



TECHNISCHE UNIVERSITÄT MÜNCHEN
Fakultät für Chemie
Fachgebiet Industrielle Biokatalyse

**Genetic Strategies to Improve Folding and Secretion of
Recombinant Proteins in *Escherichia coli***

Marc Alexander Büttner

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität
München zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. rer. nat. Tom Nilges
Prüfer der Dissertation: 1. Univ.-Prof. Dr. rer. nat. Thomas Brück
2. Univ.-Prof. Dr.-Ing. Kai-Olaf Hinrichsen

Die Dissertation wurde am 01.06.2016 bei der Technischen Universität München
eingereicht und durch die Fakultät für Chemie am 17.10.2016 angenommen.

Note of restriction of access

This thesis contains internal confidential information of Boehringer Ingelheim RCV GmbH & Co KG. Disclosure of any content of this thesis and attached drawings and data, or parts thereof, is strictly prohibited for two years after submission. It is also not allowed to make copies or transcriptions - including in digital form. Exceptions require prior written approval by Boehringer Ingelheim RCV GmbH & Co KG. The thesis may only be made accessible to university examiners, to the extent that they are bound by respective confidentiality obligations.

Sperrvermerk

Die vorliegende Dissertation beinhaltet interne vertrauliche Informationen der Firma Boehringer Ingelheim RCV GmbH & Co KG. Die Weitergabe des Inhaltes der Arbeit und eventuell beiliegender Zeichnungen und Daten im Gesamten oder in Teilen ist für zwei Jahre nach Einreichung strikt untersagt. Es dürfen keinerlei Kopien oder Abschriften - auch nicht in digitaler Form - gefertigt werden. Ausnahmen bedürfen der vorherigen, schriftlichen Zustimmung (Einwilligung) der Firma Boehringer Ingelheim RCV GmbH & Co KG. Die Dissertation darf nur den jeweiligen Korrektoren - soweit diese durch eine entsprechende Verschwiegenheitsverpflichtung gebunden sind - zugänglich gemacht werden.

Danksagung

Ich bedanke mich herzlich bei ...

... meinem Doktorvater, Thomas Brück, für Feedback, Diskussionen und die Ermöglichung einer professionellen Industrie-Promotion. Hoffentlich werden wir auch in Zukunft weiter gut zusammenarbeiten.

... meinem Betreuer bei BI, Simon, für ehrliches Feedback zu Versuchsplanung, Berichten, Präsentationen und dieser Arbeit. Du hast die Struktur und Qualität meiner Arbeiten stark positiv beeinflusst. Die Zusammenarbeit war, auch persönlich, bereichernd. Danke an Caterina für einen erfolgreichen Start und an Daniela für die Unterstützung beim strukturierten und fokussierten Arbeiten.

... allen weiteren Kollegen bei Boehringer Ingelheim in Wien. Die angenehme Arbeitssphäre bei „den Fermentierern“ war ein Grund, warum ich immer gerne zur Arbeit gegangen bin. Ein besonderer Dank gilt Julia, Mark, Artur und Dominik - ein Großteil dieser Arbeit beruht auf eurem Engagement. Danke an die BI-Molekularbiologen (Carin, Ruth, Bernhard, Johanna) für Feedback, angenehme Zusammenarbeit und Labororganisation. Vielen Dank an Alma, Matthias, Andreas und die Techniker des Upstream-Teams für die professionelle Durchführung von Fermentationen. Ohne euch hätte ich die entwickelten Stämme nicht testen können und die Arbeit wäre nur halb so viel wert. Vielen Dank an Michaela für zahlreiche SDS PAGE-Analysen.

... meinen Freunden. Helen und Sebastian, danke, dass ihr darauf geachtet habt, dass mein Privatleben nicht zu kurz kommt. Stefan und Christoph (und die anderen BI-Fußballer), unser montaglicher Jour fixe auf dem Fußballplatz war eines meiner Wochen-Highlights.

... meinen Eltern für ihr Interesse an einem ihnen fremden Thema. Danke, dass ihr mich dabei unterstützt, meine Ausbildung und meinen Lebensweg selbst zu bestimmen.

... meiner Freundin Dani für ihre Unterstützung, ihr Verständnis für unmögliche Arbeitszeiten und ihr Durchhaltevermögen in mehr als drei Jahren Wochenendbeziehung. Dass du mir nach anstrengenden Tagen im Labor und Büro zugehört hast, hat mich immer wieder aufgebaut.

„Glaube denen, die die Wahrheit suchen,
und zweifle an denen, die sie gefunden haben.“

~ André Gide

Contents

| | |
|---|------------|
| Abstract | I |
| Zusammenfassung | III |
| List of Abbreviations | V |
| Introduction | 1 |
| 1.1 Antibody Fragment Biopharmaceuticals | 1 |
| 1.2 Production of Biopharmaceuticals in <i>Escherichia coli</i> | 4 |
| 1.3 Soluble Production of Biopharmaceuticals in the <i>Escherichia coli</i> Periplasm | 6 |
| 1.3.1 Issues Associated with Inclusion Body Processes Suggest Soluble Protein Production | 6 |
| 1.3.2 The <i>Escherichia coli</i> Periplasm: a Valuable Compartment for Pro- duction of Recombinant Proteins | 8 |
| 1.3.3 Potential Issues in Soluble Recombinant Protein Production | 19 |
| 1.3.4 Host Engineering Approaches | 22 |
| 1.3.5 Process Development Approaches | 28 |
| 1.3.6 Current Benchmarks and Conclusion | 31 |
| 1.4 Aim and Objectives | 33 |
| Materials and Methods | 35 |
| 2.1 Model Proteins | 35 |
| 2.2 Generation of Examined Strains | 37 |
| 2.2.1 Creation of Plasmid Constructs | 37 |
| 2.2.2 Genomic Integration Employing Homologous Recombination | 45 |
| 2.2.3 Creation of Plasmid-Bearing Cells | 51 |
| 2.2.4 Preparation of Cryo Cultures | 53 |
| 2.3 Promoter Strength Analysis | 54 |
| 2.3.1 Shake Flask Experiments | 54 |
| 2.3.2 Fluorescence Measurement | 55 |
| 2.3.3 Data Analysis | 57 |
| 2.4 Production Experiments | 58 |
| 2.4.1 Shake Flask Experiments | 58 |
| 2.4.2 Fed-Batch Fermentations for FabZ and scFv-Z-1 Production | 60 |
| 2.4.3 Fed-Batch Fermentations for Production of FabX | 65 |
| 2.5 Examination of Gene Expression Cassette Stability | 66 |
| 2.6 Protein Analytics | 67 |
| 2.6.1 FabZ Standard Preparation | 69 |
| 2.6.2 Sample Preparation | 71 |
| 2.6.3 Enzyme-Linked Immunosorbent Assay | 72 |
| 2.6.4 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis | 75 |
| 2.6.5 Western Blot | 76 |
| 2.6.6 N-terminal sequencing | 79 |

| | |
|---|------------|
| Results and Discussion | 81 |
| 3.1 Genomic Integration of Gene Expression Cassettes Enables Significantly Improved Product Yields | 81 |
| 3.1.1 Green Fluorescent Protein | 83 |
| 3.1.2 Fragment Antigen Binding Z | 87 |
| 3.1.3 Fragment Antigen Binding X | 101 |
| 3.1.4 Single-Chain Fragment Variable Z 1 | 103 |
| 3.1.5 Conclusion | 106 |
| 3.2 Genomic Integration Addresses Important Bottlenecks in Soluble Periplasmic Protein Production | 109 |
| 3.2.1 Expression Cassette (In)Stability | 110 |
| 3.2.2 Influence of Multi-Copy Plasmid Presence | 115 |
| 3.2.3 Prevention of Target Gene Basal Expression | 119 |
| 3.2.4 Harmonization of Expression and Secretion Rate | 121 |
| 3.2.5 Conclusions | 128 |
| 3.3 Beyond Single-Copy Genomic Integration; Rationale Behind and Preparation for Further Strain Engineering | 133 |
| 3.3.1 Process Development to Improve Periplasmic Production of Recombinant Proteins | 133 |
| 3.3.2 Establishing a Small Scale Format for Increased Throughput During Strain Evaluation | 141 |
| 3.4 Establishment of Different Systems for Co-Synthesis of Periplasmic Folding Modulators | 146 |
| 3.4.1 Rationale | 146 |
| 3.4.2 Co-Synthesis of Folding Modulators Under Control of the T7 RNA Polymerase | 155 |
| 3.4.3 Constitutive Promoter Systems for Co-Synthesis of Folding Modulators | 172 |
| 3.4.4 Conclusions | 190 |
| Conclusion and Outlook | 197 |
| Confidential Supplement | 203 |
| 5.1 Information on Model Proteins | 205 |
| 5.1.1 FabX | 205 |
| 5.1.2 FabZ | 205 |
| 5.1.3 scFv-Z-1/2 | 205 |
| 5.2 Modular Vector System | 206 |
| 5.3 Oligonucleotides | 207 |
| 5.4 DNA Sequences of Modules | 210 |
| 5.4.1 Origins of Replication | 210 |
| 5.4.2 LacI Module and Spacer | 211 |
| 5.4.3 Spacer Upstream of Promoter Modules | 211 |
| 5.4.4 Promoters | 211 |
| 5.4.5 Multiple Cloning Site | 215 |
| 5.4.6 Model Protein Genes | 215 |

| | | |
|--------|-----------------------------------|------------|
| 5.4.7 | Folding Modulator Genes | 217 |
| 5.4.8 | Terminator Modules | 219 |
| 5.4.9 | <i>cer</i> module | 219 |
| 5.4.10 | Resistance Modules | 219 |
| 5.5 | Media | 221 |
| 5.5.1 | Media Components | 221 |
| 5.5.2 | T7 Pre-Culture Medium | 221 |
| 5.5.3 | Shake Flask Media | 221 |
| 5.5.4 | Fermentation Media | 222 |
| | References | 223 |

Abstract

Antibody fragments are important protein tools for research, therapeutic, diagnostic and purification applications. Due to their reduced size and lack of post-translational modifications, they often can be produced in non-mammalian expression systems, *e.g.* *Escherichia coli*. Production as protein aggregates (inclusion bodies) and subsequent refolding is often cumbersome. Thus, numerous techniques for production of soluble and correctly folded recombinant antibody fragments have been reported. However, approaches which address several bottlenecks simultaneously are still rare.

Several techniques were tested and combined within this study. The use of plasmid-free strains with a genome-integrated expression cassette utilizing the T7 RNA polymerase system successfully addressed issues of plasmid-based approaches (plasmid loss, negative influence of the basic plasmid, high level basal expression, growth cessation and secretion machinery overload). In fed-batch fermentations, soluble titers of three antibody fragment model proteins produced in the periplasm of *E. coli* BL21(DE3) were substantially increased (FabX: 3.3-fold, $53 \rightarrow 174 \text{ mg l}^{-1}$; FabZ: 4.4-fold, $83 \rightarrow 363 \text{ mg l}^{-1}$; scFv-Z-1: 8.6-fold, no absolute quantification). An overload of the periplasmic folding machinery was the major remaining issue. Unexpectedly, plasmid-based co-expression of folding modulator genes under control of condition-inducible or T7 RNA polymerase systems significantly interfered with overall and soluble expression of the target gene. The use of constitutive promoters with different activity levels allowed screening of periplasmic folding modulators and variation of their cellular levels. Soluble product titers were improved substantially in shake flasks (FabZ: $> 300 \%$, FabX: $> 600 \%$) and moderately in fed-batch fermentations (FabZ: $+ 50 \%$, FabX: $+ 24 \%$). FkpA, Skp and SurA appeared relevant, while DsbA, DsbC and PpiD exerted no or a negative effect. Eventually, soluble product titers of 450 and 576 mg l^{-1} were obtained for FabX and FabZ, respectively.

In conclusion, the established toolbox comprises three basic production strains (gene of interest on a multi-copy plasmid or forward/reverse genome-integrated). PpiD and the strongest constitutive promoter seemed expendable in co-expression approaches. Thus, each of the three basic strains can, if necessary, be enhanced with one of five folding modulators on two different cellular levels, yielding up to 33 different production strains. This flexible set of tools will be of value for production of biopharmaceuticals by enabling addressing frequently encountered issues of secretion and folding of recombinant proteins in *E. coli*.

Zusammenfassung

Antikörper-Fragmente werden in den Bereichen Forschung, Diagnostik und Aufreinigung und als Therapeutika eingesetzt. Sie können aufgrund ihrer geringen Größe und nicht vorhandenen post-translationalen Modifikationen auch in nicht-Säuger-Systemen, z.B. *Escherichia coli*, produziert werden. Die Produktion als aggregierte Proteine (Einschlusskörper) gestaltet sich vor allem aufgrund der anschließenden Rückfaltung oft als schwierig. Um Zielproteine direkt korrekt gefaltet herzustellen, wurden bereits diverse Techniken entwickelt. Dennoch gibt es wenige Ansätze, die gleichzeitig mehrere Probleme bei der löslichen Produktion adressieren.

In dieser Arbeit wurden verschiedene Techniken untersucht und kombiniert. Typische Probleme Plasmid-basierter Expression (Plasmidverlust, negativer Einfluss des Basisplasmids, Basalexpression, Wachstumsprobleme und Überlastung der Sekretionsmaschinerie) wurden durch Genomintegration der Zielgen-Expressionkassette mit starkem T7-RNA-Polymerase-System adressiert. Dies steigerte die löslichen Ausbeuten von ins Periplasma von *E. coli* BL21(DE3) sekretierten Antikörper-Fragmenten in Fed-Batch-Fermentationen erheblich (FabX: 3.3-fach, $53 \rightarrow 174 \text{ mg l}^{-1}$; FabZ: 4.4-fach, $83 \rightarrow 363 \text{ mg l}^{-1}$; scFv-Z-1: 8.6-fach, keine Absolutwerte). Jedoch wurde eine Überlastung der periplasmatischen Faltungsmaschinerie festgestellt. Die Plasmid-basierte Co-Synthese periplasmatischer Faltungshelfern mit induzierbaren Promotoren beeinflusste die Expression des Zielgens. Konstitutive Promotoren verschiedener Stärke ermöglichten schließlich die Untersuchung verschiedener Helfer mit verschiedenen zellulären Konzentrationen. Lösliche Produkttiter zweier Fab-Fragmente wurden in Schüttelkolben-Experimenten deutlich (FabZ: $> 300 \%$, FabX: $> 600 \%$) und in Fed-Batch-Fermentationen moderat gesteigert (FabZ: $+ 50 \%$, FabX: $+ 24 \%$). FkpA, Skp and SurA erschienen förderlich, wobei DsbA, DsbC und PpiD keine oder negative Effekte zeigten. Die besten Expressionssystemen resultierten in löslichen Produkttiter von 450 (FabX) und $576 \text{ mg l}^{-1} \text{ (FabZ)}$ in Fed-Batch-Fermentationen.

Das System besteht aus drei Basisstämmen mit Plasmid-basierter oder „forward“-/„reverse“-Genom-integrierter Expressionskassette. Da PpiD und der stärkste konstitutive Promotor verzichtbar erschienen, können diese jeweils mit fünf Faltungshelfern mit je zwei zellulären Konzentrationen kombiniert werden. Diese 33 Stämmen werden die lösliche Produktion von therapeutischen Proteinen erleichtern, indem bekannte Probleme bei der Sekretion und Faltung von rekombinanten Proteinen in *E. coli* gezielt adressiert werden können.

List of Abbreviations

A_{wavelength}, absorption at given wavelength; **AA**, amino acid; **BI (RCV)**, Boehringer Ingelheim Regional Center Vienna GmbH & Co. KG; **CBB**, coomassie brilliant blue; **cml(R)**, chloramphenicol (resistance module); **DCW**, dry cell weight; **DO**, dissolved oxygen; **ds**, double stranded; **FDA**, U.S. Food and Drug Administration; **FRT**, flippase recognition target; **FS**, feed start; **GFP**, green fluorescent protein; **GoI**, gene of interest; **HC**, C_H1V_H (Fab) or V_H (scFv); **HPLC**, high pressure liquid chromatography; **HRP**, horseradish peroxidase; **IB**, inclusion body; **IEX**, ion exchange chromatography; **IM**, inner membrane; **kan(R)**, kanamycin (resistance module); **LC**, C_LV_L (Fab) or V_L (scFv); **mAb**, monoclonal antibody; **MW**, molecular weight; **o/n**, over night; **OD**_{wavelength}, optical density at given wavelength; **OM**, outer membrane; **ori**, origin of replication; **PFCF**, plasmid-free cell fraction; **PCN**, plasmid copy number; **PCR**, polymerase chain reaction; **PoI**, protein of interest; **PPIases**, peptidyl prolyl *cis-trans* isomerases; **qPCR**, quantitative PCR; **REN**, restriction endonuclease; **RFU**, relative fluorescence unit; **RO-H₂O**, water purified by reverse osmosis; **RT**, room temperature; **S-LB**, soy-lysogeny broth; **SEC**, size exclusion chromatography; **SOC**, super optimal broth with catabolite repression; **SOD**, superoxide dismutase; **sol**, soluble; **SOP**, standard operating procedure; **spc(R)**, spectinomycin (resistance module); **stm**, streptomycin; **sup**, supernatant; **tet**, tetracycline; **tot**, total; **VBNC**, viable but non-culturable cells; **WC**, whole cell.

Introduction

1.1 Antibody Fragment Biopharmaceuticals

The term “biopharmaceuticals” was coined in the 1980s and describes biotechnologically produced pharmaceuticals, which were not extracted from biological sources. Rather, biopharmaceuticals are produced utilizing modern biological methods including recombinant DNA techniques (Walsh, 2014). Recombinant human insulin (Humulin, Eli Lilly) was the first biopharmaceutical to receive FDA approval (1982). During the 1980s, eight further biopharmaceutical products received approval. Approval rates increased afterwards and have constantly ranged above 50 approvals per five year time period, starting with 1995. In July 2014, 212 biopharmaceutical drugs were approved and marketed in the United States and the EU. Their indications are, for instance, related to treatment of cancer, inflammation-related diseases, hemophilia, metabolic disorders and diabetes. In 2014, the biopharmaceutical market accounted for approximately 40 % of the total sales of the global pharmaceutical market (Spadiut et al., 2014). 37 biopharmaceuticals held “blockbuster” status, *i.e.* reached sales volumes of above US \$1 billion, in 2013. In the same year, the total sales volume achieved by biopharmaceutical drugs was approximately US \$140 billion. Monoclonal antibodies (mAbs) and related products represent the most lucrative biopharmaceuticals product class. The first monoclonal antibody drug for clinical use in humans (Orthoclone OKT-3, Janssen-Cilag) was approved by the FDA in 1986 (Smith, 1996). Nowadays, several antibody products (*e.g.* Avastin, Herceptin, Remicade, Rituxan, Humira, and Erbitux) have reached blockbuster status (Spadiut et al., 2014). Major indications are inflammatory and autoimmune diseases (US \$41 billion sales in 2013) and cancer (US \$26 billion sales in 2013). Unless otherwise stated, this paragraph is based on a 2014 article (Walsh, 2014).

The structure of immunoglobulin G (IgG), which is the most abundant antibody isomer type in humans, was published in 1971 (see Figure 1.1) (Edelman, 1971). The tetrameric full-length mAb molecules consist of two identical heavy and light chains and have an approximate molecular weight of 150 kDa. Two heterodimeric Fab (fragment antigen binding, C_H1-V_H and C_L-V_L) and the homodimeric Fc (fragment crystallizable, C_H2-C_H3) regions form the Y-shaped structure. Heavy and light chain are connected *via* disulfide bonds in the Fab region. In each Fab fragment, variable regions of both chains (V_H and V_L) form two variable fragments (Fv), which enable antigen binding with high

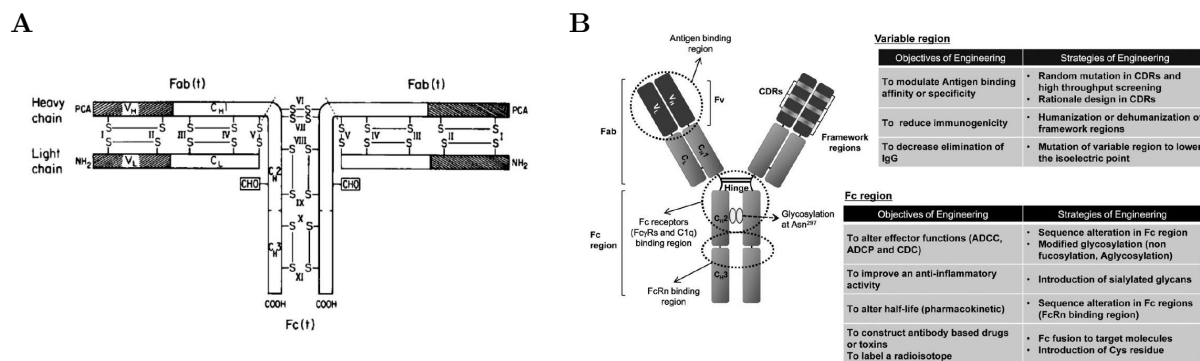


Figure 1.1: Overall arrangement of chains and disulfide bonds of γ G1 immunoglobulin Eu (**A**) and antibody (IgG) structure and engineering (**B**). **A**: Half-cystinyl residues are numbered I to XI beginning from the N-terminus. Numbers I to V designate corresponding residues in light and heavy chains. PCA: pyrrolidone carboxylic acid. CHO: carbohydrate. Fab(t) and Fc(t) refer to fragments produced by trypsin, which cleaves the heavy chain as indicated by dashed lines above half-cystinyl residues VI. V_H , V_L : homologous variable regions of heavy and light chains. C_{H1} , C_{H2} , C_{H3} : homology regions comprising C_H , the constant regions of the heavy chain. C_L : constant region of the light chains which is homologous to C_{H1} , C_{H2} and C_{H3} . The figure and description were copied from the original publication (Edelman, 1971). **B**: Antibody (IgG) structure and engineering. Immunoglobulin G consists of three domains (Fab, Fc and Hinge). In general, the Fab region is responsible for antigen-binding specificity and affinity and the Fc part is responsible for antibody stability and effector functions through Fc receptor bindings. In antibody engineering, various strategies can be employed for each domain. The figure and description were copied from the original publication (Jeong et al., 2011).

specificity and affinity. Fc domains, which bear the Hinge region and carbohydrate side chains (“glycosylation”), stabilize the antibody and are important for antibody-dependent cell-mediated cytotoxicity (ADCC). Native antibodies are part of the immune system. Due to their antigen-binding characteristics, they can bind to foreign substances (*e.g.* viruses or bacteria). By antibody binding, these foreign substances are neutralized or marked for actions of alternative defence systems. To utilize this high specificity and affinity, antibody-based compounds have been successfully developed and marketed for therapeutic, research and diagnostic applications. Unless otherwise indicated, this paragraph is based on text book information or review articles (Janeway et al., 2001; Jeong et al., 2011; Nelson and Reichert, 2009).

Derived from full-length antibodies, different parts or compounds, *e.g.* Fabs (Better et al., 1988) or scFvs (Bird et al., 1988), have been developed and used (see Figure 1.2). While binding specificity of these “antibody fragments” is similar to that of full IgGs, removal of antibody parts has important consequences. Partly due to their decreased size, antibody fragments are often characterized by rapid and deep tissue penetration, a lower retention time in non-target tissues and a rapid blood clearance, which results in shorter circulating half lives (Jain, 1990; Larson et al., 1983; Ward et al., 1989; Yokota et al., 1992). However, “half life extension”, *e.g.* by means of PEGylation (Jain and Jain, 2008) or albumin con-

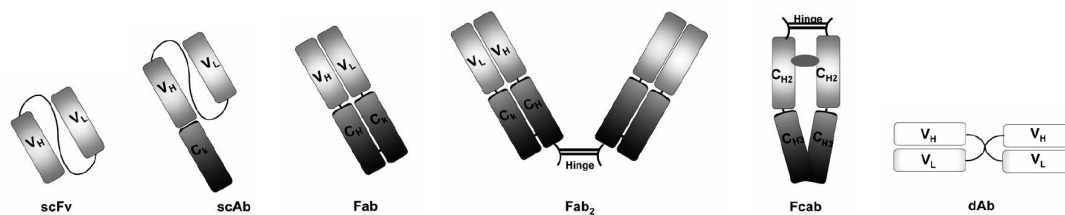


Figure 1.2: Schematic structure of various antibody fragments. In scFv, scAb and dAb, heavy and light chains are connected by flexible Gly-Ser linkers. In Fab, Fab₂ or Fcab molecules, both chains are covalently linked *via* disulfide bonds. Fcab requires glycosylation at Asn²⁹⁷ position. The figure and description were copied from a review article (Jeong et al., 2011).

jugation (Holt et al., 2008), is possible. Absence of Fc region and glycosylation in many Ab fragments prevents complement-dependent (CDC) or antibody-dependent cellular cytotoxicity (ADCC). Thus, certain fragments are characterized by reduced immunogenic potential (Raghavan and Bjorkman, 1996). Furthermore, the lack of the Fc region can lead to reduced stability (Bird et al., 1988). Unless specified otherwise, this paragraph is based on a review article (Nelson and Reichert, 2009) and the introductory section of a research article (Ahmad et al., 2012).

Considering their antigen-binding properties and the described characteristics, antibody fragment can constitute relevant alternatives for full-length antibodies. In accordance, many antibody fragments have become important tools in research, diagnostic, therapeutic and purification applications (de Marco, 2011; Li and Zhu, 2010; Reichert, 2012). For example, bivalent Fab₂ and monovalent Fab fragments have shown good therapeutic results in both animal and clinical studies (Goldenberg, 2002). Abciximab (ReoPro, Centocor/Eli Lilly), Ranibizumab (Lucentis, Genentech) (Ferrara et al., 2006) and Certolizumab pegol (Cimzia, UCB) are FDA-approved Fab fragments (Nelson and Reichert, 2009). Furthermore, since production of small, glycosylation-free antibody fragments requires less complex synthesis machineries, their production in non-mammalian expression systems is often feasible (Jeong et al., 2011).

1.2 Production of Biopharmaceuticals in *Escherichia coli*

Chemical synthesis of complex biological products, especially proteins, is hardly possible (Lee and Bang, 2010). They are usually produced in genetically modified host organisms as “recombinant proteins”. A plethora of production systems is available for production of biopharmaceuticals. Depending on target product characteristics, these are characterized by advantages and drawbacks (Brondyk, 2009). This variety is also represented in the expression systems that are used for manufacturing of newly approved products (see Figure 1.3). As many biopharmaceutical products require post-translational modifications to fulfill demands of activity and/or stability, mammalian expression systems are becoming more important. Currently, chinese hamster ovary (CHO) cells are most commonly used. Developments concerning this expression system were comprehensively reviewed in 2012 (Kim et al., 2012). In 2010, 32 % of all manufactured active pharmaceutical ingredient (API) (a total of approximately 26400 kg) were derived from mammalian systems. The remaining amount (approximately 18000 kg) was produced in microbial systems. Eukaryotic microbial systems (*Saccharomyces cerevisia* and *Pichia pastoris*) also play major roles in production of biopharmaceuticals. On the other hand, non-*E. coli* procaryotic microbial systems (e.g. *Vibrio cholera* and *Bordetella pertussis*) are of less importance. Thus, the gram-negative bacterium *Escherichia coli* remains the most common non-mammalian production system. Very few approved products are produced in transgenic animals, insect or plant cells. In conclusion, CHO cells and the gram-negative bacterium *E. coli* currently represent the most important expression systems. Unless stated otherwise, this paragraph is based on a 2014 review article (Walsh, 2014).

E. coli in particular has a long history of successful application as a production organism in the biopharmaceutical industry. In 1977, the first recombinant human therapeutic protein (Somatostatin) was successfully synthesized in *E. coli* (Itakura et al., 1977). Five years later, in 1982, recombinant human insulin for treatment of diabetes (Humulin), produced in *E. coli*, received FDA approval. Since then, various recombinant hormones, thrombolytics, growth factors, interferons and antibody fragments, which are produced in *E. coli*, have received approval (EMEA, 2007; Walsh, 2010, 2014). Up to 2014, 19 % of all approved biopharmaceuticals were produced in *E. coli* (see Figure 1.3) (Walsh, 2014).

Application and characteristics of the production organism *E. coli* have been reviewed comprehensively elsewhere (Frenzel et al., 2013; Green and Sambrook, 2012; Huang et al., 2012; Jeong et al., 2011; Kamionka, 2011; Meyer and Schmidhalter, 2012; Nelson and

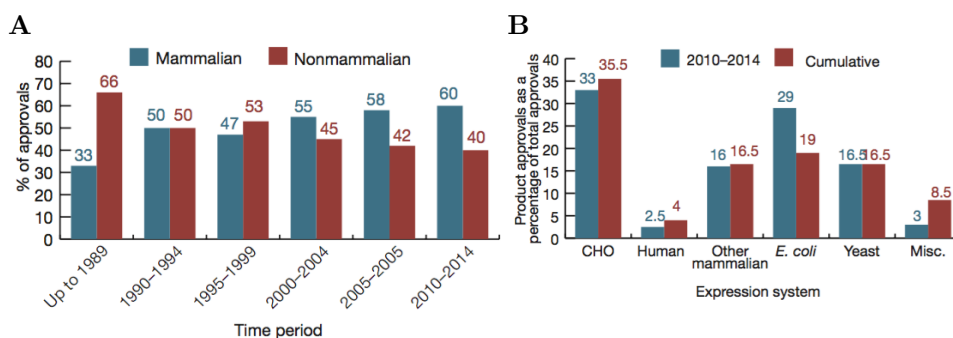


Figure 1.3: Expression systems used to manufacture biopharmaceutical products. **A**: Relative application of mammalian versus non-mammalian expression systems in the production of biopharmaceuticals approved over the indicated time periods. Each data set is expressed as a percentage of total biopharmaceutical product approvals for the period in question. **B**: Product approvals, cumulative (1982 - 2014) and during the considered study (2010 - July 2014). The relative distribution among the employed expression system is shown. Figures and descriptions were copied from a review article (Walsh, 2014).

Reichert, 2009; Overton, 2014; Spadiut et al., 2014; Walsh, 2014) and will be covered briefly. Due to decades of research and use in laboratories, the genetics of this, in biochemical and physiological terms, relatively simple organism are characterized well. For example, *E. coli* was one of the first organisms to have its entire genome sequenced (Blattner et al., 1997). A plethora of tools for genetic manipulation and examination has been developed. In terms of biotechnological applications, *E. coli* cells can be grown rapidly to high cell densities on inexpensive media. Furthermore, they allow more robust and shorter processes than, for instance, CHO cells. Growing numbers of expression vectors, different production strains, protein refolding and fermentation technologies for industrial application are available. Combinations of several of these technologies allowed for production of numerous approved biopharmaceuticals. This, in itself, constitutes an advantages over alternative production systems in the drug approval process. However, while a broad variety of biologics has been successfully produced in *E. coli*, certain product characteristics may prevent its use. For instance, certain posttranslational modifications (such as glycosylation), proteolytic protein maturation and/or disulfide bond formation are mandatory for full function of some biopharmaceuticals (*e.g.* mAbs with blockbuster status). While engineering approaches currently address these topics, most commonly used *E. coli* strains are not or very insufficiently able to perform such modifications. Furthermore, the presence of endotoxins that need to be removed from therapeutic products, has to be considered prior to choosing the expression system.

In conclusion, several relevant production hosts for recombinant biopharmaceuticals exist. Under circumstances that allow exploiting its many advantages, *E. coli* is an attractive production organism.

1.3 Soluble Production of Biopharmaceuticals in the *Escherichia coli* Periplasm

1.3.1 Issues Associated with Inclusion Body Processes Suggest Soluble Protein Production

In most cases, biopharmaceuticals require a defined conformation for achieving full biological activity. Thus, they are administered as dispersed soluble entities (Ferrer-Miralles et al., 2009). Obtaining target proteins in the native and active form in *E. coli*-based bioprocesses is possible in two different ways (Sahdev et al., 2008). Formation of aggregated proteins, described as “inclusion bodies” (IBs) when located in the cytoplasm, often occurs as an effect of a high level of overproduction. It can be induced deliberately and aggregated protein can, subsequently, be isolated, solubilized, refolded and purified by chromatographic techniques (Burgess, 2009; Lee et al., 2006; Singh et al., 2015; Vallejo and Rinas, 2004). On the other hand, production of soluble and active target protein in the upstream part of the bioprocess, *i.e.* directly in the expression host, is another feasible synthesis route. For instance, the Fab fragments Ranibizumab (Lucentis) and Certolizumab (Cimzia) are produced as soluble entities in the periplasm of *E. coli* (EMEA, 2007; Jalalirad, 2013). The “tissue-type plasminogen activator” (tPA, 527 amino acids, 17 disulfide bonds) is a relevant biopharmaceutical product, which has been produced by both IB-based and soluble protein production processes (Vallejo and Rinas, 2004).

Advantages and disadvantages associated with IB processes were covered in several review articles (Arakawa and Ejima, 2014; Baneyx and Mujacic, 2004; Frenzel et al., 2013; Garcia-Fruitos et al., 2012; Huang et al., 2012; Lee et al., 2006; Sahdev et al., 2008; Singh and Panda, 2005; Sorensen and Mortensen, 2005a; Sugiki et al., 2014; Vallejo and Rinas, 2004) and introductory sections of a research article (Wyre and Overton, 2014). Unless otherwise specified, this paragraph is based on these articles. High-yield production and feasibility to also produce toxic proteins are advantages associated with IB-based production. Furthermore, IBs are stable protein aggregates and are resistant to protease activities *in vivo*. In addition, proteomic analysis showed that IBs are relatively homogeneous in composition. The recombinant protein can account for more than 90 % of all embedded polypeptides (Valax and Georgiou, 1993). During downstream processing, IBs can be easily isolated after cell disruption, yielding relatively pure recombinant protein material. The resultant IB paste can be frozen and stored for several months, providing manufacturing flexibility (Garcia-Fruitos et al., 2009). In conclusion, IBs can be produced

at high yield and isolated and purified with minimal effort. However, IBs require full denaturation and subsequent refolding to regain biological functionality. Identification of conditions for efficient recovery of soluble and functional biopharmaceuticals from IBs produced in the bacterial cytoplasm is often cumbersome and time-consuming. Furthermore, the success of these approaches is highly target protein-dependent. The conventionally used “one-step-denaturing and refolding” method works for approximately 40 % of inclusion body proteins (Yang et al., 2011). It was reported that *in vitro* refolding efficiency is often far too low for economically feasible IB-based production processes (Lee et al., 2006). In conclusion, severe drawbacks affiliated with IB-based *E. coli* bioprocesses make soluble biopharmaceutical protein production a valuable alternative.

1.3.2 The *Escherichia coli* Periplasm: a Valuable Compartment for Production of Recombinant Proteins

The *Escherichia coli* periplasm

The gram-negative bacterium *E. coli* consists of two separate compartments, the cytoplasmic and the periplasmic space (Choi and Lee, 2004; Duong et al., 1997; Goemans et al., 2014; Madigan and Martinko, 2009; Merdanovic et al., 2011; Schlegel et al., 2013). The cytoplasm is highly crowded and contains an aqueous solution with a large number of macromolecules (proteins, genomic and plasmid DNA and RNA), small organic molecules and anorganic ions (Cayley et al., 1991; Zimmerman and Trach, 1991). Transcription, translation and folding of non-secreted protein occur in this compartment. The thioredoxin system (composed of the thioredoxin reductase TR and two thioredoxins, Trx1 and -2) and the glutaredoxin system (including the glutathione reductase Gor, the tripeptide glutathione and three glutaredoxins, Grx1, -2 and -3) maintain the cytoplasm in a reducing state (Prinz et al., 1997; Ritz and Beckwith, 2001; Stewart et al., 1998). The cytoplasm is limited by the inner membrane (IM), a symmetric bilayer consisting of phospholipids and integral membrane proteins. The outer membrane (OM) is an asymmetric lipid bilayer with phospholipids in the inner leaflet and lipopolysaccharides (LPS) in the outer leaflet (Ruiz et al., 2006). It constitutes the border to the extracellular space. The aqueous and viscous oxidizing periplasmic space is located in between the two membranes. This compartment represents approximately 10 to 20 % of the total cellular volume (Van Wielink and Duine, 1990). More than 300 proteins reside in the periplasm (Weiner and Li, 2008). These perform a variety of physiological functions, such as protein folding (*e.g.* chaperones), uptake and transport of nutrients (*e.g.* hydrolases and binding proteins) and detoxification of harmful substances. Furthermore, a peptidoglycan meshwork provides cell shape and anchors the outer membrane lipoprotein.

Overview of the periplasmic folding machinery

The periplasmic folding machinery promotes folding of secreted polypeptides. Comprehensive review articles on the topic are available (Baneyx and Mujacic, 2004; Denoncin et al., 2012; Herrmann and Riemer, 2014; Kolaj et al., 2009). Furthermore, the periplasmic chaperone network has been examined using differential proteomics techniques (Denoncin et al., 2012). Functions of the folding machinery are performed by i) folding catalysts, *i.e.* peptidyl-prolyl *cis-trans* isomerases (PPIases) (Pliyev and Gurvits, 1999)

and thiol-disulfide oxidoreductases (the Dsb family), and ii) general chaperones. PPIases are subdivided into the cyclophilin group (*e.g.* PpiA), the FKBP family (*FK506 binding protein, e.g.* FkpA) and the parvulin group (*e.g.* PpiD and SurA). They catalyze the isomerization of peptidyl-prolyl bonds. Oxidoreductases catalyze the formation and isomerization of disulfide bonds. While the folding catalyst group is actively involved in the folding process, general chaperones prevent aggregation and keep unfolded proteins soluble and folding-competent. Also, these chaperones keep misfolded protein soluble and degradable. This is achieved by binding of exposed hydrophobic moieties in unfolded, partially folded or misfolded polypeptides. For instance, chaperones interact with unfolded β -barrels of outer membrane proteins (OMPs) which have to cross the hydrophilic periplasmic space before insertion into the outer membrane (OM). Note that overlapping functions of components of the periplasmic folding machinery have previously been reported (see Figure 1.7 (B), p. 14) (Duguay and Silhavy, 2004).

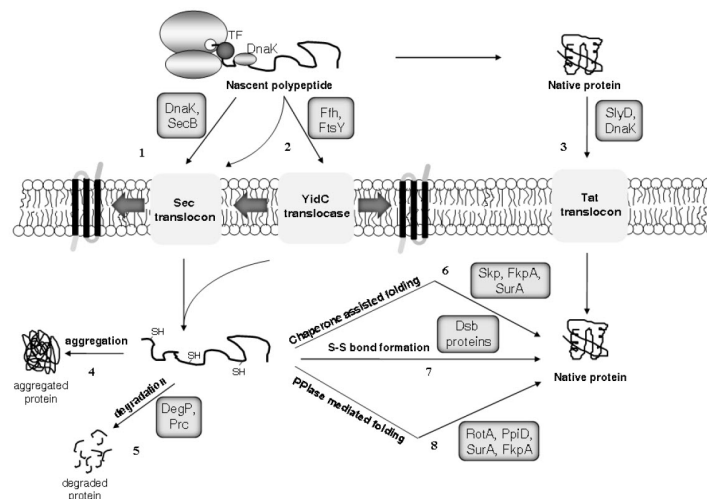


Figure 1.4: Membrane translocation and periplasmic folding in *E. coli*. Most polypeptides are delivered to SecA at the inner surface of the membrane by DnaK or SecB chaperones, prior to crossing the cytoplasmic membrane in an unfolded conformation *via* the Sec translocon (1). Polypeptides with highly hydrophobic signal sequences or transmembrane domains may, however, be recognised by Ffh, which, together with its FtsY receptor, can target the polypeptide to either the Sec machinery or to the YidC translocase (2). Alternatively, the twin-arginine translocation (Tat) machinery is responsible for the translocation of folded proteins (3). Upon crossing the membrane and after cleavage of the leader peptide, partially folded proteins may (4) aggregate, (5) be degraded by periplasmic proteases, or fold into their native state. Folding is often assisted by periplasmic chaperones (6) and/or folding catalysts such as disulfide bond-forming enzymes (7) or peptidyl-prolyl *cis-trans* isomerases (8). Figure and description were copied from a previously published review article (Kolaj et al., 2009).

Oxidoreductases

Structure and function of periplasmic oxidoreductases (“disulfide bond” or “Dsb” proteins) have been reviewed extensively (Arredondo and Georgiou, 2011; Cho and Collet, 2012;

de Marco, 2009; Denoncin and Collet, 2012; Depuydt et al., 2010; Goemans et al., 2014; Ito and Inaba, 2008; Kadokura and Beckwith, 2010; Messens and Collet, 2006; Nakamoto and Bardwell, 2004). One representative, the potent oxidase DsbA (Huber-Wunderlich and Glockshuber, 1998) is a soluble monomeric protein of 21 kDa (189 amino acid residues) with a C-P-H-C catalytic motif and a thioredoxin-fold domain (Baneyx and Mujacic, 2004; Denoncin and Collet, 2012; Pan and Bardwell, 2006). DsbA is secreted to the periplasm *via* the SRP pathway (Schierle et al., 2003). It catalyzes the formation of disulfide bonds in post- or co-translationally secreted proteins in a co-translocational and vectorial (*i.e.* in between consecutive cysteines) manner (Bardwell et al., 1991; Kadokura and Beckwith, 2009; Kadokura et al., 2004). 15 DsbA substrates have been identified, but more were predicted (Hiniker and Bardwell, 2004; Kadokura et al., 2004; Leichert and Jakob, 2004). After disulfide transfer to the substrate *via* a mixed-disulfide complex, the catalytic motif is re-oxidized *in vivo* by the sulfhydryl oxidase DsbB (Joly and Swartz, 1997). It resides in the inner membrane, bears four transmembrane segments and exposes two small hydrophilic loops to the periplasm. Each loop contains a pair of conserved cysteine residues (Baneyx and Mujacic, 2004), which are maintained in an oxidized state. After shuttling of disulfide bonds from DsbB to DsbA (Bardwell et al., 1993; Kadokura and Beckwith, 2002; Kadokura et al., 2003), DsbB delivers accepted electrons to ubiquinone or menaquinone under aerobic and anaerobic conditions, respectively, to maintain its own oxidized state (Bader et al., 1999). These reside in close proximity to the four-helix bundle structure (Bader et al., 1999) and transfer electrons to the electron transport chain. Under aerobic conditions, molecular oxygen constitutes the terminal electron acceptor. DsbB structure and mechanism have been comprehensively reviewed (Denoncin and Collet, 2012; Herrmann and Riemer, 2014).

If non-vectorial disulfide bonds are required in proteins containing more than two cysteines, oxidation by DsbA might promote formation of incorrect variants. Disulfide bond isomerization is catalyzed by the isomerase DsbC, a soluble V-shaped homodimer with 23.3 kDa and four cysteines per monomer (Missiakas et al., 1994). Both monomers bear an N-terminal dimerization domain, required for stability, and a C-terminal catalytic domain, which bears the thioredoxin-fold and the C-G-Y-C motif (McCarthy et al., 2000; Sun and Wang, 2000). DsbC is thought to capture folding intermediates within the 38 Å wide, uncharged and partially hydrophobic cleft formed by its dimerization interface (Baneyx and Mujacic, 2004; Chen et al., 1999; McCarthy et al., 2000). This might support interaction with misfolded proteins due to their exposed hydrophobic core. Disulfide bond isomerization relies on a process involving mixed disulfide intermediates (Baneyx and Mujacic, 2004). Comprehensive review articles for details on two feasible *in vivo* models for the

DsbC-catalyzed reaction mechanism are available (Berkmen, 2012; Denoncin and Collet, 2012; Herrmann and Riemer, 2014). *In vivo*, DsbC is maintained in a reduced state by the inner membrane-bound DsbD at the expense of NADH oxidation in the cytoplasm (Hiniker and Bardwell, 2003; Joly and Swartz, 1997; Kadokura et al., 2003). DsbD consists of the three domains α (N-terminal, exposed to the periplasm), β (membrane-bound) and γ (C-terminal, exposed to the periplasm, thioredoxin-like) (Goemans et al., 2014). Each domain contains a pair of redox-active cysteines (Kadokura et al., 2003), which participate in reduction of DsbC (Herrmann and Riemer, 2014). Electrons flow from the cytoplasmic NADPH pool to the thioredoxin system of the cytosol, and *via* β , γ and α domains of DsbD towards the target protein DsbC (Herrmann and Riemer, 2014; Rietsch et al., 1997).

Other periplasmic oxidoreductases are DsbE, which is required for cytochrome *c* biogenesis, and DsbG, which is homologous to DsbC and has a similar architecture but is characterized by a larger and less hydrophobic cleft and a different substrate specificity (Andersen et al., 1997; Berkmen, 2012; Heras et al., 2004; Kolaj et al., 2009).

Note that Figure 1, 2 B and 7 A of a previously published review article (Kadokura and Beckwith, 2010) provide a good overview of the DsbA/C system. However, due to licensing conditions, the images could not be used in this thesis.

Peptidyl-prolyl *cis-trans* isomerases

FkpA, a representative of the FKBP (*FK506 binding protein*) family (Horne and Young, 1995) of peptidyl-prolyl *cis-trans* isomerases (PPIases), is a V-shaped homodimer (see Figure 1.5 (B)). Structure and function have been studied extensively (Arie et al., 2001; Missiakas et al., 1996; Ramm and Plückthun, 2000; Saul et al., 2004). Each monomer has 245 residues and two domains. N-terminal segments with their three helices are responsible for dimerization and general chaperone activity by forming a hydrophobic cleft. C-terminal domains bear the PPIase function (Arie et al., 2001; Bothmann and Plueckthun, 2000; Missiakas et al., 1996; Ramm and Plückthun, 2000; Saul et al., 2004). FkpA was described to accommodate partially folded substrates within the hydrophobic cleft, allowing the flexible C-terminal domains to access prolyl bonds (Baneyx and Mujacic, 2004).

PpiA (RotA), a cyclophilin, is a globular protein with 164 residues and structurally similar to human cyclophilin A (Clubb et al., 1995; Liu and Walsh, 1990). Like all PPIases, it catalyzes the *cis-trans* isomerization of proline peptide bonds.

The parvulin PpiD contains an N-terminal helix anchored in the inner membrane and three soluble domains facing the periplasm. Two of these bear the enzyme's chaperone activity, while the third bears a parvulin-like fold (Weininger et al., 2009). As PpiD is localized close to the Sec translocon and interacts with exiting proteins, it was assumed to assist in folding initiation (Antonoaea et al., 2008). Contrary to a previous report (Dartigalongue and Raina, 1998), PpiD was more recently reported to function in a network of periplasmic chaperones without directly interacting with outer membrane proteins (Matern et al., 2010). Its substrate specificity overlaps with that of SurA (Stymest and Klappa, 2008).

SurA is another member of the parvulin group. It consists of a large N-terminal domain, two parvulin-like PPIase domains (Baneyx and Mujacic, 2004) and a short C-terminal helix (see Figure 1.5 (A)). N-terminal, first parvulin and C-terminal domain build a core structural module with a cleft responsible for substrate binding (Bitto and McKay, 2002) and chaperone activity (Behrens et al., 2001). The enzyme's PPIase activity is mediated by the second PPIase domain (Behrens et al., 2001). It was reported to be involved in survival during the stationary growth phase (Tormo et al., 1990). Its chaperone activity supports maturation of outer membrane proteins (Behrens et al., 2001; Behrens-Kneip, 2010; Lazar and Kolter, 1996; Missiakas et al., 1996; Rouviere and Gross, 1996; Sklar et al., 2007). Furthermore, SurA was found to improve the folding of unstable or aggregation-prone proteins in the periplasm (Missiakas et al., 1996). It shows a preference for Ar-X-Ar motifs (Ar: aromatic residue, X: any residue), which are common in OMPs but rare in other polypeptides (Bitto and McKay, 2003; Hennecke et al., 2005; Xu et al., 2007).

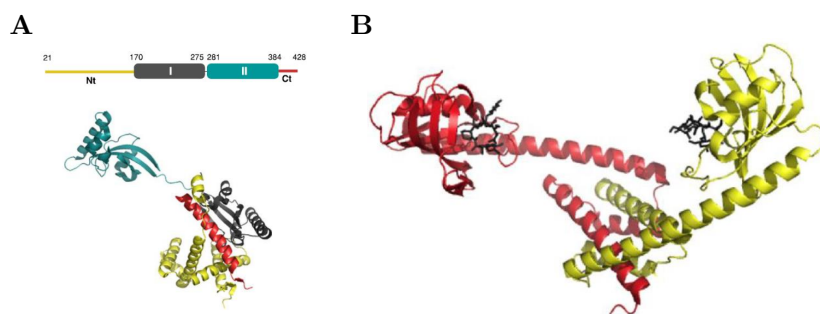


Figure 1.5: Structure of *E. coli* PPIases SurA and FkpA. **A:** Structure of *E. coli* SurA. Top: Schematic diagram of the domains of the mature SurA. The numbers refer to the amino acid positions. Bottom: Ribbon representation of SurA. **B:** The crystal structure of *E. coli* FkpA. Figures and description were copied from a previously published review article (Goemans et al., 2014).

General chaperones

General periplasmic chaperones are involved in maturation of outer membrane proteins and protection of periplasmic proteins under stress conditions (Goemans et al., 2014). They recognize surface-exposed hydrophobic areas of proteins in non-native conformations. Chaperones were proposed to function in a major (*via* SurA) and a secondary redundant (*via* Skp and DegP) folding pathway (Denoncin et al., 2012; Rizzitello et al., 2001; Sklar et al., 2007). Note that several PPIases also mediate chaperone function (see Figure 1.7).

DegP (see Figure 1.6 (B)), for instance, is a periplasmic protease with general chaperone activity (Goemans et al., 2014; Sklar et al., 2007; Strauch et al., 1989). DegP functions as a protease at high and a chaperone at low temperatures (Lipinska et al., 1990; Price and Raivio, 2008; Skorko-Glonek et al., 2008; Spiess et al., 1999).

The “seventeen kilodalton protein” Skp (PDB: 1SG2) is a soluble homotrimeric protein with an inner substrate-binding cavity defined by α -helical tentacles diverging from a β -barrel core domain (Schlupschy et al., 2005; Walton and Sousa, 2004) (see Figure 1.6 (A)). The cavity can accommodate modules of up to approximately 20 kDa in size (Walton and Sousa, 2004) and prevent parts of target proteins from aggregation, while other parts are correctly folded (Walton et al., 2009). Skp’s primary function was found to be assisting the folding and membrane insertion of outer membrane proteins (Chen and Henning, 1996; Schafer et al., 1999; Sklar et al., 2007). Furthermore, it was shown to capture unfolded proteins as they emerge from the Sec translocation apparatus (Baneyx and Mujacic, 2004). In summary, it was shown to interact with over 30 proteins of the cell envelope (Jarchow et al., 2008).

Further general periplasmic chaperones are the “spheroblast protein Y” Spy (see Figure 1.6 (C)), which prevents aggregation of a wide range of substrates (Quan et al., 2011; Schwalm et al., 2013), HdeA/B, which is active during acid survival (Gajiwala and Burley, 2000; Goemans et al., 2014; Hong et al., 2005, 2012; Kern et al., 2007) and LolA, which assists during transport of outer membrane lipoproteins across the periplasm (Tajima et al., 1998).

Proteolytic degradation in the periplasm

Proteolytic activities in the periplasm were previously reported to be lower than in the cytoplasm. However, as previously reviewed, numerous proteases are present in this com-

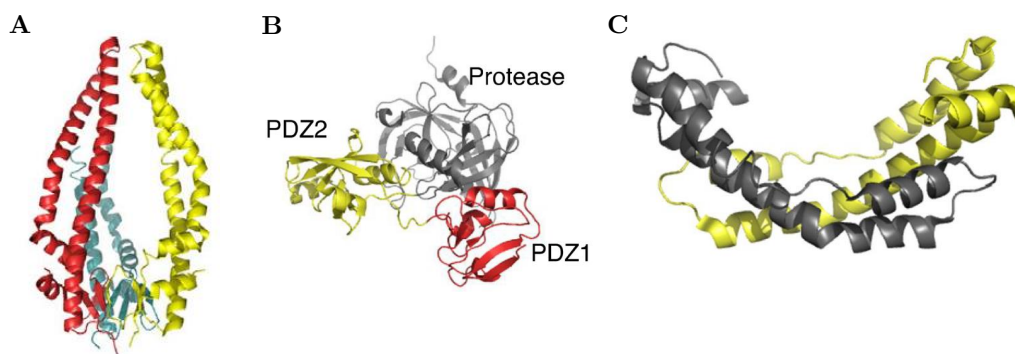


Figure 1.6: Structures of *E. coli* periplasmic general chaperones. **A**: Crystal structure of the trimeric chaperone Skp. **B**: DegP monomer with its protease domain and two PDZ domains (parts of the protein that can interact with other proteins; named after the three proteins which were first discovered to bear this domain). **C**: Crystal structure of the *E. coli* Spy chaperone. Figures were copied from a review article (Goemans et al., 2014).

| A | | | B | | | | | |
|---|-----------------------|--|----------------------------|------------|-------------------------------------|-----------------------|------------|--------------|
| Classification | Protein | Substrates | Peptidyl-prolyl isomerases | Chaperones | Disulfide bond catalysts/isomerases | | | |
| Generic chaperones | Skp (OmpH) | Outer membrane proteins and misfolded periplasmic proteins | SurA FkpA | Skp | DsbG | | | |
| | | FkpA | | | | Broad substrate range | | |
| Specialized chaperones | SurA | Outer membrane proteins | | | | PpiD PpiA | Chaperones | DsbC DsbA |
| | LolA | Outer membrane lipoproteins | | | | | | |
| | PapD (and its family) | Proteins involved in P pili biosynthesis | | | | | | |
| | FimC | Proteins involved in type 1 pili biosynthesis | | | | | | |
| PPIases | SurA | Outer membrane β -barrel proteins | PpiD PpiA | Chaperones | DsbC DsbA | | | |
| | PpiD | Outer membrane β -barrel proteins | | | | | | |
| | FkpA | Broad substrate range | | | | | | |
| | PpiA (RotA) | Unknown | | | | | | |
| Proteins involved in disulfide bond formation | DsbA | Reduced cell-envelope proteins | PpiD PpiA | Chaperones | DsbC DsbA | | | |
| | DsbB | Reduced DsbA | | | | | | |
| | DsbC | Proteins with nonnative disulfides | | | | | | |
| | DsbG | Proteins with nonnative disulfides | | | | | | |
| | DsbD | Oxidized DsbC, DsbG and CcmG | | | | | | |
| | DsbE (CcmG) | Cytochrome <i>c</i> biogenesis | | | | | | |
| | CcmH | Cytochrome <i>c</i> biogenesis | | | | | | |

Figure 1.7: Overview of periplasmic folding modulators (**A**) (Baneyx and Mujacic, 2004) and visualization of overlapping functions (**B**) (Duguay and Silhavy, 2004).

partment (Herrmann and Riemer, 2014; Merdanovic et al., 2011). According to Merdanovic *et al.*, nine are serine (DegQ, DegP, Tsp, YdgD, PBP4, PBP5, PBP6, PBP6B and PBP7), six are metallo (PtrA, MepA, IAP, YebA, YfgC and YhjJ) and three are cysteine proteases (NlpC, YafL and YdhO). Two potential proteases (YdcP and YhbU) are unclassified. Proteolysis eliminates damaged proteins, thereby contributing to protein quality control. Moreover, proteases regulate stress pathways *via* signaling protein activation or inactivation.

Folding mechanisms of disulfide bond-containing (recombinant) proteins

Many biopharmaceutical proteins bear disulfide bonds. These are essential for stability (Elksne and Rasmussen, 1996; Laminet and Plueckthun, 1989; Zhang et al., 1994) or

catalytical activity (Berkmen, 2012; Creighton, 1986). As reported in a 2014 review article (Herrmann and Riemer, 2014) and articles cited therein (Bartlett and Radford, 2009; Braakman and Bulleid, 2011; Hartl et al., 2011; Kang and Kini, 2009), disulfide bond-containing proteins assume their native conformation (“oxidative protein folding”) in several steps. Partial folding (step 1) occurs prior to cysteine oxidation (step 2). In the first part of the step 1, formation of secondary structures occurs within milliseconds to seconds. Secondary structure elements assume a three-dimensional orientation based on non-covalent interactions on a similar time scale. Afterwards, *cis-trans* isomerization of proline residues occurs. This happens on the seconds to minutes time scale and often constitutes a rate-limiting step in protein folding (Lu et al., 2007). When thiol groups of two cysteine residues are in proximity, formation of covalent disulfide bonds (step 2) can occur. Like proline isomerization, this can take seconds to minutes and may also constitute a rate-limiting step. As cysteines have to be sufficiently close for disulfide bond formation, this step rather accompanies than supports the folding process (Kosuri et al., 2012).

Relevance of the periplasm for recombinant protein production

The suitability of the two *E. coli* compartments, *i.e.* cytoplasmic and periplasmic space, for production of disulfide bond-bearing recombinant proteins has previously been reviewed (Baneyx and Mujacic, 2004; Choi and Lee, 2004; de Marco, 2012; Denoncin and Collet, 2012; Depuydt et al., 2010; Huang et al., 2012; Jeong et al., 2011; Messens and Collet, 2006; Rosano and Ceccarelli, 2014; Schlegel et al., 2013; Sorensen and Mortensen, 2005b). While the cytoplasm of conventional *E. coli* production strains is the compartment of choice for IB formation, it is characterized by severe disadvantages for synthesis of soluble, disulfide bond-bearing proteins. For instance, reducing conditions in the cytoplasm prevent disulfide bond formation. The *E. coli* alkaline phosphatase did not acquire disulfide bonds and was inactive when retained in the cytoplasm (Derman and Beckwith, 1991). The oxidizing periplasmic space, on the other hand, allows formation of disulfide bonds. Progress in the development of strains with an oxidative cytoplasmic space, *e.g.* by mutation of the *trxB* gene or introduction of a *trxB/gor* double mutation, has been made. In the resulting strains (Novagen AD494 and Novagen Origami, respectively) soluble production of disulfide bond-containing proteins in the cytoplasm was reported (Cassland et al., 2004; Levy et al., 2001; Sonoda et al., 2010; Sorensen and Mortensen, 2005b). However, the periplasmic space offers further advantages for soluble protein production. Its folding machinery and the fine-tuned quality control mechanisms can constitute an advantage (Duguay and Silhavy, 2004; Merdanovic et al., 2011). The

proteolytic activity in the periplasm is decreased compared to the cytoplasm (Swamy and Goldberg, 1982). This can result in decreased target protein degradation (Talmadge and Gilbert, 1982). Purification of proteins from the periplasm (as opposed to isolation from whole cell lysates) is supposed to be simplified due to fewer contaminating proteins in the periplasm. Upon translocation, signal peptide cleavage generates an authentic N-terminus of the recombinant target protein, which is often crucial for biopharmaceuticals. In summary, synthesis of disulfide bond-containing proteins, such as antibody fragments and peptide hormones in the periplasm constitutes the preferred route in case of “conventional” *E. coli* production strains (Huang et al., 2012).

The mechanism of secretion

In order to take advantage of the periplasm’s characteristics, recombinant proteins have to be translocated across the inner membrane (IM). Translocation of native *E. coli* proteins has already been reviewed comprehensively (Choi and Lee, 2004; Driessen and Nouwen, 2008; Goemans et al., 2014; Herrmann and Riemer, 2014; Merdanovic et al., 2011; Mergulhão et al., 2005; Natale et al., 2008). Unless otherwise specified, this section is based on these articles.

Proteins destined for secretion are synthesized in the cytoplasm with a specific N-terminal amino acid sequence (“signal sequence”). This sequence determines which secretory machinery catalyzes the translocation. Signal sequences are often composed of a hydrophobic domain (ten to 20 amino acids) and a short, positively charged N-terminal domain (two to ten amino acids). During the protein translocation process, the signal sequence is cleaved by a signal peptidase (*e.g.* at the Ala-X-Ala box in case of signal peptidase I). One of three pathways (Sec, SRP or TAT pathway) is typically involved in secretion.

The twin-arginine translocation (TAT) pathway allows for secretion of folded proteins from the cytoplasm. It utilizes signal sequences with the characteristic twin-arginine motif. Approximately 30 native *E. coli* proteins are secreted *via* this pathway. The vast majority of polypeptides are secreted in an un- or loosely folded conformation across the IM *via* the Sec pathway under involvement of the SecYEG apparatus/Sec translocon (Dalal and Duong, 2011; du Plessis et al., 2011; Mergulhão et al., 2005; Papanikou et al., 2007; Rusch and Kendall, 2007; Valent et al., 1998). Polypeptides can be translocated co-translationally and depend on interaction with SRP (signal recognition particle). Alternatively, they are secreted post-translationally and depend on interaction with SecB. SRP recognizes the protein’s signal sequence as it emerges from the ribosome, causing stalling

of the translation process. Afterwards, the complex of SRP, ribosome and nascent protein binds to the Sec translocase, where translation and translocation occur simultaneously. Most proteins residing in the cytoplasmic membrane are targeted to the Sec-translocon in this fashion (Luirink et al., 2012). However, the majority of Sec-secreted proteins is translocated across the IM *via* the SecB-dependent pathway (Driessen, 2001). In this case, nascent amino acid chains bind the trigger factor (TF) protein, which prevents interaction with SRP. Subsequently, the produced proteins are maintained in an unfolded and translocation-competent state by the chaperone SecB. When the complex of SecB and the fully synthesized amino acid chain reach the translocase, the protein is secreted post-translationally. The SecB-dependent pathway was reported to be approximately 10 times faster than the SRP-dependent pathway in review articles (Delic et al., 2014; Kadokura and Beckwith, 2010).

Typically, SecB-related signal sequences (*e.g.* derived from the native *E. coli* proteins OmpA or PelB) are used to target recombinant proteins to the bacterial periplasm (Samuelson, 2011a).

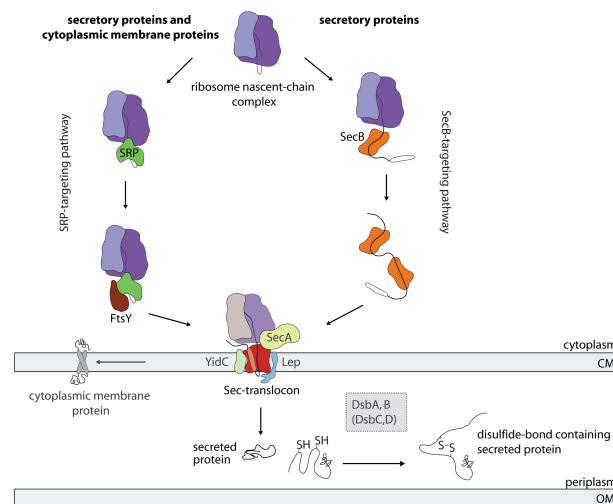


Figure 1.8: The biogenesis of Sec translocon-dependent secretory and cytoplasmic membrane proteins in *E. coli*. Most secretory and cytoplasmic membrane proteins require the Sec-translocon for their biogenesis. The Sec-translocon is a protein-conducting channel in the cytoplasmic membrane (CM in the figure equals IM in the text), which mediates the vectorial transfer of secretory proteins across and the biogenesis of membrane proteins in the cytoplasmic membrane. Secretory proteins are equipped with a cleavable N-terminal signal sequence. The signal sequence determines whether a secretory protein is targeted to the Sec-translocon via the post-translational SecB-targeting pathway or the co-translational signal recognition particle (SRP)-targeting pathway, which is comprised of the SRP and its receptor FtsY. Upon translocation, the signal sequence is cleaved off by the leader peptidase Lep and the secretory protein is released into the periplasm. In this compartment, the Dsb system can catalyze the formation of disulfide bonds. The disulfide oxidoreductase DsbA catalyzes the *de novo* formation of disulfide bonds in polypeptide chains. The disulfide bond formation protein B (DsbB) is essential to maintain DsbA in an oxidized state. Incorrectly formed disulfide bonds can be corrected by DsbC/D. Cytoplasmic membrane proteins are targeted to the Sec-translocon via the SRP-targeting pathway. SecA: peripheral membrane ATPase associated with the Sec-translocon. OM: outer membrane. YidC: cytoplasmic membrane protein translocase/insertase. The figure and description were copied from a research article (Schlegel et al., 2013).

1.3.3 Potential Issues in Soluble Recombinant Protein Production

Many disulfide bond-bearing recombinant proteins are secreted to the oxidizing periplasmic space, where they assume their final and active conformation. The underlying pathway is depicted in Figure 1.9.

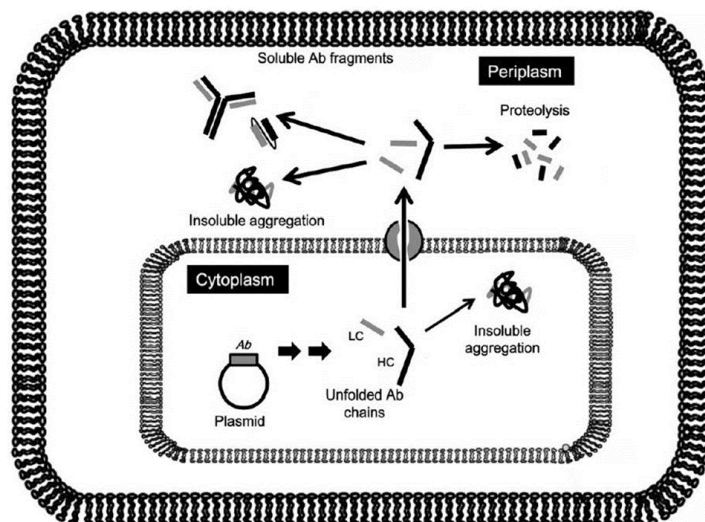


Figure 1.9: Overview of the synthesis route of soluble antibody fragments in the *Escherichia coli* periplasm. After transcription and translation in the cytoplasm, precursors are secreted to the periplasm. For diverse reasons, aggregation in the cytoplasm can occur. After secretion, amino acid chains can fold properly, form insoluble aggregates or be subject to periplasmic proteolysis. The figure was copied from a previously published review article (Jeong et al., 2011) and modified.

Pharmaceutically relevant proteins derive from various classes of proteins (see Section 1.1) and have different origins. Proteins below 100 kDa, which do not require post-translational modifications are often considered suitable for production in *E. coli* (Sugiki et al., 2014). Recombinant protein production relies on expression of the foreign gene in the microenvironment of a host cell (pH, osmolarity, redox potential, cofactors and folding mechanisms) (Rosano and Ceccarelli, 2014). Products of eukaryotic origin might be incompatible with a bacterial host strain's production machinery, which is characterized by transcription-translation coupling and lacks suitable chaperones, components for introducing post-translational modifications and compartmentalization (Lebendiker and Danieli, 2014). In conclusion, several unscheduled events can hamper soluble protein production *via* the pathway depicted in Figure 1.9. Events which might occur after translation (*i.e.* during translocation and folding in the periplasm) are presented in Figure 1.10. Numerous review articles concerning potential issues and improvements in soluble protein production have been published (Arredondo and Georgiou, 2011; Berlec and Strukelj, 2013; Carneiro et al.,

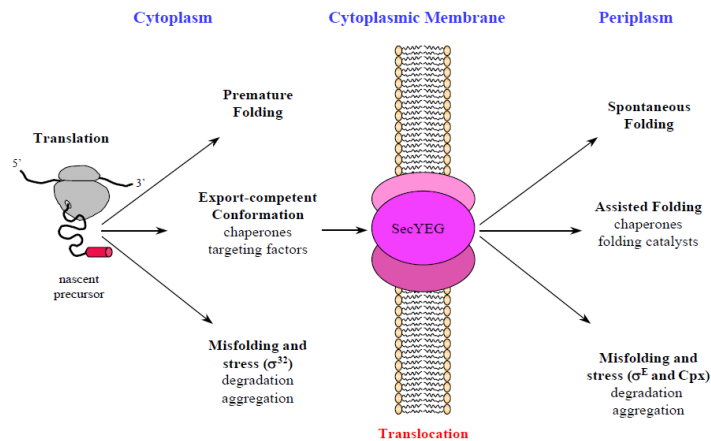


Figure 1.10: Alternative folding pathways for periplasmic proteins in *E. coli*. This scheme illustrates the present discussion and tries to emphasize the alternative and competing process between folding and misfolding of periplasmic proteins before and after translocation across the cytoplasmic membrane. Figure and description were copied from a previously published review article (Miot and Betton, 2004).

2013; Chou, 2007; de Marco, 2013; Francis and Page, 2010; Frenzel et al., 2013; Hoffmann and Rinas, 2004; Huang et al., 2012; Jeong et al., 2011; Lebendiker and Danieli, 2014; Overton, 2014; Rosano and Ceccarelli, 2014; Salinas et al., 2011; Spadiut et al., 2014). Unless otherwise stated, this section is based on these articles.

Recombinant protein production poses a “metabolic burden” on the host cell. This has been defined as the amount of resources like raw material, cellular energy (in the form of ATP), reducing power (NADH) and metabolites (especially amino acids), that is withdrawn from the host’s metabolism for maintenance of the foreign DNA and expression of foreign genes (Bentley et al., 1990; Carneiro et al., 2013; Hoffmann and Rinas, 2004). Specifically, presence of multicopy plasmids (Bentley et al., 1990; Birnbaum and Bailey, 1991; Rozkov et al., 2004; Silva et al., 2012; Spadiut et al., 2014; Wang et al., 2006), potential toxicity of gene products, high-level gene expression and protein misfolding might be underlying reasons (Berlec and Strukelj, 2013; Chou, 2007). Extracellular factors that pose a metabolic burden on cells include accumulation of toxic metabolites, limitation of nutrients and limitation of oxygen (Berlec and Strukelj, 2013; Chou, 2007). Metabolic consequences comprise growth inhibition and modification of catabolism or anabolism. Replication of chromosomal and/or plasmid DNA might be impaired. In addition, competition for resources can lead to cellular stress responses (Battesti et al., 2011; Jin et al., 2012). All these aspects might influence the host cell’s potential for recombinant protein production. Hosts for recombinant protein production might also be characterized by a low total productivity, potentially caused by weak promoters, a high mRNA degradation rate, high proteolytic activity, expression cassette loss (Rodriguez-Carmona et al., 2012;

Striedner et al., 2010) or an impaired cellular state. Protein translocation across the cytoplasmic membrane can be a bottleneck. When productivity is high, overload of the translocation and/or the holding machinery, which prevents cytoplasmic aggregation and premature folding, is feasible (de Marco, 2013; Samuelson, 2011a; Wagner et al., 2008). This can severely reduce the amount of translocated protein. In addition to hampering soluble product yield, the resulting formation of cytoplasmic aggregates has previously been reported to be toxic to the host cell (de Marco, 2013; Samuelson, 2011a; Wagner et al., 2008). Even after successful secretion of the recombinant protein to the periplasm, amino acid chains are still prone to aggregation in the periplasmic space (Hunke and Betton, 2003). A high aggregation propensity of the assembled recombinant protein itself or the appearance of labile folding intermediates can result in aggregate formation in the periplasm. Intermediates with incompletely formed disulfide bonds, which are prone to nonspecific association, are of particular relevance in this context (Baneyx and Mujacic, 2004; Schlapschy and Skerra, 2011). Note that many other limitations can impede soluble recombinant protein yield. An overview is presented in Table 1.2 (p. 32).

1.3.4 Host Engineering Approaches

Due to *E. coli*'s long history as a production organism for recombinant proteins, several tools for improving target yield are available. An overview is presented in Table 1.2 (p. 32).

Choice of strain

Several strains are (commercially) available for heterologous protein production and have previously been reviewed (Samuelson, 2011b). B and K-12 strains are frequently used in recombinant protein production approaches. B strains were reported to be less susceptible to acetate accumulation under high glucose concentration (Shiloach et al., 1996). Furthermore, they are deficient in specific proteases and are characterized by a high outer membrane permeability. *E. coli* B strains were also reported to be more resistant to environmental changes (Huang et al., 2012; Marisch et al., 2013b; Phue et al., 2005; Shiloach et al., 1996; Son et al., 2011; Terpe, 2006). A review on the history of B strains is available (Daegelen et al., 2009). BL21, a B strain representative, and its λ DE3 lysogen BL21(DE3) are widely used for recombinant protein production. They are naturally deficient in i) the ATP-dependent serine peptidase Lon, which degrades foreign proteins (Gottesman, 1996), and ii) the aspartyl protease OmpT, which is found on the outer membrane and degrades extracellular proteins (Grodberg and Dunn, 1988). BL21(DE3) bears a single expression cassette for synthesis of the T7 RNA polymerase under control of the *lacUV5* promoter in its genome (Moffatt et al., 1984; Studier and Moffatt, 1986; Studier et al., 1990). The T7 bacteriophage-derived polymerase is used to steer transcription of heterologous genes which are under control of a T7 promoter.

The “Walker” strains C41(DE3) and C43(DE3) were identified and isolated in a screen for BL21(DE3)-derivatives with improved membrane protein production characteristics (Miroux and Walker, 1996). Their superior characteristics were reported to result from mutations in the *lacUV5* promoter controlling the T7 RNA polymerase gene (Dumon-Seignovert et al., 2004). BL21(DE3)-derivative Lemo21(DE3) allows for rhamnose-induced production of varying levels of T7 lysozyme (lysY), the natural inhibitor of T7 RNA polymerase (Wagner et al., 2008). K-12 strains were initially used for analysis of biochemical genetics (Gray and Tatum, 1944; Tatum, 1945; Tatum and Lederberg, 1947). Strain MG1655 was the first strain to have its whole genome sequenced (Blattner et al., 1997) and was obtained by modification of the original K-12 wild-type strain. To date, representatives (*e.g.* RV308 and HMS174) are also frequently used as production hosts for

recombinant proteins (Choi et al., 2006; Huang et al., 2012; Marisch et al., 2013a,b). The genome of the K-12 strain W3110, a widely spread protein production host, is similar to that of MG1655 (Hayashi et al., 2006). RV308 and HMS174 are derivatives of MG1655 and W3110, respectively (Marisch et al., 2013a). Many K-12 strains are also available as DE3-bearing derivatives. B and K-12 strains were compared in numerous studies (*e.g.* BL21/RV308/HMS174 in batch cultivations (Marisch et al., 2013b), BL21(DE3)/JM109 under different feeding strategies (Shiloach et al., 1996) and B/K-12 in genome comparison studies (Studier et al., 2009)).

Codon usage optimization

Use of codons in heterologous genes, which are used by *E. coli* at a frequency below 1 % (“rare codons”) was shown to negatively affect expression levels of the heterologous genes (Kane, 1995). Codon optimization was previously suggested as a tool to improve recombinant protein production in *E. coli* (Jeong et al., 2011; Overton, 2014; Spadiut et al., 2014). Tools for codon optimization are available commercially (GeneOptimizer[®] by Thermo Fisher Scientific or GeneGPS[®] by DNA 2.0) or free of charge online (*e.g.* Puigbo et al. (2007))

Engineering of the cytoplasm

E. coli strains with an oxidizing cytoplasmic space are available (*e.g.* Origami or SHuffle (New England Biolabs) (Lobstein et al., 2012)) (Berkmen, 2012; Delic et al., 2014). Additional co-synthesis of DsbC or Skp in the cytoplasm was reported to constitute a valuable strategy in case of these strains (Frenzel et al., 2013; Kong and Guo, 2014). Co-synthesis of cytoplasmic folding chaperones (*e.g.* DnaK/DnaJ/GrpE (Calloni et al., 2012; Gragerov et al., 1992; Huang et al., 2012; Perez-Perez et al., 1995; Samuelson, 2011b) or GroES/EL (Goloubinoff et al., 1989; Gragerov et al., 1992; Huang et al., 2012; Kerner et al., 2005; Maeng et al., 2011)) also seems reasonable to improve of recombinant protein production in the cytoplasm (de Marco et al., 2007; Frenzel et al., 2013; Kolaj et al., 2009). In case of periplasmic protein production, several measures for circumventing the limited secretion capacities of the Sec translocon have successfully been applied. For thermodynamically stable proteins which fold rapidly in the cytoplasm, switching to SRP-dependent (co-translational) secretion was previously suggested (Delic et al., 2014; Schierle et al., 2003; Steiner et al., 2006; Thie et al., 2008). Furthermore, increasing the level of stabilizing cytosolic chaperones appears suitable to avoid aggregation and ensure

the translocation-competent state of target proteins. For instance, co-synthesis of disaggregating/holding chaperones ClpB, IbpA or IbpB (de Marco, 2013; Huang et al., 2012) or components of the synthesis pathway (*e.g.* the ribosome-associated trigger factor (TF) (Hoffmann et al., 2010; Samuelson, 2011b) and SecB (Baars, 2007; Baars et al., 2006)) appear relevant in this context.

Adjusting the target gene's expression rate

Fine-tuning of target gene expression was reviewed as a profound technique for harmonization with the Sec-translocon or folding capacity (Rosano and Ceccarelli, 2014; Schlegel et al., 2013). Roughly, bacterial gene expression can be divided into transcription, translation, translocation (for periplasmic proteins) and protein folding (Snyder, 2013). The transcription process consists of transcription initiation, elongation and transcription termination. Afterwards, mRNA molecules are translated into an amino acid chain in the translation process. Roughly, this process consists of three steps. Initiation occurs at the translational initiation region (TIR) or ribosomal binding site (RBS), which bears the initiation codon, Shine-Dalgarno (SD) sequence and translational enhancers (Vimberg et al., 2007). During elongation along the mRNA in 5'-to-3' direction, the peptidyltransferase within the ribosome complex catalyzes formation of peptide bonds between new amino acids and the growing peptide chain. When one of three "stop" or "nonsense" codons, for which no tRNAs are available, enters the ribosome, three release factors (RF) catalyze disintegration of the ribosome and translation termination occurs. Along these complex processes, which are exploited in recombinant protein synthesis, many targets can be used for adjusting the target protein's synthesis rate. Details are presented in a comprehensive review article from 2013 (Arpino et al., 2013), which is the basis for the following section.

Gene expression can be modified by choice or modification of promoters, which bear the sequences for recruitment of the transcriptional machinery (Balzer et al., 2013). One prominent promoter system for regulation of heterologous gene expression is the T7 system based on transcription by a T7 phage RNA polymerase (Studier and Moffatt, 1986). Several further IPTG- or lactose-inducible systems (*e.g.* *tac/trc* (Amann et al., 1983, 1988; Brosius et al., 1985; de Boer et al., 1983; Huang et al., 2012) or *lac* (Gronenborn, 1976; Hirschel et al., 1980; Huang et al., 2012)) are available. These are characterized by different possibilities of titration, promoter leakiness and level of expression. Due to several drawbacks of the T7 promoter system, a variety of non-T7 promoters has emerged in the field of heterologous gene expression. For instance, a down-regulated version of the arabinose-inducible P_{araBAD} system was shown to improve yields of membrane protein

production in *E. coli* (Nannenga and Baneyx, 2011). The *XylS/Pm* expression system constitutes an alternative system (Aune, 2008; Aune et al., 2010). The phage λ pR, pL system is temperature shift-inducible, titratable, rather leaky and enables high-level expression (Huang et al., 2012; Valdez-Cruz et al., 2010). The P_{lux} system enables high-level expression, tight control and inexpensive induction using homoserine lactone.

The translational level can, for instance, be adjusted by modifying the SD sequence or its distance from the translation initiation codon (Chen et al., 1994; Osterman et al., 2013). Tools for prediction and experimental data concerning strength of different SD sequences are available (Salis et al., 2009; Seo et al., 2013).

The gene expression rate can also be influenced by the number of expression cassettes present in the host cell. For instance, different types of replication origin (*ori*), *e.g.* pUC-based (hundreds of copies), pBR322/ColE1 (tens of copies) (Tolia and Joshua-Tor, 2006), pSC101 (few copies) or p15A/pACYC184 (few copies) (Chang and Cohen, 1978; Tolia and Joshua-Tor, 2006) enable plasmid copy number (PCN) adjustment (de Marco, 2013). Using different *oris* was also suggested for adjusting different levels of periplasmic chaperones in co-expression approaches (Schaefer and Plueckthun, 2010). Selection markers are necessary to maintain high copy plasmids, which opposes the FDA's demand to omit the use of antibiotics (Huang et al., 2012). An appropriate strategy is the use of low copy-number plasmids (Samuelson, 2011a,b). Furthermore, the *cer* determinant has been shown to effectively control copy number as a stabilizing plasmid-borne genetic module (Summers and Sherratt, 1984). Genomic integration of a single gene expression cassette, *i.e.* reduction of the gene copy number to one, was shown to be advantageous in case of production of recombinant proteins in the *E. coli* cytoplasm (Chen et al., 2008; Mairhofer et al., 2013; Striedner et al., 2008, 2010).

Co-synthesis of periplasmic folding modulators

Due to their broad substrate spectrum, native components of *E. coli*'s periplasmic folding machinery constitute a target for up-regulation approaches to improve host cell performance in soluble recombinant protein production (Kolaj et al., 2009). An overview of the periplasmic folding machinery is presented in Section 1.3.2. As numerous combinations are feasible, several review and method articles have already been published (see, *e.g.*, Kolaj et al. (2009) and Schlapschy and Skerra (2011)). Several (combinations of) components have successfully been applied to enhance soluble and active/functional recombinant target protein yields during periplasmic expression in *E. coli*. An overview is

presented in Table 1.1. In conclusion, success of folding modulator gene co-expression is barely predictable (Fahnert, 2012; Kolaj et al., 2009; Overton, 2014). However, a guideline for deciding which folding modulator to test was introduced by Kolaj *et al.* in 2009 (see Figure 1.11).

Table 1.1: Overview of studies which report successful application of periplasmic folding modulators for enhanced soluble and active recombinant target protein yield during periplasmic expression.

| Folding Modulators |
|---|
| Products and References |
| DegP Lipase B (Xu et al., 2008), penicillin acylase (Lin et al., 2001a,b; Pan et al., 2003) |
| DsbA Cyclohexanone monooxygenase (Lee et al., 2004), heat-labile enterotoxins (Wulfing and Rappuoli, 1997), human superoxide dismutase (Mao et al., 2010), human leptine (Jeong and Lee, 2000), insulin-like growth factor-I (Joly et al., 1998), lipase B (Xu et al., 2008), T-cell receptor fragments (Wulfing and Pluckthun, 1994) |
| DsbC Full-length IgG (Lee et al., 2013), horseradish peroxidase (Kurokawa et al., 2000), insulin-like growth factor-I (Joly et al., 1998), <i>Ragi</i> bifunctional inhibitor (Maskos et al., 2003), reteplase (Zhuo et al., 2014), scFv (Hu et al., 2007; Sun et al., 2014), tissue plasminogen activator (Qiu et al., 1998) |
| FkpA Maltose-binding protein (Arie et al., 2001), penicillin acylase (Wu et al., 2007), scAb (Zhang et al., 2003), scFv (Bothmann and Plueckthun, 2000; Ow et al., 2010; Padiolleau-Lefevre et al., 2006; Sonoda et al., 2011), single chain T cell receptors (Gunnarsen et al., 2010) |
| Skp Fab (Lin et al., 2008), scAb (Hayhurst et al., 2003; Hayhurst and Harris, 1999; Mavrangelos et al., 2001), scFv (Bothmann and Pluckthun, 1998; Ow et al., 2010; Sonoda et al., 2011; Wang et al., 2013), T-cell receptors (Maynard et al., 2005) |
| Combinations of Dsb proteins Brain-derived neurophilic factor (Hoshino et al., 2002), glutamate racemase (Kohda et al., 2002), horseradish peroxidase (Kurokawa et al., 2000), human nerve growth factor (Kurokawa et al., 2001), salmon C-type lectin receptor C (Soanes et al., 2008), scFv (Sandee et al., 2005; Sun et al., 2014) |
| DsbA, DsbC, FkpA, SurA (pTUM4) β -scorpion toxin (O'Reilly et al., 2014), DC-SIGN domain (Schlapschy et al., 2006), Fab (Friedrich et al., 2010), lipase B (Xu et al., 2008), lipocalins (Breustedt et al., 2006), Pfs48/45 protein (Outchkourov et al., 2008), plasma retinol-binding protein (Schlapschy et al., 2006) |
| FkpA, Skp scFv (Ow et al., 2010) |

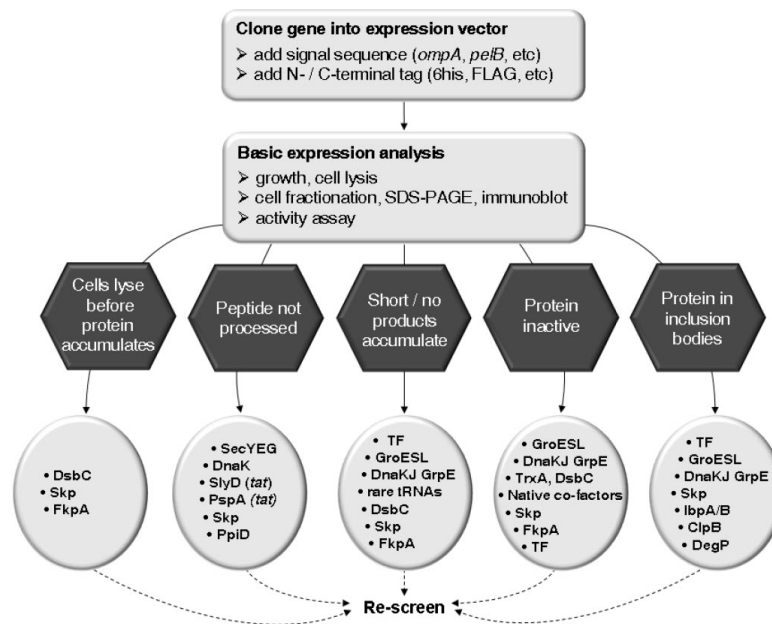


Figure 1.11: Strategy for selection of molecular chaperones and folding catalysts for co-synthesis approaches. Following production of a recombinant protein in *E. coli*, analysis of cell growth, protein solubility and subcellular location, macromolecular state and activity provide some insight into the limiting step in the folding and production process. This figure shows the major issues typically encountered (in hexagons) during production of a difficult-to-produce recombinant target and identifies the co-production strategies that have been most successful in overcoming these bottlenecks to date (corresponding ovals). Image and text were copied from a review article (Kolaĳ et al., 2009).

1.3.5 Process Development Approaches

In addition to genetic host engineering approaches, fermentation process development strategies bear the potential to improve soluble product yield of recombinant proteins produced in *E. coli*. This section is mainly based on a review article from 2014 (Overton, 2014).

Selection of growth medium

The choice of growth medium has a substantial influence on the production process outcome (Cote and Gherna, 2002). Growth media typically contain nutrients like carbon and nitrogen sources, amino acids and micronutrients (*e.g.* vitamins or metal ions). Glucose and glycerol are the most commonly used carbon/energy sources in bacterial growth media. In case of glucose as the main energy source, fed-batch fermentations ensuring limited glucose supply can avoid overflow metabolism phenomena and resulting acetate secretion (Doelle et al., 1982; Shiloach and Fass, 2005). Glycerol, on the other hand, usually does not lead to acid generation but is more expensive than glucose and is associated with lower biomass yields. Amino acids and other biochemicals are often added to prevent limitations during rapid growth. Both chemically defined media or media, for which certain components, *e.g.* yeast hydrolysates, are chemically undefined, can be used in bioprocesses. Both types are characterized by advantages and drawbacks concerning costs, preparation, growth behavior and regulatory requirements.

Choice of temperature

Process temperature is one of the main influential factors for the product yield and is, thus, of special interest in process development approaches (Rodriguez-Carmona et al., 2012). A reduced temperature after inducer addition in order to enable improved yield of soluble product was mentioned previously (Samuelson, 2011a). Amongst other reasons, a decreased protein synthesis rate presumably prevents accumulation of folding intermediates in the cytoplasm (Battistoni et al., 1992; Huang et al., 2012; Jeong et al., 2011; Thieringer et al., 1998; Vasina and Baneyx, 1997). In case of periplasmic Fab production in *E. coli* HMS174(DE3) in carbon-limited continuous cultures, product yield increased as temperature decreased (Rodriguez-Carmona et al., 2012). Furthermore, reduced temperature has also been reported to amplify “secretion” effects (Dragosits et al., 2011; Rodriguez-Carmona et al., 2012), which might be desired. On the other hand, negative effects of

rapidly lowering temperature before induction are also possible (Tolia and Joshua-Tor, 2006). In conclusion, a suitable temperature strategy has to be identified for each target product and host strain (considering the target protein’s aggregation tendency, production rates and potential physiological stress due to temperature shifts) (Fahnert, 2012).

Induction strategy

In dependence of the target protein, different time points to induce recombinant gene expression are feasible (see Figure 1.12) (Overton, 2014). Depending on the host cell system, different types of induction of recombinant protein production are necessary. For the following introduction, an isopropyl- β -D-1-thiogalactopyranoside (IPTG)-inducible system and bolus addition of the inducer are assumed. In other cases (*e.g.* induction by arabinose), inducer metabolization has to be considered.

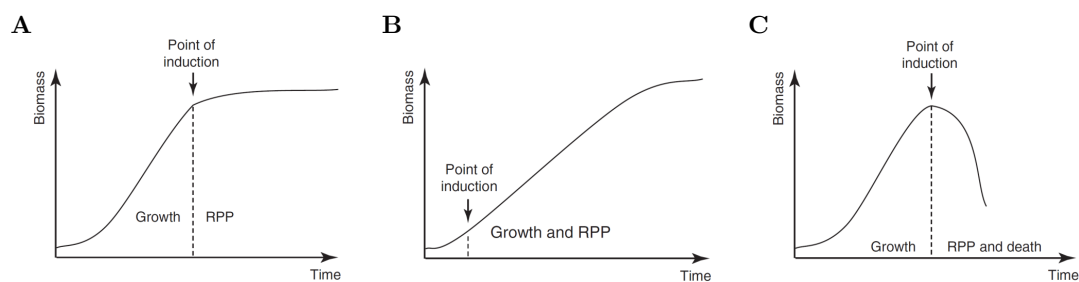


Figure 1.12: Different induction strategies. Induction at a high biomass (**A** and **C**) enables separation of growth and recombinant protein production (RPP) phases. Following induction, growth often slows as metabolic resources are channelled to protein production. Induction at low biomass (**B**) can only proceed if metabolic resources are evenly split between biomass formation and recombinant protein production; if this does not occur, then growth and/ or recombinant protein production will be impaired. When “toxic” proteins are produced (**C**), growth is severely inhibited or even cell lysis occurs. Figure and description were copied from a review article (Overton, 2014).

The specific growth rate at induction is influenced by the time point chosen for inducer addition. Growth rate has been reported to influence production of the recombinant protein (Curless et al., 1990). If the inducer is added to the culture in the middle of logarithmic growth phase, *i.e.* prior to biomass accumulation (see Figure 1.12 (B)), cells can provide sufficient levels of energy and metabolic precursors for the recombinant biosynthesis. The cells are at their maximum catabolic capacity (Carneiro et al., 2013). This strategy was recommended for target proteins that do not cause growth inhibition (Overton, 2014). Inducer can also be added during late-logarithmic growth phase (see Figure 1.12 (A) and (C)). This uncouples growth and production phase and leads to higher cell densities, *i.e.* increases the amount of cells for production. However, cells might be in an unfavorable metabolic state. This comprises the presence of stress-related proteins like proteases,

which can reduce the yield of the target protein (Carneiro et al., 2013). This strategy was previously recommended for recombinant proteins that are known to inhibit growth (Overton, 2014). Apart from induction time point, the amount of inducer might also influence the product formation. In this regard, lower inducer concentrations were previously suggested to be advantageous for soluble product synthesis (see Table 1.2, p. 32). Note that inducer titration during IPTG induction of plasmid-bearing production systems was previously reported to be unpredictable (Rosano and Ceccarelli, 2014). Reasons were identified in heterogeneous distribution of the lac permease enzyme, which actively transports IPTG into the cell (Fernandez-Castane et al., 2012). Thus, inducer titration is rather applicable for systems that do not rely on IPTG-based induction. On the contrary, Striedner *et al.* reported successful titration of the IPTG-inducible T7 system, albeit for a plasmid-free expression system with a single genome-integrated copy of the heterologous gene.

1.3.6 Current Benchmarks and Conclusion

Numerous techniques have been applied in attempts to improve soluble product yields during recombinant antibody fragment production in *E. coli*. In review articles by Jeong et al. (2011) (Jeong et al., 2011, Table 2) and Frenzel et al. (2013) (Frenzel et al., 2013, Table 1) soluble product titers of 1.1 to 2.5 g l⁻¹ (F(ab')₂), 14 mg l⁻¹ to 2 g l⁻¹ (Fab), 0.3 to 1.3 g l⁻¹ (scFv), 3.0 g l⁻¹ (dimeric scFv), 0.5 g l⁻¹ (Fv) and 35 to 65 mg l⁻¹ (VL dAb) were reported. Note that product titers were obtained with different strains (*e.g.* RV308, UL635, W3110, BL21(DE3), XL1-Blue or DH5 α), promoters (*e.g.* PhoA, Lac, Trc, XylS/Pm or T7) and in various fermentation formats (*i.e.* different culture volumes and process parameters). Furthermore, different genetic strategies (*e.g.* protease gene knock out or co-synthesis of helper proteins) were applied. In a review article by Meyer and Schmidhalter (2012), significant variability in soluble product titers (0.5 - 10 g l⁻¹) obtained with the commercial Lonza XS Technologies™ platform was reported, depending on the product (Fab, Fab fusion proteins, scFv, growth factors, enzymes and various formats of amphipathic proteins). In accordance, it was previously reviewed that no rational approach for *in silico* identification of bottlenecks and/or solutions exists (Fahnert, 2012; Frenzel et al., 2013). The potential for soluble production in *E. coli* greatly depends on the respective recombinant product and techniques at hand. Techniques for improvement have been suggested in several review articles (Jeong et al., 2011; Lebendiker and Danieli, 2014; Overton, 2014; Rosano and Ceccarelli, 2014; Spadiut et al., 2014). A summary overview of genetic host engineering and process development approaches depending on the observed bottlenecks is presented in Table 1.2. Although all mentioned solutions are relevant for soluble protein production in *E. coli*, the optimization of synthesis rate and periplasmic folding capacity are of major relevance (Delic et al., 2014; Denoncin and Collet, 2012).

Table 1.2: Overview of common issues during plasmid-based soluble protein production in the periplasm of *E. coli*. The observed issue, its analysis methods and proposed solutions are listed. The content of this table is based on several review articles (Berlec and Strukelj, 2013; Chou, 2007; Huang et al., 2012; Jeong et al., 2011; Lebendiker and Danieli, 2014; Overton, 2014; Rosano and Ceccarelli, 2014; Spadiut et al., 2014).

| Synthesis step | Issue | Analysis | Solutions |
|-------------------------------|-------------------------------------|---|---|
| Cell viability | Growth cessation prior to induction | Biomass monitoring, product concentration measurement, RT-PCR/qPCR | PCN, T_{low} , glucose addition (for <i>lac</i> promoters), P_{opt} , pLys plasmids |
| | Growth cessation after induction | Monitoring biomass progress | P_{opt} , T_{low} , Ind_{low} |
| Expression cassette stability | Plasmid loss | Replica plating | PCN, T_{low} , Ind_{low} , different/no marker, plasmid-free strains |
| | Runaway effect | PCN determination | Stabilizing elements, PCN, plasmid-free strains |
| Transcription | Basal expression | Monitoring biomass progress, product concentration measurement or RT-PCR/qPCR | P_{opt} , PCN, T_{low} , glucose addition (for <i>lac</i> promoters), pLys plasmids |
| | GoI mRNA abundance | Northern blot, RT-PCR/qPCR | P_{opt} , RBS_{opt} , P_{opt} |
| | Read-through effects | <i>in vitro</i> transcription assay, RT-PCR/qPCR | Terminator change |
| Translation | mRNA degradation | Northern blot, RT-PCR/qPCR, bioinformatic analysis | Removal of degradation-triggering features, protective elements, down-regulation of RNase genes |
| | Translation stalling | Bioinformatic analysis, RT-PCR/qPCR | Codon optimization, provision of rare tRNAs, new media formulations |
| Translocation | Cytoplasmic aggregation | N-terminal sequencing, MS, subcellular fractionation, stress monitoring | Ind_{low} , T_{low} , PCN, good aeration, prevention of foaming, cytoplasmic FM, secretion pathway switch |
| Periplasmic folding | Folding machinery overburdening | Cell fractionation, N-terminal sequencing, MS, stress monitoring | T_{low} , periplasmic FM, Ind_{low} , P_{opt} , RBS_{opt} , chemical chaperones and co-factors |
| | Unbalanced monomer ratio | MS, SDS PAGE | LC-HC switch, monocistronic organization with different promoters |
| | Proteolysis | MS, Omics | Protease gene regulation |

FM, folding modulators. IB, inclusion body. Ind_{low} : lowered inducer concentration. MS, mass spectrometry. qPCR, quantitative PCR. PCN, reduced or stabilized plasmid copy number. P_{opt} : optimized (weaker, tight and tunable) promoters. RBS_{opt} : optimized ribosomal binding sites. RT-PCR, reverse-transcriptase PCR. T_{low} : reduced temperature (growth/production phase).

1.4 Aim and Objectives

A plethora of host engineering techniques for improving soluble protein production in the *E. coli* periplasm is available. A synergistic combination of single approaches was previously suggested (Delic et al., 2014) but has, so far, rarely been applied. The aim of this work was to create and test a robust and lean toolbox for improving recombinant soluble product yield of therapeutically relevant proteins. Commercial demands (*e.g.* limited time and resources for development) were to be considered. The developed toolbox was supposed to enable a quick screening of several host strains/systems prior to start of process development.

Different techniques were supposed to be tested concerning their suitability for improving soluble protein production. Specifically, genomic integration of target gene expression cassettes and co-synthesis of cytoplasmic or periplasmic folding modulators appeared relevant for this project. During this work, construction and testing of a variety of genetic elements was intended.

The suitability of techniques and combinations thereof was supposed to be verified in production experiments with different products. Four model proteins (GFP and three antibody fragments) were available for this purpose. Production experiments with conventional and newly developed systems were supposed to be performed side-by-side. In order to obtain meaningful results, parameters of fermentation experiments were supposed to be defined according to current industrially relevant process formats. To enable a pre-evaluation, appropriate small scale approaches had to be designed.

During production experiments, results for assessing why applied techniques did or did not improve the soluble product yields were supposed to be obtained. For this objective, analysis procedures necessary for detailed examination of the impact of host engineering techniques were to be established and/or applied.

A guideline on how to use the host engineering toolbox elements (see, *e.g.*, Figure 1.11) was considered a crucial part of the final toolbox. Based on the findings of production experiments and the detailed analyses, such a guideline for using the developed and tested techniques was supposed to be established.

Materials and Methods

2.1 Model Proteins

Model proteins GFP, FabZ, FabX and scFv-Z-1 (see Table 2.1) were used within this study. Protein characteristics were determined using ExPASy ProtParam (Gasteiger et al., 2005) using the amino acid sequence without N-terminal methionine and leader peptide.

Table 2.1: Model proteins used during the course of this work. Protein monomer parameters were calculated using ExPASy ProtParam (Gasteiger et al., 2005).

| Protein | Type | MW [kDa] | AA | -S-S- | II | AI | GRAVY ³ |
|----------|-------------|------------------------|--------------------|-------|--------------------------------------|------------------------|--------------------------|
| FabX | Heterodimer | HC: ~ 25 LC: ~ 23 | n.a. | 5 | n.a. | n.a. | n.a. |
| FabZ | Heterodimer | HC: 23.20 LC: 23.14 | HC: 219 LC: 213 | 5 | HC: 43.52 (u) LC: 54.61 (u) | HC: 73.38 LC: 75.07 | HC: -0.171 LC: -0.288 |
| GFP | Monomer | 26.78 | 238 | 0 | 28.05 (s) | 75.71 | -0.509 |
| scFv-Z-1 | Monomer | 24.88 | 236 | 2 | 47.44 (u) | 73.56 | -0.171 |

MW: molecular weight. AA: number of amino acid residues. -S-S-: number of disulfide bonds in the correctly folded product. II: insolubility index. s/u: classified as stable (II < 40)/unstable (II ≥ 40) (Guruprasad et al., 1990). AI: aliphatic index (positive factor for the increase of thermostability of globular proteins) (Ikai, 1980). GRAVY: Grand average of hydropathicity (hydropathy values of all amino acids, divided by residue number) (Kyte and Doolittle, 1982).

The monomeric protein GFP used in this study has a molecular weight of 26.78 kDa (238 residues). It was originally introduced as a marker for gene expression in 1994 (Chalfie et al., 1994). The version used in this study derives from a previously published optimized type of GFP (GFPmut3b) (Cormack et al., 1996). It was also used in initial studies for examination of plasmid-free strains (Striedner et al., 2010).

Antibody fragments FabZ (fragment antigen-binding) and scFv-Z-1 (single-chain fragment variable) are derived from a therapeutically relevant monoclonal antibody. Due to confidentiality reasons, details are presented in Section 5.1 (Confidential Supplement). FabZ has 432 residues and a molecular weight of approximately 46 kDa (LC: 23.14, HC: 23.20). Each monomer of this heterodimer bears two intra-chain disulfide bonds. An additional disulfide bond is formed between C_{H1} and C_L. scFv-Z-1 has 236 residues and a molecular weight of 24.88 kDa. It has a HC-linker-LC structure and bears one intra-chain disulfide bond in each of its monomers. Note that the LC-linker-HC version scFv-Z-2 was used for

establishing a Protein L-based method for scFv detection (see Section 2.6.5). However, in initial shake flask experiments, an increased total productivity was determined for the HC-linker-LC-type (scFv-Z-1). Thus, this scFv variant was used in this study.

FabX has a molecular weight of approximately 48 kDa (LC: 23, HC: 25). Each monomer bears two 2 intra-chain disulfide bonds. A fifth disulfide bond between C_H1 and C_L connects the two monomers. Due to confidentiality reasons, further details are provided in Section 5.1.1 (Confidential Supplement).

In conclusion, the model proteins include therapeutically relevant (mAb-derived), disulfide bond-bearing and rather hard-to-produce (see II and GRAVY values) proteins of different type (Fab and scFv).

2.2 Generation of Examined Strains

During standard molecular biology procedures, cells were grown in S-LB Medium (5 g l⁻¹ bacto yeast, 5 g l⁻¹ NaCl, 10 g l⁻¹ soy flour peptone). All components were dissolved in RO-H₂O, heat-sterilized and stored for a maximum of three months at RT prior to use. For obtaining single colonies, cells were spread on S-LB agar (5 g l⁻¹ bacto yeast, 5 g l⁻¹ NaCl, 10 g l⁻¹ soy flour peptone, 15 g l⁻¹ bacto agar). S-LB agar was sterilized, heated, poured into Petri dishes at approximately 60 °C and used when cooled to RT. In accordance with previous articles, strains bearing genome-integrated genes are identified with smaller- and greater-than signs (“<” and “>”), whereas plasmid-bearing strains are identified by parentheses (Mairhofer et al., 2013; Striedner et al., 2010). B strains (BL21/BL21(DE3)) are identified by an initial “B”, K-12 strain HMS174(DE3) by an initial “K” in strain short identifiers.

Table 2.2: Basic strains used in the course of this work. BL21, BL21(DE3) and HMS174(DE3) were used for target protein production and promoter tests. DH5 α was used for plasmid propagation.

| Strain | Genotype | Supplier |
|--------------|---|---------------------------------------|
| BL21 | B F ⁻ <i>fhuA2 ompT gal dcm lon</i> <i>hsdS_B(r_B⁻m_B⁻) gal [malB⁺]_{K-12}(λ^S)</i> | Novagen (Merck, Darmstadt, Germany) |
| BL21(DE3) | like BL21, λ (DE3 [<i>lacI lacUV5-T7</i> gene 1 <i>ind1 sam7 nin5</i>]) | Novagen (Merck, Darmstadt, Germany) |
| DH5 α | F- Φ 80 <i>lacZ</i> Δ M15 Δ (<i>lacZYA-argF</i>) U169 <i>recA1 endA1 hsdR17</i> (rK ⁻ , mK ⁺) <i>phoA</i> <i>supE44</i> λ - <i>thi-1 gyrA96 relA1</i> | Life Technologies (Carlsbad, CA, USA) |
| HMS174(DE3) | K-12 F ⁻ <i>recA1 hsdR</i> (r _{K-12} ⁻ m _{K-12} ⁺) (Rif R) λ (DE3 [<i>lacI lacUV5-T7</i> gene 1 <i>ind1</i> <i>sam7 nin5</i>]) | Novagen (Merck, Darmstadt, Germany) |

2.2.1 Creation of Plasmid Constructs

A modular approach for cloning was developed within another project at BI RCV. This modular plasmid system was used for creation of almost all constructs. Exceptions are mentioned in the respective results section. Due to confidentiality reasons, module restriction sites and DNA sequences are only presented in Sections 5.2 and 5.4 (Confidential Supplement). Enzymes used during plasmid creation were supplied by New England Biolabs (Ipswich, MA, USA), Fermentas (Waltham, MA, USA) or Promega (Fitchburg, WI, USA). DNA preparation kits (QIAprep[®] MiniPrep Kit, QIAquick[®] PCR Purification Kit, QIAquick[®] Gel Extraction Kit) were supplied by Qiagen (Stockach, Germany). Single-strand DNA oligo molecules were supplied by Sigma-Aldrich (Vienna, Austria).

Table 2.3: Basic plasmids used in this study. Note that certain plasmids were not generated as basic plasmids but only as plasmid with target genes in their MCS. Note also that plasmid 707-FLPe, pSIM5 and pSIM9 bear genes. However, to allow an overview of the components of these plasmids, they are listed. Plasmids are arranged in alphabetical order.

| Name | P | T | <i>lacO</i> | <i>lacI</i> | ori | <i>cer</i> | Marker | Reference |
|---------------|---|------------------------|-------------|-------------|-------------------------------------|------------|---------------|---------------------|
| 707-FLPe | λ R cI857 ^{ts} | n.a. | – | – | pSC101 <i>repA</i> ^{ts} | – | tet | GeneBridges |
| pACIB-int | T7 | Zenit | + | – | pBR322 | – | kan(F) amp | ACIB |
| pBI1Kara.1 | araBAD | T7 | – | – | pBR322 | – | kan | BI, pre-work |
| pBI1KT7.1 | T7 | T7 | + | + | pBR322 | – | kan | BI, pre-work |
| pBI1KiT7.1 | T7 | T7 | + | + | pBR322 | – | kan(F) | this work |
| pBI1KiT7.3 | T7 | Zenit | + | + | pBR322 | – | kan(F) | this work |
| pBI1ST7.1 | T7 | T7 | + | + | pBR322 | – | stm | Schuller (2015) |
| pBI1ST7.2 | T7 | T7 | + | + | pBR322 | + | stm | Schuller (2015) |
| pBI4iSC(n).2 | C(n) | T7 | – | – | p15A | + | stm | Huber (2015) |
| pBI4iSC(n).3 | C(n) | Zenit | – | – | p15A | – | stm | Huber (2015) |
| pBI4iST7.1 | T7 | T7 | + | + | p15A | – | stm | Schuller (2015) |
| pBI4iST7(n).2 | T7(n) | T7 | + | + | p15A | + | stm | Schuller (2015) |
| pBI4iST7(n).5 | T7(n) | T7 | + | – | p15A | + | stm | Schuller (2015) |
| pBI4iST7(n).7 | T7(n) | T7 | – | – | p15A | + | stm | Schuller (2015) |
| pBI4iST7.10 | T7 | Zenit | – | – | p15A | – | stm | Huber (2015) |
| pET30a(+) | T7 | T7 | + | + | pBR322 | – | kan | Novagen |
| pSIM5 | <i>p</i> _L cI857 ^{ts} | <i>t</i> _{L3} | – | – | pSC101 <i>repA</i> ^{ts} | – | cml | Datta et al. (2006) |
| pSIM9 | <i>p</i> _L cI857 ^{ts} | <i>t</i> _{L3} | – | – | RK2 <i>trfA</i> ^{ts} | – | cml | Datta et al. (2006) |

P: Promoter. T: Terminator. Marker: Antibiotic resistance gene. ACIB: Austrian Centre of Industrial Biotechnology (Graz, Austria). BI: Boehringer Ingelheim RCV GmbH & Co KG. GeneBridges: GeneBridges (Heidelberg, Germany). Novagen: Novagen (Merck, Darmstadt, Germany).

Plasmid constructs were propagated in *E. coli* DH5 α cells (Life Technologies, Carlsbad, CA, USA). Basic plasmids (see Table 2.3), genes of interest or folding modulator genes were used to create expression plasmids (see Table 2.4 and 2.5, respectively).

Standard cloning and cloning of PCR amplicates

Unless stated otherwise, modules used in this study were present in plasmids from previous projects. Folding modulator genes were amplified from the *E. coli* HMS174(DE3) genome in the course of a master thesis prior to this work (Heistingner, 2013) and were present in different plasmids with appropriate restriction sites. FabZ genes (HC-LC and LC-HC-

Table 2.4: Plasmids for expression of target genes. Products are in alphabetical order. Due to confidentiality reasons only selected plasmids for FabX production are presented.

| Product (<i>gene</i>) | Comments | Expression plasmid |
|------------------------------|--|---|
| FabX (<i>fabX</i>) | dicistronic LC-HC expression cassette, both monomers genetically fused to SP _{ompA} | pBI1Kara.1-fabX pBI1KT7.1-fabX |
| FabZ (<i>fabZ</i> , HC-LC) | dicistronic HC-LC expression cassette, both monomers genetically fused to SP _{ompA} | pACIB-in-fabZ pBI1KT7.1-fabZ |
| FabZ (<i>fabZ</i> , LC-HC) | dicistronic LC-HC expression cassette, both monomers genetically fused to SP _{ompA} | pBI1KT7.1-fabZ |
| GFP (<i>gfp</i>) | original GFPmut3b | pACIB-int-gfp pET30a(+)-gfp |
| GFP (<i>gfp.1</i>) | <i>NdeI</i> binding site removed from <i>gfp</i> gene | pBI1KT7.1-gfp.1 pBI4iSC(n).2-gfp.1 pBI4iSC(n).3-gfp.1 pBI1ST7.1-gfp.1 pBI1ST7.2-gfp.1 pBI4iST7.1-gfp.1 pBI4iST7.2-gfp.1 pBI4iST7(n).2-gfp.1 pET30a(+)-gfp.1 |
| scFv-Z-1 (<i>scfv-Z-1</i>) | <i>scfv-Z-1</i> gene genetically fused to SP _{ompA} | pBI1KiT7.3-scfv-Z-1 |

variants) with appropriate restriction sites were obtained from Genearth/Life Technologies (Regensburg, Germany) and cloned from the supplied vector to the target plasmids.

Basic cloning was performed using restriction endonuclease (REN) digest, purification by agarose gel electrophoresis (AGE) and subsequent ligation (Sambrook and Russell, 2001). For REN digest, reaction solutions with different volumes (35 - 50 µl) containing different amounts of the plasmid that carried the target DNA (1 - 2 µg), 10x reaction buffer (1/10 of the reaction solution volume), 10x BSA (1/10 of the reaction solution

Table 2.5: Plasmids for folding modulator co-synthesis. Basic plasmids are presented in order of chronological appearance. Selected plasmids were used to transform strains B<FabZ-kanR>(02) and B<FabX-kanR>(01). Resulting strains were tested in production experiments.

| Basic plasmid | Promoters | Folding modulators |
|---------------|--|-----------------------------------|
| pBI4iST7(n).2 | T7xxiii, T7xx, T7xxix, T7 ¹ | DsbA, DsbC, FkpA, PpiD, Skp, SurA |
| pBI4iST7(n).7 | T7xxiii, T7xx, T7xxix, T7 | DsbA, DsbC, FkpA, PpiD, Skp, SurA |
| pBI4iSC(n).3 | Ci, C8, C2 ² | DsbA, DsbC, FkpA, PpiD, Skp, SurA |
| pBI1KC(n).1 | bglA, dnaK, ibpA | DsbA |

¹Measured strength: 3, 20, 61 and 100 %, see Section 3.4.2

²Measured strength: 32, 78 and 100 %, see Section 3.4.3

volume, if required) and different amounts of REN enzyme solutions (0.5 - 1.2 μl) were prepared. Solutions were incubated at 37°C for 1 - 2 h. RENs from different suppliers (Fermentas/Thermo Scientific™ (FastDigest™ in certain cases) or NEB) were used during this work. Parameters were subject to optimization experiments in several cases. After addition of loading dye, DNA molecules were separated by agarose gel electrophoresis. Agarose gels contained agarose in 1x TAE (tris base, acetic acid, EDTA) buffer and three to four drops of a 0.025 % ethidium bromide solution (Roth) for each 200 ml of gel were added. The amount of agarose (0.8 - 1.5 %), voltage (60 - 110 V) and separation duration (60 - 100 min.) were varied depending on gel and construct size. Size of DNA sequences was assessed by comparison with signals from MassRuler™ DNA Ladder Mix (Thermo Fisher). Relevant pieces of gels were excised under UV light. DNA was purified using the QIAquick® Gel Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. For purification of DNA fragments with less than 200 bp in size, the MinElute™ Spin Columns (Qiagen, Hilden, Germany) were used. For ligation, reaction solutions of 10 μl containing different volumes of solution of excised "insert" (2.0 - 4.0 μl) and vector "backbone" (0.5 - 1.0 μl), 0.5 μl T4 DNA Ligase (1 - 3 U μl^{-1} , Promega, Fitchburg, WI, USA) and 5 μl 2x Rapid Ligation Buffer (Promega, Fitchburg, WI, USA) were prepared. The ratio of volumes of insert to backbone solution was chosen to obtain a minimum 3-fold excess of insert and was subject to optimization in case of troubleshooting. Solutions were incubated for 30 - 45 min. at 25°C without agitation. 50 μl of chemically competent *E. coli* DH5 α cells (Invitrogen/Life Technologies or prepared in-house) were thawed on ice for 15 min. 5 μl of the ligation reaction were added to cells. Afterwards, the suspension was incubated on ice for 10 min., at 42°C for 45 sec and, again, on ice for 2 min. After addition of 900 μl of S-LB medium, the cell suspension was incubated at 37°C and gentle agitation for 50 min. Subsequently, the cells were sedimented by centrifugation (3220 x g, 30 sec). The supernatant was discarded by decanting. Sedimented cells were resuspended in the residual volume and plated on S-LB plates containing an appropriate antibiotic. Plates were incubated in inverted position at 37°C over night. Colonies were used to inoculate 5 ml of S-LB medium supplemented with an appropriate antibiotic in 15 ml reaction tubes. After incubation at 37°C and 250 rpm over night, cells were sedimented by centrifugation (3220 x g, RT, 10 min). Plasmids were purified using the QIAprep® MiniPrep Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. DNA concentration was determined using a NanoDrop photometer (Thermo Fisher).

In initial folding modulator co-synthesis approaches, condition-inducible promoters were used (see Section 3.4). Plasmids were created in a master thesis supervised by the au-

thor of this work (Dürkop, 2014). To create these constructs, a *bglA* promoter sequence (see Section 5.4.4, Confidential Supplement) synthesized by GeneArt/Life Technologies (Regensburg, Germany) was used to replace the T7 promoter sequence in pBI1KT7.1 by standard cloning methods described above. This step yielded pBI1KbglA.1. Subsequently, native promoter sequences (see Table 3.8) were amplified from the genome of the *E. coli* strain HMS174(DE3) using Phusion[®] High-Fidelity DNA Polymerase (NEB). Purified genomic DNA was available from the work of a previous project (Heistingner, 2013). Genomic DNA (1 μ l of a 5 ng μ l⁻¹ solution), 4 μ l of 5x Phusion HF buffer, 0.4 μ l of dNTPs (10 mM each), 1 μ l of forward and reverse primer (0.5 μ M) and 0.2 μ l of polymerase (2 U μ l⁻¹) were used in a reaction solution of 20 μ l. The initial denaturation (98°C, 30 sec.) was followed by a second denaturation (98°C, 30 sec.), an annealing (50 - 72°C, depending on the primer, 15 sec.) and an elongation (72°C, 30 sec. per kb) step. Second denaturation, annealing and elongation step were performed 30 times in total. Final elongation was performed at 72°C for 10 sec. After the reaction, the solution was incubated at 10°C until further use and, subsequently, purified using the QIAquick[®] PCR Purification Kit according to the manufacturer's protocol. Promoter sequences, folding modulator genes and plasmid backbone were ligated in one step. PCR amplicates, plasmids bearing folding modulator genes and the plasmid for insertion of both promoter and folding modulator gene (pBI1KbglA.1) were digested with appropriate enzymes as described above. After purification by agarose gel electrophoresis, the "insert" share of the ligation reaction solution volume was divided into "promoter" and "folding modulator" share. Shares were varied in optimization approaches. The remaining cloning procedure was performed as described above.

The *cer* module, which is located downstream of the terminator, was present in plasmids from previous projects and introduced into this work in the course of a master thesis (Schuller, 2015). One DNA sequence consisting of *lacI*, spacer and transcription/translation module (see Table 5.1, Confidential Supplement) was cloned from basic plasmid pBI1KT7.1 and used to replace the araBAD promoter module in basic plasmids pBI4Sara.2 (p15A ori, *cer* module), pBI4Sara.1 (p15A ori, no *cer* module) and pBI1Sara.2 (pBR322 ori, *cer* module). This resulted in basic plasmids pBI1ST7.2, pBI4iST7.2 and pBI4iST7.1, respectively (.2: with *cer* sequence, .1: without *cer* sequence).

QuikChange site-directed mutagenesis

An *NdeI* restriction site was removed from the *gfp* gene by a single base exchange using the QuikChange II Site Directed Mutagenesis Kit (Agilent, Santa Clara, CA, USA) with

PfuUltra[®] High-Fidelity DNA polymerase and primer pair qc-GFP.1-1-fwd/qc-GFP.1-1-rev (see Table 5.2 in the confidential supplement). A reaction solution of 20 μ l containing 7.5 ng of pET30a(+)-gfp template, 1.25 μ l of both primer solutions (10 μ M), 0.5 μ l of dNTP mix (10 μ M per dNTP), 2 μ l 10x reaction buffer and 0.2 μ l of polymerase (2.5 U μ l⁻¹) was prepared. For the PCR reaction, an initial denaturation step (95°C, 30 sec) was followed by 18 cycles of denaturation (95°C, 30 sec), annealing (55°C, 1 min) and elongation (68°C, 13 min) and a final elongation step (68°C, 10 min) prior to cooling to 10°C (hold step). After PCR reaction the solution was transferred to 1.5 ml microcentrifuge vials. 0.5 μ l of *DpnI* was added and the solution was incubated for 1 h at 37°C. 50 μ l of XL1-Blue Supercompetent Cells (Agilent, same kit) were thawed on ice for 10 min. 1 μ l of the PCR reaction solution after *DpnI* digest was added to the cells. The remaining steps (transformation, incubation, plating and plasmid purification) were performed as described above. The introduced modification yielded plasmid pET30a(+)-gfp.1 bearing the *gfp.1* gene.

Oligo cloning

T7 promoter modules as well as constitutive promoter modules without insulation and weaker RBS (see Section 3.4.3) were supplied as single-strand DNA (ssDNA) oligos from Sigma-Aldrich (Vienna, Austria). Oligos were designed to resemble a double-strand (dsDNA) molecule with appropriate promoter restriction sites after annealing (see Table 5.1 in the confidential supplement). After they were dissolved with H₂O, the resulting solution (100 μ M) were stored at -20°C until further use. Both oligonucleotides necessary for cloning were phosphorylated separately. A reaction solution of 50 μ l containing 3 μ l of the ssDNA solution (100 μ M), 5 μ l of 10x T4 DNA Ligation Buffer (NEB, Ipswich, MA, USA) and 1 μ l of T4 Polynucleotide Kinase solution (NEB, Ipswich, MA, USA) was prepared. The solution was incubated for 30 min at 37°C, and 20 min at 65°C prior to cooling to 10°C and incubation until further use. After phosphorylation, both solutions were mixed for hybridization. The mixture was incubated at 95°C for 3 min, 72°C for 1 min and 50°C for 1 min prior to cooling to 10°C and incubation until further use. The target backbone plasmid was subjected to preparative restriction digest as described above. Ligation of dsDNA promoter modules and digested backbone plasmids, transformation and plasmid purification were also performed as described above. Oligo cloning was performed in the course of two master theses supervised by the author of this study (Huber, 2015; Schuller, 2015).

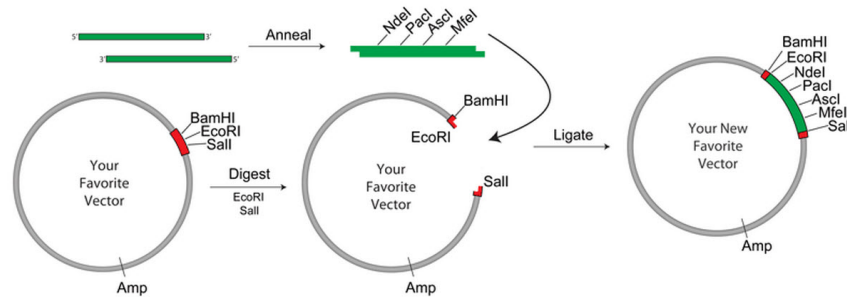


Figure 2.1: Annealed oligo cloning: schematic representation. The figure was copied from the Addgene website (<https://www.addgene.org/plasmid-protocols/annealed-oligo-cloning/>, Addgene (2014)).

Cloning with GeneArt™ Strings™

Constitutive promoters with insulation and strong RBS (see Section 3.4.3) were obtained as GeneArt™ Strings™ bearing appropriate promoter restriction sites (see Table 5.1, Confidential Supplement). Upon delivery of freeze-dried DNA fragments, these were dissolved with H₂O and inserted into pJET1.2/blunt using the CloneJET PCR Cloning Kit (Thermo Scientific) according to the manufacturer's instruction. Modules were cloned into plasmids of choice by standard cloning (see above). GeneArt™ Strings™ cloning was performed within the course of a master thesis supervised by the author of this work (Huber, 2015).

Addition of FRT sites to the kanamycin resistance gene

For addition of FRT sites to the kanamycin resistance gene, a specific primer pair (mv-frt-insertion_fwd and mv-frt-insertion_rev, see Table 5.2 in the confidential supplement) and Phusion® High-Fidelity DNA Polymerase (NEB) were used to amplify the kanamycin resistance gene from pBI1KT7.1. A reaction solution of 50 µl containing 1 ng of pBI1KT7.1 template, 2.5 µl of both primer solutions (10 µM per dNTP), 1.0 µl of dNTP mix (10 µM per dNTP), 10 µl 5x phusion HF buffer and 0.5 µl of polymerase was prepared. For the PCR reaction, an initial denaturation step (98 °C, 30 sec) was followed by 18 cycles of denaturation (98 °C, 10 sec), annealing (48 - 72 °C (12 step gradient), 15 sec) and elongation (72 °C, 30 sec per kb) and a final elongation step (72 °C, 10 min) prior to incubation at 10 °C (hold step) until further use. The PCR amplicate bore FRT sites and resistance gene module restriction sites. Success of the PCR reaction was examined by AGE. Selected reaction solutions were pooled and purified using the QIAquick® PCR Purification Kit. Standard cloning (see above) was performed to digest PCR amplicate and original plasmid backbone and replace the FRT site-free resistance module in pBI1KT7.1 by the new resistance module, yielding plasmid pBI1KiT7.1 (see Table 2.3).

Verification of correct construct composition

Sequences were examined by analytical control REN digest and sequencing after cloning in all cases. REN enzymes used for cloning were, in most cases, also used for control digests. For control digests, a reaction solution of 20 μl containing 200 ng of DNA, 2 μl of 10x buffer and 10x BSA solution (if required) and 0.2 μl of solutions for each enzyme were prepared. The solution was incubated for 1 - 2 h at 37°C. Subsequently, size of DNA molecules was assessed by agarose gel electrophoresis as described above. After all cloning procedures, the inserted module including restriction sites was subjected to sequencing. Periodically, entire empty basic plasmids were subjected to sequencing. Plasmid-containing solutions (20 μl , 40 ng μl^{-1}) and corresponding sequencing primers (20 μl , 10 μM) were prepared. Sequencing was performed by GATC Biotech (Cologne, Germany) or VBC-Biotech Service GmbH (Vienna, Austria).

2.2.2 Genomic Integration Employing Homologous Recombination

Genomic integration of gene expression cassettes was performed using pSIM plasmids (Datta et al., 2008) as previously reported (Sharan et al., 2009) (see Figure 2.2). Details on the molecular mechanisms are reported elsewhere (Murphy, 2012). Linear dsDNA cartridges were generated and integrated into the transposon Tn7's target site *attTn7* (Lichtenstein and Brenner, 1982) in the genome of strains HMS174(DE3) and BL21(DE3).

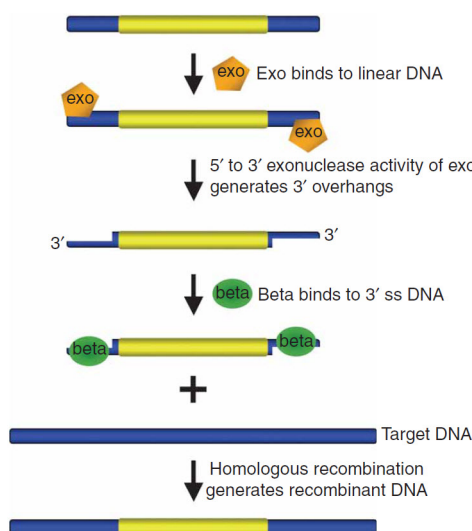


Figure 2.2: Overview of the bacteriophage λ recombination system used for recombineering. Exo has a 5'-to 3'-dsDNA exonuclease activity, which can generate 3'-overhangs on linear DNA. Beta binds the single-stranded DNA (3'-overhangs), promotes ss-annealing and generates recombinant DNA. An additional protein, Gam (not shown here), which prevents RecBCD nuclease from degrading double-stranded linear DNA fragments, is also required for dsDNA recombineering. Figure and description were copied from a previously published protocol (Sharan et al., 2009).

Preparation of the dsDNA integration cassette

Plasmid templates bearing the gene of interest were linearized using REN enzymes *KpnI* and *PciI* in initial experiments. Later in the project *KpnI* and *SphI* were used to increase the size difference of the resulting plasmid fragments. A reaction solution of 50 μ l containing 1 μ g of template plasmid DNA, 0.5 μ l of solution of each enzyme, 5 μ l of 10x buffer and 5 μ l of 10x BSA was prepared. The solution was incubated over night at 37°C. DNA fragments were separated by AGE. The relevant fragment was purified as described in Section 2.2.1. Subsequently, dsDNA integration cassettes containing sequences with 50 bp of homology to the *attTn7* site of the *E. coli* genome on both ends were generated. Linearized and gel purified plasmid from the previous step were used as the PCR

Table 2.6: Strains created by genomic integration used in this work. Strains are ordered by product.

| Product | Producer strain ² | Plasmid origin | Primer pair ³ | Strand ¹ | Resistance ⁵ | Origin |
|----------|------------------------------|------------------------|--|---------------------|-------------------------|--------------|
| GFP | B<kanR-GFP> | pACIB-int-gfp | TN7_HO1_back, TN7_HO2_for | Plus/Plus | kanR, FRT | pre-work |
| | B<GFP> | pACIB-int-gfp | TN7_HO1_back, TN7_HO2_for | Plus/Plus | - | B<kanR-GFP> |
| FabX | B<FabX-kanR>(01) | pBI1KT7.1-fabX | 095_TN7_for, 095_TN7_back | Plus/Minus | kanR | this work |
| | B<FabX-kanR>(02) | pBI1KT7.1-fabX | 095_TN7_for_fwd-int, 095_TN7_back_rev-int | Plus/Plus | kanR | this work |
| FabZ | B<kanR-FabZ> | pACIB-int-fabZ (HC-LC) | TN7_HO1_back, TN7_HO2_for | Plus/Plus | kanR, FRT | this work |
| | B<FabZ> | pACIB-int-fabZ (HC-LC) | TN7_HO1_back, TN7_HO2_for | Plus/Plus | - | B<kanR-FabZ> |
| | B<FabZ-kanR>(01) | pBI1KT7.1-fabZ (HC-LC) | 095_TN7_for, 095_TN7_back | Plus/Minus | kanR | this work |
| | B<FabZ-kanR>(02) | pBI1KT7.1-fabZ (LC-HC) | 095_TN7_for, 095_TN7_back | Plus/Minus | kanR | this work |
| | K<FabZ-kanR> | pBI1KT7.1-fabZ (LC-HC) | 095_TN7_for, 095_TN7_back | n.a. ⁴ | kanR | this work |
| | K<kanR-FabZ> | pACIB-int-fabZ (HC-LC) | TN7_HO1_back, TN7_HO2_for | n.a. ⁴ | kanR, FRT | this work |
| scFv-Z-1 | B<scFv-Z-1>(01) | pBI1KiT7.3-scfv-Z-1 | int_Tn7-rev_mv-fwd, int_Tn7-fwd_mv-rev | Plus/Minus | kanR, FRT | this work |
| | B<scFv-Z-1>(02) | pBI1KiT7.3-scfv-Z-1 | int_Tn7-fwd_mv-fwd, int_Tn7-rev_mv-rev | Plus/Plus | kanR, FRT | this work |

¹BLAST nomenclature; Plus/Plus: forward; Plus/Minus: reverse²Nomenclature according to previous articles (Mairhofer et al., 2013; Striedner et al., 2010). B: BL21(DE3). K: HMS174(DE3).³see Table 5.2 in the confidential supplement⁴Genome not available on BLAST website.⁵FRT: flippase recognition target site

template. Different primer pairs (see Table 2.6) and Phusion[®] High-Fidelity DNA Polymerase (NEB) were used for the PCR reaction. Reaction solutions of 50 μ l containing 1 ng of template, 2.5 μ l of primer solutions (10 μ M), 1 μ l of dNTPs (10 μ M per dNTP), 10 μ l 5x Phusion HF buffer and 0.5 μ l of polymerase were prepared. Annealing temperature gradient PCR was performed as follows. After the initial denaturation step (98°C, 30 sec), 35 cycles of denaturation (98°C, 10 sec), annealing (48 - 72°C (12 step gradient), 15 sec) and elongation (72°C, 30 sec per kb) were conducted prior to the final elongation step (72°C, 10 min). Afterwards, the solution was stored at 10°C (hold step) until further use. PCR success was examined by AGE of 5 μ l of the solutions. PCR reaction solutions corresponding to annealing temperatures with the most intense bands were pooled. In case of visible impurities after AGE analysis of PCR solutions, pools were purified by AGE and subsequent excision of relevant bands. After purification using the QIAquick[®] Gel Extraction Kit (Qiagen), dsDNA cassettes were used for genomic integration. In case of absence of visible impurities, *DpnI* was added to PCR reaction solutions and incubation was performed for 2 h at 37°C. After PCR purification using the QIAquick[®] PCR Purification Kit (Qiagen), PCR products were used for genomic integration.

Genomic integration

Plasmids pSIM5 (pSC101 ori, *clm*^R) or pSIM9 (RK2 ori, *clm*^R) were introduced into electrocompetent *E. coli* BL21(DE3) or HMS174(DE3) target strains by electroporation. 50 μ l of electrocompetent cells (see Section 2.2.3) were thawed on ice. 50 ng of plasmid was added. Cells with DNA were transferred to electroporation cuvettes (0.1 cm gap) and electroporation was performed (1.8 kV, 25 μ F capacitance and 200 Ω resistance).

Immediately afterwards, 950 μ l of S-LB medium was added to cells. Cells were incubated for 1 h at 32°C in a bench-top incubator with 550 rpm agitation. After centrifugation (1 min., 13000 x g) and discarding the supernatant, the suspension was plated on pre-warmed S-LB agar containing 20 μ g ml⁻¹ chloramphenicol. Agar plates were incubated at 30°C until colonies were detected (approximately 48 h). Afterwards, pSIM-bearing cells were either used directly for recombineering or stored as glycerol stocks (see Section 2.2.4) at -80°C until further use.

Single colonies or different volumes of thawed cryo cultures were used to inoculate 5 ml of S-LB medium containing 20 μ g ml⁻¹ chloramphenicol in a 15 ml reaction tube. Incubation was performed over night at 32°C and 250 rpm. 500 μ l of cultures were transferred to 35 ml of sterile S-LB medium with 20 μ g ml⁻¹ chloramphenicol and grown until OD₆₀₀ values of

0.4 to 0.6 were reached. For induction of λ Red gene expression, 17 ml were transferred to a sterile 100 ml unbaffled Erlenmeyer flask and incubated at 42°C for 15 min with manual agitation in a water bath. The remaining 17 ml of suspension were further incubated at 32°C. These cells were treated similarly to induced cells and subsequently functioned as a negative control for assessing colony numbers. Preparation of electrocompetent cells was performed as previously reported (Sharan et al., 2009) (see p. 51). Afterwards, 100 or 300 ng of dsDNA integration cassette (see above) were added to 50 μ l of electrocompetent cells. The mixture was transferred to ice-cold electroporation cuvettes (1 mm gap, 732-1135, VWR, Leuven, Belgium). Electroporation was performed as previously reported (1.75 kV, 25 μ F capacitance and 200 Ω resistance) (Sharan et al., 2009). Immediately after electroporation, 1 ml of S-LB medium was added to electroporation cuvettes. Cells were, then, transferred to 2 ml microcentrifugation tubes and incubated for a minimum of 3 h at 32°C and 450 rpm. Afterwards, cells were sedimented by centrifugation for 30 s at 4000 x g at RT. After aspirating and discarding approximately 900 μ l of supernatant, cells were resuspended in the remaining solution and plated on S-LB agar plates containing 30 μ g ml⁻¹ kanamycin. Plates were incubated at 30 to 32°C until colonies appeared (up to 48 h).

Integration events were verified by colony PCR with colonies as template, *Taq* DNA Polymerase (NEB) and two different primer pairs (see Table 5.2). One primer of each set was homologous to a sequence within the integration cassette, the other was homologous to both sides of the native *E. coli* genome. A small amount of cells was picked up from selected colonies using a pipet tip. After saving colonies by touching a plate with agar containing 30 μ g ml⁻¹ kanamycin, cells attached to the pipet tip were resuspended in 50 μ l of RO-H₂O by aspirating and extruding the water three times. The resulting cell suspension was used as the colony PCR template. Plates were incubated at 30°C until colonies became detectable. Reaction solutions of 50 μ l containing 5 μ l of template solution, 2.5 μ l of primer solutions (10 μ M), 1 μ l of dNTPs (10 μ M per dNTP), 10 μ l of 5x *Taq* Standard Buffer and 0.3 μ l of polymerase (5 U ml⁻¹) were prepared. The PCR reaction was performed as follows. After the initial denaturation/cell lysis step (94°C, 10 min), 30 cycles of denaturation (94°C, 10 sec), annealing (50°C, 15 sec) and elongation (68°C, 1 min per kb) were conducted prior to the final elongation step (68°C, 10 min). Afterwards, the solution was incubated at 10°C (hold step) until further use. PCR reactions with positive and negative controls were examined by AGE. Colonies from the sample preparation step, which were associated with expected band patterns, were subjected to the pSIM plasmid curing procedure and, subsequently, stored at -80°C as cryo cultures.

Induction of pSIM plasmid loss

Colonies with verified integration events were subjected to a pSIM plasmid loss procedure. Plasmid loss was achieved based on a previously reported procedure (Datta et al., 2006) with slight adaptations. Colonies were transferred to 15 ml tubes containing 5 ml of S-LB with 20 $\mu\text{g ml}^{-1}$ chloramphenicol and 30 $\mu\text{g ml}^{-1}$ kanamycin. Cells were grown over night at 32°C and 300 rpm. 15 μl of the suspension were transferred to 15 ml of S-LB medium containing 30 $\mu\text{g ml}^{-1}$ kanamycin. Note that chloramphenicol, resistance to which is mediated by pSIM presence, was omitted in this step. Cells were grown for a minimum of 8 h at 37°C and 300 rpm to trigger temperature-mediated pSIM plasmid loss. The suspension was diluted using sterile 0.9 % (w/v) NaCl. Dilutions were spread on S-LB agar plates containing 30 $\mu\text{g ml}^{-1}$ kanamycin and grown over night at 37°C. Using inoculation loops, colonies were transferred to plates containing S-LB agar with 20 $\mu\text{g ml}^{-1}$ chloramphenicol and, subsequently, to agar containing 30 $\mu\text{g ml}^{-1}$ kanamycin. Colonies with chloramphenicol sensitivity and kanamycin resistance were considered plasmid-free. These colonies were, again, examined for correct genomic integration by colony PCR as described above. In addition to primers for verification of integration, primer pairs for examination of original template plasmid absence were employed (see pBI_absence_fwd and pBI_absence_rev in Table 5.2). Colonies saved during the procedure were used for preparation of cryo cultures (see Section 2.2.4) and stored at -80°C until further use.

Excision of the kanamycin resistance gene after genomic integration

For selected strains, the integrated resistance marker cassettes flanked by FRT sites were removed by a FLP-FRT-recombination system (Gene Bridges, Heidelberg, Germany) according to a protocol slightly adapted from the manufacturer's protocol (GeneBridges, 2013). 100 μl of cryo stocks of cells with genome-integrated expression cassette were used to inoculate 10 ml of S-LB medium containing 30 $\mu\text{g ml}^{-1}$ kanamycin in a 100 ml unbaffled shake flask. Cells were incubated over night at 37°C and 250 rpm. 40 ml of S-LB medium with 30 $\mu\text{g ml}^{-1}$ kanamycin in a 100 ml unbaffled shake flask were inoculated with 400 μl of the suspension. Cells were grown at 37°C and 250 rpm until an OD_{600} value of 0.4 to 0.6 was reached. Afterwards, cells were subjected to the procedure for preparation of electrocompetent cells described on page 51. 50 μl of the resulting cell suspension were transferred to pre-cooled 1.5 ml reaction tubes. After addition of 25 ng of plasmid 707-FLPe (GeneBridges, Heidelberg, Germany) (GeneBridges, 2013), the mixture was transferred to pre-cooled electroporation cuvettes and electroporation was performed (25 μF , 2.5 kV and 200 Ω). After addition of 1 ml S-LB medium to the

electroporation cuvettes and transfer of the suspension to 2 ml reaction tubes, cells were incubated at 30°C under gentle agitation for a minimum of 3 h. Afterwards, cells were sedimented by centrifugation (4000 x g, 30 sec, RT). Subsequently, cells were spread on S-LB agar plates containing 3 µg l⁻¹ tetracycline and incubated at 30°C until colonies became visible (approximately 48 h). Single colonies were transferred to 2 ml reaction tubes containing 1 ml of AB-free S-LB medium and incubated at 30°C for 3 h under agitation before switching the temperature to 37°C and growing cells over night. Afterwards, the suspension was diluted and spread on AB-free S-LB agar plates. Plates were incubated over night at 37°C. Colonies were replica plated on S-LB agar plates containing 30 µg l⁻¹ kanamycin, 3 µg l⁻¹ tetracycline and no antibiotics (in this order). Colonies with the correct phenotype (tet- and kan-sensitive) were examined again for correct integration of the GoI expression cassette by colony PCR as described above. Furthermore, kanamycin resistance gene removal was examined with primer pair pACIB-int_kan-removal_fw/pACIB-int_kan-removal_rev. Colonies associated with marker-free cells were used for preparation of cryo cultures (see Section 2.2.4).

2.2.3 Creation of Plasmid-Bearing Cells

Strains that were created for plasmid-based expression of target genes are presented in Table 2.7. Production strains transformed with plasmids for folding modulator co-synthesis are presented in the respective results sections.

Table 2.7: Plasmid-bearing strains for GoI expression. In this table, strains that were used for production experiments are presented and ordered by the protein of interest.

| Product | Plasmid-free strain | Plasmid for transformation | Strain identifier |
|----------|---------------------|---|--------------------------------------|
| FabX | BL21 | pBI1Kara.1-fabX | B(FabX) |
| FabZ | BL21(DE3) | pBI1KT7.1 pBI1KT7.1-fabZ (HC-LC) pBI1KT7.1-fabZ (LC-HC) | B(pBI) B(FabZ)(01) B(FabZ)(02) |
| GFP | BL21(DE3) | pET30a(+)-gfp | B(GFP) |
| scFv-Z-1 | BL21(DE3) | pBI1KiT7.3-scFv-Z-1 | B(scFv-Z-1) |

Preparation of chemically competent cells

5 ml of S-LB medium (selective, if possible) were inoculated with 100 μ l of thawed cryo stocks in a 15 ml reaction tube. Cells were incubated over night at 37°C and 250 rpm. 2 ml of the over night culture were used to inoculate 200 ml of S-LB medium (selective, if possible) in a 1 l un baffled Erlenmeyer shake flask. Cells were grown at 37°C and 250 rpm until an OD₅₅₀ of 0.6 to 0.8 was reached. The suspension was aliquoted into four sterile 50 ml reaction tubes and cooled on ice for 10 minutes. Subsequently, cells were sedimented by centrifugation (3220 x g, 10 min, 4°C) and the supernatant was discarded. Cells were resuspended in 10 ml of ice-cold 100 mM CaCl₂ solution containing 10 % (v/v) glycerol. After another centrifugation step (3220 x g, 10 min, 4°C), the sedimented cells were resuspended in 1 ml of ice-cold 100 mM CaCl₂ solution containing 10 % (v/v) glycerol. Aliquots of 110 μ l of resuspended competent cells were shock frozen using dry ice and stored at -80°C until further use.

Preparation of electrocompetent cells

Vessels with cells from growth steps were cooled on ice for 15 min and, periodically, agitated manually. Aliquots of 20 ml were, then, transferred to 50 ml reaction tubes and centrifuged (7 min, 2°C, 4600 x g). After discarding the supernatant, the cell pellet was resuspended with 1 ml of ice-cold RO-H₂O. After addition of another 30 ml of ice-cold RO-H₂O, centrifugation was carried out (4600 x g, 7 min, 2°C). After discarding the

supernatant, the cell pellet was resuspended with 1 ml of ice-cold RO-H₂O, transferred to a 2 ml centrifugation tube and centrifugation was carried out (10000 x g, 30 sec, 4°C). The supernatant was discarded and the cell pellet was resuspended in 1 ml of ice-cold sterile RO-H₂O. A final centrifugation step was performed (30 sec, 4°C, 10000 x g) and the supernatant was discarded. The cell pellet was resuspended in 200 µl of ice-cold sterile RO-H₂O and used for electroporation

Transformation of competent cells

Unless stated otherwise, plasmid-containing strains were created by heat-shock transformation using chemically competent *E. coli* cells. Competent cells of different strains were thawed on ice for 15 min. 10 ng of the respective plasmid were added to the cells. Afterwards, the mixture was incubated on ice for 10 min, at 42°C for 45 sec and on ice for another 2 min. 900 µl of S-LB medium were added. Afterwards, cells were incubated at 37°C and 450 to 600 rpm for 45 to 90 min. To obtain appropriate colony numbers, 100 µl of a ten-fold dilution and the undiluted suspension were plated on selective S-LB agar plates. Subsequently, remaining cells were sedimented by centrifugation (3220 x g, 30 sec, RT) and 700 µl of the supernatant were discarded. Sedimented cells were resuspended in the residual volume. The resulting suspension was spread on selective S-LB plates containing an appropriate antibiotic. Cells were incubated on plates in inverted position at 37°C over night.

2.2.4 Preparation of Cryo Cultures

After strain creation by genomic integration or transformation, colonies were transferred to 50 ml reaction tubes containing 15 ml of selective (if possible) S-LB medium using sterile inoculation loops. Cells were grown at 37°C and 250 rpm until OD₅₅₀ values of 0.2 to 0.5 were reached. After addition of 2.7 ml of 15 % (v/v) glycerol and thorough mixing, the suspension was divided into aliquots of approximately 1 ml and stored at -80°C until further use.

2.3 Promoter Strength Analysis

2.3.1 Shake Flask Experiments

Shake flask experiments during promoter strength analysis were similar to the ones in production experiments (see Section 2.4.1). Briefly, they consisted of preculture (37°C, different vessels) and main culture in 1 l shake flasks. After reduction of the incubation temperature from 37 to 25°C in the main culture phase, recombinant protein production was induced by addition of 1 mM IPTG.

In experiments for testing T7 promoters under the influence of regulatory elements *lacI* and *lacO* (see Section 3.4.2), genome-integrated FabZ production strain B<FabZ-kanR>(02) (see Table 2.6) was transformed with pBI4iST7(n).2-gfp.1 plasmids bearing one of several T7 promoters. The medium used in pre- and main culture was supplemented with 50 µg l⁻¹ streptomycin. Cell samples were prepared by centrifugation immediately before and 4 h after IPTG addition (see Section 2.4.1).

In experiments for testing T7 promoters without the influence of regulatory elements, basic strain BL21(DE3) was transformed with pBI4iST7(n).7-gfp.1 plasmids. Cryo cultures were prepared and used for inoculation of preculture medium. The medium used in pre- and main culture was supplemented with 50 µg l⁻¹ streptomycin. Cell samples were prepared immediately before and 1, 4, 19 and 24 h after addition of 1 mM of IPTG.

In experiments for testing constitutive promoters, basic strain BL21(DE3) was transformed with pBI4iSC(n).3-gfp.1 plasmids bearing different constitutive promoter modules. Cryo cultures were prepared and used for inoculation of preculture medium. The medium used in pre- and main culture was supplemented with 50 µg l⁻¹ streptomycin. Cell samples were prepared immediately before and at different time points (up to 24 h) after addition 1 mM of IPTG.

2.3.2 Fluorescence Measurement

For GFP synthesis rate analysis, RFU (relative fluorescence units) of samples of interest were determined by emission measurement (wavelength: 512 nm) after excitation of samples (wavelength: 475 nm). The fluorescence measurement procedure is depicted in Figure 2.3. 5/OD cell samples obtained from promoter test shake flask experiments (see Section 2.3.1) were thawed at RT for at least 15 min. Subsequently, cells were resuspended using 5 ml of “OD Buffer” (20.7 g l⁻¹ Na₂HPO₄ dodecahydrate, 5.7 g l⁻¹ KH₂PO₄, 11.6 g l⁻¹ NaCl). Note that the resulting suspension was characterized by the same amount of cells for all samples. 200 µl were transferred to wells of a 96-well flat bottom transparent UV-plate (Costar[®], Sigma-Aldrich). During examination of T7 promoter with regulatory elements, duplicate measurements in neighboring columns were performed. However, a significant influence of the row position on the measurement plate was observed later (Huber, 2015). Thus, in the part of the project, which dealt with constitutive promoters, resuspended cells subjected to fluorescence measurement were applied in all eight positions of an entire plate column. Fluorescence of samples was measured using a Safire^{2™} fluorescence microplate reader (Tecan, Life Technologies, Carlsbad, CA, USA) (Excitation wavelength: 475 nm, Emission wavelength: 512 nm, Excitation and emission bandwidth: 20.0 nm, Gain: 200, Number of reads: 10, Flashmode: High sensitivity, Integration time: 500 µs, Lag time: 60 µs, Z-Position (Manual): 12400 µm, Shake duration (Orbital Low): 30 s, Shake settle time: 10 s). The measured fluorescence was displayed in relative fluorescence units (RFU) using the software Magellan V7.1 (Tecan).

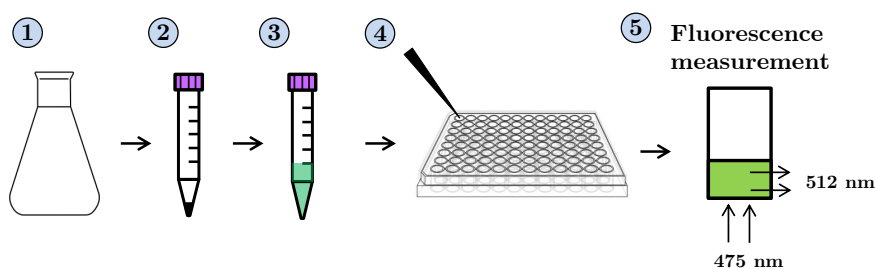


Figure 2.3: Schematic overview of the fluorescence measurement procedure. Shake flask experiments (1) were performed as described in Section 2.3.1. Cell samples were prepared by centrifugation (2) and stored. Cells were thawed and resuspended (3). The suspension was transferred to a 96-well plate (4) and the fluorescence of the samples was measured (5). Resulting fluorescence values were used to calculate the promoter strength of the tested module. Figure and description were copied from a master thesis supervised by the author of this work (Huber, 2015) and adapted.

In a pre-experiment, resuspended cell samples associated with tests of ten different promoters with strong and no fluorescence (depending on the sampling time point), were mixed in six different ratios. By fluorescence measurement, a linear range ($R^2 \geq 0.97$)

between 0 and approximately 45000 RFU was determined for all ten promoters. As the greatest share of samples from later project parts resulted in maximum values of below 45000 RFU, promoter strength calculations are, mostly, based on raw data from the linear range of the applied method.

2.3.3 Data Analysis

Strength of T7 promoter variants under influence of regulatory elements *lacO* and *lacI* was calculated based on the acquired raw data according to Equation 2.1.

$$\text{Relative Promoter Strength}_n = \frac{RFU_{n,T4} - RFU_{n,T0}}{RFU_{native,T4} - RFU_{native,T0}} \quad (2.1)$$

For both the examined (n) and the native reference promoter (native), RFU values obtained from samples prepared prior to induction were subtracted from the RFU values obtained after 4 h of GFP production. The difference calculated for the examined was divided by the one calculated for the reference promoter, yielding the relative promoter strength of the examined promoter. Standard deviations of the GFP measurement at each time point (*i.e.* T0 and T4) were calculated based on technical replicates. These values were used to calculate the standard deviation of the determined promoter strengths using the law of propagation of uncertainty.

For calculation of strength of constitutive and T7 promoters without regulatory elements, mean value and standard deviation of eightfold measurements (an entire column of a 96 well plate) for each sample were calculated. Fluorescence of non-GFP-producing strains was considered autofluorescence and subtracted from fluorescence of all GFP-producing strains at each time point. Resulting GFP fluorescence values were used for relative promoter strength calculation. Depending on the experimental setup, strength of a selected promoter was set to 100 %. Relative promoter module strength was determined by dividing the GFP fluorescence of a respective strain by the value determined for the reference. Details on the procedure, *i.e.* sampling time points, choice of reference and calculation of relative promoter strength, are presented in the respective results sections.

2.4 Production Experiments

2.4.1 Shake Flask Experiments

A 10-fold concentrated stock solution (without “Sterile Addition” components) of the basic shake flask medium was prepared, sterile filtered and stored at 4°C for up to two months. Likewise, “Shake Flask Sterile Additions” were prepared and also stored for up to two months. Prior to use, stock solution of the basic medium was diluted 10-fold, transferred to shake flasks and heat-sterilized. Immediately prior to shake flask experiments, sterile additions were added. Due to confidentiality reasons, media composition is presented in the confidential supplement (see Section 5.5.3). The majority of strain screening experiments were performed according to the following procedure. Deviations are described in the respective results section.

In all shake flask experiments, cells from cryo cultures (thawed for at least 15 min at RT) or single colonies were used to inoculate T7 Preculture Medium selective for the plasmid or genome-integrated cassette borne by the cell (if applicable). In initial experiments, 500 ml unbaffled Erlenmeyer Flasks, 100 ml of medium and 100 µl of cryo cultures (or single colonies) were used in precultures. As the procedure itself was subject to optimization, 30 ml of medium, 50 ml reaction tubes and 50 µl of cryo cultures (or single colonies) were used in later stages of the project. Precultures were incubated at 37°C and 250 (initial experiments) to 300 (later experiments) rpm for a minimum of 9 h.

Cells corresponding to 40 ml of a suspension with an OD₅₅₀ value of 1 were transferred to a sterile 50 ml reaction tube and sedimented by centrifugation (3200 x g, 10 min., 22°C). Subsequently, the supernatant was discarded and sedimented cells were resuspended using 5 ml of a 0.9 % (w/v) NaCl solution. Resuspended cells were transferred to 200 ml of “T7 Shake Flask Medium, Version 3” (5 g l⁻¹ glycerol, see Table 5.8, Confidential Supplement) supplemented with 30 µg ml⁻¹ kanamycin in unbaffled 1 l Erlenmeyer flasks. This resulted in an initial OD₅₅₀ of 0.2. Incubation was carried out at 37°C and 300 rpm until an OD₅₅₀ value of at least 0.5 was reached. Subsequently, shake flasks were transferred to a shaker incubator set to 25°C and incubated at 300 rpm for at least 15 min. After cells reached the induction criterium (OD₅₅₀ of 1.0 ± 0.3), 1 mM of the inducer IPTG was added. Incubation was continued under the same conditions for 12 h. OD₅₅₀ values were determined at the end of the experiment (EoF). Samples for protein analytics were prepared immediately prior to addition of IPTG (T0) and at EoF. For preparation of cell samples 5 ml (T0) or a volume corresponding to 5 ml of a suspension with OD₅₅₀

of 1 (EoF) were transferred to 2 or 15 ml reaction tubes, depending on the calculated volume. Samples were centrifuged for 10 min. at 4°C (2 ml tubes: 16000 x g, 15 ml tubes: 3220 x g). The centrifugation supernatant was decanted, filtrated (pore diameter 0.2 µm, Merck Millipore, Darmstadt, Germany) and used as supernatant samples. All samples were stored at -20°C until analysis was performed.

2.4.2 Fed-Batch Fermentations for FabZ and scFv-Z-1 Production

Basic procedure

Computer-controlled stirred-tank 6.9 l BIOSTAT[®] Cplus bioreactors (Sartorius Stedium Biotech, Göttingen, Germany) were used for fed-batch fermentations. The temperature was maintained at 37°C during the “growth phase” of all fermentation process. In certain cases, the temperature was reduced prior to induction of recombinant protein production (see below). The pH was maintained at 6.8 ± 0.2 by addition of 25 % ammonia solution and 3 M phosphoric acid as required. The dissolved oxygen level was stabilized to range at or above 20 %. In the major share of fed-batch fermentations, DO stabilization was achieved by stirrer speed adjustment (step 1, 400 to 1200 rpm) and aeration variation (step 2, constant total aeration rate: 5 l min^{-1} , air flow: $2.5 - 5 \text{ l min}^{-1}$, oxygen flow: $0 - 2.5 \text{ l min}^{-1}$). The pressure was stabilized at 1000 mbar. Foaming was suppressed by addition of 1 ml l^{-1} PPG2000 (Wacker-Chemie, Munich, Germany) to batch medium prior to start of the experiment and further addition throughout the fermentation procedure as required.

For preculture incubation, 300 ml of selective T7 Preculture Medium (see Section 5.5.2) in a 1 l unbaffled shake flask were inoculated with 100 μl of a glycerol culture. These were previously thawed ($-80^\circ\text{C} \rightarrow \text{RT}$, 30 min). Incubation was carried out at 33.5°C and 250 rpm until an $\text{OD}_{550} 2 \pm 1$ was reached. For batch phase initiation (“inoculation”), cell numbers equivalent to 35 ml of a solution with OD_{550} value of 2.0 were transferred to the bioreactor containing 2500 ml of Batch Medium under monoseptic conditions using a syringe. Unless indicated otherwise, non-selective conditions were employed after start of batch phase. Glucose addition (“feeding”) with Feed 500 or 600 Solution (see confidential supplement, Section 5.5.4) was initiated after occurrence of a sharp increase in the pO_2 signal (“DO peak”), which indicated glucose depletion.

Feeding strategy and handling of glucose accumulation

The “feeding phase” following batch phase termination consisted of a maximum of four phases (1: Exponential Feed, 2: Constant Feed Rate, 3: Feed Rate Reduction (linear, optional), 4: Constant Feed Rate (optional)). Strategies for feeding and handling of glucose accumulation are presented below. Fermentations were carried out under glucose-limited conditions ($\leq 0.5 \text{ g l}^{-1}$). Initially, fermentation experiments were carried out using Feed 500 Solution (500 g l^{-1} glucose, see Section 5.5.4). After start of the second year of

Table 2.8: Feeding strategies applied in fed-batch fermentations. Strategies relied on different feed solutions and exponential and post-exponential feeding phase parameters. In the “exponential phase” column, durations, initial and final feeding rate as well as calculated growth rate are presented. Parameters (durations and progress of feeding rates) of the phase after the exponential phase are also presented (“Post-exponential phase”).

| Strategy | Solution ² | Exponential phase | Post-exponential phase |
|----------------|-----------------------|---|---|
| 1 | 600 | 12 h, 13.00 → 126.7 g h ⁻¹ $\mu = 0.190 \text{ h}^{-1}$ | DO-stat |
| 2 | 600 | 12 h, 15.29 → 127.91 g h ⁻¹ $\mu = 0.177 \text{ h}^{-1}$ | 1 h constant, 127.91 g h ⁻¹ 0.5 h ramp, 127.91 → 35.00 g h ⁻¹ constant, 35.00 g h ⁻¹ |
| 3 | 600 | 12 h, 15.29 → 127.91 g h ⁻¹ $\mu = 0.177 \text{ h}^{-1}$ | constant, 127.91 g h ⁻¹ |
| 4 ¹ | 600 | 9.5 h, 15.12 → 162.55 g h ⁻¹ $\mu = 0.250 \text{ h}^{-1}$ | 2 h ramp, 162.55 → 100.00 g h ⁻¹ constant, 100.00 g h ⁻¹ |
| 5 | 600 | 12 h, 13.00 → 126.70 g h ⁻¹ $\mu = 0.190 \text{ h}^{-1}$ | constant, 126.70 g h ⁻¹ |
| 6 | 500 | 12 h, 14.00 → 154.70 g h ⁻¹ $\mu = 0.200 \text{ h}^{-1}$ | constant, 154.70 g h ⁻¹ |

¹Strategy applied in FabX production experiments.

²See Section 5.5.4 for composition of Feed 500/600 Solution.

the project, a feed solution with a higher nutrient concentration (600 g l⁻¹ glucose, see Section 5.5.4 in the Confidential Supplement) was used. Six different feeding strategies were applied during fermentation experiments (see Table 2.8).

The following rules were applied in case of glucose accumulation events ($> 0.5 \text{ g l}^{-1}$) during feeding phases. If moderate accumulation ($> 0.5; \leq 2.0 \text{ g l}^{-1}$) was observed, the feeding rate was reduced by 50 g h⁻¹ and glucose concentration was determined after another 30 min. When a value of below 0.5 g l⁻¹ glucose was measured after 30 min of cultivation, the initial feed rate was adjusted. If the glucose concentration was still above 0.5 g l⁻¹, the feed rate was reduced by another 50 g h⁻¹ and a measurement after another 30 min was conducted. In case of a greatly increased glucose concentrations ($> 2 \text{ g l}^{-1}$), glucose feeding was stopped. After occurrence of a sharp increase in pO₂ signal (“DO peak”), a feeding rate of 50 g h⁻¹ was set and the glucose concentration was determined at the next regular sampling time point. In case of glucose-limited conditions ($< 0.5 \text{ g l}^{-1}$), the feeding rate was increased by 20 g h⁻¹ steps. In case of repeated glucose accumulation, glucose feeding was stopped again.

In case of DO-stat approaches during fermentations with 16 °C production phase temperature, DO was stabilized by glucose feed rate adjustment. To enable maintenance of a DO value of approximately 20 %, total aeration rate, agitation and pressure had to be

adjusted to 3 l min^{-1} , 400 rpm and 700 mbar, respectively. Under the applied conditions, a linear development of added feed solution was observed. Gradient calculation in the linear range yielded an approximate feed rate of 35 g h^{-1} . This feed rate was set for the following fermentation experiment with a 16°C production phase.

Induction strategies

Different IPTG-based induction strategies were employed in fed-batch fermentation experiments. These are presented in Table 2.9. Note that induction strategies resulted in production phases of different duration. Details and deviations are described in the respective results sections.

Table 2.9: Induction strategies employed in fed-batch fermentations. Four different induction strategies were employed. These relied on types of induction, amounts of inducer and induction time points. Values in the T0 column indicate the duration after start of feeding before recombinant protein production was initiated.

| Strategy | Type | T0 | IPTG |
|----------------|-------|--------|------------------------|
| 1 | bolus | 14.0 h | 894 mg |
| 2 | bolus | 14.0 h | 298 mg |
| 3 ¹ | feed | 11.5 h | 50 mg h^{-1} |
| 4 | bolus | 14.0 h | 596 mg |

¹Strategy applied in FabX production experiments.

Temperature strategy

During the growth phase, *i.e.* prior to induction of recombinant protein production, a temperature of 37°C was employed. In most cases, the production phase temperature was set to 37°C , as well. When reduced production phase temperatures were employed, temperature shifts to 16, 25, 28 and 30°C , respectively, were performed shortly prior to the production phase. In these cases, different temperature ramps (**1**: ramp of 0.75 h duration, initiated 13 h after start of feeding; **2**: 1 h duration, 13 h after feed start; **3**: 2 h duration, 9.5 h after feed start) were employed.

At-line analytics

OD₅₅₀, dry cell weight (DCW) and glucose concentration were determined periodically. Samples were taken from the fermenter and stirred at RT during the sampling procedure.

For determination of optical density, the cell suspension was diluted with “OD Buffer” (20.7 g l⁻¹ Na₂HPO₄ dodecahydrate, 5.7 g l⁻¹ KH₂PO₄, 11.6 g l⁻¹ NaCl) to suit the linear range of the GENESYS™ 10S UV-Vis Spectrophotometer (Thermo Fisher, Waltham, MA, USA). The extinction of this dilution before and after filtering (0.2 µm pore size, blank) was determined. Under consideration of blank and dilution, the OD₅₅₀ of the fermentation cell suspension was calculated. For DCW determination, 10 ml of cell suspension were transferred to a 50 ml reaction tube. Cells were sedimented by centrifugation (15 min, 9200 x g, 4°C), a share of the supernatant was used for supernatant sample preparation for off-line analytics (see below) and the remaining volume discarded. Cells were resuspended with 10 ml of RO-H₂O and, again, centrifuged (15 min, 9200 x g, 4°C). After discarding the supernatant, cells were resuspended with 5 to 10 ml of RO-H₂O and transferred to a weighing bowl, which was previously used for taring. The suspension was dried at 125°C and DCW was determined by the Moisture Analyzer HG63 (Mettler-Toledo, Giessen, Germany). At the end of the fermentation (EoF), the wet cell weight (WCW) was determined. 35 ml of cell suspension were transferred to a 50 ml reaction tube and centrifugation was conducted (15 min, 15100 x g, 4°C). The weight of the empty reaction tube and the weight after discarding the supernatant after centrifugation were used to calculate the WCW. For glucose concentration, the cell suspension was filtrated (0.2 µm pore diameter). Glucose concentration was determined with a Biochemistry Analyzer (YSI Inc., Yellow Springs, OH, USA).

Sampling for off-line analysis

To obtain cell pellets for subsequent offline analysis, a fermentation suspension volume equivalent to 10 ml of a suspension with an OD₅₅₀ of 1.0 (“10/OD samples”) was transferred to 0.9 % (w/v) NaCl solution in several 2 ml reaction tubes. After centrifugation (13200 x g, 4°C, 10 min.), the supernatant was discarded and samples were stored at -20°C until further use. Centrifugation supernatant from the first centrifugation step of the DCW value determination procedure (see above) was used to obtain supernatant samples. Several aliquots of approximately 2 ml were filtrated (pore diameter: 0.2 µm) and stored at -20°C until further use. Samples for culturable cell number and plasmid stability determination were collected in sterile syringes directly from the fermenter. After addition of 10 % (v/v) sterile glycerol, samples were mixed thoroughly and stored at -20°C until further use.

Titer calculations

The calculation of total soluble product concentration was performed according to Equation 2.2. It takes into account the reduced share of supernatant with increasing cell density. Specifically, the determined product concentration in the supernatant $P_{sup,t}$ was multiplied by the estimated fraction of supernatant from the total suspension. The WCW value determined for a certain cell density (DCW_{EoF}) at the end of a fermentation (WCW_{EoF}) was recalculated for all time points under consideration of the cell density at that time (DCW_t). Addition to the intracellular volumetric concentration ($P_{IC,t}$) yielded the total soluble product concentration ($P_{total,t}$)

$$P_{total,t} = P_{IC,t} + P_{sup,t} \left(1 - WCW_{EoF} \frac{DCW_t}{DCW_{EoF}}\right) \quad (2.2)$$

Product content of cells was examined according to Equation 2.3. Different values were inserted as “Product per volume” (*i.e.* intracellular soluble or total product titer) or “Cells per volume” (*i.e.* OD_{550} or DCW values) into the equation. Details are presented in the respective results section.

$$Y_{P/X} = \frac{\text{Product per volume}}{\text{Cells per volume}} \quad (2.3)$$

For calculation of specific product formation rates, an increase in the concentration was divided by the product of the mean cell density and the time difference (see Equation 2.4 (Striedner et al., 2010)).

$$q_P = \frac{conc_{t2} - conc_{t1}}{\frac{DCW_{t1} + DCW_{t2}}{2} \times (t_2 - t_1)} \quad (2.4)$$

2.4.3 Fed-Batch Fermentations for Production of FabX

Note that, due to confidentiality reasons, the experimental set up of FabX production experiments in the fed-batch fermentation format cannot be disclosed in detail.

For production of FabX with arabinose-driven plasmid-bearing expression systems, a fed-batch fermentation process was developed and modified for improved soluble product yield. A defined salt medium was employed. The feeding was performed using the Feed 600 Solution (see Section 5.5.4). For induction of recombinant gene expression, a second feed, which was supplemented with the inducer arabinose, was applied. The end of the fermentation batch phase was indicated by a sharp increase in dissolved oxygen (DO) caused by glucose exhaustion. This DO peak triggered the feed start. The feeding started with an exponential feeding rate, followed by a negative linear feed ramp (accompanied by a decrease in temperature) and ended with a constant feeding of glucose/inducer feed. FabX location was analyzed thoroughly and the vast majority of the properly folded target protein was found within the cell-free culture broth at the end of the cultivation process. Thus, two product related volumetric yields were determined. The total amount of Fab chains within the cells was analyzed by SDS PAGE. The amount of intact, soluble product within the culture supernatant was determined by both ELISA and 2D-HPLC, which constitute two independent, product-specific methods.

For FabX synthesis based on BL21(DE3) strains bearing a genomically integrated target gene, a slightly different production process was established. The batch medium as well as the glucose feeding solution was the same that was used for plasmid-bearing expression systems. For the induction of recombinant protein production, IPTG was added to a second glucose-containing feed. The end of the batch phase was also marked by the sharp DO increase due to carbon source limitation. A feeding strategy similar to the one described above was applied. In the first stage, the feeding was conducted in exponential mode, followed by a negative linear ramp, accompanied by a temperature drop. Finally, a constant feeding of the inducer-containing feed solution was performed. However, the inducer feeding was performed over a three-time prolonged period of time. Since the distribution between soluble/total and intra-/extracellular product was similar in comparison to experiments described above, the same types of product analyses were conducted.

To induce increased product titers in the supernatant at the end of the process, a hold time with increased pH value was conducted after the fermentation process. However, details cannot be disclosed due to confidentiality reasons.

2.5 Examination of Gene Expression Cassette Stability

During analysis of the expression cassette stability, samples from production experiments were diluted with sterile 0.9 % (w/v) NaCl solution. Dilutions from samples from FabZ fed-batch fermentation production experiments were prepared as triplicates. 100 μl of different dilutions were plated on non-selective S-LB agar. Cells were grown over night at 37°C in inverted position of the plates. Colony numbers of plates with less than 1000 colonies were used to approximate the culturable cell count in the suspension (in colony-forming units per ml). Plates were stored at 4 to 8°C for a maximum of two weeks.

Different numbers of colonies (196 for FabZ production experiments in fed-batch format, 300 for FabX production experiments in fed-batch format and 47 for shake flask experiments) were subjected to replica plating (see Figure 2.4). Specifically, colonies were transferred to S-LB agar with the respective antibiotic concentration and, subsequently, to agar without antibiotics using inoculation loops. Selective S-LB agar in the first step of replica plating contained 30 $\mu\text{g l}^{-1}$ kanamycin, 50 $\mu\text{g l}^{-1}$ kanamycin and 50 $\mu\text{g l}^{-1}$ streptomycin in case of samples from plasmid-free expression in fed-batch fermentations, plasmid-based expression in fed-batch fermentations and co-expression experiments in shake flasks, respectively. Plates were incubated at 37°C over night. The number of colonies on the first plate(s) was divided by the number of colonies on the second (non-selective) plate. This number was considered as the ratio of expression cassette-bearing cells.

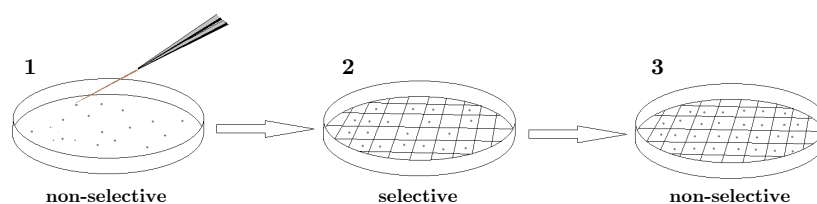


Figure 2.4: Illustration of the replica plating step during the expression cassette stability analysis procedure. 100 μl of an appropriate dilution of cell suspension was plated on non-selective agar plates (**1**). Colonies were transferred to a selective plate (**2**) and, subsequently, to a plate containing no antibiotic (**3**). Comparison of colony numbers of plates of steps **2** and **3** were used to calculate the expression cassette stability. The figure was adapted from a figure created and used in a master thesis supervised by the author of this work (Schuller, 2015).

2.6 Protein Analytics

During protein analysis, target products were analyzed with various methods (see Table 2.10), depending on target product and aim of the experiments.

Table 2.10: Overview of product analysis approaches depending on the target product. Different analysis methods were employed for analysis of products in different fractions.

| Product | Intracellular | | | Extracellular | | |
|----------|---------------------------------------|------------------------------|----------|-------------------------------------|-----------|----------|
| | soluble | insoluble | total | soluble | insoluble | total |
| GFP | SDS PAGE Fluorescence ² | SDS PAGE | SDS PAGE | n.a. | n.a. | n.a. |
| FabX | ELISA Western Blot | SDS PAGE ¹ NTS | SDS PAGE | ELISA Western Blot 2D-HPLC | n.a. | SDS PAGE |
| FabZ | ELISA Western Blot | SDS PAGE ¹ NTS | SDS PAGE | ELISA | n.a. | n.a. |
| scFv-Z-1 | Western Blot | SDS PAGE NTS | SDS PAGE | n.a. | n.a. | n.a. |

¹SDS PAGE performed for N-terminal sequencing only.

²Promoter test procedure (see Section 2.3).

n.a.: no analysis; NTS: N-terminal sequencing.

Note that a high ratio of soluble product does not necessarily represent high protein activity. For instance, incomplete folding could result in a stable soluble conformation, whereas the active site might not mediate activity (Gonzalez-Montalban et al., 2007; Martinez-Alonso et al., 2008).

However, due to several considerations, soluble product was chosen as the target value for examination of optimization efforts. Purification of FabX resulted in an overall yield of approximately 50 %. Purified product was examined concerning its activity and solubility. Within these examinations, no striking differences in the amounts of both were observed (data not shown). Potentially, soluble but inactive FabX, which was present after fermentation, was removed by downstream activities. However, the share seems to be insignificant, as, otherwise, downstream-related target product reduction would have been more pronounced than the plausible observed 50 %. Furthermore, during extensive upstream process development within the project associated with FabX (data not shown), efforts for soluble product improvement yielded a robust process with sufficient active product yield (after purification). Thus, the strategy of using soluble target product as the main target value for optimization appears promising for host engineering approaches. As FabX and FabZ are identical in approximately 80 % of their amino acid

sequence, considerations for FabX were assumed to apply for FabZ.

Furthermore, the ELISA method used for quantification of soluble model proteins FabX and FabZ was shown to detect heterodimers but not monomers or homodimers in a separate BI RCV project (data not shown). Under consideration of the other arguments, FabX and FabZ are likely to be active if they are soluble and heterodimeric.

The Protein L-based method for quantification of soluble scFv-Z-1 was shown to detect scFv product with intact disulfide bonds (see Section 2.6.5). As i) V_L and V_H of this monomeric antibody fragment are, *per se*, present, ii) disulfide bonds in both domains are presumably intact (due to successful detection by Protein L) and iii) the product is soluble, activity seems probable.

Establishing an affinity-based immuno-assay would have been possible. However, as genetic strategies were the main focus of this work and resources were considered thoroughly, initial ideas were not pursued. This study was about the direct comparison of engineered hosts rather than absolute values. Discussion and conclusions relied on significant improvements and overall trends, in parts, across multiple products. Significant soluble product titer improvements were expected to correspond, at least to a certain extent, to increased activity. In conclusion, the improvements in soluble recombinant protein amount determined by ELISA, 2D-HPLC or Western Blot was assumed to also represent an improvement in the amount of active product.

2.6.1 FabZ Standard Preparation

After the first third of the project, the first batch of soluble FabZ reference molecules for analytical procedures was consumed. New material in sufficient quantity was produced using plasmid-free FabZ producer B<FabZ-kanR>(02) (see Table 2.7). The strain was employed in fermentation process described in Section 2.4.2 with a production phase temperature of 37°C. Feeding strategy 6 (see Table 2.8) and induction strategy 4 (see Table 2.9) were applied during the process. After the fermentation process, cells were sedimented by centrifugation (45 min, 4°C, 11344 x g). The supernatant was used for FabZ purification. Sedimented cells were discarded.

Binding of FabZ to MabSelect beads (GE Healthcare Life Sciences) and effective elution by pH down-shift was tested prior to actual purification in small scale experiments using supernatant from a fermentation, which was performed under the same parameters that were used for the later production fermentation. Purification of FabZ from the supernatant was performed by the BI RCV Process Science Downstream department. Comprehensive details cannot be disclosed due to confidentiality reasons. After filtration (0.2 µm pore diameter), the solution was loaded on a MabSelect Protein A column using the same resin, that was also used for pre-experiments. After a wash step, FabZ product was eluted employing a pH gradient (high → low). Fractions were collected and analyzed for product and impurities by SDS PAGE. Selected elution fractions were pooled. Prior to cation exchange chromatography (CEX), conductivity and pH of the solution were adjusted by dialysis and dilution with appropriate buffers. Due to high resulting volume, the solution was aliquoted and CEX purification was performed separately for all aliquots. After the loading and wash step, target product was eluted by a pH gradient. Fractions were collected and analyzed for product and impurities by SDS PAGE. Selected elution fractions were pooled. After concentration of target product, the buffer was exchanged by storage buffer (50 mM NaH₂PO₄, 50 mM NaCl, pH 6.0 (adjusted with NaOH)) by dialysis. After a final filtration step, the protein concentration was determined by four-fold measurement of absorption at 280 nm wavelength using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and dividing the resulting values by 1.47 mg ml⁻¹ (A₂₈₀ value of a 1 mg ml⁻¹ FabZ solution). An average value of 5.91 mg ml⁻¹ (SD: 0.02 mg ml⁻¹) was determined. The protein solution was aliquoted in 100 µl aliquots and stored at -80°C until further use.

Purified FabZ molecules showed the same product peak in SEC-HPLC analysis when compared to the standard previously used. Furthermore, the new standard was compared to the existing one by SDS PAGE under reducing and non-reducing conditions. Linearity

of the purified FabZ molecule was verified by SDS PAGE ($R^2 > 0.99$). An impurity (8 - 21 % of total product) in the low molecular weight range (10 - 15 kDa) was observed in case of the old standard. Considering this impurity, same amounts of new and old standard resulted in similar densities ($\pm 6\%$). Under reducing conditions, similar amounts of protein resulted in lower band intensities (up to -20 %) for the new standard (also under consideration of impurities). However, this might be due to the sample dilution process. In conclusion, standard molecules appeared sufficiently similar. Furthermore, most conclusions of this work depend on relative and not absolute values. For direct comparisons, the same standard molecule was used during analysis procedures. The new reference molecule was used during analytical procedures during two thirds of the project (all shake flask experiments and 59 of 77 fermentations for FabZ production).

2.6.2 Sample Preparation

Prior to protein analysis, supernatant samples from fermentation or shake flask experiments were thawed for a minimum of 15 min at RT and used for product analytics within 90 min. Cell samples were thawed at RT for at least 15 min. Afterwards, 250 and 500 μ l of BugBuster™ Protein Extraction Reagent (Merck KGaA, Darmstadt, Germany) were added to pellets from shake flask and fermentation experiments, respectively. Subsequently, cells were incubated at 22°C and 1400 rpm for 15 min. The resulting solution was used for analysis of the total intracellular protein. If required, separation into soluble and insoluble fraction was performed by centrifugation (16000 x g, 22°C, 15 min). The supernatant after the centrifugation step was used as the soluble intracellular protein fraction. The sediment was considered the aggregated fraction. 500 μ l of reducing 1x LDS solution (4-fold dilution of NuPage® 4x LDS Sample Buffer (Life Technologies, Carlsbad, CA, USA) with 10 % β -mercaptoethanol) were added to the aggregated fraction. Subsequently, reaction tubes were incubated at RT under vigorous agitation for 15 min.

2.6.3 Enzyme-Linked Immunosorbent Assay

Amounts of soluble and correctly folded products FabX and FabZ were quantified by sandwich Enzyme-Linked Immunosorbent Assay (ELISA). Buffers, solutions and material used for ELISA analysis were prepared prior to analysis or supplied by external suppliers (see Table 2.11).

Table 2.11: Material used during ELISA analysis.

| Step | Material | Composition/Details | Supplier |
|--------------------|------------------------|--|--------------------------------------|
| General | Wash Buffer | from 20x PBST 1x PBST + 0.10 % (v/v) Tween20 | this work |
| | 1x TBS | 10x dilution of 10x TBS | this work |
| | 1x PBST | 20x dilution of 20x PBST | this work |
| | 10x TBS | 1.12 g l ⁻¹ Tris 4.8 g l ⁻¹ Tris-HCl 17.6 g l ⁻¹ NaCl | other projects |
| | 20x PBST | Concentrated PBS Tween20 | Cell Signaling, Cambridge, UK |
| Plate coating | Coating Antibody | Ab7497 | Abcam, Cambridge, UK |
| | Coating Buffer | NaHCO ₃ /Na ₂ CO ₃ pH 9.5 | this work |
| | Microtiter well plates | Polystyrene clear flat bottom (Greiner 655061) | Sigma Aldrich, St. Louis, MO, USA |
| Sample preparation | Diluent Buffer | 1x TBS 0.1 % (v/v) Tween20 | this work |
| | Microtiter plate | NUNC F shape | Thermo Fisher, Waltham, MA, USA |
| Blocking | Blocking Solution | 1x TBS 1 % Casein Blocker | Biorad, Hercules, CA, USA |
| Detection | Secondary Antibody | Peroxidase-conjugated anti-human IgG (AP.003M) | The Binding Site, San Diego, CA, USA |
| Color reaction | Substrate Solution | TMB One Component Microwell Substrate | SouthernBiotech, Birmingham, AL, USA |
| | Hydrochloric acid | 1 M HCl | this work |

Solutions with different concentrations (25, 10, 7.5, 5, 3.75, 2.5 and 1 ng l⁻¹) of reference molecule (purified soluble FabZ or FabX product) were prepared using Diluent Buffer. To obtain results of samples within the standard curve range, different dilutions of soluble intracellular and supernatant fractions were prepared, depending on the expected soluble product content. Prior to analysis, 100 µl of a 4000-fold dilution of the Coating Antibody in Coating Buffer were transferred to each well of a polystyrene clear flat bottom microtiter

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| A | Ref 25 | Ref 10 | Ref 7.5 | Ref 5 | Ref 3.75 | Ref 2.5 | Ref 1 | Ref 0 | S06 Dil1 | S06 Dil2 | S06 Dil2 | S06 Dil2 |
| B | Ref 25 | Ref 10 | Ref 7.5 | Ref 5 | Ref 3.75 | Ref 2.5 | Ref 1 | Ref 0 | S06 Dil1 | S06 Dil3 | S06 Dil3 | S06 Dil3 |
| C | Ref 25 | Ref 10 | Ref 7.5 | Ref 5 | Ref 3.75 | Ref 2.5 | Ref 1 | Ref 0 | S06 Dil1 | S07 Dil1 | S07 Dil1 | S07 Dil1 |
| D | S01 Dil1 | S01 Dil1 | S01 Dil1 | S01 Dil2 | S01 Dil2 | S01 Dil2 | S01 Dil3 | S01 Dil3 | S01 Dil3 | S07 Dil2 | S07 Dil2 | S07 Dil2 |
| E | S02 Dil1 | S02 Dil1 | S02 Dil1 | S02 Dil2 | S02 Dil2 | S02 Dil2 | S02 Dil3 | S02 Dil3 | S02 Dil3 | S07 Dil3 | S07 Dil3 | S07 Dil3 |
| F | S03 Dil1 | S03 Dil1 | S03 Dil1 | S03 Dil2 | S03 Dil2 | S03 Dil2 | S03 Dil3 | S03 Dil3 | S03 Dil3 | S08 Dil1 | S08 Dil1 | S08 Dil1 |
| G | S04 Dil1 | S04 Dil1 | S04 Dil1 | S04 Dil2 | S04 Dil2 | S04 Dil2 | S04 Dil3 | S04 Dil3 | S04 Dil3 | S08 Dil2 | S08 Dil2 | S08 Dil2 |
| H | S05 Dil1 | S05 Dil1 | S05 Dil1 | S05 Dil2 | S05 Dil2 | S05 Dil2 | S05 Dil3 | S05 Dil3 | S05 Dil3 | S08 Dil3 | S08 Dil3 | S08 Dil3 |

Figure 2.5: Position of samples on ELISA 96 well measurement plates. Ref: purified soluble FabX or FabZ molecules. Numbers indicate concentrations in ng l^{-1} . S: samples. Up to eight samples were measured on one plate. Dil: different dilutions that were prepared for each sample. All dilutions were applied in triplicates (separated by dotted lines in the graphic).

plate wells (Greiner 655061, Sigma Aldrich, St. Louis, MO/USA). Wells were sealed and the plate was stored over night at 4°C without agitation. Afterwards, a washing procedure was performed (aspiration of buffer from the previous step, washing of wells three times with $250\ \mu\text{l}$ of Wash Buffer, final Wash Buffer aspiration and tapping of plate on paper towel to remove residual buffer). $200\ \mu\text{l}$ of Blocking Solution were added to washed empty wells. Subsequently, the plate was incubated at RT and 450 rpm for 1 h and subjected to the washing procedure described above. $100\ \mu\text{l}$ of reference molecule, blank (Sample Diluent Buffer) and sample solutions were added to the empty microtiter plate wells in triplicates (see Figure 2.5). Subsequently, the plate was incubated at RT for 1 h at 450 rpm. The washing procedure described above was applied after this step. $100\ \mu\text{l}$ of a 4000-fold dilution of the Secondary Antibody in Blocking Solution were added to wells. Subsequently, the plate was incubated at RT and 450 rpm for 1 h. Wells were washed according to the washing procedure described above. $100\ \mu\text{l}$ of Substrate Solution were added to microtiter plate wells prior to incubation for 10 min at RT in the dark. Reaction was aborted by addition of $100\ \mu\text{l}$ 1 M HCl to wells. OD values for wave lengths 620 and 450 nm were determined within 30 min after substrate reaction abortion. OD_{620} values were subtracted from OD_{450} values, resulting in ΔOD values. They were exported and used for product concentration calculations.

Based on the obtained ΔOD values and applied concentrations, a standard curve for the reference molecules was calculated by linear regression. The coefficient of variation (CV) was calculated for triplicate measurements. If the CV was above 10 %, single or all values associated with one concentration were not considered in the calculation. A minimum coefficient of determination (R^2) of 0.99 was defined as acceptance criterium. 0 and $25\ \text{ng l}^{-1}$ were inserted into the calculated equation, yielding lower and upper borders of acceptable ΔOD values, respectively. The value of the “lower border” was arbitrarily increased by

0.1. For samples, values out of the standard curve calculation, *i.e.* below or above the calculated values, were removed from calculations. With the remaining values, volumetric concentrations in the original fermentation suspension (intracellular/extracellular) under consideration of dilutions and sample preparation steps were calculated. A maximum of 9 values was obtained for each sample. Mean and CV were calculated using all available values. Mean values associated with CV values of above 10 % were interpreted with caution or not at all.

2.6.4 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Samples from selected time points were analyzed by Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS PAGE) analysis (see Table 2.10). Appropriate dilutions of 4x NuPage[®] LDS Sample Buffer (Life Technologies, Carlsbad, CA, USA) were used for dilution of solutions resulting from sample preparation (see Section 2.6.2). To establish reducing conditions, 10 % β -mercaptoethanol was added prior to dilution of LDS buffer. Different dilutions of reference molecules were prepared using the same buffer that was used for sample dilution. Samples and reference molecules were incubated for 5 min at 80°C and 450 rpm.

In case of stain-free gel preparations, Precision Plus (Unstained) MW standard (BioRad, Hercules, CA/USA), reference molecule solutions and samples were loaded on Criterion TGX AnyKD Stain-Free Precast Gel (Bio-Rad, Hercules, CA/USA) gels. Protein separation was conducted in 1x Tris/Glycin/SDS buffer applying a voltage of 300 V for approximately 20 min.

In case of Coomassie Brilliant Blue (CBB)-stained gels, samples, the MW standard Precision Plus Protein[™] Dual Color Standard (Bio-Rad, Hercules, CA, USA) and appropriate references were loaded on a NuPAGE[®] 4 - 12 % Bis-Tris Midi Gel (Life Technologies, Carlsbad, CA/USA). Separation was performed in NuPAGE[®] MES SDS Running Buffer (Life Technologies, Carlsbad, CA/USA) NuPAGE[®] applying 200 V for approximately 45 min. After protein separation, gels were washed and stained using the SimplyBlue[™] SafeStain Coomassie[®] G-250 solution (Life Technologies, Carlsbad, CA/USA) according to the manufacturer's protocol. Gel images were acquired after over night destaining in RO-H₂O using a Gel Doc device (Bio-Rad, Hercules, CA/USA).

After acquisition, images were analyzed for quantification purposes using Image Lab 5.0 software. After assigning lanes, relevant bands were assigned by comparison to the MW standard applied on the gel. Quantification of the amount loaded on the gel was carried out using the respective reference molecule's standard curves. A coefficient of determination of above 0.99 was set as the acceptance criterium. During SDS PAGE analysis of Fab fragments, bands at the product MW range were detected in the pre-induction sample (T0). These product amounts were subtracted from the remaining values prior to calculation of volumetric concentrations. Target product amounts detected on the gel were used to calculate volumetric concentrations under consideration of sample preparation and previous dilution.

2.6.5 Western Blot

After electrophoresis according to the CBB-based method described in Section 2.6.4 (non-reducing conditions), proteins were transferred to a nitrocellulose membrane (20 V, 10 min) using an iBlot[®] Gel Transfer Device and Stacks (Life Technologies, Carlsbad, CA, USA). Afterwards, the membrane was incubated with 40 ml of Blocking Buffer for approximately 4 h at RT under gentle agitation. Remaining steps differed depending on the target product (see below). Materials, buffers and solutions used during Western Blot analyses are presented in Table 2.12.

Table 2.12: Material used during Western Blot analysis.

| Step | Item | Composition/Details | Supplier |
|--------------------|---------------------------|---|---|
| General | 20x PBST Buffer | Concentrated PBS Tween20 | Cell Signaling Technol- ogy Cambridge, UK |
| | 1x PBST buffer | 20x dilution of 20x PBST Buffer with RO-H ₂ O | this work |
| | iBlot [®] system | Gel Transfer Device and Stacks | Life Technologies Carlsbad, CA, USA |
| Blocking | Blocking Solution | 3 % (w/v) Blotting Grade Blocker in 1x PBST Buffer | Bio-Rad Hercules, CA, USA |
| Fab detection | Primary Antibody | Anti-Human IgG anti- body [2A11], ab7497 | Abcam Cambridge, MA, USA |
| | Detection Antibody | Anti-Mouse IgG (Fab specific)-Alkaline Phos- phatase, A2179 | SIGMA Aldrich St. Louis, MO, USA |
| | Color Reaction Substrate | Western Blue [®] Stabi- lized Substrate for Alka- line Phosphatase | Promega Madison, WI, USA |
| scFv-Z-1 detection | Protein L-HRP Conjugate | Pierce [®] Protein L, Peroxidase Conjugated, Product Number 32420 | Thermo Scientific Waltham, MA, USA |
| | Color Reaction Substrate | 1-Step Ultra TMB- Blotting Solution | Thermo Scientific Waltham, MA, USA |

Quantification of soluble and correctly folded FabZ and FabX molecules

In selected cases, samples associated with FabX and FabZ production experiments were analyzed by Western Blot analysis. After the blocking step, Blocking Buffer was replaced by 40 ml of a 4000-fold dilution of the Primary Antibody in Blocking Solution. The membrane was incubated over night at RT under gentle agitation. After incubating

the membrane three times for 5 min. each with 1x PBST Buffer at RT under gentle agitation, a 2000-fold dilution of the Detection Antibody in Blocking Buffer was added. The membrane was incubated for 1 h at RT under agitation. After discarding the detection antibody solution, the membrane was washed three times for 5 min. each with 1x PBST Buffer at RT under agitation. The color reaction was started by addition of approximately 10 ml Color Reaction Substrate.

Quantification of soluble and correctly folded scFv-Z-1 molecules

All steps including the Blocking Buffer incubation step were performed as described above. Afterwards, the Blocking Buffer was replaced by 40 ml of a 2000-fold dilution of the Protein L-HRP Conjugate in Blocking Buffer. The membrane was incubated over night at RT under gentle agitation. After discarding the Protein L-HRP Conjugate solution, the membrane was washed three times for 5 min. each with fresh 1x PBST Buffer at RT under agitation. The color reaction was started by addition of approximately 10 ml of Color Reaction Substrate.

Note that, initially, scFv-Z-2 (LC-linker-HC composition) was planned to be used as a scFv model protein. Thus, it was used during establishment of the Protein L-based Western Blot approach. No signals were detected when reducing conditions were applied during sample preparation and dilution procedure (data not shown). Furthermore, sufficient linearity of ($R^2 > 0.99$) between 0.02 and 0.50 μg of loaded material was observed. An increased overall productivity was detected for the alternative scFv-Z variant scFv-Z-1 (HC-linker-LC) in shake flask experiments (data not shown). Thus, scFv-Z-1 was used as a model protein in further approaches. Chain order in scFv molecules was previously shown to influence expression yield and antigen-binding (Buehler et al., 2010). However, while certain characteristics might be influenced, Protein L-binding was expected not to be influenced. Protein L is known to bind to κ light chains of antibodies and their fragments (*e.g.* scFv molecules) and is frequently used in purification approaches (Bjoerck, 1988; Nilson et al., 1993; Rodrigo et al., 2015; Thermo Fisher Scientific Inc., 2015). As active ingredients are desired in these applications, its widely spread use suggests that it merely binds to soluble and correctly folded product variants selectively and independent of the construct. The widely used $(\text{G}_4\text{S})_3$ linker, which was used in both scFv-Z molecules, was reported to not influence ordered secondary structure or interfere with domain folding (Freund et al., 1993; Huston et al., 1988). Thus, the conformation of the target domain of Protein L-binding (κ light chain) was assumed not to be significantly influenced by domain re-organization. It was reported that binding of Protein L to the κ light chain is

dependent on the 3D structure of the target molecule and does not function after reduction and alkylation of cysteines (Nilson et al., 1992). A similar observation was made for scFv-Z-2. No evidence was found that suggests differences in case of scFv-Z-1. If signals are detected by Protein L-based Western Blot, a correct 3D structure was assumed to be present, independent of chain order. In conclusion, chain arrangement seems to strongly influence expression of the target product but not domain folding. Thus, the Protein L-based Western Blot method described was assessed as suitable for specific detection of soluble and correctly scFv-Z-1 fragments.

Quantification of protein concentration using acquired Western Blot images

For all Western Blot procedures, the color reaction was stopped at desired band intensity (5 - 30 min of incubation at RT) by discarding the substrate solution and washing the membrane with RO-H₂O several times for approximately 5 min. each at RT under agitation. Image acquisition and analysis was performed using a GS-900™ Calibrated Densitometer and Image Lab 5.0 (Bio-Rad, Hercules, CA, USA), respectively. Values obtained by densitometric analysis were used to quantify target product content. Specifically, signal intensities from different amounts of reference molecule were used to calculate a standard curve using linear regression with respect to the acceptance criterion ($R^2 \geq 0.99$). This equation was used to calculate the amount of target product on the nitrocellulose membrane in case of samples with unknown product content. The calculated amount was used to determine the volumetric concentration in the suspension of the production experiment under consideration of sample preparation and dilution steps. If no values for samples were below the standard curve, absolute quantification was not performed and values from densitometric measurement of samples were compared directly.

2.6.6 N-terminal sequencing

The cellular location (cytoplasm/periplasm) of aggregated target product was determined by N-terminal sequencing (see Section 3.2.4). The aggregated intracellular fraction obtained during sample preparation (see Section 2.6.2) was diluted using reducing LDS buffer. Samples were subjected to CBB-based SDS PAGE analysis (see Section 2.6.4). Bands affiliated with product aggregates were excised from gels after SDS PAGE analysis using clean scalpels and stored at -20°C until further analysis. After shipping samples on dry ice, purification and Edman degradation for determination of N-terminal amino acid chain sequence was performed by Eurosequence B.V. (Groningen, The Netherlands).

Results and Discussion

3.1 Genomic Integration of Gene Expression Cassettes Enables Significantly Improved Product Yields

Most approaches for bacterial production of recombinant proteins rely on multi-copy plasmids as the location for the gene of interest's expression cassette. Especially for production of soluble and correctly folded products, this technique is often characterized by severe drawbacks, such as plasmid loss, basal expression, uncontrollable increase in plasmid copy number or uncontrolled overload of cellular machineries involved in transcription, translation, secretion or folding (Bentley et al., 1990; Birnbaum and Bailey, 1991; Hoffmann and Rinas, 2004; Rodriguez-Carmona et al., 2012; Terpe, 2006). In a previous study, *E. coli* strains generated by stable integration of the GoI expression cassette into the host genome were successfully employed in bioprocesses for production of recombinant superoxide dismutase (SOD) (Kramer et al., 1996) and green fluorescent protein (GFP) (Cormack et al., 1996). Cytoplasmic soluble product yields were improved up to 3.4-fold. Furthermore, Striedner et al. (2010) reported various advantages of plasmid-free expression systems (*e.g.* increased expression cassette stability, minimized system leakiness, improved process stability and increased host cell capacity for recombinant protein production). A patent application for applying the technique for production of recombinant proteins on a manufacturing scale was filed (WO2008142028, Striedner et al. (2008)). However, plasmid-free expression systems have, so far, not been utilized for production of soluble antibody fragments in the periplasmic space.

During this work, ten *E. coli* strains for production of model proteins GFP, FabX, FabZ and scFv-Z-1 were created by genomic integration. Additionally, two selection marker-free strains for production of GFP and FabZ were created by application of the FRT-FLP technique. Comparable colony numbers after using cells with and without recombinering enzymes for electroporation of the integration cassette during the first approaches made intense colony PCR approaches for identification of correct integration events necessary. By adjusting the dsDNA preparation method the rate of successful integration attempts was increased and the number "false-positive" colonies was reduced.

E. coli strains generated by genomic integration were examined concerning productivity of therapeutically relevant soluble and correctly folded target proteins in the periplasm.

The performance of conventional plasmid-bearing and plasmid-free systems with genome-integrated target gene expression cassette were directly compared in fed-batch fermentation production experiments. Selected product concentration as well as at- and on-line data were evaluated. The aim of this chapter is to present results concerning the main goal of upstream development efforts, increased volumetric soluble product titers.

3.1.1 Green Fluorescent Protein

In this work, genetic toolbox elements for production of soluble and correctly folded antibody fragments in the periplasm were developed and examined. Prior to this work package, results for production of GFP in the cytoplasm as reported by Striedner *et al.* were supposed to be reproduced employing local standard operating procedures (SOPs) and equipment. Three strains were engineered (see Table 3.1) and employed in fed-batch fermentation experiments. Strain B<kanR-GFP> (*gfp* expression cassette genome-integrated into the forward strand of the *attTn7* locus) was prepared by recombineering (Datta *et al.*, 2006; Sharan *et al.*, 2009) using plasmid pACIB-int-gfp for integration cassette creation. This strain was created prior to initiation of work for this thesis. For creation of selection marker-free strain B<GFP>, the B<kanR-GFP> strain’s kanamycin resistance gene module was removed by application of the FRT-FLP technique, utilizing the FRT sites flanking the module. BL21(DE3) cells bearing plasmid pACIB-int-gfp were characterized by growth impairment in basic molecular biology procedures, complicating cryo culture preparation and determination of preculture duration for fermentations. This was probably due to elements present on the pACIB-int plasmid. To avoid these issues, strain B(GFP), bearing the expression cassette on multi-copy plasmid pET30a(+), was employed as a reference strain in fermentation experiments. In all cases, GFP gene expression was controlled by a T7 promoter. Translational control was exerted by a strong, *i.e.* purine-rich, ribosomal binding site (RBS) in both cases. The tZenit and T7 terminators were used for transcription termination in plasmid-free and plasmid-bearing strains, respectively.

Table 3.1: Strains employed in GFP production experiments. Gene location: position of the GFP gene expression cassette. Marker: antibiotic resistance gene used for selection. kanR: gene mediating kanamycin resistance.

| Strain | Gene location | Promoter | Terminator | Marker |
|-------------|---------------------------|----------|------------|--------|
| B(GFP) | pET30a(+) | T7 | T7 | kanR |
| B<kanR-GFP> | <i>attTn7</i> locus (fwd) | T7 | tZenit | kanR |
| B<GFP> | <i>attTn7</i> locus (fwd) | T7 | tZenit | none |

Fed-batch fermentation processes were performed according to the procedure described in Section 2.4.2. The temperature was set to 37°C throughout the production process. Feeding phase was initiated upon glucose depletion indicated by appearance of a sharp increase in the dissolved oxygen level (“DO peak”). A 500 g l⁻¹ glucose feed (Feed 500 Solution, see Section 5.5.4 of the Confidential Supplement) was employed during feeding phases. Initially, the rate of feed solution addition was increased exponentially from 14.0 to 154.7 g h⁻¹ within 12 h. Afterwards, feed solution was added at a constant rate of

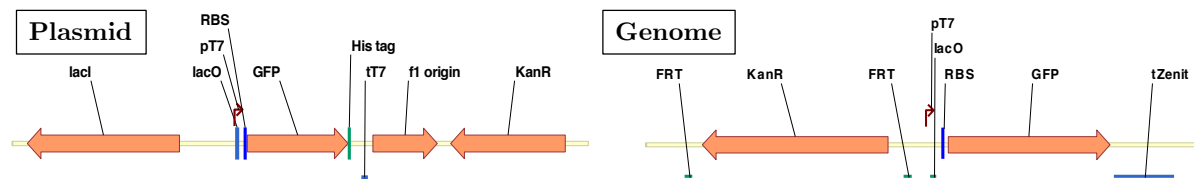


Figure 3.1: Schematic representation of GFP expression cassette composition. Left: plasmid-based expression cassette. The expression cassette was cloned into the MCS of a pET30a(+) plasmid. Right: genome-integrated expression cassette. The expression cassette was integrated into the forward strand of the BL21(DE3) genome's attTn7 site. FRT: flippase recognition target site. KanR: kanamycin resistance module. lacO: binding site for inhibitor LacI (lac operator sequence). pT7: T7 promoter. RBS: ribosomal binding site/Shine-Dalgarno sequence. tT7: T7 terminator. tZenit: tZenit terminator.

154.7 g h⁻¹ until the end of the fermentation process (“EoF”, 27 h after start of feed phase). Two induction strategies were employed during the experiments. In one case (strain B<kanR-FabZ>), IPTG was constantly added at 43.6 mg h⁻¹ starting 5 h after start of exponential glucose feeding. This resulted in an increasing IPTG concentration (0.8 mM at end of fermentation, EoF). In all other cases, IPTG addition was initiated 19 h after start of glucose feeding. In these cases, the addition rate was constantly increased towards EoF, resulting in an increasing IPTG concentration throughout the production phase (mean: 0.9 mM). Cell samples were analyzed concerning their soluble and insoluble GFP content by SDS PAGE (Coomassie Brilliant Blue staining procedure). BSA was used as a reference protein for quantification.

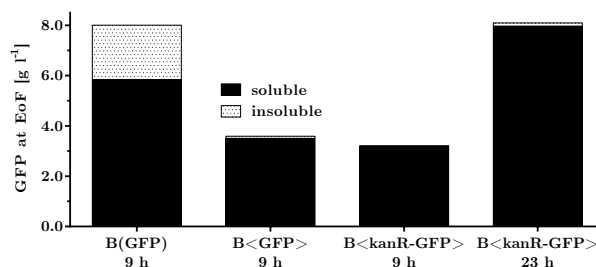


Figure 3.2: Summary of results from 5 l Fed-Batch GFP production experiments employing 4 different production strains. Fermentations differed in strain and/or induction strategy. Volumetric GFP concentrations are presented in dependence of the strain employed in fermentation experiments. Values are based on measurements of soluble and insoluble intracellular GFP concentrations by SDS-PAGE analysis (CBB staining and quantification using a BSA standard solution). Solid black bars represent the soluble, dotted bars the insoluble intracellular product fraction. Category axis labels identify strains used for fermentations. “9 h”/“23 h”: production phase duration.

When the plasmid-bearing strain B(GFP) was employed in fermentation experiments, a total intracellular GFP concentration of 8.0 g l⁻¹ was obtained at EoF. Soluble and insoluble fraction constituted 72.5 (5.8 g l⁻¹) and 27.5 % (2.2 g l⁻¹) of the total product, respectively. Under the same conditions (9 h production phase duration), strain B<kanR-GFP> (plasmid-free) yielded approximately 3.2 g l⁻¹ of total product. The

ratio of insoluble product was determined to be below 1 %. A total product concentration of 3.6 g l^{-1} (98 % soluble product) was measured at EoF in case of marker-free strain B<GFP> (genome-integrated expression cassette) under these conditions. When the production phase was prolonged to 23 h by earlier start of inducer addition, strain B<kanR-GFP> yielded a total product titer of 8.1 g l^{-1} . In this case, the soluble product share was 98 %. Note that the amount of glucose feed solution prepared for this fermentation experiment was not sufficient for the entire feed duration. This resulted in abortion of the fermentation process 1.6 h prior to planned EoF. However, as the production phase was shortened by merely 7 %, a significant influence on resulting data was not assumed.

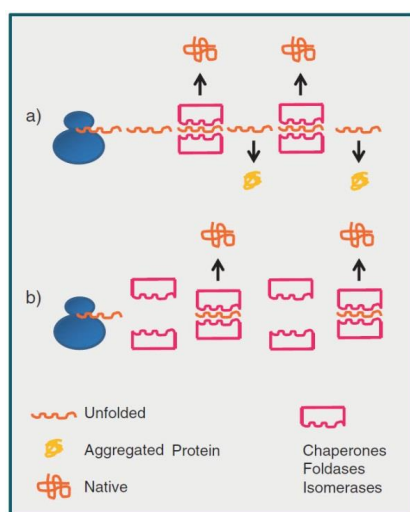


Figure 3.3: Schematic overview of cytoplasmic protein folding. A recombinant polypeptide synthesized by the ribosomal machinery is mostly dependent on one/more chaperones and foldases to reach its native conformation. Disequilibrium between expression rate and folding capacity (a) will result in the accumulation of aggregates. This fate can be prevented by both reducing the target gene's expression rate and favoring the excess of chaperones and foldases (b). The figure was copied from a review article (de Marco, 2013).

In conclusion, total GFP yields were comparable. However, an approximately 1.4-fold increase in soluble GFP concentration was observed after genomic integration of the *gfp* expression cassette and appropriate adaptation of the recombinant protein production phase (9 h \rightarrow 23 h). As expected, a significant fraction of the total product was produced as insoluble aggregates when multi-copy plasmid pET30a(+) was used for expression. Apparently, one of the steps for soluble protein formation in the cytoplasm (transcription, translation, protein folding) is overburdened (see Figure 3.3) (Balzer et al., 2013; de Marco, 2013; Kolaj et al., 2009). A decelerated synthesis rate and prolonged production time enabled improved soluble product formation. Thus, application of the genomic integration technique resulted in expected improvements of soluble product concentration. However, the calculated 1.4-fold improvement is not as significant as previously reported

by Striedner *et al.* (3.4-fold improvement), while a higher volumetric yield of approximately 8.0 g l^{-1} was obtained in this work (*cf.* 6.69 g l^{-1}). Varying absolute titer values might derive from differences in fermentation set up and target product measurement method. GFP concentration was determined by fluorescence and ELISA measurement by Striedner *et al.*, whereas SDS PAGE analysis was applied in this work. Furthermore, Striedner *et al.* initiated the protein production phase 7 h after start of feed addition for both plasmid-bearing and plasmid-free strains, whereas different processes were used for comparison in this study. In addition, crucial information about the integration direction of the *gfp* expression cassette in plasmid-free strains, which were shown to influence the expression rate (see Section 3.2.5), were not disclosed by Striedner *et al.*. Possibly, a higher specific productivity (compared to the values presented above) might be achievable by integration of the expression cassette into the reverse strand of the *attTn7* site (see Sections 3.2.4 and 3.2.5). In summary, a direct comparison of results of the work performed by Striedner *et al.* and this work is not meaningful. However, judging by the results presented above, the initial aim of reproducing the general trend observed in the previous study was fulfilled.

3.1.2 Fragment Antigen Binding Z

Improvements of soluble product yield were observed for cytoplasmic expression using plasmid-free strains within this study and in a previous publication (Striedner et al., 2010). However, no data concerning plasmid-free expression systems have been published for production of proteins in the bacterial periplasm. Expression cassettes for production of model protein FabZ in the periplasm were designed to yield dicistronic mRNA molecules, *i.e.* both monomers were encoded on a single mRNA molecule. Both monomers were genetically fused to DNA sequences encoding the leader peptide of outer membrane protein A (SP_{ompA}) (see Figure 3.4). In relevant strains, the T7 promoter and terminator were employed to initiate and terminate transcription, respectively. Two different alternatives for expression cassette composition, characterized by a different order of the DNA sequences encoding single monomers (HC-LC and LC-HC), were employed during production of FabZ target protein. The codon usage of HC and LC gene differed slightly between the two resulting expression cassettes (611/657 bp identical for HC monomer, 608/639 bp identical for LC monomer).

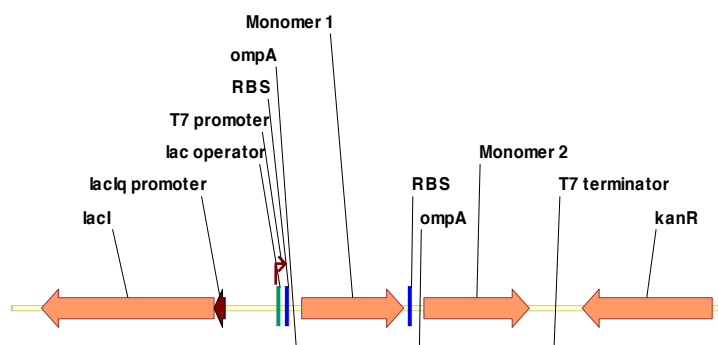


Figure 3.4: Schematic representation of FabZ expression cassettes and their DNA environment. *lacI*: LacI repressor gene. *lacIq* Promoter: strong constitutive promoter for *lacI* expression. *lac* operator: binding site for inhibitor LacI. RBS: ribosomal binding site/Shine-Dalgarno sequence. *ompA*: DNA sequence encoding the OmpA leader peptide. Monomer 1/2: DNA sequence coding for a monomer (LC or HC). *kanR*: kanamycin resistance module.

Both expression cassettes were cloned into a pBI1KT7.1 plasmid. The resulting plasmids were used to create plasmid-bearing production strains by transformation of BL21(DE3) cells. Furthermore, they were used as templates to generate plasmid-free FabZ-producing strains by integration of the expression cassette into the BL21(DE3) genome. A reference strain for assessing the influence of the empty plasmid was created by transforming BL21(DE3) with basic plasmid pBI1KT7.1. All 5 strains were used in fermentation experiments (see Table 3.2) with a constant temperature of 37°C. Fermentation process 1 (feeding strategy 6, induction strategy 4, see Section 2.4.2) was initially applied (see sub-

Table 3.2: Overview of FabZ production experiments with focus on comparison of plasmid-bearing and plasmid-free strains. GoI location: position of the FabZ gene. Cassette: organization of the FabZ gene expression cassette. Fermentation processes 1, 2 and 3 represent different fermentation processes (see text). Numbers in columns represent the number of replicate fermentations for each strain and fermentation set up.

| Strain | <i>fabZ</i> location | Cassette | Fermentation process | | |
|------------------|------------------------------|----------|----------------------|----|----|
| | | | 1 | 2 | 3 |
| B(pBI) | pBI1KT7.1 (no GoI) | none | - | - | 3x |
| B(FabZ)(01) | pBI1KT7.1 | HC-LC | 2x | - | - |
| B(FabZ)(02) | pBI1KT7.1 | LC-HC | 2x | - | 3x |
| B<FabZ-kanR>(01) | <i>attTn7</i> locus, reverse | HC-LC | 2x | - | - |
| B<FabZ-kanR>(02) | <i>attTn7</i> locus, reverse | LC-HC | 3x | 3x | 3x |

sequent sections). Strains bearing the LC-HC or no (reference) expression cassette were compared again using fermentation set up 3 (feeding strategy 3, induction strategy 1, selective conditions). Fermentation process 2 (feeding strategy 3, induction strategy 1) was employed in experiments in the context of co-synthesis of periplasmic folding modulators (see Section 3.4).

HC-LC expression cassette organization

The production performance of plasmid-free strain B<FabZ-kanR>(01) and its plasmid-bearing counterpart B(FabZ)(01) was evaluated in fed-batch fermentation experiments. Both strains are based on strain BL21(DE3) and carry a “HC-LC-type” FabZ gene expression cassette. After cloning of the gene into the multi-copy plasmid pBI1KT7.1, the resulting plasmid was either used to transform BL21(DE3) cells or as a template for recombineering. Fed-batch fermentation processes comprised a batch, an exponential and a constant feed phase. Feeding was initiated after DO peak appearance. Feed 500 Solution (see Section 5.5.4 of the Confidential Supplement) was employed during feeding phases. During the initial 12 h of feeding, a constant specific growth rate ($\mu = 0.2 \text{ h}^{-1}$) was set by exponential increase of the feed rate ($14.0 \rightarrow 154.7 \text{ g h}^{-1}$). Afterwards, feed solution was added at a constant rate of 154.7 g h^{-1} until the end of the fermentation process (“EoF”, 27 h after start of feeding). Glucose feeding was adjusted in case of glucose accumulation according to the rules presented in Section 2.4.2. The temperature was set to 37°C throughout the whole processes. A constant DO value of 20 % was ensured by stirrer speed adjustment and addition of pure O_2 to the inlet air. Recombinant protein production was induced by bolus addition of 894 mg ($\approx 3.75 \text{ mmol}$) IPTG after 14 h of feeding (T0). This bolus resulted in a mean IPTG concentration of 0.6 mM between T0 and EoF. Samples for OD_{550} , DCW, SDS PAGE and ELISA analysis were prepared throughout the process. For each strain, two duplicate fermentation experiments were performed. Different product- and growth-related target parameters were analyzed (see Figure 3.5). Culturable cell numbers and plasmid stability were also examined (see Section 3.2.1).

When the plasmid-bearing strain was employed in fermentation experiments, total intracellular FabZ concentration increased for approximately 8 h after induction, peaking at a maximum of approximately 10 g l^{-1} (see Figure 3.5 (E)). Afterwards, the volumetric concentration decreased towards EoF, resulting in a final total intracellular FabZ concentration of approximately 6 g l^{-1} . The plasmid-free strain showed a constant increase of total FabZ towards a maximum of 1.1 g l^{-1} at EoF. An intracellular soluble product concentration maximum of approximately 50 mg l^{-1} was reached in fermentations of the plasmid-bearing strain (see Figure 3.5 (A)). In case of strain B<FabZ-kanR>(01), which bears the FabZ gene expression cassette in its genome, intracellular soluble FabZ concentrations increased until EoF. This resulted in a mean maximum concentration of 131 mg l^{-1} . Independent of the used strain, less than 50 mg l^{-1} of soluble product were measured in the supernatant (see Figure 3.5 (B)). The total soluble FabZ concentration (see Figure 3.5 (C)) remained constant after 10 h of production phase in case of plasmid-

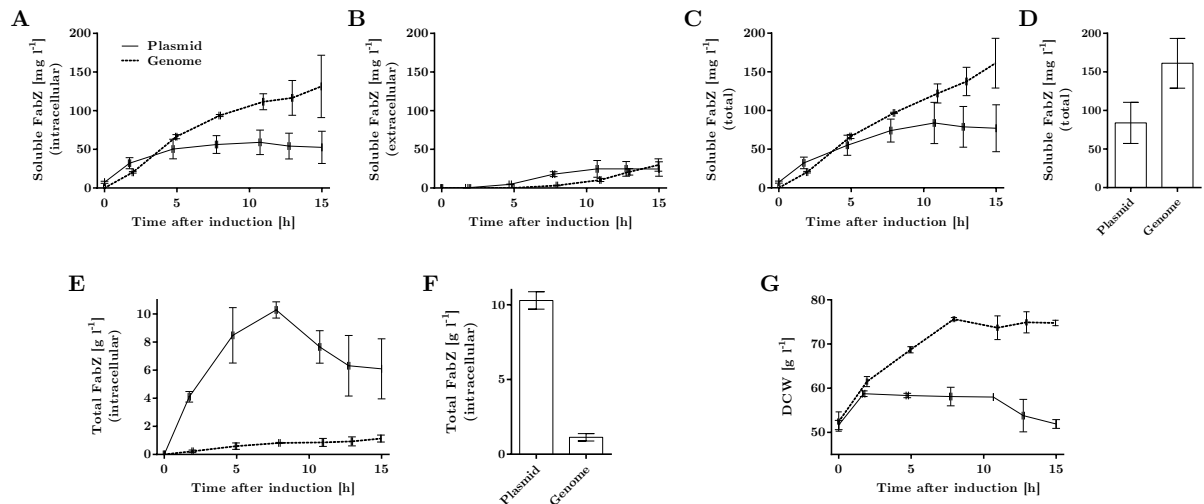


Figure 3.5: Plasmid-bearing vs. plasmid-free FabZ production systems (HC-LC expression cassette): volumetric product concentration, growth and maximum product titers of fermentation experiments. Plasmid-bearing strain B(FabZ)(01) and plasmid-free strain B<FabZ-kanR>(01) were employed in 5 l fed-batch fermentations according to procedures described in the text. Samples for protein analytics (sedimented cells and supernatant fraction) were analyzed by ELISA (soluble FabZ) or SDS PAGE (total FabZ). Dry cell weight (DCW) was determined throughout fermentations. In scatter plots, results referring to plasmid-bearing and plasmid-free strains are represented with solid and dashed lines, respectively. Progress of volumetric soluble FabZ titers in the intracellular (A) and extracellular (B) fraction as well as the sum of soluble products in both fractions (C) are presented. In D, the mean maximum total soluble FabZ concentration in dependence of the location of the expression cassette is presented. Furthermore, total (soluble + insoluble) intracellular FabZ progress (E) and mean maximum (F) as well as progress of DCW values (G) are presented. Based on the results of two replicate fermentations, mean values and standard deviations (depicted as error bars) for values and time points (if applicable) were calculated. Mean of maximum values of multiple fermentations (D and F) are used throughout the text.

bearing strains, resulting in a mean maximum value of 84 mg l^{-1} (see Figure 3.5 (D)). The plasmid-free strain yielded a maximum total soluble FabZ concentration of 161 mg l^{-1} as a mean for both replicate fermentations. In case of this strain, increasing values were measured until the EoF time point. Approximately 2 h after inducer addition, a similar DCW value of approximately 60 g l^{-1} was determined for both strains (see Figure 3.5 (G)). In case of the plasmid-bearing strain, values remained constant for further 8 h and decreased subsequently towards 52 g l^{-1} at EoF. Growth of plasmid-free strains continued for further 6 h, yielding DCW values of approximately 75 g l^{-1} . During the remaining fermentation time, values remained constant. Glucose concentration values of above 1 g l^{-1} were determined in experiments employing plasmid-bearing strains after approximately 4.5 h of induction with IPTG (data not shown). As mitigation measure, an initial feed rate reduction and subsequent continuous adjustments were performed. However, glucose concentrations of above 0.5 g l^{-1} were determined, once again, approximately 2 h prior to EoF.

In summary, application of the multi-copy plasmid resulted in a maximum total soluble

FabZ concentration of 84 mg l^{-1} and a maximum total intracellular product concentration of approximately 10 g l^{-1} as mean of both fermentation runs. Apparently, the high total productivity did not lead to a high soluble protein production. Excessive recombinant protein production might have induced overload of cellular machineries, leading to product aggregation. This might also have led to the observed growth cease. It is likely that plasmid-based expression exerted a selective pressure in favor of cells that have lost the expression plasmid (see Section 3.2.1). Glucose accumulation, a result of feeding glucose exceeding the enzymatic uptake limit of the cells (Varma and Palsson, 1994), was observed in case the of plasmid-bearing strain. Plasmid-based expression was previously shown to cause metabolic breakdown for diverse reasons (Mairhofer et al., 2013). By application of strains bearing a single expression cassette copy in their genome, the maximum total soluble FabZ concentration was improved approximately 1.9-fold to 161 mg l^{-1} . An increase in soluble product concentration by application of this technique is in accordance with previously published results (Striedner et al., 2010). The maximum total FabZ concentration in the cell was decreased approximately 9-fold (max. 1 g l^{-1}). Under the applied conditions, the genome-integrated strain is the system of choice for soluble FabZ production and process convenience. Furthermore, since total soluble product concentrations were still increasing towards EoF for cultivations of this strain, a prolonged production phase might lead to further increased soluble product yields. Due to cell viability impairment (see Section 3.2.1), this strategy seems not suitable for the plasmid-based expression system.

LC-HC expression cassette organization

In case of the “HC-LC” type expression cassette, significant differences between plasmid-bearing and plasmid-free production strains were observed. It was reported that the gene order of Fab fragment monomers on an expression cassette significantly influenced product yields (Humphreys et al., 2002; Jeong et al., 2011; Kirsch et al., 2005). Thus, a dicistronic expression cassette with “LC-HC” monomer order was designed and used for creation of reduction strains. By accident, slightly different monomer sequences (99 % sequence identity) were employed in the new expression cassette. However, due to the minimal variation, it seems reasonable to use this strain to draw general conclusions. After cloning the expression cassette into the multi-copy plasmid pBI1KT7.1, the plasmid was used to either transform BL21(DE3) cells or as a template for recombineering. The resulting strains were compared in 5 l fed-batch fermentation experiments, which were performed as described in the previous section. For each strain, two duplicate fermentation experiments were performed. An additional fermentation experiment employing the plasmid-free strain was performed for the production of FabZ as a reference molecule for analytical procedures. Different product-, growth-, cell viability- and expression cassette stability-related target parameters were examined. Results are presented in Figure 3.6 and Section 3.2.1.

With the “LC-HC” expression cassette, the plasmid-based expression system yielded a maximum total (soluble + insoluble) intracellular FabZ concentration of 6.4 g l^{-1} (see Figure 3.6 (F)). A maximum value of 4.6 g l^{-1} was determined when employing the genome-integrated strain, while values reached a plateau approximately 2.5 h prior to EoF. In fermentation experiments associated with plasmid-based “LC-HC” expression, intracellular soluble product concentration values remained below 90 mg l^{-1} (see Figure 3.6 (A)) (*cf.* values at or below 50 mg l^{-1} for HC-LC expression). Values obtained when employing plasmid-free strains reached a peak at approximately 11 h and of recombinant protein production (270 mg l^{-1}) and, subsequently, decreased towards EoF (200 mg l^{-1}). Soluble product concentration in the supernatant remained below 40 mg l^{-1} (per total volume) for the plasmid-bearing expression system over the entire fermentation duration (see Figure 3.6 (B)). In samples prepared after 9 h of recombinant protein production, soluble product was first detected in the supernatant of fermentations associated with genomic integration strains. During further cultivation, concentrations increased constantly and resulted in a maximum value of 179 mg l^{-1} at EoF. For the plasmid-bearing strain, total soluble FabZ concentrations stayed below 100 mg l^{-1} (see Figure 3.6 (C)), resulting in a maximum value of 83 mg l^{-1} as mean value of the duplicate fermentations (see Figure 3.6 (D)). A maximum of 363 mg l^{-1} was obtained as mean of the cultivations of

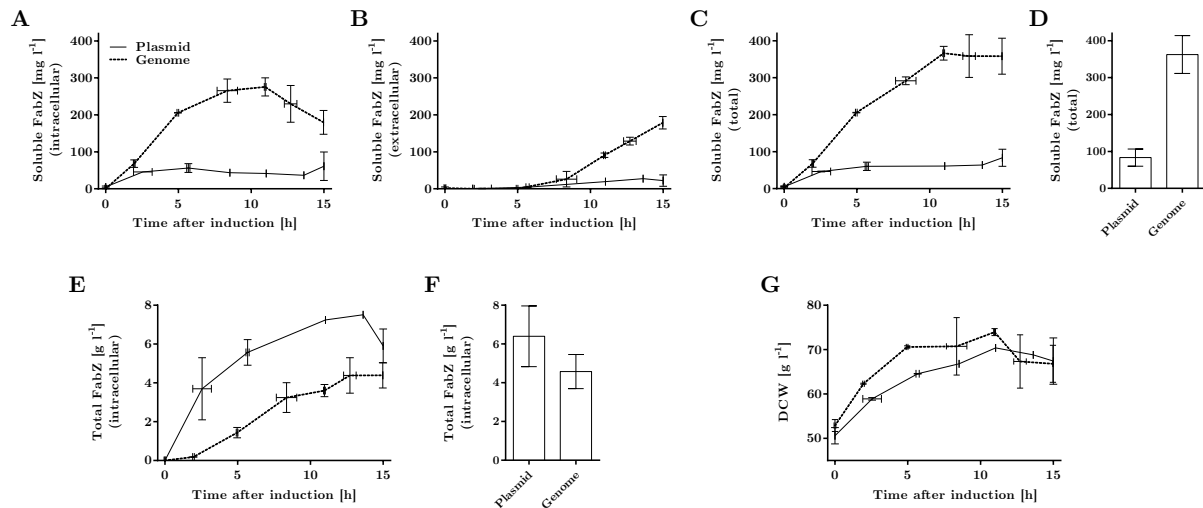


Figure 3.6: Plasmid-bearing vs. plasmid-free FabZ production systems (LC-HC expression cassette): volumetric product concentration, growth and maximum product titers of fermentation experiments. Plasmid-bearing strain B(FabZ)(02) and plasmid-free strain B<FabZ-kanR>(02) were employed in 5 l fed-batch fermentations according to procedures described in the text. Samples for protein analytics (sedimented cells and supernatant fraction) were analyzed by ELISA (soluble FabZ) or SDS PAGE (total FabZ). Dry cell weight (DCW) was determined throughout fermentations. In scatter plots, results referring to plasmid-bearing and plasmid-free strains are represented with solid and dashed lines, respectively. Progress of volumetric soluble FabZ titers in the intracellular (**A**) and extracellular (**B**) fraction as well as the sum of soluble products of both fractions (**C**) are presented. In **D**, the mean maximum total soluble FabZ concentration in dependence of the location of the expression cassette is presented. Furthermore, total (soluble + insoluble) intracellular FabZ progress (**E**) and mean maximum (**F**) as well as progress of DCW values (**G**) are presented. Based on the results of two (plasmid-bearing) or three (plasmid-free) replicate fermentations, mean values and standard deviations (depicted as error bars) for values and time points (if applicable) were calculated. Mean of maximum values of multiple fermentations (**D** and **F**) are used throughout the text.

genome-integrated strains. These cultivations reached the maximum at approximately 11 h after induction, whereupon values remained constant until EoF. The batch phase duration for plasmid-bearing and plasmid-free strains was 18 and 8 h, respectively (see Section 3.2.3 for details and discussion). DCW courses were similar for both expression systems. Differences in culturable cell count and expression cassette stability were observed (see Section 3.2.1 for details and discussion). Glucose accumulation ($\geq 0.5 \text{ g l}^{-1}$) was not detected in any of the fermentation experiments. To prevent anticipated glucose accumulation in an unattended process phase, the feeding rate was reduced from 154.7 to 80 g h^{-1} after approximately 6 h of Fab production in case of one fermentation employing the plasmid-bearing strain.

When plasmid-bearing LC-HC strain was employed in fermentation experiments, approximately 1 % of the total intracellular product (5.9 g l^{-1}) at EoF was soluble and correctly folded (61 mg l^{-1}). Furthermore, a low mean maximum total soluble value of 83 mg l^{-1} was determined. As observed in HC-LC experiments described above, limitations in steps

of the synthesis path (*e.g.* system overburdening or plasmid loss) appear to prevent efficient formation of native FabZ. By application of genomic integration of the target gene, the maximum total soluble FabZ concentration was increased approximately 4-fold to 363 mg l^{-1} , while the maximum total intracellular FabZ concentration decreased approximately 1.5-fold to 4.6 g l^{-1} . Soluble and correctly folded FabZ constituted approximately 4 % of the total intracellular FabZ content at EoF ($180 \text{ mg l}^{-1}/4.4 \text{ g l}^{-1}$). While target protein (soluble FabZ) content was improved significantly, high total intracellular FabZ concentrations suggest further optimization potential to improve soluble product yields. Feed rate reduction and reduced sampling during one of the fermentations employing the plasmid-bearing strain and the observed plasmid loss phenomenon (see Section 3.2.1) complicate interpretation. However, results unambiguously support the assumption that genomic integration constitutes a valuable tool for improving soluble product yields.

Obtained results were further examined concerning differences in expression system performance caused by arrangement of monomers (HC-LC vs. LC-HC order). Considering plasmid-bearing strains, the peak in total intracellular FabZ concentration shifted from 8 (HC-LC) to 11 h (LC-HC) after inducer addition. The intracellular soluble FabZ concentration for plasmid-bearing strains remained below a maximum of 90 mg l^{-1} at all sampling time points, independent of expression cassette organization. Likewise, progress of these values throughout the processes did not show significant differences. Soluble FabZ concentrations in the supernatant ranged between 0 and 40 mg l^{-1} at all time points for both expression cassettes, when plasmid-bearing strains were employed. Progress of total soluble FabZ concentrations was comparable for both expression cassette types. Maximum total soluble FabZ concentrations were comparable (84 (HC-LC) and 83 mg l^{-1} (LC-HC)). Similar DCW values were determined prior to induction, employing either HC-LC (52 g l^{-1}) or LC-HC-type (51 g l^{-1}) expression cassettes. After approximately 2 h of induction, DCW values for the HC-LC expression cassette remained constant at approximately 59 g l^{-1} for 8 h and, subsequently, decreased towards EoF (52 g l^{-1}). On the other hand, DCW values measured in fermentations of the LC-HC-expression cassette-bearing strain increased further, yielding EoF values of approximately 67 g l^{-1} . By changing the expression cassette organization from HC-LC to LC-HC, the batch phase was prolonged from approximately 9 to 18 h. While the culturable cell count decreased rapidly after inducer addition when the HC-LC cassette was used, usage of LC-HC resulted in nearly constant values throughout the fermentation process. The plasmid stability did not differ much between strains comparing either one of the expression cassettes. In conclusion, an influence of the expression cassette on certain fermentation parameters (*e.g.* culturable cell number, growth after induction, batch phase duration) was detected in case of

plasmid-bearing strains. Product-related parameters were not significantly impacted.

Considering plasmid-free expression strains, the maximum total intracellular FabZ concentration was increased 4-fold from 1.1 to 4.6 g l⁻¹ by application of the LC-HC cassette (compared to the HC-LC variant). For both alternatives, overall FabZ content increased until EoF. While intracellular soluble FabZ content increased towards EoF in case of the “HC-LC” expression cassette, values associated with LC-HC strains increased more rapidly, peaked after 9.5 h of recombinant protein production and, subsequently, decreased towards EoF. The maximum soluble intracellular FabZ concentration was increased approximately 2-fold from 132 to 269 mg l⁻¹ by application of LC-HC-type organization in the genome-integrated expression cassette. Soluble FabZ concentrations in the supernatant of fermentations associated with genome-integrated HC-LC-type expression cassettes ranged below 40 mg l⁻¹ at all process time points. Re-organization of the expression cassette to LC-HC led to increased extracellular FabZ concentrations of 179 mg l⁻¹ at EoF. This constitutes a 6-fold increase in the extracellular titer. Using the “HC-LC” expression cassette, total soluble FabZ concentration increased towards EoF, yielding maximum values of 161 mg l⁻¹. When using the LC-HC alternative, a 2.2-fold increased maximum total soluble FabZ concentration of 363 mg l⁻¹ was reached. No significant influence of expression cassette organization on the DCW and its progress was determined. Changing the expression cassette from HC-LC to LC-HC led to a reduction of culturable cells at EoF (10⁹ vs. 10⁵ cfu ml⁻¹). Both genome-integrated strains showed an expression cassette stability of 100 % for all analyzed samples (see Section 3.2.1).

Differing effects of expression cassette re-organization were observed in case of plasmid-bearing and genome-integrated strains. In plasmid-free strains, product-related values, *i.e.* intra- and extracellular soluble and total intracellular product concentration, were substantially affected, while other process/growth parameters, except for the culturable cell count, remained unchanged. It appears that expression cassette organization has the highest impact on the target value “soluble product yield” when remaining noise factors affiliated with plasmid-based expression (*e.g.* plasmid loss, metabolic burden) are eliminated.

Changing the expression cassette organization from HC-LC to LC-HC and *vice versa* might influence several factors like mRNA secondary structure and stability. This might affect the gene expression levels through differences in mRNA availability or low translation rates due to structural hindrance (Carrier and Keasling, 1997; Goodman et al., 2013; Grunberg-Manago, 1999; Gu et al., 2010; Lenz et al., 2011; Plotkin and Kudla, 2011). Furthermore, different mRNA stabilities within a polycistronic operon were previously found

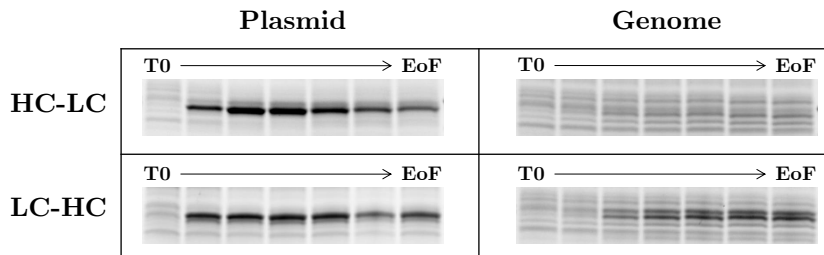


Figure 3.7: SDS PAGE analysis (reducing conditions) of FabZ samples from different fermentation experiments. Plasmid-bearing (**Plasmid**) and plasmid-free (**Genome**) *E. coli* expression systems, bearing either **HC-LC** or **LC-HC** expression cassettes were subjected to fed-batch fermentations. Processes were conducted as described in the text. Comparable amounts of cells of seven different time points (prior to (T0) and throughout the recombinant protein production phase until EoF) were analyzed concerning their total intracellular product content. In previous experiments associated with another project, it was shown by SDS PAGE analysis of monomers that the two distinct bands on the SDS PAGE gel refer to FabZ LC (upper) and HC (lower band), respectively. Note that the glucose feed rate was reduced in case of the plasmid-based HC-LC expression cassette (upper left gel section).

to influence relative gene expression (Newbury et al., 1987). This suggests the possibility of different monomer chain ratios in the cases described above. The development of a system with optimized expression ratio of monomers improved Fab production (Humphreys et al., 2002). It appears that changing the order of the expression cassette from HC-LC to LC-HC shifts the production from HC to LC as it can be seen for the plasmid-bearing expression systems in Figure 3.7. However, it was also observed within this project that periplasmic and cytoplasmic aggregates form separate bands on SDS PAGE gels under reducing conditions (data not shown). Thus, observed double bands might also have a different origin. Despite more than 99 % sequence homology, small differences in codon usage within monomer DNA sequences, might also be influential (Fredrick and Ibba, 2010; Sorensen et al., 1989; Tuller et al., 2010a,b). While identifying minor influential factors is hardly possible based on the obtained results, changing the monomer order from HC-LC to LC-HC unambiguously increased total intracellular FabZ content in recombinerred strains. Both bands visible in the bottom right section of Figure 3.7 (Genome, LC-HC) were later shown to originate from the periplasm (see Section 3.2.4). Thus, they are not overlaid by double bands caused by improper leader peptide cleavage and results suggest a balanced LC to HC ratio in the bacterial periplasm. In conclusion, a combination of higher total productivity and a more balanced monomer ratio probably caused improvement of soluble product formation in case of a genomically integrated LC-HC expression cassette. Due to its higher total productivity and improved soluble product titers, the LC-HC expression cassette was used for further experiments.

Repetition of fermentation experiments employing LC-HC expression cassettes for more detailed analysis

In this part of the study, strains B<FabZ-kanR>(02) and B(FabZ)(02) were, again, employed in fed-batch fermentation experiments. Results concerning comparison of plasmid-bearing and genome-integrated expression systems (see previous section) were supposed to be reproduced. In order to evaluate the influence of the basic plasmid during production experiments, BL21(DE3) cells bearing the empty plasmid pBI1KT7.1 without FabZ genes were also employed in production experiments. Fermentations were performed as described in Section 2.4.2 and were adapted to the most recent process parameters at that state of the project. Parameters employed in previous sections are subsequently presented in parentheses. Fermentation processes consisted of batch, exponential and constant feed phase. Experiments were terminated after 27 h (previous: 29 h) of feeding. Feed 600 Solution (see Section 5.5.4, previous: Feed 500 Solution) was used for feeding. Due to an increased glucose concentration in the feed solution, the feed rate was reduced. The feeding rate was set to increase exponentially from 15.29 to 127.91 g h⁻¹, yielding a constant growth rate within the first 12 h of feeding ($\mu = 0.177$ h⁻¹, previous rate: 14.0 to 154.7 g h⁻¹). Afterwards, a constant feeding rate of 127.91 g h⁻¹ (previous: 154.7 g h⁻¹) was set (feeding strategy 3, see Section 2.4.2). After 14 h of feeding, 894 mg (≈ 3.75 mmol) of inducer IPTG were added to the fermenter (induction strategy 1), resulting in a 13 h recombinant protein production phase with IPTG concentrations of 1.1 and 0.9 mM prior to induction and at EoF, respectively. A constant cultivation temperature of 37°C was applied throughout the process. In order to avoid plasmid loss, batch and feed media were supplied with 30 and 50 $\mu\text{g l}^{-1}$ kanamycin for plasmid-free and plasmid-bearing strains, respectively (previous: non-selective conditions). Standard procedures described in Section 2.4.2 were applied in case of glucose accumulation. Selected at-, off- and on-line parameters were determined as described above. Results are presented in Figure 3.8.

As expected, no product was identified in any of the analyzed samples associated with the non-productive mock strain carrying the empty plasmid. When the plasmid-bearing expression system was employed, total intracellular FabZ concentrations increased rapidly for the first 4 h after induction up to 7 g l⁻¹. Afterwards, they remained nearly constant. A maximum value of 7.7 g l⁻¹ was determined for the plasmid strain. Titters for the plasmid-free strain remained at a plateau after 10 h, yielding a maximum of 4.5 g l⁻¹. Intracellular soluble FabZ concentrations of below 80 mg l⁻¹ were determined for all samples in case of the plasmid-bearing system. For the genomic integration strain, values peaked after approximately 8.5 h (269 mg l⁻¹) and, subsequently, decreased towards EoF. Soluble

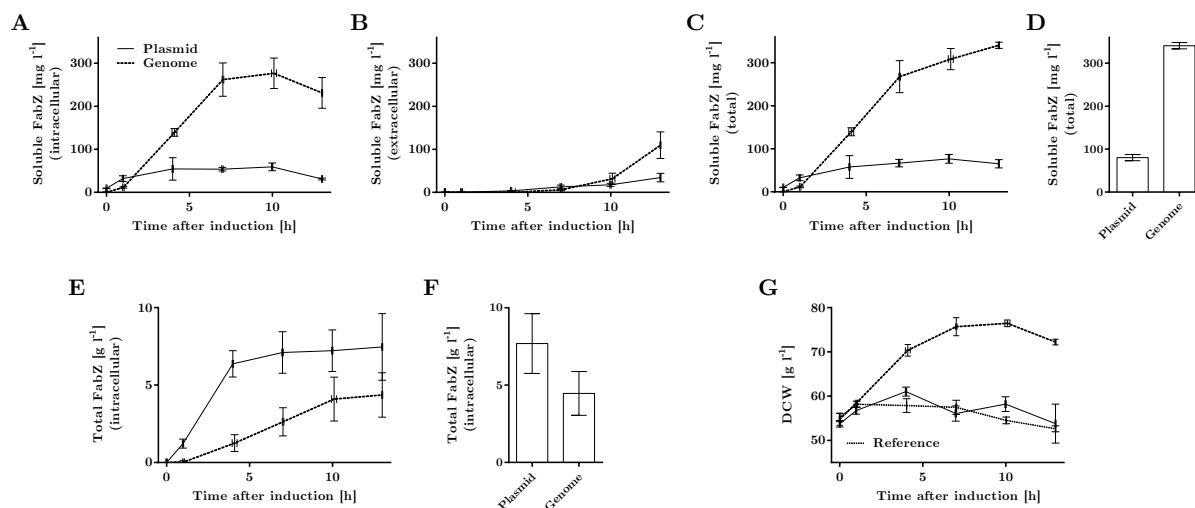


Figure 3.8: Plasmid-bearing vs. plasmid-free FabZ production systems (LC-HC expression cassette): volumetric product concentration, growth and maximum product titers of repeated fermentation experiments. Strains B(FabZ)(02) (plasmid-bearing), B<FabZ-kanR>(02) (plasmid-free) and B(pBI) (basic plasmid-bearing) were employed in 5 l fed-batch fermentations according to procedures described in the text. Samples for protein analytics (sedimented cells and supernatant fraction) were analyzed by ELISA (soluble FabZ) or SDS PAGE (total FabZ). Dry cell weight (DCW) was determined throughout fermentations. In product-related scatter plots, results referring to plasmid-bearing and plasmid-free strains are represented with solid and dashed lines, respectively. As no product concentration was measured, no product-related values for the non-producing reference strain are presented. Progress of volumetric soluble FabZ titers in the intracellular (**A**) and extracellular (**B**) fraction as well as the sum of soluble products (**C**) are presented. In **D**, the mean maximum total soluble FabZ concentration in dependence of the location of the expression cassette is presented. Furthermore, total (soluble + insoluble) intracellular FabZ progress (**E**) and mean maximum (**F**) as well as progress of DCW values (**G**) are presented. DCW values for the reference strain are presented as well (dotted line in **G**). Based on the results of three replicate fermentations, mean values and standard deviations (depicted as error bars) for values and time points (if applicable) were calculated. Mean of maximum values of multiple fermentations (**D** and **F**) are used throughout the text.

product appeared in the supernatant after approximately 4 h when the plasmid-bearing strain was employed. Titers remained below 50 mg l^{-1} throughout the entire process. Genome-integrated expression cassettes facilitated the appearance of soluble product in the supernatant after approximately 7 h. Furthermore, a steady increase was detected, reaching a maximum of 110 mg l^{-1} . Total soluble product concentrations for the plasmid-bearing strain at all time points ranged between 0 and 90 mg l^{-1} , resulting in a mean maximum value of 80 mg l^{-1} . In case of recombiner strains, total soluble product concentration increased constantly for 7 h (268 mg l^{-1}). Subsequently, a less rapid increase was detected. A mean maximum value of 341 mg l^{-1} was determined. The plasmid-free production strain and the mock strain bearing the empty plasmid pBI1KT7.1 resulted in batch phases of similar duration (8.5 - 8.6 h). Batch phases in experiments employing plasmid-bearing FabZ producer B(FabZ)(02) lasted longer (14.6 h on average). Similar DCW values were determined for all examined strains at 1 h after inducer addition (see

Figure 3.8 (G)). Fermentation experiments employing the plasmid-bearing strains (productive or non-productive) were characterized by nearly constant DCW values during the remaining “induction phase”. DCW values associated with genome-integrated strains continued to increase for further 9 h (76 g l^{-1}) before decreasing towards EoF (72 g l^{-1}). Accumulation of glucose ($> 0.5 \text{ g l}^{-1}$), was observed at 4 h and 7 h after inducer addition in case of the mock and the production plasmid-bearing strains, respectively. Despite feed rate adjustment, glucose accumulation was, once again, detected for both plasmid strains at EoF. Glucose remained limiting in fermentations employing plasmid-free strains. DO values oscillated between 0 and 60 % (standard deviation (SD): 13 %) during the first 5 h after inducer addition when the plasmid-bearing expression system was employed in fermentation experiments. Similar oscillation phenomena (SD: 11 %) were observed between 7 and 12.5 h after inducer addition in case of the plasmid-free strain. For comparison, in periods without apparent oscillation phenomena, SD values ranged below 2 %. Stirrer speed increased less rapidly during exponential feed phase in case of the productive plasmid-bearing system. Moreover, for the productive plasmid strain, stirrer speed decreased and increased repeatedly during the recombinant protein production phase. Stirrer speed remained at the maximum of 1200 rpm for both other strains. CO_2 exhaust values were highly similar for all strains and all replicate fermentations prior to inducer addition. In case of the plasmid-bearing systems, post-induction exhaust CO_2 values increased more rapidly compared to the recombinant strain. CO_2 values decreased at 4.5 and 7.5 h after induction in case of mock and the production plasmid-bearing strain, respectively. Furthermore, different amounts of ammonia were consumed in fermentation of plasmid-free production, plasmid-bearing production and mock strain (in order of descending consumed base amount).

In summary, the plasmid-bearing FabZ producer strain was, again, characterized by previously observed drawbacks (*i.e.* prolonged batch phase, growth cease after induction and a low maximum total soluble FabZ concentration). In contrast to previous experiments with this expression system, accumulation of glucose was observed. As selective conditions were applied, plasmid loss was probably prevented. Thus, unlike previously, all cells remained productive and metabolically burdened. The reduced glucose metabolization potential might, thus, have led to accumulation. Comparable to previous experiments during this study, genomic integration resulted in a significantly increased maximum total soluble FabZ concentration. Still, a maximum value of several g l^{-1} of total intracellular FabZ indicates unaddressed bottlenecks during synthesis of soluble and, thus, correctly folded FabZ product and optimization potential for further expression system development. In the course of these experiments, a non-Fab-producing mock strain bearing the

basic plasmid pBI1KT7.1 was employed in fermentation experiments as well. Mere plasmid presence resulted in growth cease and glucose accumulation after induction. These phenomena are typically associated with high-rate recombinant protein production using multi-copy plasmid. Furthermore, on-line data (*e.g.* exhaust CO₂ values and stirrer speed progress) resembled on-line data of fermentations of the plasmid-based production system more closely than those obtained from experiments employing the plasmid-free strain. The influence of a multi-copy plasmid as such (as opposed to effects of recombinant protein production) has been examined in several previous research articles (Andersson et al., 1996; Bentley et al., 1990; Birnbaum and Bailey, 1991; Carneiro et al., 2013; Mairhofer et al., 2013; Ow et al., 2006; Wang et al., 2006). Results of this work in the light of these previous studies are discussed below (see Section 3.2.5). In conclusion, despite process changes (selective media, shortened production phase, different feed conditions), previous results concerning advantages of plasmid-free strains employing the LC-HC expression cassette variant could be reproduced.

3.1.3 Fragment Antigen Binding X

Results described above showed that soluble recombinant protein yields were significantly improved by application of the plasmid-free expression systems. To verify the technique's suitability as a general toolbox element, a plasmid-free strain for production of the therapeutically relevant Fab fragment FabX was generated and employed in fed-batch fermentation experiments. Subsequently, results were compared with product titers obtained in experiments employing plasmid-bearing FabX producers. Apart from being scientifically relevant, application of genomic integration constituted a promising strategy for improving soluble FabX yields. This was scope of a different project conducted at BI RCV. In preceding approaches during this project, several plasmid-bearing strains were employed in production experiments. Furthermore, production process developments were conducted. In order to generate a plasmid-free expression system, the *fabX* genes were cloned into the basic plasmid pBI1KT7.1. The resulting plasmid was used as a template for integration of the expression cassette into the reverse strand of the *E. coli attTn7* site. For FabX production using plasmid-bearing expression systems, a fed-batch fermentation process was developed (see Section 2.4.3) and optimized for improved soluble product yield. A modified process was established for plasmid-free expression systems. Thus, as opposed to previous experiments, results from two different fermentation processes were compared. Since both processes were intensively optimized for the use in combination with the employed expression system, a direct comparison of both systems seemed feasible. Soluble product concentrations in the supernatant and total intracellular FabX content were determined by 2D-HPLC and SDS PAGE, respectively (see Figure 3.9).

The mean total intracellular FabX concentration was 3.9 g l^{-1} (SD: 25 %) when plasmid-bearing strains were employed in fermentation experiments. For these strains, a mean maximum concentration of 66 mg l^{-1} (SD: 34 %) was determined for soluble FabX in the supernatant. Mean intracellular FabX concentrations of 2.4 g l^{-1} (SD: 24 %) were determined when plasmid-free strains were applied. On the other hand, increased mean values for soluble product in the supernatant of 205 mg l^{-1} (SD: 36 %) were determined.

Examination of consolidated results (Figure 3.9) showed that genomic integration of the *fabX* genes improved soluble product concentrations in the supernatant at the expense of a slightly reduced total (soluble + insoluble) intracellular FabX content. In conclusion, expression systems with genomically integrated target genes performed significantly better than their plasmid-bearing counterparts under the applied conditions. In this regard, results are in accordance with the previous studies concerning production of soluble FabZ and reinforce genomic integration's value for expression system development. Due to

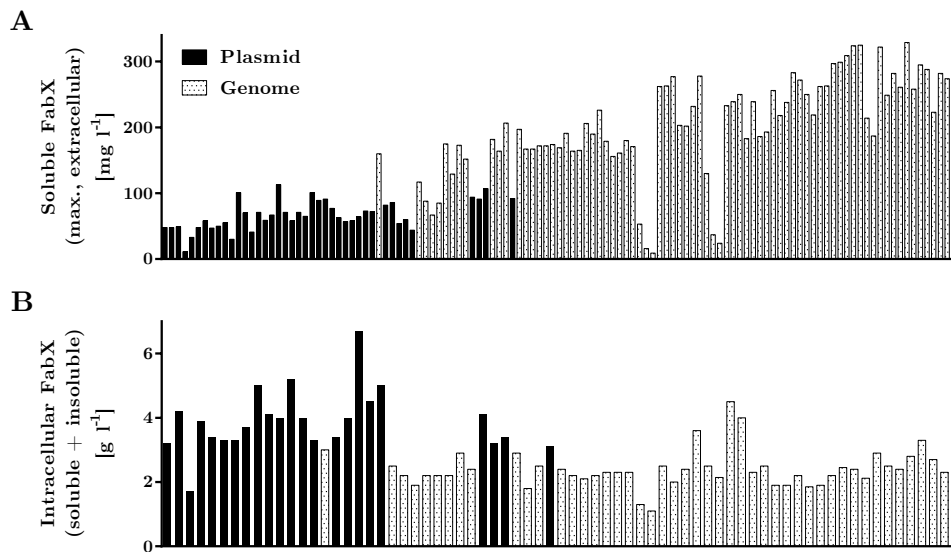


Figure 3.9: Summary of results obtained during FabX production experiments of more than 2 years. As the majority of correctly folded FabX product accumulated in the supernatant, different expression systems and process formats were employed in order to improve soluble extracellular FabX concentration. **A**: soluble FabX titers in the culture supernatant, determined by 2D-HPLC analysis. **B**: total intracellular FabX titer, determined by SDS PAGE analysis. Maximum titers of different fermentations are represented as bars. From the left to the right the results of the first to the most recent fermentation are depicted. Filled black and dotted bars represent approaches, in which plasmid-bearing and genome-integrated expression systems were employed, respectively.

its beneficial production characteristics, the new plasmid-free expression system was employed in the further course of the above-mentioned project. Comparing the mean values presented above, only small fraction of the total product (below 10 %) seems to be correctly assembled and reached its native conformation. As reported above for plasmid-free FabZ production systems, synthesis steps in soluble FabX formation seem to constitute inhibitory bottlenecks, even after genomic integration of target genes.

3.1.4 Single-Chain Fragment Variable Z 1

During previous parts of this study, it was shown that plasmid-free strains improved production of soluble GFP (cytoplasm), FabZ and FabX (periplasm). Aim of this part of the work was to examine single-copy, genome-based expression during periplasmic production of the monomeric scFv fragment “single-chain fragment variable Z 1” (scFv-Z-1), which bears two disulfide bonds. The *scfv-Z-1* gene was cloned into the basic plasmid pBI1KiT7.3, which bears the tZenit terminator and a FRT-site-flanked kanamycin resistance gene module. The resulting plasmid was used to transform BL21(DE3) cells, yielding production strain B(scFv-Z-1). Furthermore, it was used as a template for genomic integration of the expression cassette into the reverse and the forward strand of the *attTn7* site of the BL21(DE3) strain, yielding strain B<scFv-Z-1>(01) (reverse integration) and B<scFv-Z-1>(02) (forward integration), respectively. All three strains were cultivated in equally designed fed-batch fermentation processes (see Section 2.4.2 for details). Briefly, processes consisted of batch and exponential and constant feed phase. Experiments were terminated after 27 h of feed phase duration. Appearance of a dissolved oxygen (DO) peak triggered start of Feed 600 Solution addition. The feeding rate was set to increase exponentially from 15.29 to 127.91 g h⁻¹ within 12 h after start of feeding to enable a constant growth rate ($\mu = 0.177$ h⁻¹). Afterwards, a constant feeding rate of 127.91 g h⁻¹ was set. Standard actions described in Section 2.4.2 were applied in case of glucose accumulation. The dissolved oxygen (DO) level was kept constant at above 20 % by agitation and addition of pure O₂ to the inlet air. After 14 h of feeding, 894 mg (≈ 3.75 mmol) of the inducer IPTG were added to the fermenter, resulting in 13 h of recombinant protein production with IPTG concentrations of 1.16 and 0.85 mM immediately after induction and at EoF, respectively. In the initial growth phase of the process, a temperature of 37°C was applied. Prior to inducer addition, the temperature was decreased to 30°C within one hour using a linear ramp and kept constant throughout the recombinant protein production phase. Selected at-, off- and on-line parameters were determined as described above. The scFv production potential was assessed by analysis of intracellular soluble scFv-Z-1 concentration using a Protein L-based Western Blot procedure developed in the course of this project. Selected results are presented in Figure 3.10.

Fermentation experiments, in which the multi-copy plasmid-bearing scFv-Z-1 production strain was employed, resulted in 8.7 and 19.0 g l⁻¹ of insoluble intracellular product after 8 and 13 h of recombinant protein production, respectively. Employing plasmid-free strain B<scFv-Z-1>(01), created by genomic integration into the *attTn7* reverse strand, resulted in roughly 8 g l⁻¹ insoluble intracellular scFv-Z-1 at both sampling time points.

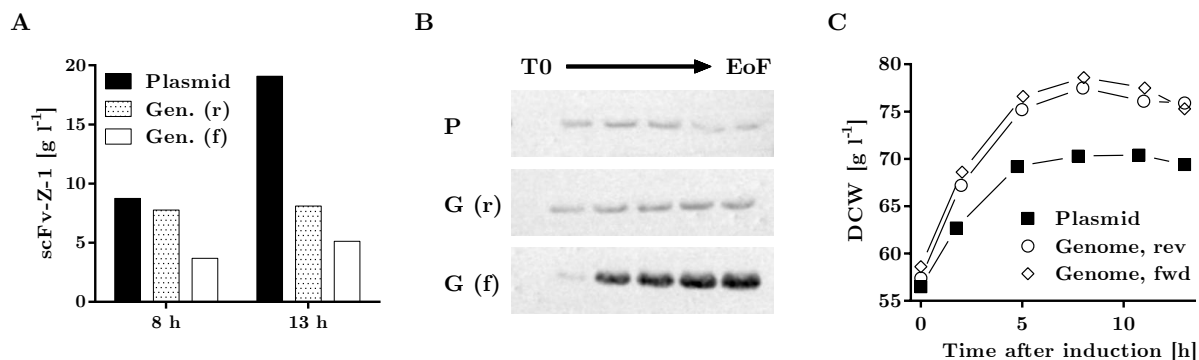


Figure 3.10: Plasmid-bearing vs. plasmid-free scFv-Z-1 production systems: analysis of insoluble and soluble product concentration and growth in fermentation experiments. A plasmid-bearing (pBI1KiT7.3) and two plasmid-free (reverse and forward integration at *attTn7* site) scFv-Z-1 production strains were employed in fed-batch fermentations in 5 l bioreactors according to the procedure described in the text. “Genome, rev/Gen. (r)/G (r)”: plasmid-free expression system generated by integration into the reverse strand. “Genome, fwd/Gen. (f)/G (f)”: plasmid-free expression system generated by integration into the forward strand. Cell samples of sedimented cells were analyzed by SDS PAGE (insoluble product, **A**) and Western Blot (soluble product, **B**) analyses. Dry cell weight (DCW, **C**) was determined throughout the fermentation experiments.

In case of the second plasmid-free expression system (forward integration), an insoluble intracellular product titer of approximately 4 g l^{-1} was determined after 8 h of recombinant protein production phase. The value increased to approximately 5 g l^{-1} at EoF. Western Blot analysis for examination of the soluble intracellular product content resulted in faint scFv-Z-1 bands below the quantification limit for both the plasmid-bearing strain and the plasmid-free strain bearing the “reverse-integrated” expression cassette. In case of the “forward-integrated” cassette, the sample of the first “post-induction” time point showed a barely detectable signal in Western Blot analysis. Bands associated with subsequent samples showed clearly detectable product-related bands exceeding the intensities of bands associated with the other tested strains by far. A maximum soluble product titer of approximately 1 g l^{-1} was calculated for this strain based on these bands. Prior to induction, DCW values ranged between 57 and 59 g l^{-1} and the batch phase duration was approximately 8.6 h for all strains. After induction, DCW values for the plasmid-bearing strain increased to approximately 70 g l^{-1} 5 h after inducer addition and remained constant until EoF. DCW increased during the first 8.5 h of recombinant protein production to a peak of 78 g l^{-1} and remained constant afterwards during fermentations associated with either of the plasmid-free strains. A glucose concentration of 1.7 g l^{-1} was determined at 3.5 h after inducer addition in case of the plasmid-bearing strain. Subsequently, the feed rate was repeatedly adjusted. However, glucose accumulation occurred again after approximately 15 h of recombinant protein production phase. Increased glucose values were not detected in fermentations associated with either of the plasmid-free strains. However, since glucose concentration increased constantly during the first 5 h of recom-

binant protein production, when the forward integration strain was employed, glucose accumulation was expected. Thus, the feed rate of this fermentation was reduced from 127.91 to 120 g l⁻¹ 5.5 h after inducer addition. Most on-line parameters did not differ in between fermentation experiments. Changes that were observed (*e.g.* exhaust and aeration mass flow values) corresponded to the feed rate adjustments.

In conclusion, plasmid-based expression of the *scFv-Z-1* gene resulted in high concentrations of insoluble product. On the other hand, amounts of soluble scFv-Z-1 were barely detectable. Furthermore, drawbacks associated with plasmid-based expression (*i.e.* glucose accumulation and growth deficiencies after inducer addition) were observed. As expected, reverse genomic integration, which was previously applied during production of FabX and FabZ resulted in decreased amounts of insoluble product, especially at the end of the fermentation. However, unlike described for FabZ and FabX (see Sections 3.1.2 and 3.1.3), reverse genomic integration did not significantly increase the soluble product concentration. Integration of the expression cassette into the forward strand resulted in further decreased insoluble intracellular product titers. While the other expression systems yielded no quantifiable amounts of soluble intracellular product, concentrations obtained with this system were clearly detectable. Genomic integration of the expression cassette into the forward strand of the *attTn7* site of BL21(DE3) seems to yield the expression system of choice for the production of soluble scFv-Z-1 under the applied conditions. Arguably, the reduced feed rate for this strain might have positively influenced soluble product formation. However, the first sample with clearly detectable bands during Western Blot analysis was already prepared 5 h after induction of recombinant protein production with IPTG. At that time point (prior to feed rate reduction), the soluble product-related band was already more intense than bands from all time points of fermentations associated with the alternative expression systems. As band intensities and, thus, volumetric concentration increased towards EoF, a prolonged fermentation process employing this strain might yield even higher soluble product titers. Due to glucose accumulation, this strategy does not appear promising in case of the two alternative expression systems.

3.1.5 Conclusion

Plasmid-free expression was shown to be beneficial in a previously published study (Striedner et al., 2010), suggested in a review article (Spadiut et al., 2014, Table 1) and was basis for a patent application (WO2008142028, Striedner et al. (2008)). In order to assess the technique’s potential for improving soluble protein production in the periplasm, expression systems were created applying the “complex cloning procedure” (Striedner et al., 2010) in the course of this study. Subsequently, plasmid-free production strains were tested in fed-batch fermentation processes and compared to their plasmid-bearing counterparts.

Initial recombineering approaches yielded “false-positive” colonies, indicating impurities in the expression cassette preparations and increasing efforts for identification of correct integration events. False-positive results after chromosomal integration attempts have been reported before, especially for large integration cassette constructs. They are believed to result from incompletely digested template vectors, spontaneous formation of resistance gene cassette-bearing vectors or non-directed integration of the expression cassette (Albermann et al., 2010; Datsenko and Wanner, 2000; Kuhlman and Cox, 2010). After method adjustment regarding dsDNA preparation, all integration attempts were successful and barely any “false-positive” colonies were detected. However, low numbers of recombinant (“correct”) colonies suggest a low integration efficiency. Low viable cell numbers prior to electroporation, inadequate outgrowth conditions, not enough dsDNA used for electroporation or a low time constant during electroporation were previously determined as possible reasons (Murphy et al., 2000; Sharan et al., 2009; Yu et al., 2000). Furthermore, it was reported that increased expression cassette length negatively influences the integration efficiency (Maresca et al., 2010). Examination of single steps of the procedure during the course of this study support these assumptions (data not shown). Primers containing homologous sequences to regions of both the forward and reverse strand of the *attTn7* site of the *E. coli* genome were used for integration cassette creation. As these are compatible with different vector templates frequently used at BI, standardized creation of GoI-bearing template plasmids, plasmid-bearing and plasmid-free production systems for the respective product is easily possible. To enable FRT-FLP-based removal of resistance gene modules after genomic integration, “FRT sites” were added to the kanamycin resistance module. Using cryo stocks of pSIM-bearing cells and an expression cassette-bearing plasmid and integration primers at hand, genomic integration (including cryo culture preparation) required approximately two weeks. Antibiotic resistance gene removal starting from cryo cultures required approximately an additional week. For establishing a highly efficient routine integration method, the above-mentioned bottlenecks should be

Table 3.3: Summary of maximum soluble product titers in relevant fractions obtained in fed-batch fermentation experiments using plasmid-bearing and plasmid-free expression systems. Improvements by using plasmid-free expression are presented as fold changes of maximum soluble product titers.

| Product | Conditions | System | Sol. Product | Improvement |
|----------|---|---|------------------------|-----------------|
| GFP | 37°C; inducer feed for 9 h (plasmid) or 23 h (genome) | Plasmid | 5.8 g l ⁻¹ | 1.4-fold |
| | | Genome | 8.0 g l ⁻¹ | |
| FabZ | HC-LC cassette; 37°C; inducer bolus after 14 h of feeding; * | Plasmid | 84 mg l ⁻¹ | 1.9-fold |
| | | Genome | 161 mg l ⁻¹ | |
| | | LC-HC cassette; 37°C; inducer bolus after 14 h of feeding | Plasmid | |
| Genome | 363 mg l ⁻¹ | | | |
| FabX | LC-HC cassette; 37°C; inducer bolus after 14 h of feeding; selective conditions | Plasmid | 80 mg l ⁻¹ | 4.3-fold |
| | | Genome | 340 mg l ⁻¹ | |
| | | Temperature shift prior to induction; inducer: arabinose (plasmid)/IPTG (genome); inducer feed; supernatant concentration | Plasmid | |
| Genome | 174 mg l ⁻¹ | | | |
| scFv-Z-1 | plasmid vs. forward integration; 37°C growth, 30°C production temperature; * | Plasmid | n.a. | 8.6-fold |
| | | Genome | n.a. | |

*Feed rate adjustment due to glucose accumulation in some of the fermentations.

FabX: Note that a limited set of data (direct comparison of two triplicate sets of fermentations performed in side-by-side experiments) is the basis for presented fold-changes. scFv-Z-1: improvement calculated based on densitometric analysis of Western Blot results.

addressed. Several alternative protocols, which might be characterized by higher integration efficiencies, already exist (Hillyar, 2012; Sawitzke et al., 2013; Sharan et al., 2009; Turan et al., 2013; Zucca et al., 2013). In total, twelve *E. coli*-based expression systems were created by recombineering and FRT-FLP technique prior to and throughout the course of this project (see Table 2.6).

Nine strains were employed in fed-batch fermentations for cytoplasmic (cyt) and periplasmic (PP) production of four different products (GFP (cyt)/FabX (PP)/FabZ (PP)/scFv-Z-1 (PP)). These products have different characteristics (mono- vs. multimeric, varying numbers of disulfide bonds). Different genetic constructs (LC-HC/HC-LC expression cassette) were employed during production of heterodimer FabZ. Results with focus on improvement induced by application of plasmid-free expression are presented in Table 3.3.

Application of genomic integration improved maximum soluble product titers between 1.4- and 8.6-fold, depending on the product. In this regard, expected improvements based on the report by Striedner *et al.* (Striedner et al., 2010) were proven. Production of soluble proteins in the periplasm constitutes a complex process (see, *e.g.*, bottlenecks presented in Table 1.2). It seems plausible that a lowered expression rate, which is reliably ensured by

genomic integration, is appropriate for efficient synthesis and secretion. Note that soluble product titers reported for GFP and FabX production using plasmid-bearing and -free strains are based on different cultivation processes. Furthermore, in two cases, glucose accumulation occurred in fermentations associated with plasmid-bearing strains and made unscheduled feed rate adjustments necessary. In conclusion, a dilemma arises when a plasmid-free strain is supposed to be compared to its plasmid-bearing counterpart due to their different demands concerning optimal process design. On the one hand, strains can be compared in equal processes, which might bias results towards the one expression system that fits better to the applied process. On the other hand, processes can be optimized separately for each expression system type. This would demand significant resources even for a small set of varied process parameters. Furthermore, such attempts, independent of the size of the experimental approach, do not guarantee an equally optimized process for each expression system. As the results presented here are based on both strategies, it is assumed that a profound basis for evaluation of the plasmid-free expression systems was guaranteed. While the presence of an expression plasmid resulted in process irregularities (*i.e.* glucose accumulation or unscheduled batch phase duration prolongation), fermentations of plasmid-free strains showed good reproducibility. This indicates an improved process stability and suggests that using plasmid-free expression systems allows “simple and stable” (Striedner *et al.*, 2010) upstream process design. Further advantages, *e.g.* increased expression cassette stability, reduced system leakiness and no negative influence of multi-copy plasmid maintenance, which were previously reported by Striedner *et al.* and Mairhofer *et al.* (Mairhofer *et al.*, 2013; Striedner *et al.*, 2010), were proven within this study. Detailed examination of the bottlenecks targeted by genomic integration are presented in Section 3.2. In conclusion, genomic integration unambiguously constitutes a valuable item in a toolbox for soluble recombinant protein production.

3.2 Genomic Integration Addresses Important Bottlenecks in Soluble Periplasmic Protein Production

When the recombinant protein is targeted to the *Escherichia coli* periplasm, several issues can impair the soluble product yield. A plethora of host engineering techniques for addressing these bottlenecks is available (see Section 1.3.4 and Table 1.2). As suggested in a review article from 2014, genomic integration of the gene of interest is a suitable option for generation of “stable and efficient” production strains (Spadiut et al., 2014, Table 1). Within the course of this study, it was shown that plasmid-free expression improved soluble product yields of the model proteins GFP, FabZ, FabX and scFv-Z-1 (see Section 3.1). In previous articles, plasmid-free systems were compared with their plasmid-bearing counterparts and reasons for improved production performance were proposed (Mairhofer et al., 2013; Striedner et al., 2010). Based on in-depth characterization of producer strains, Mairhofer *et al.* identified several relevant differences between plasmid-free and plasmid-bearing expression systems (Mairhofer et al., 2013). High-rate transcription of multiple copies of the product gene paired with transcriptional read-through due to inefficient transcription termination by the T7 terminator led to the presence of long heterologous mRNA molecules. This constituted a major metabolic burden in plasmid-based systems. In consequence, translation of host-intrinsic mRNA molecules was significantly reduced. Selective pressure against productive cells resulted in an impaired growth behavior and/or plasmid loss (Striedner et al., 2010). Plasmid-free strains, on the other hand, showed a reduced stress response on the transcriptional level and decreased cell growth impairment due to recombinant protein production. Moreover, decreased system leakiness and no lethal consequences for the host metabolism despite high rates of recombinant protein production characterized the T7 production systems. In summary, a higher capacity for recombinant protein production was reported for plasmid-free systems in previous studies (Mairhofer et al., 2013; Striedner et al., 2010) and above (see Section 3.1). In this section, plasmid-free and plasmid-bearing strains were compared considering the bottlenecks of the synthesis path of recombinant proteins depicted in Figure 1.9 (p. 19). Expression cassette stability, influence of the plasmid as such and system leakiness were examined. Harmonization of expression rate and secretion capacity was also assessed. As previous studies compared expression systems for cytoplasmic products, the impact of genomic integration concerning secretory production has not yet been examined.

3.2.1 Expression Cassette (In)Stability

The presence of the heterologous gene expression cassette is fundamental for recombinant protein production. Formation of potentially toxic gene products might impose a major burden on productive cells, constituting an advantage for non-productive cells. The resulting plasmid loss of plasmid-based expression systems interfered with efficient cytoplasmic production of proteins in a previous study (Striedner et al., 2010). Reasons for plasmid loss and its kinetics were reviewed by Summers in 1991 (Summers, 1991). Several methods to prevent overgrowth of non-productive cells exist (Friebs, 2004, Section 4.2). In this part of the study, it was examined if plasmid loss constituted a relevant bottleneck under the applied conditions. Furthermore, plasmid-free systems were examined concerning expression cassette stability.

During fed-batch fermentations for FabZ production (see Table 3.2, HC-LC/LC-HC expression cassette, fermentation process 1), samples were prepared to assess expression cassette stability. During the analysis procedure (see Section 2.5), cells were plated on non-selective medium. Resulting colony numbers were used to determine the culturable cell count. The ratio of resistant cells was determined by replica plating of these colonies to kanamycin-containing (step 1) and non-selective (step 2) agar plates. As the resistance gene is part of the expression cassette, this ratio was considered as the fraction of productive cells.

Independent of the used expression system, similar culturable cell counts were determined prior to induction of recombinant protein production (approximately 10^{10} cfu ml⁻¹, see Figure 3.11). In case of the genome-integrated HC-LC expression cassette, culturable cell numbers were constant until 11 h after induction and decreased, subsequently, towards EoF. In case of the plasmid-bearing counterpart, values decreased rapidly until 5 h after induction (10^6 - 10^7 cfu ml⁻¹). Subsequently, increasing values were determined towards EoF (10^8 - 10^9 cfu ml⁻¹). The fermentation of bacteria containing LC-HC expression cassette-bearing plasmids were characterized by constant culturable cell numbers during the whole time of induction (10^9 - 10^{10} cfu ml⁻¹). Values for plasmid-free counterparts showed a decrease towards approximately 10^5 cfu ml⁻¹ at EoF. Experiments employing plasmid-bearing cells were characterized by a significant number of expression plasmid-free cells already prior to induction (30 - 40 % plasmid-free cells). Within 6 h after induction, the fraction of plasmid-bearing cells even decreased to below 10 % and remained only slightly above or at 0 % throughout the process. In accordance with a previous study (Striedner et al., 2010), expression cassette stability of genome-integrated strains was determined to be 100 % for all time points.

In conclusion, dramatic overgrowth by non-productive plasmid-free cells (“plasmid loss”) was detected under non-selective conditions in case of cells that carried the GoI expression cassette on a multi-copy plasmid. Apparently, presence of the expression cassette exerts a highly negative selective pressure. In case of the plasmid-borne HC-LC expression cassette, increasing culturable cell numbers in combination with the lack of plasmid-bearing cells after approximately 5 h of protein production suggest overgrowth of unproductive cells. The LC-HC expression cassette probably led to an increased basal expression, which resulted in a prolonged batch phase (see Section 3.2.3). When the LC-HC expression cassette was applied, no glucose accumulation was observed. This suggested that productive cells might have been overgrown by non-productive cells earlier. Due to non-productive cells’ increased glucose metabolization capabilities, glucose accumulation might have been prevented. For cytoplasmic production of recombinant proteins, genomic integration was reported to prevent loss of expression cassettes (Striedner et al., 2010). Comparable results were obtained during fermentation experiments for production of FabZ in the *E. coli* periplasm. Furthermore, culturable cell numbers decreased more significantly if stable genomic integration was combined with the higher-yielding expression cassette (LC-HC). This suggested that decreased cell viability can be directly correlated with increased cellular stress due to the production of recombinant protein.

Note that, in case of the LC-HC expression cassette, the feed rate was reduced to approximately half of the planned rate after the 6 h of sampling time point to avoid glucose accumulation during unattended phases of the fermentation process (see “4” in Figure 3.11). Further, no sampling could be performed in the unattended process phase, resulting in the lack of one sampling time point for this strain (see “2” in Figure 3.11). In a repeated fermentation without glucose feed reduction, similar progress of off-, at- and on-line values was observed. Thus, a similar value for the culturable cell count and the ratio of retained expression cassettes was expected. When the plasmid-based HC-LC expression system was used, glucose accumulation was detected approximately 5 h after induction (see “3” in Figure 3.11). The feed rate was reduced in regular intervals according to the procedure described in Section 2.4.2. Apparently, glucose accumulation was a result of the combination of employed strain and process parameters. It was also observed in a replicate fermentation. Thus, it was considered a strain-inherent result rather than an unexpected deviation. Note also that, in previous reports, it was stated that the number of cells, which are able to form colonies on agar plates (“culturable cells”), does not necessarily correlate with the concentration of viable and productive cells in the fermenter. The occurrence of “viable but non-culturable” cells (VBNC) appears to be a response of non-sporeforming bacteria in case of (environmental) stress (Colwell, 2000; Oliver, 2005;

Roszak and Colwell, 1987; Sundström et al., 2004). This also seems to be the case in these experiments.

For example, the observed drop in culturable cells for certain strains did not correlate with DCW courses (*cf.* Figures 3.11 and 3.5/3.6). It was previously reported that increases in biomass despite constant cell numbers can derive from increasing cell size during production of recombinant proteins (Hoffmann and Rinas, 2004). This means that culturability potentially decreased with increasing productivity. Thus, mainly cells without expression cassette retain the ability to be cultivated in the first step of the plasmid retention analysis. Accordingly, this procedure was previously considered not optimal for quantifying plasmid loss, since the plating method requires cells to be able to grow on agar plates (Hoffmann and Rinas, 2004; Patkar et al., 2002). In conclusion, the analysis method might be biased towards increased numbers of expression cassette-free cells. However, all strains were treated similarly and the discussion was based on trends rather than exact values. Thus, the results obtained from the selective plate count procedure appear valid for the drawn conclusions. Only one colony on one of the nine plates was obtained during culturable cell number determination at EoF for the strain bearing a genome-integrated LC-HC expression cassette (see “1” in Figure 3.11). Accordingly, a calculation of the culturable cell number is not reliable. However, since two of the replicate plates for this dilution did not yield any colonies, actual culturable cell numbers are expected to be even lower than the indicated value of 10^5 cfu ml⁻¹. Due to lack of colonies, ratio of cells bearing the expression cassette could not be determined.

Expression cassette stability was expected to be dependent on the recombinant target protein. Thus, plasmid loss was also assessed during production of model protein FabX. At the end of a fermentation, in which a plasmid-bearing FabX production system (araBAD promoter, T7 terminator, pBR322 ori, kanamycin resistance module) was employed, the plasmid stability was examined. Strain and process parameters were the results of process and host optimization for FabX production within the corresponding development project. Likewise, expression cassette stability was analyzed at EoF time points of various experiments, in which plasmid-free expression systems for FabX production (T7 promoter, expression cassette integrated into the *attTn7* reverse strand of *E. coli* BL21(DE3)) were employed. For analysis of expression cassette stability, 300 colonies grown on non-selective agar-plates were transferred to agar plates with and without kanamycin. The division of colony numbers from selective and non-selective plates yielded the “plasmid stability” value.

At the end of the fermentation, in which the plasmid-based system was employed, a

plasmid stability of 73 % was determined. In case of two alternative plasmid-bearing expression systems tested in fermentation experiments during host improvement efforts, plasmid stability values of 46 and 73 % were determined. When plasmid-free production systems were used, a constant expression cassette stability of 100 % was determined in all cases. In conclusion, loss of the expression cassette was observed in case of plasmid-bearing systems and could successfully be prevented by application of plasmid-free host strains. However, the extent of plasmid loss was not as pronounced as observed during FabZ experiments. Possibly, the reduced production temperature and the weaker araBAD promoter prevented further decreased plasmid stability values.

Plasmid loss constituted a relevant drawback of the employed plasmid systems. Genomic integration technique increased the expression cassette stability to 100 % for all strains. In conclusion, genomic integration provides an important tool for production of recombinant proteins, whose synthesis might exert negative selective pressure for productive host cells.

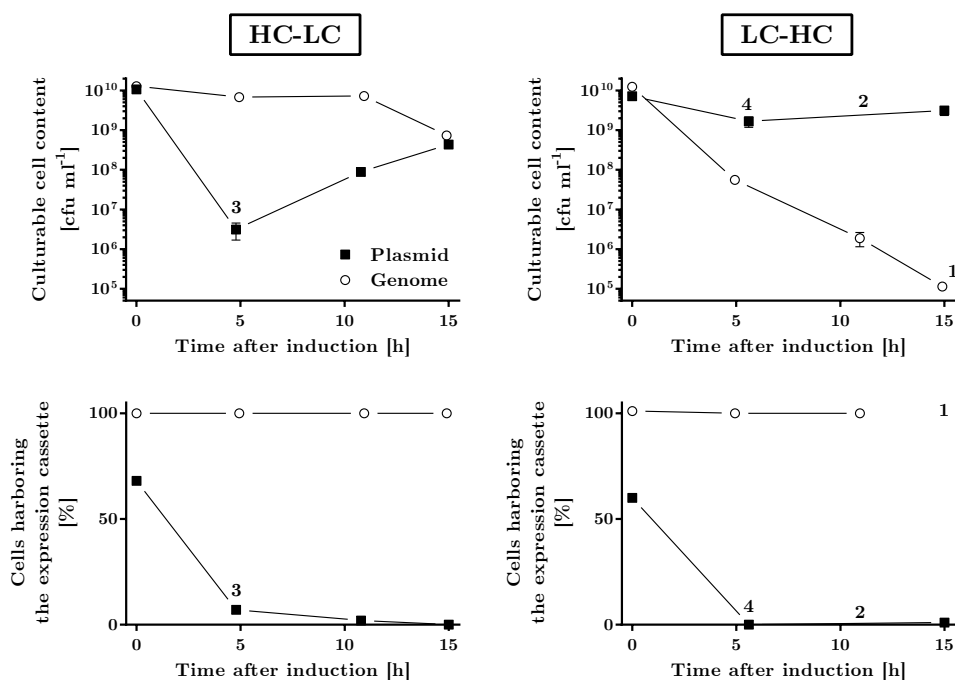


Figure 3.11: Culturable cell content (top) and fraction of cells harboring the expression cassette (and, thus, the GoI) (bottom) plotted against time after induction of recombinant protein production. Four different strains subjected to fed-batch fermentation experiments. Two FabZ gene expression cassette alternatives (HC-LC and LC-HC, see headings) were used in both the plasmid-based (filled squares) and the genome-integrated (blank circles) expression system. Cultivations were performed at 37°C under non-selective conditions. Recombinant protein production was induced by bolus addition of IPTG after 14 h of feeding, resulting in 15 h of recombinant protein production. Culturable cell counts were determined by plating different dilutions of fermentation suspension on non-selective S-LB agar plates and subsequent colony counting. The percentage of cells with expression cassette was calculated based on replica plating on kanamycin-containing and non-selective S-LB agar plates. Culturable cell numbers represent mean values of colony counts from 1 to 3 dilutions per sampling time point and 3 replicates per dilution. Error bars represent the standard deviations of the indicated values. Numbers in the figure are associated with unexpected incidents. Their significance is discussed in the text. 1: Culturable cell count was approximately 10^5 cfu ml⁻¹. Due to lack of colonies, no determination of expression cassette stability was possible. 2: Due to prolonged batch phase, the process was unattended during scheduled sampling time points. Thus, no samples could be prepared. 3: Due to glucose accumulation, the feed rate was reduced at this time point. 4: To avoid expected glucose accumulation, feed rate was reduced after 6 h of recombinant protein production.

3.2.2 Influence of Multi-Copy Plasmid Presence

Production of recombinant proteins can have a significant influence on the *E. coli* host cell (see Section 1.3.3). However, the mere presence of multi-copy plasmids has previously been reported to also significantly affect the host cell's metabolism (Ow et al., 2009, 2006; Wang et al., 2006). Perturbations could, in part, be associated with the expression of the antibiotic resistance gene (Rozkov et al., 2004) or presence of other regulatory molecules (Pasini et al., 2016). As discussed in Section 3.2.1, genomic integration of the target gene reliably ensures the presence of the expression cassette. It is also expected to prevent negative effects mediated by the plasmid backbone. This is because, in addition to the expression cassette, only a single copy of the resistance gene (and no other plasmid-associated elements) is integrated into the host genome.

Within this part of the study, the influence of basic plasmid pBI1KT7.1 during fermentation experiments was examined. Furthermore, the influence of a single copy of the kanamycin resistance gene was assessed in fermentations for GFP production.

Plasmid presence affects cell behavior after addition of the inducer IPTG

In a previous part of this study, fermentations were performed with the goal to compare plasmid-bearing and plasmid-free FabZ production systems (see Section 3.1.2). In one of the experimental set ups, BL21(DE3) cells bearing the empty plasmid pBI1KT7.1 were employed in addition to the plasmid-free and plasmid-bearing production strains (B<FabZ-kan>(02) and B(FabZ)(02)). Plasmid pBI1KT7.1 is comparable to pET30(+) and carries a pBR322 ori, a kanamycin resistance gene cassette and the T7 transcription regulation module (T7 promoter, lac operator binding site, *lacI* gene, T7 terminator). The multiple cloning site (MCS) was engineered to be compatible with local cloning strategies. It is representative for the characteristics of the pET plasmid series (Novagen). These plasmids are commonly used for recombinant protein production in *E. coli*. Details on the comparison of plasmid-bearing and plasmid-free production systems are presented in the respective section. In this section, elaborations focus on observations related to the influence of the basic plasmid. As no product-related values were obtained for the empty “mock” strain, dry cell weight (DCW), optical density at 550 nm (OD₅₅₀) and glucose concentration measurements were compared (see Figure 3.12).

The duration of the batch phases was similar for plasmid-free and the plasmid-bearing mock reference strain. In comparison, the feed phase started approximately 6 h later in

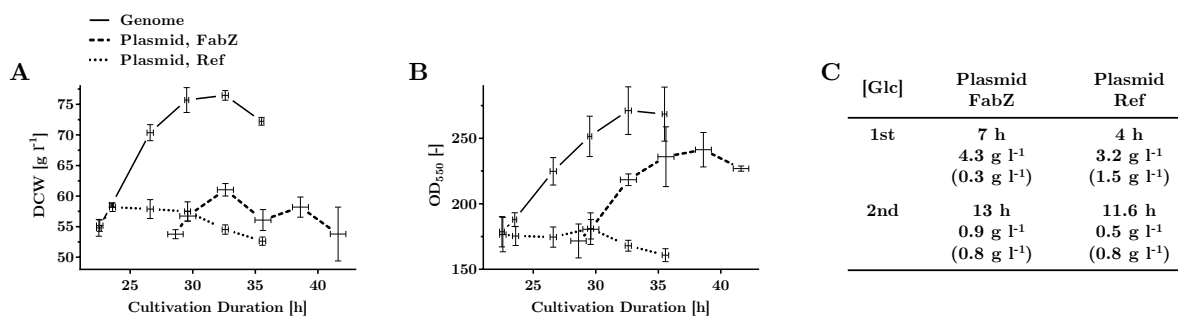


Figure 3.12: Dry cell weight (DCW, **A**), OD_{550} (**B**) and glucose concentrations (**C**) were measured during fermentations of plasmid-free and plasmid-bearing FabZ production strains and a strain bearing the empty plasmid. Fed-batch fermentations were performed as described in Sections 2.4.2 and 3.1.2. Values were determined immediately prior to and during the phase of recombinant protein production. “Genome” (solid lines): plasmid-free FabZ production strain B<FabZ-kanR>(02). “Plasmid, FabZ” (dashed lines): plasmid-bearing FabZ production strain B(FabZ)(02). “Plasmid, Ref” (dotted lines): mock reference strain B(pBI), bearing the empty basic plasmid pBI1KT7.1. Mean values are plotted in dependence of cultivation duration. Error bars represent standard deviations based on measurements from three fermentation experiments for each strain. **C**: Increased glucose concentrations were observed twice in fermentations of plasmid-bearing FabZ production and mock reference strain. Time points after inducer addition (in h), mean glucose concentrations and standard deviations of triplicate measurements (in g l⁻¹, without and with parentheses, respectively) are presented.

case of the plasmid-bearing FabZ producer (see Figure 3.13 in Section 3.2.3). Based on DCW values (see Figure 3.12 (A)), growth in fermentation experiments continued after inducer addition in case of the plasmid-free production strain. Growth ceased after induction for both the producing and reference empty plasmid-bearing strain. When the plasmid-free strain was employed, OD_{550} values increased during until 3 h prior to EoF and remained constant afterwards (mean EoF value: 269) (see Figure 3.12 (B)). In case of the plasmid-bearing FabZ-production strain (LC-HC expression cassette), values increased, peaked ($OD_{550} = 240$) and decreased towards EoF (mean at EoF: 227). When the empty plasmid-bearing reference strain was employed in fermentation experiments, OD_{550} values remained approximately constant for several hours after inducer addition and, subsequently, decreased (mean EoF value: 161). For the productive plasmid-bearing strain, OD and DCW values did not correlate with each other (compare lines for “Plasmid, FabZ” in left and center graph of Figure 3.12). No glucose accumulation was detected in fermentations associated with the plasmid-free strain. However, when one of the two plasmid-bearing strains was employed, glucose concentrations above a threshold of 0.5 g l⁻¹ were measured, albeit at different time points after induction (see Figure 3.12 (C)). On-line data obtained from fermentation experiments (*e.g.* dissolved oxygen, stirrer speed, aeration data) did not reveal unambiguous trends concerning the used strains (data not shown).

Negative effects of multi-copy plasmids are usually related to plasmid maintenance (Ow

et al., 2009, 2006; Wang et al., 2006) or excessive heterologous gene expression. This can result in drainage of host resources (Mairhofer et al., 2013), uncontrolled increase in plasmid copy number (Striedner et al., 2010; Teich et al., 1998; Wrobel and We, 1998; Yavachev and Ivanov, 1988) and loss of the plasmid due to negative selective pressure (Popov et al., 2011). In the experiments described above, the presence of the empty plasmid did not show any effect prior to inducer addition (compared to the plasmid-free strain). After IPTG addition, parameters of fermentations of plasmid-bearing FabZ producer and reference strain (*e.g.* growth determined by DCW progress and glucose concentration) were similar. Strikingly, in fermentations of the unproductive empty plasmid-bearing strain, glucose accumulation occurred 3 h earlier compared to fermentation of the GoI-carrying counterpart. Several effects might explain the unexpected observation. As reported previously, limited T7 terminator efficiency can allow read-through by the T7 RNA polymerase (Mairhofer et al., 2014). Under induction conditions, GoI transcription is initiated and the corresponding mRNA is formed. This occurs independently of the DNA sequence which is located downstream of the transcription start. Thus, similar negative effects caused by read-through are supposed to occur in case of empty expression plasmid presence. Since the length of the DNA between transcription start and terminator is shortened by approximately 1.5 kb in case of an empty MCS, the RNA polymerase might reach the terminator sequence with higher probability. Thus, assuming a non-optimal termination efficiency, the percentage of read-through events might be increased in the strain bearing the empty plasmid. A reason for delayed glucose accumulation in fermentations of productive plasmid-bearing strains might be an increased demand of the host strain caused by formation of the recombinant protein. For the plasmid-bearing production strain, progress of DCW and OD₅₅₀ values did not correlate with each other. An influence of the formed inclusion bodies on the measurement of optical density in cell suspensions was previously reported (Sundström, 2007). In case of the plasmid-bearing FabZ production strain, OD₅₅₀ values increased for approximately 10 h after induction (39 h of cultivation), while DCW values remained constant after 4 h (33 h after start of cultivation). Inclusion body formation constitutes a potential reason for this observation. Exact reasons for the behavior of the empty plasmid-bearing strain after induction remain elusive at this stage. Examinations of, *e.g.*, plasmid copy number or the extent of read-through events, would potentially reveal relevant information concerning influence of the basic plasmid. In summary, the generated data indicated a negative influence on growth and productivity due to the mere presence of an empty plasmid in plasmid-based expression systems. However, the influence of the basic plasmid became apparent only after induction of high-rate transcription.

Influence of resistance gene expression

Presence of resistance genes (“markers”) in multiple copies was previously shown to hamper plasmid production for gene therapy in *E. coli* (Mairhofer et al., 2010). Another study showed an influence on growth rate and biomass yield of *E. coli* strain DH1 (Rozkov et al., 2004). Several systems for antibiotic-free selection were developed (*e.g.* Mairhofer et al. (2010)) and this alternative selection technique was reviewed in 2013 (Oliveira and Mairhofer, 2013). Presumably, marker-free systems could be relevant for protein production as well, especially when multi-copy plasmids are used for protein production. To examine the influence of the marker gene expression on plasmid-free recombinant protein production strains, results from two previous fermentation experiments (see Section 3.1.1) were compared. Both strains bear a genomically integrated *gfp* expression cassette. While B<kanR-GFP> bears the module for expression of the aminoglycoside 3'-phosphotransferase III gene (conveying kanamycin resistance), the resistance marker module was excised from the *E. coli* genome in case of B<GFP>. Both strains were employed in identical fed-batch fermentation processes.

At EoF, a soluble product titer of 3.2 g l^{-1} and a ratio of soluble to total product of 99 % was determined for the kanamycin-resistant strain B<kanR-GFP>. The marker-free strain yielded 3.6 g l^{-1} of soluble product and a ratio of 98 %. No significant differences in growth of both strains were detectable. Glucose concentrations remained below 0.25 g l^{-1} after inducer addition for both strains. Examination of on-line fermentation data did not reveal significant differences between both strains, either.

The presence of a single copy of a resistance gene did not significantly influence strain performance. However, more extensive experiments (*e.g.* with different recombinant target proteins) would be necessary to prove this preliminary conclusion. As other issues seemed of higher importance, marker-free hosts were not further examined in this study. After stable genomic integration, expression cassette loss is unlikely (see Section 3.2.1). Thus, excising the antibiotic resistance gene after genomic integration is possible if necessary for certain applications (*e.g.* genomic integration of another expression cassette).

3.2.3 Prevention of Target Gene Basal Expression

Application of genomic integration for expression system design was shown to prevent loss of the expression cassette and drawbacks affiliated with plasmid presence (see Sections 3.2.1 and 3.2.2, respectively). Even though these plasmid-specific drawbacks can be addressed with alternative techniques, expression from multi-copy plasmids might still cause undesired basal expression of the target gene. Shortly after development of the T7 polymerase-based expression system for production of recombinant proteins in *E. coli* (Studier and Moffatt, 1986; Studier et al., 1990), basal T7 polymerase synthesis in BL21(DE3) (*i.e.* presence prior to induction) was a known problem. This issue was addressed by insertion of *lac* operator sites near the T7 promoter sequence (Dubendorf and Studier, 1991) or using the T7 lysozyme to reduce T7 RNA levels (Studier, 1991). Basal expression of the GoI during production of recombinant proteins is a commonly known drawback and was, *e.g.*, mentioned in a previously published review article (Saida et al., 2006). Assuming that certain antibody fragments or intermediates and aggregated products might exert a negative or even toxic effect on the host cell, basal expression constitutes a problematic characteristic that can make process design cumbersome. In this section, the heterodimeric model protein FabZ, which bears five disulfide bonds, was used to examine the potential of genomic integration to avoid basal expression by reducing the gene copy number to one. Relevant results from previously performed experiments (see Section 3.1.2) are presented briefly and will be discussed in the light of basal expression. FabZ-producing strains were employed in fed-batch fermentations. Specifically, BL21(DE3) cells bearing HC-LC/LC-HC gene expression cassette integrated into the *attTn7* reverse strand of their genome or on multi-copy plasmid pBI1KT7.1 were used (see Table 3.2, fermentation processes 1 and 3). The empty plasmid-bearing reference strain B(pBI) (BL21(DE3) transformed with pBI1KT7.1) was also employed. Results are presented in Figure 3.13.

Plasmid-based production of FabZ resulted in soluble intracellular product concentrations of more than 5 mg l⁻¹ prior to induction of GoI expression with IPTG. This was seen in two different process formats and for both expression cassette organizations (see Figure 3.13 (A)). Application of plasmid-free expression systems resulted in less than 1 mg l⁻¹ prior to IPTG induction. Batch phases of fermentations of plasmid-bearing strains lasted up to 18 h (see Figure 3.13 (B)). Durations of below 9 h were observed for both genome-integrated expression strains. Cultivation of the empty plasmid-bearing strain also resulted in a batch phase of approximately 9 h. While basal expression seemed to influence process performance in FabZ experiments, no batch phase prolongation was

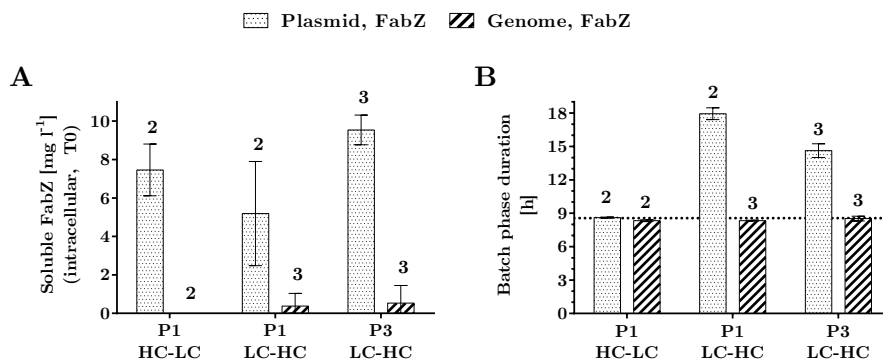


Figure 3.13: Summary of data from FabZ production experiments measured immediately prior to inducer addition. “P1” and “P3” indicate the used process format. “HC-LC” and “LC-HC” represent different expression cassette organizations. Intracellular soluble FabZ concentrations (**A**) and cultivation duration prior to start of feed addition (“batch phase duration”, **B**) are shown. Dotted and cross-hatched bars represent values associated with the plasmid-bearing and the genome-integrated FabZ producer, respectively. The dotted line in **B** represents the mean batch duration obtained from triplicate fermentations of a strain bearing the empty basic plasmid. Numbers of performed replicate fermentations are indicated above the respective bars. Error bars indicate SD values.

observed when plasmid-bearing strains were employed in scFv-Z-1 production experiments (data not shown). Furthermore, during plasmid-based FabX production, in which a reduced temperature was employed and the araBAD promoter was utilized, no basal expression was observed, either.

Thus, basal expression seems to constitute a relevant plasmid-related drawback in certain, but not all combinations of expression host and process parameters. In cases, in which basal expression issues occurred when plasmid-bearing strains were employed, using plasmid-free systems prevented it.






3.2.4 Harmonization of Expression and Secretion Rate

Thus far, bottleneck examinations in this study focused on expression cassette stability, influence of the basic plasmid and basal expression. Genomic integration mitigated issues in the context of all considered parameters. Secretion across the inner membrane can also constitute a major issue (see Section 1.3.4 and Figure 3.15). The translational level of the target gene was shown to be critical for efficient secretion (Simmons and Yansura, 1996). Thus, a transcription rate attenuated by single-copy genomic integration might be able to harmonize expression and secretion rate. The question if genomic integration facilitates de-bottlenecking of the secretion process was investigated within this part of the study. To target this question, the location of product aggregates was determined. EoF cell samples of selected fermentation experiments were lysed and soluble and insoluble product fractions were separated by centrifugation. Insoluble fractions were resolubilized under reducing and denaturing conditions by vigorous shaking. Subsequently, insoluble fractions were analyzed by SDS PAGE. Gel bands corresponding to the product were excised, purified and N-terminally sequenced using Edman degradation (Edman, 1950, 1956). Consolidated results are presented in Table 3.4. If the N-terminal amino acid sequence corresponded to the sequence of the leader peptide SP_{ompA} (M-K-K-T-A-I-A-I-A-V-A-L-A-G-F-A-T-V-A-Q-A), aggregates were assumed to originate from the cytoplasm. If, on the other hand, the N-terminal amino acid sequence corresponded to the mature product, cleavage by the periplasmic signal peptidase was assumed to have taken place and, thus, aggregates supposedly derived from the periplasm.

FabZ

In case of model protein FabZ, the maximum total soluble product titer obtained in a selected set of fermentation experiments was improved 4.4-fold by application of the genomic integration technique (see Table 3.3, FabZ, second condition). The pBI1KT7.1 plasmid-bearing and the plasmid-free strain, which were employed in these experiments, bore a dicistronic LC-HC expression cassette. In addition to examination of the soluble and correctly folded FabZ concentration, SDS PAGE analysis was carried out to determine the total FabZ content. A total product concentration in the g l⁻¹ range was observed for all selected fermentation experiments, while the concentration was reduced in case of plasmid-free strains (see Section 3.1.2, p. 92). The difference between total (several g l⁻¹) and soluble intracellular product (< 300 mg l⁻¹), presumably represents the amount of insoluble aggregates. To identify reasons for the increased soluble product amount in the

Table 3.4: N-terminal sequencing results of protein aggregates from FabX, FabZ and scFv-Z-1 production experiments. Cells from fermentation experiments were disrupted and the insoluble and soluble protein fractions were separated and analyzed by SDS PAGE. Product bands were excised, purified and subjected to Edman degradation (Edman, 1950, 1956). The determined N-terminal amino acid sequence is presented in single letter code. Amino acids in parentheses are tentative assignments or assignments of two amino acids to the same position. In case of more than one assignment per position, the amino acids are separated by a slash. Question marks indicate that no amino acid was identified. Italic letters represent minor signals. Images of excised bands are presented in the “Band(s)” column.

| Product | GoI Locus | Band(s) | Sequence |
|----------|--------------|---|--|
| FabX | Plasmid | n.a. | (D/E)-(V/I)-Q-L-(T/V)-(Q/E)-S-(G/P)-(S/G)-(S/G)-L |
| FabZ | Plasmid |  | M-K-K-T-A-I <i>(E)-V-Q-L-V-E</i> |
| | Genome (rev) |  | ?-I-V-L-T-Q (upper band) E-(V/I)-Q-L-(V/T)-(Q/E) (lower band) |
| scFv-Z-1 | Plasmid |  | M-K-K-T-A-I <i>E-V-Q-L-V-E</i> |
| | Genome (rev) |  | (M/E)-(K/V)-(K/Q)-(T/L)-(A/V)-(I/E) <i>?-G-R-G-?(P/L)</i> |
| | Genome (fwd) |  | E-V-Q-L-V-E <i>?-L-A-T-?-?</i> |

genome-integrated host strain and to generate new strategies for further host engineering approaches, aggregates were analyzed concerning their cellular location. EoF cells samples from each strain were subjected to the procedure described above. In case of the plasmid-free expression system, two bands, presumably originating from LC and HC monomer, could be separated and were analyzed individually. For the protein sample from the plasmid-based host strain, only a single broad band could be detected, which was excised from the gel and analyzed as one sample.

In case of the sample from the plasmid-bearing strain, the main N-terminal amino acid sequence was determined to be M-K-K-T-A-I. Additionally, a minor signal (E)-V-Q-L-V-E was observed. No other sequences (with signals of above 20 % with regard to the main sequence) were observed. In case of the “upper band” of the plasmid-free strain sample, one main sequence (?-I-V-L-T-Q) was determined. Analysis of the “lower band” sample revealed the mixed main signal (E-(V/I)-Q-L-(V/T)-(Q/E)). No minor sequences (above 15 % with regard to the main sequence) were observed in samples from fermentations based on plasmid-free strains.

The main signal in the sample from the plasmid-bearing strain corresponded to the SP_{ompA} sequence (M-K-K-T-A-I). A low ratio of the analyzed sample also seemed to be derived from the mature HC monomer (E-V-Q-L-V-E). The mature LC monomer (E-I-V-L-T-Q) was detected in the “upper band” sample derived from the fermentation employing

the plasmid-free strain. The “lower band” seemed to derive from a mixture of LC and HC monomer. In conclusion, the application of the plasmid-bearing expression system resulted in cytoplasmic aggregates. When the plasmid-free expression system was employed, no SP_{ompA}-related sequences were detected. This suggested that the broad majority of target protein aggregates forms in the periplasm. Apparently, a lowered protein synthesis rate induced by single-copy genomic integration prevents overburdening of the Sec translocon and enables significantly improved protein secretion. However, a major ratio of the total product still aggregates, albeit in the periplasm. Furthermore, it was confirmed that the upper band on SDS gels originated from the FabZ LC monomer. Unlike previously stated (see above), the lower band seems to be derived from a mixture of LC and HC monomers and not from the heavy chain alone.

scFv-Z-1

Different strains were employed in fermentation experiments for production of model protein scFv-Z-1. The reverse genome-integrated strain yielded decreased total scFv-Z-1 concentrations in comparison to its plasmid-bearing counterpart. However, soluble product titers were not significantly increased. A strain created by forward integration yielded further decreased total product titers. Unlike the reverse integration strain, this one yielded significantly increased amounts of soluble scFv-Z-1. In conclusion, it appeared that the expression cassette orientation of heterologous genes in the bacterial genome had an influence on the protein synthesis rate. In that regard, forward genomic integration seemed to constitute the enabling step for soluble scFv-Z-1 production. In the previous section, it was shown that single-copy genomic integration of the expression cassette could prevent secretion machinery overburdening in case of FabZ production. In this part of the study, it was examined if a similar phenomenon facilitated increased soluble scFv-Z-1 yield. A plasmid-bearing and two plasmid-free strains (forward and reverse genome-integrated expression cassette) were employed in fed-batch fermentation experiments (see Section 3.1.4). Insoluble protein fractions of EoF cell samples were subjected to N-terminal sequencing (see Table 3.4). Furthermore, samples from these fermentations were subjected to SDS PAGE analysis to examine the patterns of intracellular proteins (see Figure 3.14).

In aggregated protein fractions of the plasmid-bearing strain, the main sequence was determined to be M-K-K-T-A-I. A minor signal of E-V-Q-L-V-E was detected. No other signals (above 13 % with regard to the main sequence) were detected. When the plasmid-free strain with reverse-genome-integrated expression cassette was employed, a mixture of

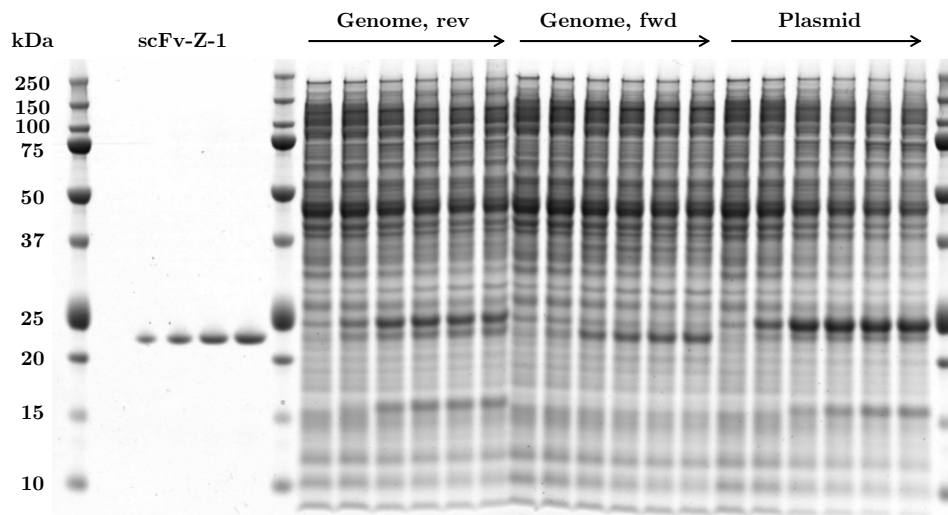


Figure 3.14: SDS PAGE analysis of samples from three scFv-Z-1 production experiments. Strains were created by genomic integration of the *scfv-Z-1* expression cassette into the forward (“Genome, fwd”) or reverse (“Genome, rev”) strand of the *E. coli* genome. Moreover, a pBI1KiT7.3-based plasmid expression strain was used (“Plasmid”). Fed-batch fermentation experiments were performed as described in the text. Samples were prepared under native conditions as target proteins with intact disulfide bonds were expected to possibly form a separate band. The gel was stained by Coomassie Brilliant Blue. Different amounts of purified scFv-Z-1 molecules (approximately 25 kDa) were applied as size standard of the product. Arrows indicate progress of the recombinant protein production phase (0 - 13 h).

two sequences ((M/E)-(K/V)-(K/Q)-(T/L)-(A/V)-(I/E)) was identified in the aggregated protein fraction. A minor signal was observed, but could not unambiguously be assigned to a specific amino acid sequence. No alternative signals (above 5 % with regard to the main sequences) were detected. A single main sequence of E-V-Q-L-V-E was determined when the sample associated with the “forward-integrated” strain was analyzed. In addition, a minor signal was obtained. However, its sequence could not be assigned to a certain amino acid sequence. No other sequences (above 3 % with regard to the main sequence) were detected. By SDS PAGE analysis, one or two bands were observed after induction of recombinant protein production (see Figure 3.14). One band appeared at the same molecular weight (MW) as the scFv-Z-1 reference molecule (“lower band”). The other band appeared above this first product-associated band (“upper band”). Samples from the plasmid-bearing strain exhibited two bands. The strength of the upper band was more pronounced compared to the lower band (EoF: upper band 7.7-fold more intense compared to the lower band). In samples from the reverse genome-integration strain, the upper and lower band intensity was decreased and increased, respectively. At EoF, densitometric measurement revealed a ratio of 1.9 in favor of the upper band. SDS PAGE analysis of samples associated with fermentations, in which the forward-integration strain was employed, revealed the most pronounced lower bands. Values determined by densitometric analysis of the upper band ranged at or below values measured in samples prepared prior

to induction of recombinant protein production.

Table 3.5: Summary of results from N-terminal sequencing and SDS PAGE analysis. Samples from fermentations employing different scFv-Z-1-producing strains were analyzed. See Table 3.4 and Figure 3.14 for separate results of the analysis procedures.

| scFv-Z-1 producer | N-terminal sequencing | SDS PAGE analysis |
|---------------------|---|-----------------------|
| Plasmid-bearing | Cytoplasmic aggregates | Pronounced upper band |
| Reverse-integration | Mixture: cytoplasmic and periplasmic aggregates | Double band |
| Forward-integration | Periplasmic aggregates | Pronounced lower band |

Considering results from N-terminal sequencing and SDS PAGE analysis (see Table 3.5) the upper band obtained on SDS gels seemed to represent the cytoplasmic fraction of the aggregated product. The lower band seemed to represent the periplasmic share (mature protein after cleavage of SP_{ompA}) of the insoluble protein fraction. In samples of the plasmid-bearing strain, N-terminal sequencing of intracellular aggregates revealed the presence of the leader SP_{ompA}. Furthermore, a strong upper band, associated with cytoplasmic aggregates, was observed. On the other hand, the lower band was less pronounced. In conclusion, the secretion machinery seemed to be overburdened, leading to aggregation of the target product already in the cytoplasm. Reverse integration of the scFv-Z-1 gene expression cassette and application of the resulting strain in fermentation experiments seemed to result in a mixture of aggregates located in the cytoplasm and in the periplasm. It appeared that the reverse genome-integrated position of the GoI expression cassette still led to overburdening of the secretion machinery and also the periplasmic folding capabilities. Genomic integration into the forward strand, on the other hand, resulted in an expression system, which yielded aggregated, but secreted product. GoI expression led to overburdening of periplasmic folding capabilities but the secretion machinery seemed to work properly. In previous sections, it was observed that the integration direction of the expression cassette influenced the synthesis rate of the recombinant target protein (see Section 3.1.4, p. 103). Obtained results concerning the location of the aggregated proteins matched the observed synthesis rate order (plasmid > genomic, reverse > genomic, forward) and are in accordance with the following assumptions from a review article (see Figure 3.15) (de Marco, 2013). In case of a high synthesis rate, the translocation machinery is overburdened, resulting in aggregation of almost the entire product in the cytoplasm. By lowering the synthesis rate (plasmid → genomic, reverse), the translocation machinery is less overburdened, reducing cytoplasmic product aggregation. However, a substantial amount aggregates in the cytoplasm, suggesting that the secretion machinery capacity is still exceeded by the synthesis rate. Furthermore, high amounts of secreted product now seem to overburden the periplasmic folding machinery. Further lowering the expression rate (genomic, reverse → genomic, forward) fully mitigates issues and enables

secretion of the entire product to the periplasm. Only for the forward genome-integration strain, significant amounts of soluble and correctly folded intracellular scFv-Z-1 amounts were unambiguously detected. This might be due to reduced translocon overburdening in answer to a decreased secretion rate.

FabX

Genomic integration was shown to prevent secretion overburdening in the context of FabZ and scFv-Z-1 production. In initial experiments in separate development project, plasmid-based expression systems employing the araBAD promoter for transcriptional control were used for FabX production. The genetic constructs were designed to target the product monomers to the periplasmic space by genetic fusion to the OmpA leader peptide SP_{ompA}. However, the greatest share of soluble, correctly folded, heterodimeric product was found in and purified from the supernatant (data not shown). Amounts of aggregated target product in the g l⁻¹ range were obtained inside of the production host cells in these experiments. To identify their location within the cell, intracellular aggregates prepared at EoF time points of two representative fermentation experiments were subjected to N-terminal sequencing. The procedure was performed as described in Sections 2.6.6 (p. 79) and the two previous sections. Results are presented in Table 3.4.

A mixture of two sequences ((D/E)-(V/I)-Q-L-(T/V)-(Q/E)-S-(G/P)-(S/G)-(S/G)-L) but not the N-terminal amino acids of the leader peptide sequence (M-K-K-T-A-I-A-I-A-V-A) was detected within the analyzed samples. Apparently, significant shares of both monomers (LC: D-I-Q-L-T-Q-S-P-S-S-L, HC: E-V-Q-L-V-E-S-G-G-G-L) are secreted to and aggregate in the periplasmic space.

In case of products FabZ and scFv-Z-1, a shift in aggregate location from the cytoplasm to the periplasm by application of genomic integration was observed. For FabX, however, aggregates obtained in fermentations, in which the plasmid-bearing strain was employed, derived from the periplasm. This suggests, that, in this case, the major issues are associated with periplasmic folding rather than the secretion machinery. Apparently, the fermentation parameters, which were optimized for soluble production (pBAD promoter and induction with arabinose and lowered production temperature), enabled efficient secretion. Except for the induction strategy (IPTG instead of arabinose), the production process within the “FabX project” was not subject to changes prior application of the genome-integrated strain (see B<FabX-kanR>(01) in Table 2.6). Furthermore, the gene expression cassette number was decreased to one by genomic integration. Accordingly,

despite utilization of the strong T7 system, a lower synthesis rate can be assumed in plasmid-free strains. This assumption is supported by a decrease of total FabX product due to usage of the genome-integrated expression strain. Secretion machinery overburdening in plasmid-free strains under the applied conditions was, therefore, considered unlikely. Thus, the majority of the total product was assumed to aggregate in the periplasm also in the genome-integrated FabX production host.

3.2.5 Conclusions

In the course of this study, it was shown that the use of plasmid-free production strains improved soluble product yields of GFP and three different antibody fragments (see Table 3.3). Several steps during the synthesis of soluble and correctly folded antibody fragments were examined concerning efficient operation.

A well known problem associated with plasmid-based recombinant protein production is the loss of the expression cassette (Bentley and Kompala, 1990; Marisch et al., 2013a; Popov et al., 2011; Summers, 1991). This drawback was reported to be efficiently addressed by genomic integration (Striedner et al., 2010). In this study, it was shown that plasmid loss also plays a critical role in recombinant antibody fragment (FabZ and FabX) production in the bacterial periplasm. Furthermore, this plasmid-related issue was, indeed, addressed by using plasmid-free expression systems.

Plasmid presence as such can already pose a significant metabolic load on host cells, depending on the choice of vector and its elements (*e.g.* size of the plasmid, plasmid copy number, origin of replication and selection marker) (Bentley et al., 1990; Birnbaum and Bailey, 1991; Carneiro et al., 2013; Ow et al., 2006; Wang et al., 2006). Especially after induction, mere plasmid presence (as opposed to high-rate recombinant protein production) can negatively influence host cell metabolism (Andersson et al., 1996; Mairhofer et al., 2013). Reasons were, *e.g.*, found in significant transcriptional read-through due to insufficient termination by the T7 terminator (Carter et al., 1981; Macdonald et al., 1994; Sousa et al., 1992; Telesnitsky and Chamberlin, 1989a). This drawback was recently targeted by generation of an improved transcription terminator (Mairhofer et al., 2014). As described in Section 3.2.2, no obvious influence of plasmid presence during the growth phase (*i.e.* prior to induction of recombinant protein production) was detected in this study. However, after induction, a significant negative influence of the basic plasmid became apparent. More specifically, the strain bearing the empty plasmid showed a behavior similar to the one of the plasmid-bearing production strain. Increased copy number of an antibiotic marker gene was assumed to affect host cells in approaches employing multi-copy plasmids (Mairhofer et al., 2010). Data generated during the course of this study suggested that single-copy presence of a resistance marker (*e.g.* in case of genome-integration strains) has negligible influence on the host cell. In conclusion, genomic integration successfully mitigated the drawbacks associated with the use of expression plasmids.

Basal expression (transcription of a recombinant gene prior to induction) has previously

been reviewed to constitute another drawback in recombinant bioprocesses (Terpe, 2006). The commonly used T7 expression system is especially susceptible to this phenomenon (Dubendorf and Studier, 1991), since small amounts of T7 RNA polymerase facilitate high-rate expression of the heterologous gene (Studier and Moffatt, 1986; Tabor and Richardson, 1985). Numerous ways to counteract this system leakiness were developed (*e.g.* *lacI* overexpression (Muller-Hill et al., 1968), T7 lysozyme co-synthesis (Studier, 1991) or arabinose-inducible T7 RNA polymerase synthesis (Thermo Fisher, 2016)). Especially for “hard-to-produce” (*i.e.* toxic) products, drawbacks of basal expression, like an increased metabolic burden and growth cease, influence the host’s production behavior, even after induction (Saida et al., 2006). As expected, basal expression was also observed in FabZ production experiments. Surprisingly, system leakiness was not apparent during examination of FabX and scFv-Z-1 production experiments. Likewise, Mairhofer *et al.* did not report noticeable T7 expression system leakiness (Mairhofer et al., 2013). Striedner *et al.* reported different levels of basal expression, depending on the basic strain used (BL21(DE3) vs. HMS174(DE3)) (Striedner et al., 2010). In this regard, the results of this study are in accordance with findings of previous studies. In cases of pronounced basal expression, leakiness could successfully be reduced or even prevented by using plasmid-free production systems. In conclusion, single-copy genomic integration of a target gene is a viable alternative to reduce system leakiness.

Recombinant protein production controlled by a plasmid-based T7 promoter can result in extraordinarily high target product synthesis rates (Studier and Moffatt, 1986). However, high-rate synthesis of recombinant mRNA and, thus, proteins can have a significant impact on the host cell (Bentley et al., 1990; Hoffmann and Rinas, 2004). For instance, high-level abundance of extrinsic mRNA species can compete with host cell RNA for elements of the translation apparatus (Mairhofer et al., 2013). This can lead to a reduced synthesis of biomass-related proteins, resulting in growth cease (Bentley et al., 1990). Recombinant genes often differ from the native host cell genes due to codon usage optimization to bias expression in favor of the recombinant gene. This might negatively affect the expression of native genes, whose products are associated with translation, transport or metabolic functions (Bonomo and Gill, 2005; Gustafsson et al., 2004; Kane, 1995). In general, reduced protein synthesis and growth arrest were reported to be consequences of metabolic stress (Dong et al., 1995; Glick, 1995). Increasing plasmid copy numbers after induction of recombinant protein production (“runaway replication”) (Grabherr et al., 2002; Hoffmann and Rinas, 2001; Teich et al., 1998) can further increase the before-mentioned effects. In this study, cessation of growth and product formation, could be prevented by application of plasmid-free strains.

In scFv-Z-1 production experiments, it was shown that the choice of the target strand for genomic integration significantly influenced the synthesis rate of the recombinant protein. In experiments with plasmid-free FabX production strains, intracellular total (soluble + insoluble) product titers of 2.6 g l^{-1} and 2.3 g l^{-1} at EoF were determined for the reverse and forward genome-integrated strain, respectively. Soluble and correctly folded FabX concentrations of approximately 230 (reverse integration) and 170 mg l^{-1} (forward integration) were measured in the culture supernatant at EoF. Results suggested a higher synthesis rate in case of reverse-integration of the FabX gene expression cassette. Based on these results, the FabX production strain with reverse-integrated expression cassette was used during the further course of the project. In conclusion, the same trend concerning the influence of the integration direction on protein synthesis was observed for two therapeutically relevant products (FabX and scFv-Z-1). Gene orientation in the bacterial genome was previously reported to influence the respective gene's expression level. It was conducted that genes on the leading strand exhibit higher expression rates (Block et al., 2012; Sabri et al., 2013). However, our results suggest that, at least in case of integration into the *attTn7* site (Lichtenstein and Brenner, 1982), the opposite is true. While reasons for this phenomenon remain elusive, targeting different strands during genomic integration constitutes a valuable item in a toolbox for soluble protein production.

By application of reverse or forward genomic integration of the recombinant gene cassettes, expression rates of two target proteins were harmonized with limited secretion capacity of the host cell. These observations are in accordance with previous findings, which suggested the importance of fine-tuning the expression rate with translocation capacities (Mergulhão et al., 2003; Schlegel et al., 2013; Simmons and Yansura, 1996) (see Figure 3.15). In summary, secretion machinery overload constituted a drawback associated with plasmid-based protein production. In relevant expression strains, a reliably lowered synthesis rate mediated by genome-borne expression cassettes enabled more efficient secretion.

In plasmid-free production strains, the capacity of the periplasmic folding instead of the translocation machinery appeared limiting. Periplasmic aggregation has previously been reported (Jeong and Lee, 2000; Lin et al., 2001b; Pan et al., 2003) and protein aggregation in general was shown to be detrimental for the cell (Lindner et al., 2007; Maisonneuve et al., 2008; Winkler et al., 2009). Periplasmic aggregation in specific can cause stress responses, which have previously been reviewed (de Marco, 2013; Duguay and Silhavy, 2004; Gasser et al., 2008; Goemans et al., 2014; Miot and Betton, 2004). Up-regulation of periplasmic proteases (see Section 1.3.2) constitutes one of the stress responses. For instance, increased DegP levels were found to be a consequence of accumulation of aggregates in the periplasm (Pan et al., 2003). For all examined recombinant products,

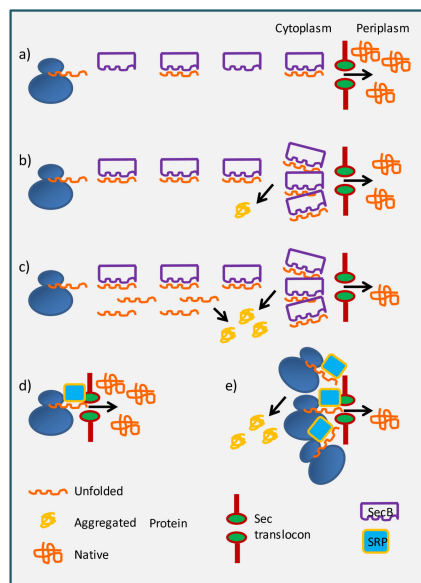


Figure 3.15: Expression rate determines the fate of proteins targeted to the periplasm. Under optimal conditions **(a)**, the newly synthesized polypeptides are successfully delivered to the translocon by the SecB co-chaperone. Subsequently, the precursor protein is translocated into the periplasm for folding. At higher expression rates, the polypeptides will progressively aggregate because either the Sec translocon **(b)** or/and SecB **(c)** will become limiting. The co-translational SRP-based secretion route could overcome SecB paucity **(d)**. However, it shares the same export machinery and, therefore, any produced in excess to the translocon capacity will remain trapped in the cytoplasm **(e)**. The figure is copied from a review article (de Marco, 2013).

strong bands (corresponding to several g l^{-1}) with a molecular weight of monomers were observed. Thus, proteolytic degradation in the periplasm was assumed not to be a major issue for soluble product formation. In accordance with previous reports concerning leakage of recombinant periplasmic proteins (Lee et al., 2001; Ukkonen et al., 2013), significant amounts of FabX and FabZ were found in the supernatant. Osmotic pressure (Hasenwinkel et al., 1997) or membrane perturbation (Slos et al., 1994) caused by accumulation of recombinant protein were identified as possible reasons. Secretion of recombinant proteins to the periplasm induced cell lysis resulting in the release of periplasmic content in other studies (Lee et al., 2001; Ukkonen et al., 2013). Decreasing values for intracellular (soluble or total) target product concentrations and a reduction of cell densities towards the end of fermentations suggested similar phenomena for FabZ and FabX production. If aggregate formation or the concentration of the soluble target protein caused cell lysis remained elusive at this stage. However, results obtained by periplasmic folding modulator co-synthesis (see Section 3.4) suggested that periplasmic aggregation has a great influence on reduced cell viability and lysis.

As previously suggested (Chen et al., 2008; Spadiut et al., 2014; Striedner et al., 2010), genomic integration of the gene of interest addressed many of the drawbacks associated

with plasmid-based recombinant protein production. Periplasmic aggregation seems to negatively influence soluble product yields. However, plasmid-free strains are a good starting point for further development. In case of strains, which already bear a plasmid for GoI expression, incompatibility groups have to be considered if the use of further plasmids is intended for host engineering. In case of a genome-integrated target gene, on the other hand, another element (*e.g.* for co-synthesis of helper proteins) could easily be introduced into the host cell by utilization any available plasmid. Moreover, when GoI expression cassettes are stably integrated into the host genome, the FLP-FRT technique can be applied to excise the integrated resistance gene module. Afterwards, also any marker gene of choice can be used for selection of cells bearing additional plasmids. Concerning the development of cultivation processes, process design appears to be simplified when plasmid-free strains are employed. For instance, it has previously been shown, that the expression rate is adjustable by inducer titration (Striedner et al., 2010). Considering the results of this study, glucose accumulation, growth deficiencies and prolonged process durations could be prevented by application of plasmid-free strains. More robust processes allow for broadening of feasible process parameters compared to processes employing plasmid-bearing strains.

3.3 Beyond Single-Copy Genomic Integration; Rationale Behind and Preparation for Further Strain Engineering

In previous sections, plasmid-free strains were employed in production experiments and compared with their multi-copy plasmid-bearing counterparts. Single-copy genomic integration of the GoI expression cassette was shown to improve soluble product concentrations for all considered products by targeting several bottlenecks. However, the predominant share of target antibody fragments was shown to aggregate in the periplasm, suggesting an overburdening of the periplasmic folding machinery. As described in the introduction both host and process engineering approaches are reasonable measures for addressing this bottleneck (see Sections 1.3.4 and 1.3.5, respectively).

3.3.1 Process Development to Improve Periplasmic Production of Recombinant Proteins

In theory, process development (as opposed to host engineering) constitutes the preferred alternative to improve recombinant product titers. This is, partly, because strain engineering and subsequent screening of performance can require a lot of resources compared to changing production process parameters. Furthermore, the use of continuous numerical factors in process design allows for the systematic application of “Design of Experiment” approaches (Haaland, 1989). More specifically, “Full/Fractional Two-Level Factorial” designs can help to identify main effects and interactions. Subsequently, “Response Surface” methodologies can be used to identify optima in the process design space for a given production system.

Determination of the induction strategy

To determine an appropriate induction strategy, strain B<kanR-FabZ> (HC-LC expression cassette, forward integration, see Table 2.6) was employed in four fermentation set ups. These differed only in the applied induction strategy. Fed-batch fermentations were performed at a constant temperature of 37°C. Feed 500 Solution (see Section 5.5.4) was used during feeding phases. In all cases, 596 mg of the inducer IPTG was supplied by bolus addition at different time points after start of feeding. An overview of the applied

strategies is presented in Table 3.6. Results are shown in Figure 3.16.

Table 3.6: Different induction strategies applied in fed-batch fermentations of strain B<kanR-FabZ>. All induction strategies were based on pulsed IPTG addition at a defined time point after feed start.

| T0 [h after feed start] | Production phase [h] | IPTG _{T0} [mM] | IPTG _{EoF} [mM] |
|-------------------------|----------------------|-------------------------|--------------------------|
| 14 | 15 | 0.755 | 0.495 |
| 15 | 13 | 0.705 | 0.493 |
| 18 | 10 | 0.634 | 0.491 |
| 22 | 6 | 0.559 | 0.495 |

T0: time point of inducer addition. Production phase: process duration after IPTG addition. IPTG_{T0}: inducer concentration immediately after inducer addition. IPTG_{EoF}: inducer concentration at EoF.

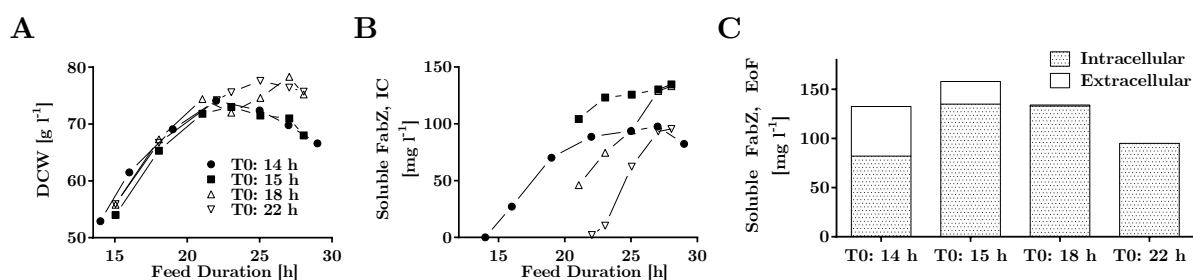


Figure 3.16: Variation of the induction time point; Influence on product formation and growth in fermentation experiments. Plasmid-free FabZ-producing strain B<kanR-FabZ> was employed in fermentation experiments. The time point of inducer addition relative to start of feeding was varied in between experiments. Samples of different time points were analyzed concerning growth (DCW) and soluble and correctly folded FabZ concentration (ELISA). **A**: DCW courses of the fermentation processes. **B**: progress of intracellular soluble FabZ concentrations throughout the fermentation processes. **C**: soluble FabZ concentration at EoF. Intracellular and extracellular fraction are indicated by differently colored bar segments. “T0: 14/15/18/22 h” abbreviations indicate feed duration prior to inducer addition.

When the inducer was added after 14 h of feeding, the intracellular soluble product concentration increased for 13 h after induction (98 mg l^{-1}) and, subsequently, decreased towards end of fermentation (EoF) (82 mg l^{-1}) (see Figure 3.16 (B)). A later induction (15 h after feed start) resulted in an increase until approximately 8 h after induction. Afterwards, values increased merely slightly, yielding 135 mg l^{-1} at EoF. When the inducer was added additional 3 h later (18 h after feed start, 10 h production phase), soluble intracellular FabZ concentration increased constantly until 133 mg l^{-1} at EoF. A further shortened recombinant protein production phase of only 6 h (inducer addition after 22 h of feed start) resulted in increasing values until 5 h after induction (feed duration: 27 h), followed by constant values until EoF (approximately 94 mg l^{-1}). At EoF, total (intracellular + extracellular) soluble product concentrations of 132, 158, 134, 95 mg l^{-1} were obtained for the different induction time points of 14, 15, 18 and 22 h, respectively. Early induction time points (14 and 15 h after feed start) resulted in detectable amounts of soluble extracellular product (50 and 23 mg l^{-1} , respectively). After late induction (18

and 22 h after feed start), amounts in the supernatant were at or below 1 mg l^{-1} . DCW progress was similar for all experiments until 22 h after feed start (latest induction time point). Due to different induction time points (14, 15, 18 and 22 h after feed start), DCW values at inducer addition varied in between the experiments (53, 54, 67 and 74 g l^{-1} , respectively). Early induction (14 and 15 h after start of feeding) resulted in DCW values of 67 g l^{-1} at EoF. Values of 75 g l^{-1} were measured when late induction (18 and 22 h after start of feeding) was applied. Amounts of consumed base were greatly increased when the early induction strategy was applied (data not shown). Furthermore, acid addition was necessary in these cases (14 h and 15 h), while for the “late induction” processes, no acid was consumed.

Differences in DCW at EoF and product appearance in the supernatant suggest cell lysis in case of early induction. Aggregation of recombinant protein was reviewed to be toxic in certain cases (de Marco, 2013) and might be a reason for the observed phenomenon. Later induction time points resulted in lowered total soluble product concentrations. Note that the production phase duration was prolonged by 1 h in the “14 h induction time point” fermentation. As this constitutes a production phase prolongation of merely 7 %, no significant effect on experiment outcome was assumed. Based on the presented results and theoretical considerations, an induction strategy for future experiments was determined. During approaches presented above, the amount of inducer (596 mg IPTG) resulted in concentrations between 0.491 and 0.755 mM. According to Striedner *et al.*, $6 \text{ }\mu\text{mol IPTG per g CDM}^{-1}$ already result in full titration of the repressor LacI (*i.e.* the maximum possible expression rate) (Striedner *et al.*, 2010). In the experiments described above, IPTG amounts per biomass ranged between 6.5 and $14.3 \text{ }\mu\text{mol IPTG per g CDM}^{-1}$. Striedner *et al.* also reported possible IPTG titration in the lower range of inducer per DCW. Contrarily, it was also reported that reliable dose-dependent expression using IPTG as the inducer is not possible due to a potentially heterogeneous response of the bacterial culture (Fernandez-Castane *et al.*, 2012; Rosano and Ceccarelli, 2014). To ensure full induction and a homogeneous response in the culture, increased IPTG amounts ($894 \text{ mg}/3.75 \text{ mmol}$) were added in subsequent fermentation experiments. With decreased liquid volume due to utilization of the more concentrated Feed 600 Solution (see Section 5.5.4), this measure resulted in increased IPTG concentrations during the second half of the project. More precisely, IPTG amounts of above 20 (T0) and $11 \text{ }\mu\text{mol IPTG per g CDM}$ (EoF) were applied. Amounts were considered sufficient to induce full LacI titration and homogeneous production behavior. As described above, the earliest induction time point resulted in observations associated with cell lysis. Increased cellular stress levels (*e.g.* induced by product aggregation) are feasible in this case (Duguay and Silhavy, 2004; Gasser *et al.*,

2008; Goemans et al., 2014; Miot and Betton, 2004). Note that a less productive FabZ production strain (forward integration, HC-LC expression cassette, see above) compared to subsequently used strain B<FabZ-kanR>(02) (reverse integration, LC-HC expression cassette) was employed in these fermentations. It was assumed that increased productivity would rather promote lysis due to promotion of aggregation. As testing engineered hosts in a process with apparent drawbacks seemed reasonable for the scope of this project, the earliest tested induction time point (14 h after feed start) was used in further experiments. To avoid night shifts in staff planning and, still, allow for unscheduled process duration adaptations, the production phase was shortened to 13 h, resulting in a process duration of 27 h after feed start.

Temperature variations slightly improved soluble product titers but hampered development potential

When the “high-performing” plasmid-free FabZ-producing strain B<FabZ-kanR>(02) was employed in production experiments at 37°C, periplasmic aggregates were observed (see Section 3.2.4). Apart from the induction strategy, temperature constitutes a major influential parameter during bioprocess development (Dragosits et al., 2011; Rodriguez-Carmona et al., 2012). To examine this parameter’s influence, fermentation experiments with varied production phase temperatures (16, 25, 28, 30 and 37°C) were performed. Fermentation experiments were performed as described in Section 2.4.2 using the strain mentioned above. Note that results of six fermentations from previous experiments (see Section 3.1) and of three fermentations from later experiments (see Section 3.4) were used for discussion in this section. During the initial two experiments employing a production temperature of 16°C, the DO value of 20 % was maintained by automated glucose feed rate adjustment. Total aeration, agitation and pressure were adjusted to 3 l min⁻¹, 400 rpm and 700 mbar, respectively, to enable feeding rate-control of DO values. Under the applied conditions, a linear course of added Feed 600 Solution was observed. Gradient calculation in the linear range yielded an approximate feed rate of 35 g h⁻¹. This feed rate was set for the following fermentation experiment with production phase temperature of 16°C. Note that different factors were varied in addition to the production phase temperature. An overview of those factors is presented in Table 3.7. Concentration of total (intracellular + extracellular) soluble and correctly folded FabZ was determined by ELISA analysis. Total (soluble + insoluble) intracellular product concentration was determined by SDS PAGE analysis of cell samples under reducing and denaturing conditions. Maximum values for these responses were determined for each fermentation approach and used for assessment.

Table 3.7: Process parameters during fermentation experiments for assessment of the influence of the production phase temperature. Different temperature, feeding and induction strategies were applied. Details are described in respective subsections of Section 2.4.2. Parameters of temperature ramp 3, feed strategy 4 and induction strategy 3 (second row of 25°C production temperature row) are not disclosed due to confidentiality reasons. n: Number of fermentations under identical conditions.

| T_{Prod} | T_{Growth} | T_{Ramp} | Feed egy | strat- | Induction strategy | n |
|-------------------------|---------------------------|-------------------------|---------------------|---------------|---------------------------|----------|
| 16°C | 37°C | 1 | 1 | | 1 | 1 |
| | 37°C | 1 | 1 | | 2 | 1 |
| | 37°C | 1 | 2 | | 1, 22 h production phase | 1 |
| 25°C | 37°C | 2 | 3 | | 1 | 1 |
| | 33.5°C | 3 | 4 | | 3 | 1 |
| 28°C | 37°C | 2 | 3 | | 1 | 3 |
| 30°C | 37°C | 1 | 5 | | 1 | 1 |
| | 37°C | 1 | 5 | | 2 | 1 |
| | 37°C | 1 | 3 | | 1 | 1 |
| 37°C | 37°C | n.a. | 6 | | 4, 15 h production phase | 3 |
| | 37°C | n.a. | 3 | | 1 | 3 |
| | 37°C | n.a. | 3 | | 1 | 3* |

*Selective conditions (30 µg ml⁻¹ kanamycin) were applied in these fermentations.

Mean soluble FabZ titers of 363, 423, 334, 305 and 35 mg l⁻¹ were determined for production temperatures of 37, 30, 28, 25 and 16°C, respectively. Mean total intracellular FabZ titers of 4.1, 1.2, 1.6 and 1.0 g l⁻¹ were determined for production temperatures of 37, 30, 28 and 25°C, respectively. Due to low soluble product titers observed for a production temperature of 16°C, no SDS PAGE analysis was performed. Note that other parameters apart from production phase temperature, *e.g.*, induction strategy, temperature during growth phase, duration of temperature ramp, feed rate, feed solution or application of antibiotics, were also varied. However, ANOVA analysis revealed no significance influence ($p < 0.5$) for any other parameter than production phase temperature. However, the influence of additional parameters might have caused increased standard deviations (see error bars in Figure 3.17). Highest soluble FabZ titers were determined for temperatures of 28°C (497 mg l⁻¹) and 25°C (436 mg l⁻¹). However, two further attempts of applying a production temperature of 28°C yielded only 179 and 327 mg l⁻¹, respectively. Apparently, the process is not very robust under the applied conditions. Promising results with a production temperature of 25°C were obtained under a more sophisticated “non-standard” process format (feed addition of inducer, adaptation of feed rate during recombinant protein production phase, lowered growth phase temperature). Under “standard conditions”, a 25°C production temperature resulted in a maximum soluble product titer of 173 mg l⁻¹. By lowering the production phase temperature from 37 to 30°C, soluble product concentration was slightly but reproducibly increased (from 363 to 423 mg l⁻¹). In case of 37°C production temperature, total intracellular FabZ

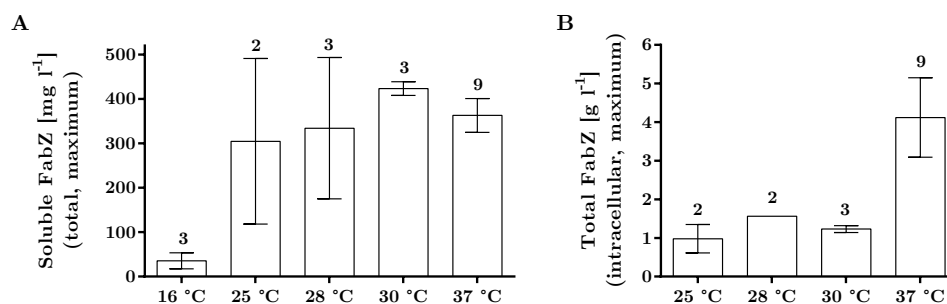


Figure 3.17: Maximum total (intra- + extracellular) soluble (A) and maximum total (soluble + insoluble) intracellular (B) FabZ titers in dependence of production phase temperature. The plasmid-free FabZ-producing strain B<FabZ-kanR>(02) (reverse genomic integration, LC-HC expression cassette) was employed in fed-batch fermentations under varying conditions (see Table 3.7). Soluble and correctly folded FabZ concentrations inside the cells and in the supernatant were determined *via* ELISA analysis. Both values were summed for each measurement time point. The maximum determined throughout the fermentation is presented in A. Furthermore, total intracellular product (soluble + insoluble) concentration was determined by SDS PAGE analysis of cell samples. Bars represent mean values determined in several experiments. Values atop bars indicate the number of conducted experiments. Error bars represent standard deviations.

concentrations ranged between 2.9 and 5.7 g l⁻¹ (see Figure 3.17 (B)). In case of all other production phase temperatures, significantly lower values (between 0.7 and 1.6 g l⁻¹) were determined.

In conclusion, a lower temperature bears the potential to increase soluble FabZ yields. However, results were ambiguous. Furthermore, a negative influence on total intracellular product titer was unambiguously observed.

Conclusion: host engineering in combination with a simple fermentation process for screening constitutes a promising de-bottlenecking strategy

At this point of the study, the influence of both host and process on soluble FabZ production had been examined. Process development attempts did not yield the expected results. Production phase temperature had a significant influence on overall intracellular but also on total soluble product titer. While soluble titers were only slightly increased compared to fermentations with 37 °C during the production phase (maximally 1.4-fold in on case of 28 °C), a significant negative influence on total productivity was observed. On the other hand, host engineering approaches (*i.e.* genomic integration, expression cassette re-organization), significantly improved soluble product yields. For instance, soluble FabZ yields were improved up to 4.4-fold (see Table 3.3). Observations from host engineering for improved production of soluble FabX and scFv-Z-1 pointed in the same direction.

In conclusion, strain engineering had more beneficial effects than the tested process de-

velopment attempts. A more profound design of experiment approach for optimizing fermentation conditions might yield higher soluble product titers compared to the concentrations obtained by temperature variation. However, as “genetic strategies” (*i.e.* host engineering) were in the focus of this project, resources to be utilized for process development were scarce. Necessary runs for meaningful design of experiment approaches were calculated using Design Expert[®]. For instance, 20 runs were calculated to be necessary for identification of the main influential of five numeric factors (inducer concentration, temperature prior to and after induction and feed rate prior to and after induction) in a fractional factorial (2^{5-1}) approach (assumptions: no replicate measurements, signal to noise ratio of 1.6, 4 blocks, one center point per block). Note that numerous further process variations (*e.g.* choice of medium, pH value or control of aeration and oxygen availability) exist (Rosano and Ceccarelli, 2014; Ukkonen et al., 2013). If only two main influential factors had been identified in this fractional factorial approach, 13 further runs would have been necessary for model optimization with a Central Composite Response Surface model (two numeric factors, 5 levels per factor, 2 blocks, 3 center points in each factorial and axial block, no replicates of factorial or axial points). A total thirty-three 5 l fed-batch fermentation runs for process development appeared out of scope of this project.

Based on the results obtained from different processes, decisions concerning important parameters were made. The more concentrated Feed 600 Solution was employed in experiments. Furthermore, the highest applied feed rate was chosen for further fermentation processes. Specifically, an exponential feed phase ($\mu = 0.177 \text{ h}^{-1}$, from 15.29 to 127.91 g h⁻¹ within 12 h) was followed by a constant feeding phase of 15 h (127.91 g h⁻¹). After temperature variation experiments described above, a production phase temperature of 37°C was applied in subsequent experiments. A lowered temperature was previously reported to improve folding by decelerating cellular processes (Berlec and Strukelj, 2013; Lebendiker and Danieli, 2014; Overton, 2014; Rodriguez-Carmona et al., 2012; Rosano and Ceccarelli, 2014; Sugiki et al., 2014). However, this could not reproducibly be verified within this study. In addition, a temperature down shift prior to induction was previously shown to have a significant influence on the host (Kim et al., 2005) and might be associated with drawbacks in recombinant protein production (Fahnert, 2012; Tolia and Joshua-Tor, 2006). These drawbacks can be prevented by a constant process temperature of 37°C. A bolus addition of 894 mg (3.75 mmol) IPTG was expected to result in “full induction” conditions and was, thus, applied for further experiments. Furthermore, the chosen induction time point (14 h after feed start) resulted in high-rate production for 13 h and ensured high total product concentrations. In conclusion, a simple process that

did not demand night shifts and enabled high intracellular total and medium to high soluble product concentrations, was applied in subsequent experiments. As high amounts of periplasmic aggregates are formed under the applied conditions, folding improvement appeared promising (see Section 3.4).

3.3.2 Establishing a Small Scale Format for Increased Throughput During Strain Evaluation

In previous experiments, the overload of the periplasmic folding machinery was not efficiently prevented by production phase temperature variation. Extensive process development approaches were evaluated to be unreasonably resource-demanding. Moreover, strain engineering techniques appeared more adequate based on previous results. Nonetheless, engineered strains require production experiments to test for performance. As performing 5 l fed-batch fermentation experiments requires a lot of resources, a different format appeared necessary at this stage of the study. Efforts to establish an appropriate small scale production format are described in this section.

Based on previous results and theoretical considerations, various parameters of the shake flask approach were determined. The underlying process was defined to consist of preculture and main culture, while the latter was divided into growth and recombinant protein production phase. To enable decreased amounts of media and facilitate handling, precultures were grown in 50 ml reaction tubes. Comparable to fermentation precultures, a salt-based medium with 10 g l⁻¹ glucose as main carbon source supplemented with 0.5 g l⁻¹ yeast extract (T7 Preculture Medium, see Section 5.5.2 in the Confidential Supplement) was used. Selective conditions (for plasmid retention, if applicable) were applied during preculture incubations. To enable rapid growth, cells were incubated at 37°C. In batch phases of fermentation experiments, a chemically defined medium with glucose as the sole carbon source was used (see Batch Medium, Section 5.5.4). After batch phase, glucose-limited conditions (< 0.5 g l⁻¹) were ensured by an appropriate feeding strategy. No feeding was possible in simple shake flask experiments, which were performed in 1 l unbaffled Erlenmeyer shake flasks. While the *lacUV5* promoter, which steers transcription of the T7 polymerase gene in DE3 strains has a reduced sensitivity to catabolite regulation, it might still be repressed under glucose excess conditions (Rosano and Ceccarelli, 2014). Furthermore, constantly high glucose loads might induce an undesired overflow metabolism leading to accumulation of acetic acid (Vemuri et al., 2006). An initial glucose concentration of 1.5 g l⁻¹ was identified to be appropriate in previous experiments (Heistingner, 2013). To simulate the glucose-limiting conditions of a fed-batch fermentation in shake flask experiments, this reduced glucose concentration was adopted and glycerol was chosen as additional carbon source. Compared to glucose, glycerol shows reduced repression of *lac* promoter-controlled gene expression. Experiments concerning the influence of different glycerol concentrations were in the scope of this part of this study. In order to accelerate bacterial growth and allow for completion of production ex-

periments in adequate time periods, shake flask medium was supplemented with 0.25 g l⁻¹ yeast extract (Heisting, 2013). An appropriate concentration of buffer components for maintaining a constant pH of 6.8 (comparable to fermentation experiments) was identified in previous experiments in the course of this study (Dürkop, 2014). The same concentrations of dipotassium hydrogen phosphate and potassium dihydrogen phosphate were added to the medium to ensure an appropriate buffering capacity. To ensure monoseptic conditions and facilitate handling, 30 µg ml⁻¹ kanamycin were added to the main culture medium when employing the genome-integrated kanamycin resistance module. To avoid deviations due to the addition of other antibiotics (resistances encoded on plasmids for, *e.g.*, folding modulator genes), no selective pressure was applied on plasmid maintenance. To accelerate bacterial growth, the main culture was started with an initial calculated OD₅₅₀ value of 0.2 (compared to approximately 0.03 in fermentation experiments). Comparable to fermentation experiments, the initial main culture phase (“growth phase”) was performed at 37 °C. In contrast to fermentation experiments, a production temperature of 25 °C was employed in shake flask experiments. A lowered temperature was evaluated to be an appropriate measure to allow for over night production phases and, thus, reasonable handling times. Furthermore, 13 h of recombinant protein production under cell growth conditions were implemented in fermentation experiments. By decreased growth due to a lowered temperature, similar conditions were supposed to be adjusted. Finally, differences concerning the influence of co-synthesis of six different periplasmic folding modulators (see below) were not expected to depend generally on production temperature. Induction was performed at OD₅₅₀ values of 0.7 to 1.3 by addition of 1 mM IPTG. Prior to induction, cells were incubated at 25 °C for at least 15 min.

To examine the appropriate glycerol concentration, different shake flask media containing 5, 10 or 15 g l⁻¹ were tested (see Table 5.8). To also assess the duration of shake flask experiments, a recombinant protein production phase of 29 h was employed. Plasmid-free FabZ production strain B<FabZ-kanR>(02), which yielded highest soluble FabZ titers and sufficient amounts of aggregates in previous experiments, was used in media composition test experiments. To simulate co-expression approaches (see Section 3.4.2) the strain was transformed with basic plasmid pBI1ST7.2. Experiments were performed using the basic setup for shake flask experiments (see Section 2.4.1) with a prolonged production phase duration. Sampling was carried out immediately prior to and 13 and 29 h after addition of IPTG. Intracellular soluble FabZ concentrations were determined by ELISA analysis. At-line measurement of OD₅₅₀, pH and glucose was performed periodically. Results are presented in Figure 3.18.

OD₅₅₀ values increased constantly towards a mean value of approximately 8 at 13 h

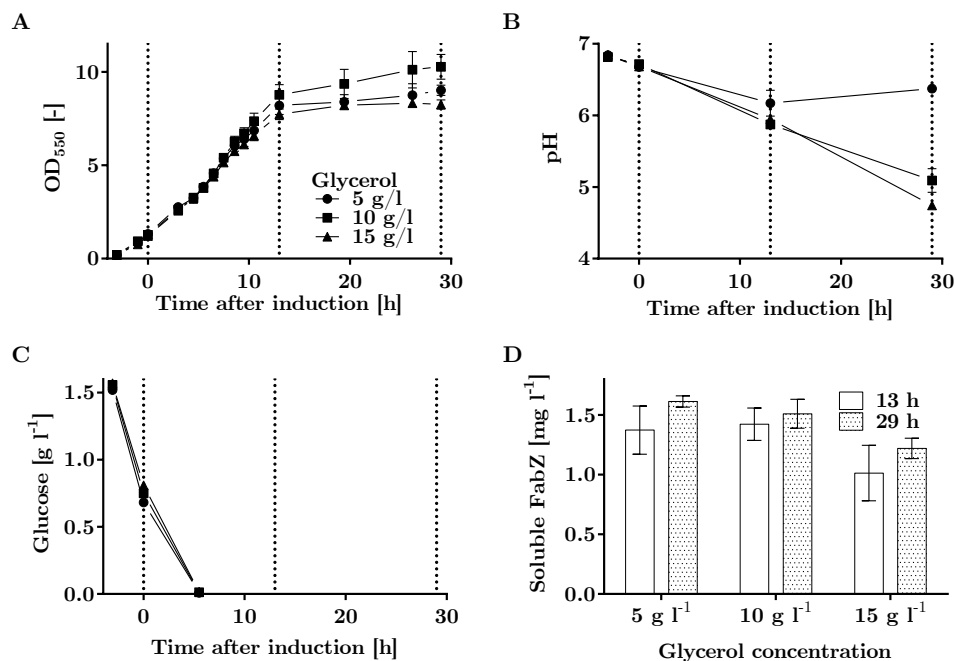


Figure 3.18: Influence of glycerol concentration and production phase duration in shake flask experiments. Plasmid-free FabZ-producing strain B<FabZ-kanR>(02) was transformed with basic plasmid pBI1ST7.2 and employed in shake flask experiments with media containing 5, 10 and 15 g l⁻¹ glycerol. OD₅₅₀ (A), pH (B) and glucose concentration in the supernatant (C) are presented in dependence of the time after IPTG addition. Vertical dotted lines represent sampling time points prior to and 13 and 29 h after induction. Intracellular soluble FabZ concentration at 13 and 29 h after induction (D) was measured by sandwich ELISA. Experiments were performed as biological triplicates for each glycerol concentration. Mean values are represented in the graphics. Error bars indicate standard deviations.

after induction for all media (see Figure 3.18 (A)). Afterwards, OD remained roughly constant until the end of the experiment (29 h after induction). Maximum OD values of approximately 9.0, 10.0 and 8.3 were determined for cells grown in medium containing 5, 10 and 15 g l⁻¹ glycerol, respectively. For all media, pH values decreased from 6.8 at inoculation to approximately 6.0 after 13 h of recombinant protein production. At the end of the experiments, differing pH values of 6.4, 5.1 and 4.7 were determined for glycerol concentrations of 5, 10 and 15 g l⁻¹, respectively. In summary, pH values remained approximately constant when 5 g l⁻¹ glycerol were employed in the medium, while values decreased in case of the other compositions. In all cases, glucose in the medium was depleted (< 0.02 g l⁻¹) after constant reduction over the first 8 hours. After 13 h of induction, soluble FabZ titers of 1.37, 1.42 and 1.01 mg l⁻¹ were determined for glycerol concentrations of 5, 10 and 15 g l⁻¹, respectively. For cell samples prepared at the end of the experiments (29 h after induction), 1.61, 1.51 and 1.22 mg l⁻¹ were measured.

Stagnant OD₅₅₀ values after 13 hours of recombinant protein production until EoF indicate the stationary phase. In all cases, glucose concentrations were slightly above 0.5 g l⁻¹ at

induction of recombinant protein production. However, after 2.5 h of recombinant protein production, glucose was limited.

Note that strains bore basic plasmid pBI1ST7.2 in the experiments in order to mimic the planned co-expression approaches. In later experiments, a significant negative influence of a comparable basic plasmid (pBI4iST7.2) on total productivity of the cell was observed (see Section 3.4.2). Strain B<FabZ-kanR>(02) without additional plasmid was later shown to result in higher total productivity and lower OD₅₅₀ values at the end of shake flask screening experiments. As shown in Figure 3.18 (C), glucose was depleted approximately 8.5 h after inoculation. Assuming that glycerol is depleted equally rapidly by the same amount of cells and considering the increasing cell density in the glucose depletion period of the experiments, glycerol will be depleted approximately 9 h after glucose depletion or 15 h after induction. pH was constant for additional 14 hours in case of the 5 g l⁻¹ glycerol medium. Probably, the buffer capacity of the medium will be sufficient in any set up until 13 h after induction. In summary, glycerol does not seem to be a limiting factor for any of the tested media.

Note that no significant aggregation of target product was observed in the used strains (data not shown). However, SDS PAGE analyses of cell samples from later shake flask experiments with plasmid-free strains revealed significant amounts of total (soluble + insoluble) intracellular product (see subsequent chapter) at 12 h after induction of recombinant protein production. Thus, an important criterium for the small scale format's suitability, *i.e.* high productivity (and aggregate formation in the periplasm) and, thus, potential for folding improvements was fulfilled.

During three of nine experiments (two replicates for 5 and one clone for 10 g l⁻¹ glycerol) OD₅₅₀ values at induction ranged between 1.3 and 1.4 and were, thus, slightly above the defined upper threshold of 1.3. However, a negligible influence on the remaining experiment was expected.

Utilization of the medium containing 5 g l⁻¹ glycerol was characterized by a minimum pH value of 6.2 after 13 h of recombinant protein production and constant values throughout the rest of the experiment. As low pH values can indicate undesired acetate formation, a pH range that was closest to the initial value of 6.8 was desired. Thus, a glycerol concentration of 5 g l⁻¹ seemed most appropriate for further experiments. Similar growth behavior of strains for all media suggested that improved pH values did not derive from carbon and, thus, growth limitation. While glycerol was previously suggested for mitigating acetate production, formation rates were also reported to be strain-dependent (Eiteman and Altman, 2006). In accordance, merely applying glycerol and low initial

glucose concentrations did not prevent acetate formation in case of media with increased glycerol concentrations (see pH values in Figure 3.18). In agreement with a previous report about co-utilization of glycerol and acetate (Martinez-Gomez et al., 2012), constant pH values in case of 5 g l⁻¹ glycerol concentration might reflect acetate consumption under low glycerol conditions. In a previous study, acetate overflow was suspected to be dependent on the degree of carbon influx (Basan et al., 2015). Here, the carbon influx might be dependent on the amount of glycerol available. Decreasing pH values (10 and 15 g l⁻¹ glycerol concentration), most probably reflect acetate production, which occurred only under glycerol excess conditions. Similar intracellular soluble FabZ concentrations were determined for the media with lower glycerol concentrations (5 and 10 g l⁻¹) 13 and 29 h after induction. As a glycerol concentration of 5 g l⁻¹ combined relevant features, *i.e.* constant pH values and sufficient growth and productivity behavior, it seemed appropriate for small scale screening experiments. Increased throughput had to be enabled by the small scale screening approach and limited resources were available for performing analytical methods. No substantial increase in product concentration was achieved by production phase prolongation. As described above, cells seem to be in a stationary phase after 13 h of recombinant protein production. In order to simulate constant growth, which is enabled by controlled fed-batch conditions, the entire shake flask experiment was supposed to be performed under growth-enabling conditions (*i.e.* prior to reaching the stationary phase). Furthermore, shorter shake flask processes supported the higher throughput. In conclusion, one final sampling point after 12 h of recombinant protein production was defined.

Under the described conditions (12 h of protein production phase using the medium with 5 g l⁻¹ of glycerol), up to 16 strains could be screened within a week (including ELISA analysis). In comparison to 4 strains per two weeks in fed-batch fermentations (one week for fermentation, one week for analytics), this constitutes a substantial improvement. The new small scale system was used in further strain engineering approaches, which yielded numerous expression systems. Intrinsic differences between shake flask and fermentation experiments (*e.g.* control of pH, temperature and oxygen supply or feeding strategies) are inevitable. To evaluate if the results from shake flask experiments were representative for the outcome in the final format of interest (fed-batch fermentations), selected strains were employed both in screening and in 5 l fed-batch fermentation approaches (see below).

3.4 Establishment of Different Systems for Co-Synthesis of Periplasmic Folding Modulators

In previous chapters, single-copy genomic integration of gene expression cassettes was shown to improve soluble target product yield by improving translocation across the inner membrane. However, great shares of the target product still aggregated in the periplasm. Process development approaches yielded only moderate improvements (see Section 3.3.1). Strain engineering constituted the most promising optimization approach. Periplasmic folding modulators have previously been used to improve soluble target product yields in *E. coli* (see Section 1.3.4). In this part of the study, different basic systems for co-expression were tested prior to examination of the influence of periplasmic folding modulators on strains with genome-integrated target gene expression cassettes.

3.4.1 Rationale

General system composition

Initially, a suitable basic set up for co-expression of folding modulator genes had to be identified based on theoretical considerations. FabZ was used as a model protein. Its genes were reverse-integrated into the *attTn7* site of the BL21(DE3) production strain. As promising results were obtained using genomic integration for target gene expression, expression of additional folding modulator genes from the host genome also appeared suitable. However, the outcome of co-synthesis of folding modulators has been described as unpredictable in several review articles and was supposed to make a trial-and-error approach necessary (Berkmen, 2012; Delic et al., 2014; Fahnert, 2012; Kolaj et al., 2009; Overton, 2014). Furthermore, several folding modulator levels and/or transcription regulation systems were planned to be applied. Assuming that six folding modulators, three intracellular levels of each factor and two transcriptional control systems were supposed to be tested, this would have required 36 genome integration events. As the integration procedure required approximately two to three weeks at that stage of the project, this approach did not seem suitable for the required trial-and-error approach. Thus, a genome-integrated gene of interest and different plasmid-based systems for folding modulator co-synthesis were combined (see Figure 3.19).

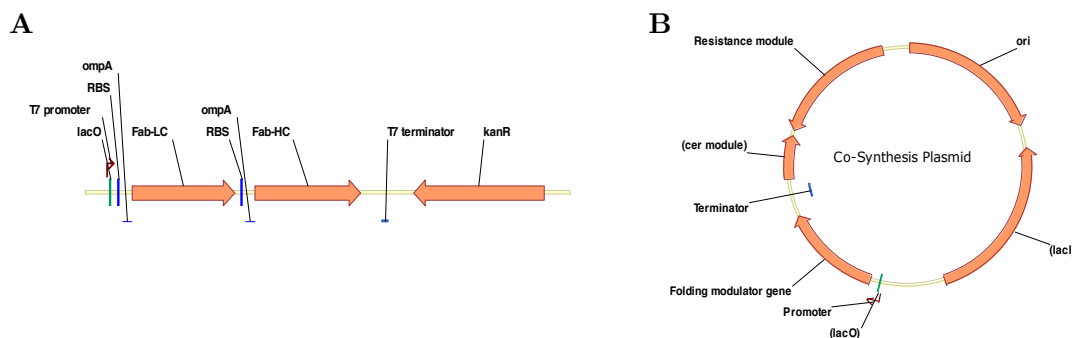


Figure 3.19: Basic set up for assessment of folding modulator co-synthesis. **A:** Dicistronic genome-integrated FabZ gene expression cassette. Transcription is initiated at a T7 promoter site (including the operator *lacO*) and terminated by a T7 terminator. Both FabZ monomers are genetically fused to an SP_{ompA} signal sequence. The genome-integrated kanamycin resistance module is located downstream of the GoI expression cassette. **B:** Plasmid for folding modulator gene co-expression. Different promoters were used to control transcription of folding modulator genes. Modules presented in parentheses were removed during the course of this study. Remaining elements of the plasmid were varied throughout the course of co-expression experiments. Details of the genetic set up of the expression plasmid employed are presented in the respective results section.

Plasmid backbone design

Different co-expression systems seemed suitable based on results of previous studies. For instance, plasmid pTUM4 was employed for co-synthesis of folding modulators in several studies (Friedrich et al., 2010; O'Reilly et al., 2014; Outchkourov et al., 2008; Schlapschy et al., 2006). It bears a p15A ori and a chloramphenicol resistance module. In different studies, the Duet plasmid system, commercially available from Novagen (Madison, WI, USA), with its components pETDuet-1 (pBR322 ori, different resistance modules), pACYCDuet-1 (p15A ori; cmlR), pCDFDuet-1 (CDF ori, stmR) and pRSFDuet-1 (RSF ori, kanR) was employed for co-synthesis (Sonoda et al., 2011; Sun et al., 2014; Wang et al., 2013; Zhuo et al., 2014). Furthermore, pKS12 (p15A ori, cmlR) (Lin et al., 2001a,b), a modified pACYCDuet-1 plasmid (p15A ori, cmlR), pBAD33 (p15A, cmlR) (Lee et al., 2013) and pAM238/pHK520 (pSC101 ori, spcR) (Mao et al., 2010) were employed in co-expression approaches. In this section, a rationale of the plasmid backbone design are presented. Components directly involved in expression of folding modulator genes are addressed in subsequent sections.

For the choice of the origin of replication (*ori*), resulting plasmid copy numbers and comparability groups were considered. Under non-selective conditions, high copy number plasmids are characterized by a higher segregational instability. This might lead to overgrowth of plasmid-free, non-productive cells and increase culture heterogeneity (Baneyx, 1999; Friehs, 2004; Huang et al., 2012; Samuelson, 2011a). In addition, an excessive folding

modulator co-synthesis was assumed to be detrimental to soluble target protein formation. While, in this study, appropriate folding modulator levels were mainly achieved by engineering the promoter and the transcription initiation module, application of mid to low copy plasmids was assumed to increase genetic stability. Plasmid copy number is regulated by the ori present on the plasmid (Snyder, 2013). Different types of ori are frequently used in recombinant protein production. For instance, pUC-, pBR322-, pSC101- and p15A-based plasmids accumulate hundreds, tens and a few (both pSC and p15A) copies per cell, respectively (Chang and Cohen, 1978; de Marco, 2013; Tolia and Joshua-Tor, 2006). Not surprisingly, ori exchange was previously suggested to be an appropriate measure for adjusting the folding modulator level (Schaefer and Plueckthun, 2010). In order to create a co-expression system with low copy numbers, only oris pSC101 and p15A were taken into consideration. Even though the current expression system bears a genome-integrated expression cassette, popular pBR322-based plasmids might be applied for gene of interest expression in future approaches. Thus, application of ori pBR322 for co-expression plasmids was ruled out. As ori p15A has been widely used in previous co-expression approaches (see, *e.g.*, Schlapschy and Skerra (2011), Schlapschy et al. (2006), Wang et al. (2013), O'Reilly et al. (2014) or Zhuo et al. (2014)), it was decided to also use this ori in the greatest part of co-expression experiments.

Chloramphenicol, kanamycin, spectinomycin and streptomycin resistance cassettes have previously been employed to select for co-expression plasmids (see above). Penicillin drugs, *e.g.* ampicillin, can trigger allergic immune responses (FDA, 2009). Due to the intended use in biopharmaceutical production, these were not considered as co-expression plasmid marker selection agents. Kanamycin resistance is the default marker when plasmid-based expression cassettes are used in the BI RCV laboratories. However, this resistance gene was integrated into the genome in several “genomic integration” strains (see Table 2.6). As no selection of plasmid-bearing strains would be possible, kanamycin resistance modules were, except for initial experiments, excluded from further considerations. As chloramphenicol and tetracycline resistance genes might be used for selection of plasmids in further strain modifications (pSIM5, 707-FLPe), they were also ruled out. Markers enabling growth under presence of the aminocyclitol aminoglycoside streptomycin are not associated with the above-mentioned drawbacks. As the respective resistance gene cassette was readily available in the laboratory and designed for fitting into the BI modular vector system, it was identified as the most suitable antibiotic for co-expression plasmid selection in “recombineered” strains.

As there were no apparent arguments against the “standard” T7 transcription terminator, it was used in initial experiments. However, in a later stage, the tZenit terminator

(Mairhofer et al., 2014; Wittwer, 2010) was employed in co-expression experiments (see Section 3.4.3).

As mentioned above, segregational instability might yield heterogeneous cultures of productive and non-productive cells under non-selective conditions. In a previous study, the *cer* determinant has been shown to effectively control the plasmid distribution during cell division (Summers and Sherratt, 1984). It was not used in initial experiments and on plasmids bearing constitutive promoters (see Section 3.4.3). However, the *cer* module was inserted into co-expression plasmids bearing T7 promoters (see Section 3.4.2).

Control of folding modulator gene expression

Inducible and constitutive promoters have been used often for steering folding modulator co-synthesis (*e.g.* the Duet plasmid system and pTUM4, respectively). Expression of folding modulator and target gene(s) under control of two different promoter systems using different plasmids has previously been suggested (de Marco and de Marco, 2004). In experiments with a two promoter, two plasmid system prior to this study, interferences of the inducers arabinose and IPTG indicated that employing different inducible promoters in one cell has to be examined cautiously (Heisteringer, 2013). Disproportionate folding modulator co-synthesis was shown to have an adverse effect on the yield of soluble and properly folded protein of interest (Martinez-Alonso et al., 2010). Furthermore, different target products might require various levels of folding modulators, making flexible regulation necessary. During this work, different modules for adjusting the intracellular folding modulator level were examined (including elements for regulation of transcriptional and translational activity).

In a master thesis, which was supervised by the author of this thesis, several “condition-inducible” expression-controlling modules were used for co-synthesis of folding modulators (Dürkop, 2014). Six promoter modules (see Table 3.8), which were PCR-amplified from the *E. coli* HMS174(DE3) genome, were used to create the basic co-expression plasmids. Specifically, the T7 promoter of plasmid pBI1KT7.1 was replaced by new modules.

To verify promoter functionality, *dsbA* expression under transcriptional control of the six modules was examined in shake flask experiments using *E. coli* strain BL21(DE3). Cell samples were examined using SDS PAGE. Under the applied conditions, promoters P_{bglA} , P_{cspA} , P_{dnaK} and P_{ibpA} showed noticeable leaky and/or condition-induced activity. On the other hand, promoters P_{phoA} and P_{fic} showed no measurable activity. If inappropriate conditions were applied during the expression experiment, incorrect sequences were

Table 3.8: Overview of different promoters that were used for controlling helper factor gene transcription during the course of the work. Constructs were designed, created and tested in a diploma thesis in the course of this study (Dürkop, 2014).

| Promoter | Native gene | Induction | Reference(s) |
|------------|----------------------------|--|--|
| P_{bglA} | 6-phospho-beta-glucosidase | constitutive this study: regular procedure | Beshay et al. (2007); Prasad et al. (1973) |
| P_{cspA} | cold shock protein | cold-shock this study: 9°C for 15 h | Giuliodori et al. (2010); Goldenberg et al. (1996); Goldstein et al. (1990); Ivancic et al. (2013); Tanabe et al. (1992); Vasina and Baneyx (1996) |
| P_{dnaK} | heat shock protein | heat-shock this study: 3x 42°C, 30 min. each | Nemecek et al. (2008) |
| P_{ibpA} | small heat shock protein | heat-stress, cytoplasmic protein misfolding/aggregation this study: 3x 42°C, 30 min. each | Arsene et al. (2000); Chuang et al. (1993); Dragosits et al. (2012) |
| P_{phoA} | alkaline phosphatase A | phosphate depletion this study: phosphate-depletion medium | Huang et al. (2012); Kikuchi et al. (1981) |
| P_{fic} | stationary phase protein | carbon depletion this study: prolonged growth | Beshay et al. (2007) |

amplified from the HMS174(DE3) genome or if the promoters were too weak to yield detectable protein yields was not further examined. Using folding modulator genes *dsbA*, *rpoE*, *surA* and *ppiD* controlled by the four active promoters, 16 co-expression plasmids were supposed to be created and tested concerning improvement of FabZ production. One of the genes, *rpoE*, encodes the RpoE sigma factor. It is involved in responses to heat-shock and misfolded peptides in the periplasmic space. This sigma factor was shown to induce activity of 20 promoters, several of which lead to an increase in levels of periplasmic folding modulators (Dartigalongue et al., 2001). Furthermore, it is essential for bacterial growth at elevated temperatures (Hiratsu et al., 1995). Information on the characteristics of DsbA, PpiD and SurA and their use in previous co-expression experiments are presented in Sections 1.3.2 and 1.3.4, respectively. Co-expression plasmid creation and subsequent transformation of recombiner and marker-free strain B<FabZ> with the resulting co-production plasmids was partly successful and yielded 14 strains. Plasmids bearing *ppiD* and *surA* genes under control of the leaky *cspA* promoter could not be created. FabZ-producing B<FabZ> cells, transformed with one of the remaining 14 co-expression plasmids, were employed in shake flask experiments. An adapted version of the T7 Preculture Medium with altered concentrations of yeast extract (increased), glucose (decreased) and glycerol (increased) (see Table 5.8, Version 2) was employed in these shake flask experiments. Cells were incubated at 37°C until an OD₅₅₀ value of 0.5 ± 0.1

was reached. After addition of IPTG to a concentration of 0.4 mM for induction of recombinant protein production, cells were incubated for 4 h under these conditions (EoF). Induction of folding modulator co-synthesis was performed following induction strategies presented in Table 3.8. Samples for protein analysis were prepared at EoF and analyzed regarding soluble and correctly folded intracellular FabZ by Western Blot. All measured concentrations ranged below the FabZ standard curve ($< 0.5 \text{ mg l}^{-1}$). However, a five-fold improvement of soluble intracellular FabZ concentration (WC-sol) caused by employing certain promoter-folding modulator combinations was tentatively identified. Co-expression prevented growth cessation and cell lysis after the first half of the protein production phase, which was observed in case of the plasmid-free reference strain. Throughout the experiments, an influence of the induction strategy applied for folding modulator co-synthesis on soluble product titer was observed. Surprisingly, this effect was independent of folding modulator which was co-produced. For instance, heat-shock treatment prior to induction of GoI expression increased the soluble product yield. In summary, obtained results did not allow for meaningful conclusions concerning suitable combinations of promoters and folding modulators. A low GoI expression rate, possibly derived from a combination of forward integration and HC-LC expression cassette composition (see above), might have prevented the detection of an influence of folding modulator co-synthesis. Furthermore, inappropriate shake flask conditions, *e.g.* too short production phases, might have exerted a negative influence on GoI expression.

The influence of DsbA, which tentatively delivered promising results in initial experiments, was also examined in a different genetic and process background. In this set of experiments, a basic plasmid with a p15A ori, a streptomycin resistance cassette, the *cer* module, the T7 terminator and different promoters (P_{bglA} , P_{dnaK} and P_{ibpA}) was employed. Plasmid-free FabZ production strain B<FabZ-kanR>(02) bears a reverse integrated LC-HC instead of a forward-integrated HC-LC expression cassette. Thus, a higher total productivity was expected. It was transformed with each of the three *dsbA* co-expression plasmids. Resulting strains were employed in shake flask and 5 l fed-batch fermentation experiments. At this stage of the study, 30°C was assumed to be the optimal production temperature for soluble FabZ. Thus, it was employed in both shake flask and fermentation approaches. As initial experiments revealed constitutive activity of the applied promoters, no particular induction conditions for folding modulator co-synthesis were applied. Shake flask experiments were performed as described in Section 2.4.1, albeit with a production phase that was conducted at 30°C. To address the expected faster growth at elevated temperatures, the production phase was shortened from 12 to 10 h. Fed-batch fermentations were performed as described in Section 2.4.2. Feed strategy 3, induction strategy 1

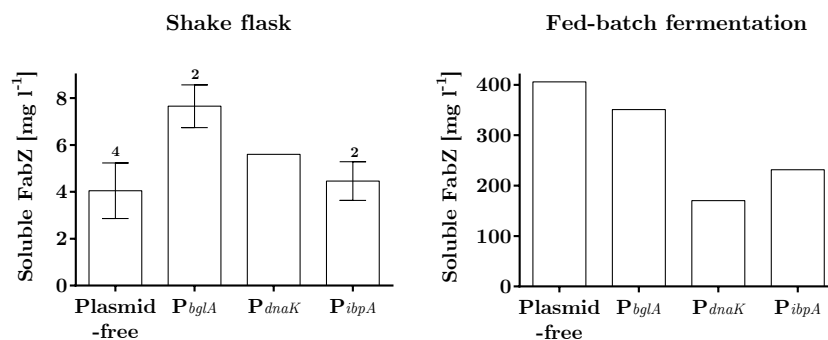


Figure 3.20: Soluble FabZ titers obtained in *dsbA* co-expression approaches with different promoters. Based on the basic FabZ production strain B<FabZ-kanR>(02) (“Plasmid-free”), three strains with increased level of the periplasmic oxidase DsbA were created. DsbA levels were controlled by different promoters (*P_{bglA}*, *P_{dnaK}* and *P_{ibpA}*), which were characterized to function in a somewhat constitutive manner under the applied conditions. Strains were employed in shake flask and fed-batch fermentation experiments (see text and Sections 2.4.1 and 2.4.2, respectively). Soluble and correctly folded FabZ content of cells and the supernatant was determined by ELISA. In case of shake flask experiments, summed values of intra- and extracellular product concentrations determined at the end of the experiment are depicted. Intracellular and supernatant product values from fermentation samples of different time points were summed. The maximum values of these progresses are presented. Numbers above the error bars (standard deviation) indicate the number of performed replicates.

and temperature ramp 1 (see Section 2.4.2) were applied in this set up. At-line (DCW, OD, glucose), off-line (soluble product titer inside of cells and in the culture supernatant, intracellular DsbA concentrations) and on-line data were collected. Soluble product titers are presented in Figure 3.20.

When the plasmid-free strain B<FabZ-kanR>(02) was employed in shake flask experiments, a soluble FabZ titer of 4.0 mg l^{-1} was determined. When DsbA was co-synthesized under control of promoters *P_{bglA}*, *P_{dnaK}* and *P_{ibpA}*, soluble FabZ titers of 7.7 , 5.6 and 4.5 mg l^{-1} were achieved, respectively. In fermentation experiments, a maximum soluble product titer of approximately 400 mg l^{-1} was determined in case of the plasmid-free expression system. DsbA co-synthesis under transcriptional control of promoters *P_{bglA}*, *P_{dnaK}* and *P_{ibpA}* resulted in soluble FabZ titers of 351 , 170 and 232 mg l^{-1} , respectively. When DsbA co-synthesis was controlled by *P_{bglA}*, *P_{dnaK}* and *P_{ibpA}*, maximum DsbA concentrations of 0.1 , 0.6 and 0.3 g l^{-1} were determined by SDS PAGE (data not shown).

In shake flask experiments, an improvement in soluble FabZ titer was observed by co-expression of the *dsbA* gene. The ranking of folding modulator co-production strains according to their soluble FabZ titer was different in shake flask ($P_{bglA}\text{-}dsbA > P_{dnaK}\text{-}dsbA > P_{ibpA}\text{-}dsbA > \text{plasmid-free}$) and fermentation ($\text{plasmid-free} > P_{bglA}\text{-}dsbA > P_{ibpA}\text{-}dsbA > P_{dnaK}\text{-}dsbA$) experiments. While co-synthesis resulted in equal or improved titers in comparison to the plasmid-free reference strain in shake flask experiments, co-synthesis of folding modulators yielded decreased soluble FabZ titers in fed-batch fermentations. Solu-

ble FabZ titers in fermentation experiments negatively correlated with the DsbA content. This supports the before-mentioned hypothesis that disproportionate overexpression of folding modulator genes might negatively impact target product titers. The fact that the ranking of strains changed in between scales suggests that different effects impact recombinant protein synthesis in the different set ups. Possibly, the applied promoters were affected differently by the culture conditions applied in the different scales. Even though a constitutive functionality of the promoters was proposed, different and quite strong overexpression of the *dsbA* gene was observed. Thus, promoters inducible by defined chemical inducers (see Section 3.4.2) or promoters defined as constitutive by literature (see Section 3.4.3) were employed in subsequent co-expression experiments.

Choice of folding modulators

An appropriate system enabling defined levels of various folding modulators was supposed to be developed within this study. In order to obtain a set of relevant co-synthesis targets, different components of the periplasmic folding machinery (see Section 1.3.2) were evaluated. Sigma factor RpoE, which is regulated as a cellular response to heat shock or outer membrane protein (OMP) misfolding (Mecsas et al., 1993; Rouviere et al., 1995), was co-synthesized with ambiguous results in a master thesis prior to this work (Heistingering, 2013). Thus, it was ruled out from further considerations.

Oxidoreductases DsbA, which catalyzes disulfide bond formation in a vectorial manner, and DsbC, involved in catalysis of disulfide bond isomerization, have been applied successfully for improving soluble product yields in co-expression experiments (see Table 1.1). As the model proteins used in this study contain only consecutive disulfide bonds, a rapid and efficient disulfide bond formation (Kadokura et al., 2004), supported by DsbA, appeared promising. No improvement concerning soluble FabX or FabZ production was expected from DsbC co-synthesis, as it was shown to be dispensable for formation of consecutive disulfide bonds (Berkmen, 2012). However, this toolbox element might become relevant for target proteins with more complex disulfide bond patterns. Thus, its effects on the host cell under recombinant protein production conditions might to reveal valuable insights and it was employed in co-synthesis approaches.

FkpA and SurA, which both combine PPIase and chaperone activity, have previously been employed to improve soluble protein production in the *E. coli* periplasm (see Table 1.1). On the contrary, a lack of effects on target protein production was reported for PPIases PpiA and PpiD in a review article from 2009 (Kolaj et al., 2009). Specifically, no beneficial

Table 3.9: Molecular weight of periplasmic folding modulators used for co-synthesis in this study. All numbers are given in kDa and represent values associated with monomers. These were calculated using ExPASy ProtParam (Gasteiger et al., 2005). The molecular weight of the precursor protein, bearing the signal peptide and N-terminal methionine, is presented in the “Precursor” column. The “Mature protein” column shows values for protein after cleavage of the signal peptide. PP: periplasmic space. IM: anchored in the inner membrane.

| Helper Factor | Precursor | Mature protein | Mono- /multimer | Position |
|---------------|--------------|----------------|----------------------|----------|
| DsbA | 23.11 | 21.13 | monomer | PP |
| DsbC | 25.62 | 23.46 | dimer | PP |
| FkpA | 28.88 | 26.22 | dimer | PP |
| PpiD | ¹ | 68.15 | ² | IM |
| Skp | 17.69 | 15.69 | monomer | PP |
| SurA | 47.28 | 45.08 | monomer ³ | PP |

¹No signal peptide predicted in the sequence (Petersen et al., 2011).

²PpiD has been found as a homodimer and -trimer in the membrane (Stenberg et al., 2005).

³In numerous articles, no evidence for SurA multimers were found.

effects of PpiA co-synthesis were observed during scFv (Bothmann and Plueckthun, 2000) and single-chain T-cell receptors (Wulfing and Pluckthun, 1994) production experiments. Furthermore, the *rotA* (*ppiA*) gene seemed dispensable and was, in accordance, removed from the co-expression vector during construction of the co-expression plasmid pTUM4 (Schlapschy et al., 2006). Like PpiA, PpiD does not seem to have beneficial effects on target protein production (Kolaj et al., 2009). However, PpiD was shown to be an important PPIase and chaperone within the periplasmic chaperone network (Matern et al., 2010). It directly interacts with proteins which enter the periplasmic space (Antonoaea et al., 2008). Due to its increased importance within the cell, PpiD was used for co-synthesis experiments within this study.

The general periplasmic chaperone Skp has a broad substrate spectrum that includes outer membrane and soluble periplasmic proteins. It was previously applied to improve soluble product titers of soluble antibody fragments and T-cell receptors (see Table 1.1). Due to its broad target protein spectrum and promising results from other studies, Skp was employed in co-expression experiments.

In conclusion, three additional folding modulators (DsbC, FkpA and Skp) were employed besides the folding modulators DsbA, PpiD and SurA, which were examined in the initial experiments. In total, six factors from all groups of the periplasmic folding machinery (see Section 1.3.2) were chosen for co-synthesis. Molecular weights of precursor and mature proteins are presented in Table 3.9.

3.4.2 Co-Synthesis of Folding Modulators Under Control of the T7 RNA Polymerase

Rationale for using and composition of the co-synthesis system

Promoters with environmental condition-dependent activities were applied for expression of folding modulator genes in a previous part of this study (see Section 3.4.1). Issues associated with induction of folding modulator gene expression by process variation were observed. Furthermore, different folding modulator gene expression levels were assumed to be, at least in part, responsible for differences observed in outcomes of shake flask and fermentation experiments. Chemically inducible and constitutive promoters were evaluated to constitute viable alternatives. Constitutive folding modulator synthesis might lead to increased metabolic burden during bacterial cultivations. Slower growth and plasmid loss were expected. Thus, induction by addition of a chemical compound appeared to be the method of choice at this stage of the study. Several systems for induction of protein production (*e.g.* the IPTG/lactose-inducible T7/P_{lac}/P_{lacUV5}/P_{tac} (Gronenborn, 1976; Studier, 1991; Studier and Moffatt, 1986), arabinose-inducible P_{araBAD} or the *m*-toluate-inducible XylS/*Pm* (Aune, 2008; Aune et al., 2010) system) were described (Arpino et al., 2013; Overton, 2014) and available. Since the T7 promoter-based and IPTG-inducible system was also used for the expression of the gene of interest, it seemed reasonable to use the same system for co-expression of folding modulator genes. As the same inducer activates the synthesis of both target and folding modulator gene, no process adjustments were necessary. High-rate expression of the gene of interest but only limited expression of the folding modulator gene was desired. However, folding modulator gene expression levels could not be adjusted by IPTG titration, since both systems were interlinked by the same inducer. Thus, different T7 promoter elements were designed.

In conclusion, a T7 RNA polymerase-controlled co-synthesis system was applied for both genome-integrated gene of interest and plasmid-borne folding modulator gene. An overview of the T7 system in general is presented in Figure 3.21, albeit for plasmid-based expression. The *E. coli* BL21(DE3)-based production strain bears a λ DE3 prophage with a gene encoding the highly active T7 RNA polymerase. Its synthesis is under control of the *lacUV5* promoter (Studier and Moffatt, 1986). This rather weak (Deuschle et al., 1986) *lac* promoter mutant reduces but does not completely eliminate sensitivity to catabolite regulation (Lanzer and Bujard, 1988; Silverstone et al., 1970). The *lacO* site within the T7 RNA polymerase promoter enables repression of transcription initiation of the phage enzyme (Dubendorf and Studier, 1991). Synthesis of the LacI repressor is under tran-

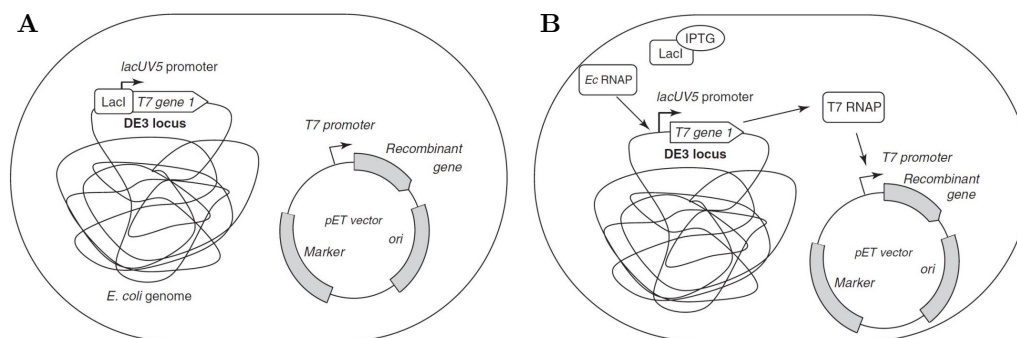


Figure 3.21: General components and function of the T7 system without (A) and with (B) the inducer IPTG. The figure was adapted from a review article (Overton, 2014). In this figure, the recombinant gene expression cassette is borne by a pET vector, whereas it was genome-integrated in this study.

scriptional control of the *lacI^q* promoter. This is a mutated version of the *lacI* promoter, which produces ten-fold more repressor compared to its native counterpart (Calos, 1978; Muller-Hill et al., 1968). Upon addition of IPTG, the T7 RNA polymerase gene is transcribed by the native *E. coli* RNA polymerase. The genome-integrated target gene is under control of a T7*lac* hybrid promoter (“T7 promoter”), which bears the LacI-binding site *lacO*. Upon addition of IPTG, LacI is inactivated by binding to the inducer and transcription can occur. Note that the genomic gene of interest expression cassette does not bear a *lacI* gene. The folding modulator gene is located on a plasmid within the host cell (see Figure 3.19, B). Different variants of the T7 promoter were employed to adjust folding modulator co-synthesis. During initial experiments, T7 promoter elements containing the *lacO* site were applied. Moreover, an additional copy of the *lacI* gene under transcriptional control of the strong *lacI^q* promoter was present on the folding modulator plasmid. Regulatory elements (*lacI*, *lacO* sites) were removed during the course of this study. Note that the *E. coli*-endogenous *lac* operon also bears one *lacO* site and one copy of *lacI*, which produces 5 to 10 molecules per cell and generation (Gilbert and Muller-Hill, 1966; Muller-Hill et al., 1968).

An initial system for T7 RNA polymerase-based folding modulator co-synthesis

A T7 promoter system for folding modulator co-synthesis was designed and tested. Substantial work was performed in the course of a master thesis under supervision of the author of this work (Schuller, 2015). The plasmid backbone consisted of a p15A ori, a streptomycin resistance gene, a *cer* module and a T7 transcription terminator. These components were selected based on considerations described in Section 3.4.1. The plasmid’s transcription regulation module was initially composed of a T7 promoter and its regulatory elements, *i.e.* the *lacI* gene and the *lacO* repressor binding site. Furthermore,

translation was initiated at a strong ribosomal binding site (RBS). The T7 promoter sequence was modified to adjust different levels of the transcribed gene.

The 250 bp *cer* site was previously reported to increase multi-copy plasmid stability (Stirling et al., 1988; Summers and Sherratt, 1984, 1988), albeit for pBR322-based plasmids. pBR322 and p15A oris are similar in sequence (Bird, 1981) and plasmid replication mechanism (Selzer et al., 1983). Thus, similar effects of the *cer* site were expected for ori p15A (ori 4 in the modular vector system). During plasmid stability evaluations in a master thesis supervised by the author of this work, the combination of p15A ori and *cer* was determined to be as stable as the combination of pBR322 ori and *cer* (Schuller, 2015). In conclusion, the plasmid pBI4iST7(n).2 backbone, with different T7 promoters (see Figure 3.22 and Table 2.3) was applied in initial helper factor gene co-expression experiments.

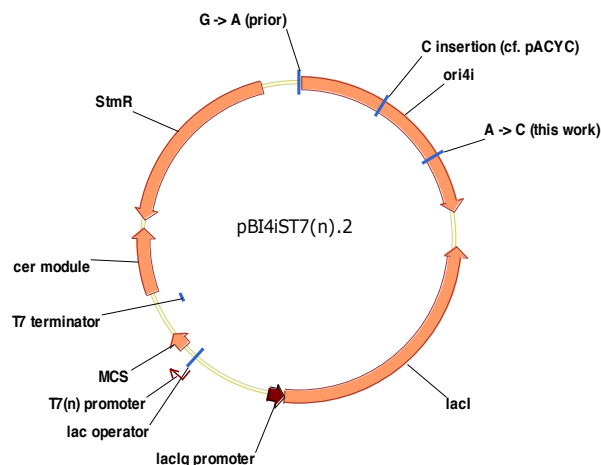


Figure 3.22: Schematic plasmid map of pBI4iST7(n).2. The plasmid bears the p15A origin of replication (ori 4i) and a streptomycin resistance gene cassette (StmR). It also harbors the plasmid elements of the T7 expression system, including the gene of the LacI repressor under control of the *lacI^q* promoter, the lac operator *lacO* and the sequences of T7 promoter and terminator. To enhance plasmid stability, pBI4iST7(n).2 harbors the *cer* module. Mutations in the origin of replication sequence are indicated. These appeared prior to this work (“prior”), existed in the original pACYC plasmid (“cf. pACYC”) or were introduced during this study (“this work”).

Determination of the promoter strength of T7 promoter mutants

Different variants of the T7 promoter sequence were designed and analyzed before (Imburgio et al., 2000). Based on the sequences described in that work, 9 promoters were chosen to replace the native T7 promoter of pBI4iST7.2-gfp.1. Plasmids were created and promoter activities were determined based on fluorescence measurement in *in vivo* experiments in a master thesis supervised by the author of this work (Schuller, 2015). A

summary of results is presented in Figure 3.23.

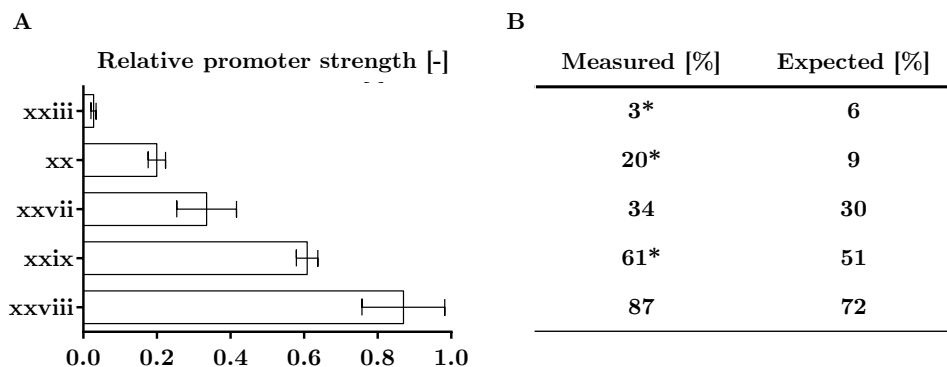


Figure 3.23: Promoter strength values of selected T7 mutants relative to the strength of the native promoter (A) and comparison of measured and expected values (B). BL21(DE3) cells with genome-integrated FabZ gene bearing pBI4iST7(n).2-gfp.1 plasmids with the indicated promoter were subjected to promoter tests as described in Section 2.3. A: Mean values from biological triplicates are presented in dependence of the T7 promoter used for steering *gfp.1* expression. Error bars indicate standard deviations, calculated by the law of propagation of uncertainty. Note that GFP fluorescence development within 20 h instead of 4 h of recombinant protein production phase was used for calculation of the weakest promoter T7xxiii. Four of the nine tested promoters did not show any activity. These are not presented in the figure. Results were generated within the course of a master thesis supervised by the author of this work (Schuller, 2015). B: Comparison of values measured in *in vivo* experiments (“Measured”) to results from previous *in vitro* experiments (“Expected”, Imburgio et al. (2000)). In addition to the native T7 promoters, promoters associated with values marked with asterisks were used in later co-expression experiments.

Test of the initial co-synthesis system in shake flask experiments

To evaluate the influence of folding modulators and their level on soluble protein production three engineered promoters in addition to the native T7 promoter sequence (T7: 100 %, T7xxix: 61 %, T7xx: 20 %, T7xxiii: 3 %) were employed for co-expression experiments. Folding modulator genes (*dsbA*, *dsbC*, *fkpA*, *ppiD*, *skp*, *surA*) were previously amplified from the *E. coli* HMS174(DE3) genome (Heistinger, 2013) and were available in plasmids derived from that work. DNA and amino acid sequences are presented in the confidential supplement (see Section 5.4.7). Genes were used to replace a *gfp.1* gene in pBI4iST7(n).2-gfp.1 plasmids, which were previously used to determine promoter strength (see above) (Schuller, 2015). To evaluate the influence of the plasmid backbone and its regulatory elements, the empty basic plasmids pBI4iST7(n).2 (see Figure 3.22 and Table 2.3) were also employed. In total, 28 plasmids were created and used to separately transform the FabZ production strain B<FabZ-kanR>(02). The plasmid-free strain was employed as a reference in production experiments. These were performed according to the standard shake flask procedure described in Section 2.4.1. Note that shake flasks were either inoculated from bacterial glycerol stocks stored at -80°C or from single colonies

from agar plates after transformation with the respective plasmid. Due to several hours of growth prior to induction (preculture and growth phase), no influence of the starting material on the experimental outcome was expected. Concentrations of soluble and correctly folded FabZ in the cells and in the supernatant were determined by ELISA. For selected samples, SDS PAGE analyses were performed to evaluate total (soluble + insoluble) intracellular FabZ concentrations. Multiple shake flask replicates were performed for each strain. Note that most of the work was performed in the course of a master thesis supervised by the author of this work (Schuller, 2015). A summary of the results is presented in Figure 3.24.

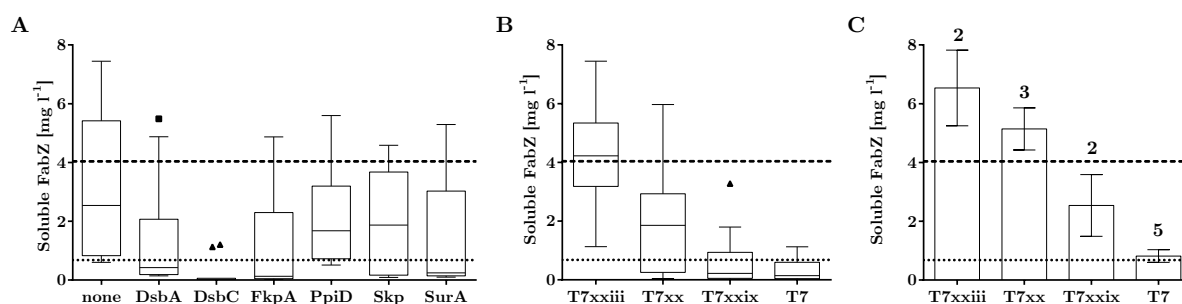


Figure 3.24: Co-synthesis using T7 promoters with regulatory elements: influence of folding modulator and promoter strength on intracellular soluble FabZ titer in shake flasks. T7xxiii, T7xx, T7xxix and T7 indicate the strength of different T7 promoters (3, 20, 61 and 100 %, respectively). Employed strains, cultivation set up and analysis procedures are described in the text and Section 2.3. **A:** Tukey box-and-whisker plot representation of titers in dependence of the co-synthesized folding modulator. **B:** Tukey box-and-whisker plot representation of titers in dependence of the promoter employed for co-synthesis. **C:** Bar chart diagram of intracellular FabZ titers produced by empty basic plasmid strains. The T7 promoter variant present on the basic plasmid is indicated on the category axis. Bar height represents mean values of soluble intracellular FabZ titers derived from multiple shake flask experiments. Error bars represent the determined standard deviations. Numbers atop bars indicate the number of performed replicates. In all figures, dotted and dashed lines represent intracellular and total (intra- and extracellular) soluble FabZ titers, respectively, which were measured when the plasmid-free reference strains was employed.

Due to errors in sample handling, selected values were excluded from the analysis. Thus, certain approaches for the empty basic plasmids bearing T7xxiii and T7xxix and plasmids encoding DsbA (controlled by T7xxiii), DsbC (T7xxiii), Skp (T7) and SurA (T7xx) were performed as duplicates only. As, at least, duplicates were performed for each combination and overall trends (as opposed to single data points) were considered, the data set was assumed to be appropriate for the conclusions presented below. For 26 of 28 combinations of folding modulator and promoter, soluble FabZ concentrations in the supernatant were examined at least once. Values of above 0.1 mg l^{-1} with an overall maximum of 0.8 mg l^{-1} were determined only for three combinations (T7xx on empty basic plasmid, T7xxix-*fkpA*, T7xxix-*skp*). As values for both further genetic constellations (lowered or increased T7 promoter strength) were significantly lower in these cases, higher supernatant titers were defined to be outliers. Thus, product FabZ titers in the supernatant seemed negligibly

low and were not further analyzed (except for the plasmid-free reference strain). For the plasmid-free reference strain, mean titers of 0.7 ($n = 6$) and 3.4 ($n = 4$) mg l^{-1} were determined in the intra- and extracellular fraction, respectively. Intracellular and total (intra- and extracellular) soluble FabZ concentrations are presented as dotted and dashed lines in Figure 3.24, respectively. Dependence of soluble intracellular FabZ concentration on folding modulators, promoters and interaction of the two factors was analyzed by ANOVA. It was calculated that effects of both factors and the interaction term were significant at the 5 % level ($p < 2.00 \times 10^{-16}$, 2.00×10^{-16} and 1.07×10^{-8} , respectively). Visual data evaluation (see Figure 3.24 (A) and (B)) supported this statistical result. Soluble FabZ concentrations of 6.5, 5.1, 2.5 and 0.8 mg l^{-1} were determined in cell samples when empty basic plasmids with promoters T7xxiii (3 % promoter strength), T7xx (20 %), T7xxix (61 %) and T7 (native) were present in the FabZ-producing strain, respectively (see Figure 3.24(C)). When folding modulator genes were present on plasmids, titers were reduced and ranged from 1.2 to 5.2 (T7xxiii), below 0.1 to 2.7 (T7xx), below 0.1 to 1.2 (T7xxix) and below 0.1 to 0.5 (T7 native). FabZ titers decreased with increasing promoter strength for all folding modulators (see Figure 3.24 (B)). Co-synthesis of DsbC yielded low soluble product titers (maximum: 1.2 mg l^{-1}), independent of the applied promoter. $\text{OD}_{550} = 2.0$ ($n = 6$, $\text{SD} = 0.2$) was determined at the end of the experiment when the plasmid-free reference strain was employed. OD_{550} ranged between 7.3 and 11.4 in case of all 28 folding modulator co-expressing strains (mean $\text{OD}_{550} = 8.4$, $n = 80$, $\text{SD} = 0.8$). SDS PAGE analysis of samples from the genome-integrated reference strain revealed two prominent product-related bands. In case of strains bearing one of the four empty basic plasmids, no product-related bands were detectable on SDS PAGE gels (data not shown).

Test of the initial co-expression system in fermentation experiments

Selected strains were employed in fermentation experiments. Results from comparative fermentation experiments employing the plasmid-free reference strain were already presented in Section 3.3.1. These served as reference values. To examine the influence of the basic plasmid, the strain bearing the empty basic plasmid with the T7 strongest promoter, which showed the highest impact on performance in shake flasks was chosen for fermentation. Furthermore, four strains, which yielded highest titers in shake flask experiments (T7xxiii promoter, folding modulators DsbA, PpiD, Skp and SurA) were employed in fermentation experiments. 5 l fed-batch fermentations were performed as described in Section 2.4.2. At that state of the project, 28°C appeared to be the most appropriate growth temperature during FabZ production. Thus, the temperature was adjusted from 37 to 28°C prior to the production phase. Temperature ramp 2, feed strategy 3 and induc-

tion strategy 1 were applied. Similar to analysis of samples from shake flask experiments, intracellular soluble and total FabZ concentrations were examined by ELISA.

Soluble intracellular FabZ titers of 232 (plasmid-free reference strain, $n = 3$, $SD = 74 \text{ mg l}^{-1}$), 4 (basic plasmid, strong T7 promoter), 30 (T7xxiii-*dsbA*), 30 (T7xxiii-*ppiD*), 30 (T7xxiii-*skp*) and 30 (T7xxiii-*surA*) mg l^{-1} were determined. Negligible product amounts of below 0.5 mg l^{-1} were determined in the supernatant when plasmids pBI4iST7.2 or pBI4iST7xxiii.2-*dsbA* were present in the production strain. As, judging by intracellular product concentrations, negligible amounts were also expected in the supernatant of fermentations of other strains, no ELISA measurements were conducted for these samples.

Conclusion: interference of gene of interest and folding modulator gene expression

In conclusion, the partially positive results from shake flask experiments could not be verified in fed-batch fermentations. The positive influence of co-expression plasmid presence in shake flasks was shown to derive from components of the basic plasmid and not from the employed folding modulator. Rather, folding modulator co-synthesis seemed to convey a negative effect on soluble FabZ production. Independent of T7 promoter strength, OD_{550} values at the end of shake flask experiments were increased and bands representing the total intracellular product were not detectable when the empty basic plasmids were present. Thus, it seems that non-promoter elements present on the plasmid negatively influence total productivity. This, in turn, might, positively affect growth behavior. The presence of the empty basic plasmid with the strongest T7 promoter reduced intracellular soluble FabZ titers about 50-fold in fermentation experiments compared to the plasmid-free host ($232 \rightarrow 4 \text{ mg l}^{-1}$). If the 5- to 10-fold increase of titers folding modulator co-producing strains compared to this value is derived from the lower strength of the used T7 promoter or from the folding modulator presence, remains elusive. In any case, a dramatic negative impact of the basic plasmid on production behavior was observed in both types of production experiments. Since the target gene expression rate was not supposed to be dependent on folding modulator co-production, this was a surprising critical issue. The plasmid-borne resistance cassette, ori, terminator and *cer* module were expected to potentially influence the host metabolism. However, an impact of this extent on production characteristics was not expected. A closer look on the further genetic elements (*lacI* and *lacO* sites as well as the chosen promoters), suggested interferences with target gene expression in a complex manner (see Figure 3.25).

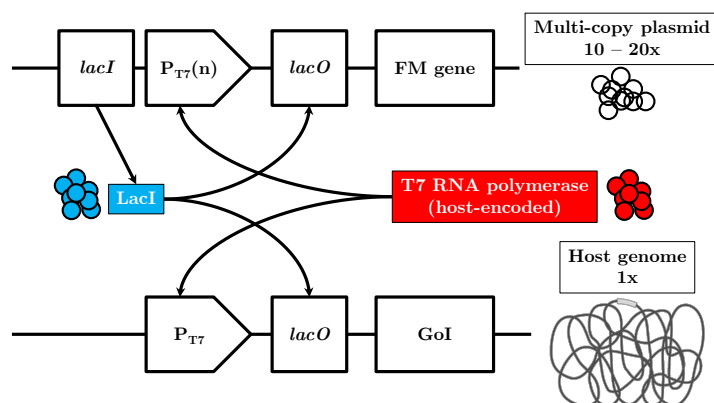


Figure 3.25: Potential interference of two T7 RNA polymerase-based expression systems (for GoI and folding modulator gene expression). Arrows indicate binding of proteins to target DNA sequences. LacI (*lacI*): lac inhibitor (gene). *lacO*: lac operator site(s). FM: folding modulator. $P_{T7}(n)$: T7 promoters variants. GoI: gene of interest. Note that elements of the lac operon native to *E. coli* are not depicted.

T7 promoter-controlled transcription is known to lead to rapid accumulation of mRNA in the cell (Studier and Moffatt, 1986). Abundance of folding modulator mRNA might compete with the target gene's mRNA for the host's translational machinery. In the performed shake flask experiments, however, the influence on protein of interest (PoI) concentration did not correlate with sequence length of the folding modulator gene (*dsbA*: 624 bp, *dsbC*: 705 bp, *fkpA*: 810 bp, *ppiD*: 1869 bp, *skp*: 483 bp, *surA*: 1284 bp; $p \geq 0.177$). If ribosome competition had had a major influence on PoI concentration, a correlation would have been expected due to increased ribosome utilization dependent on the helper mRNA's size. Furthermore, lower T7 promoter strength did not result in increased overall PoI formation (*i.e.* detectable product-related bands; data not shown). It seemed that folding modulator mRNA and utilization of ribosomes was one but not the major point of interference with target gene expression. Note that the number of regulatory elements *lacI* and *lacO* is also drastically increased in cells, which bear multi-copy plasmid pBI4iST7(n).2. Based on the results above, these two elements were identified to disturb the current system as follows (see Figure 3.25). Multiple *lacI* copies under control of the strong *lacI^q* promoter (Muller-Hill et al., 1968) resulted in high levels of the repressor LacI in the cell. Apparently, *lacO* sites provided by multi-copy plasmid presence is outnumbered by LacI repressors. Thus, even high IPTG levels could not antagonize LacI repressor levels in the cell. In detail, repressor binding to *lacO* elements is not sufficiently prevented and expression of the GoI was impaired. Thereby, formation of PoI aggregates was also prevented, which results in higher OD₅₅₀ values at the end of the experiments for strains harboring a co-expression plasmid. In addition, multiple copies of the T7 promoter of the co-expression plasmids competed for available T7 RNA polymerase molecules with the promoter of the target gene expression cassette. Thus, strong co-expression promoters

impaired target gene expression even further (see Figure 3.24 (C)). When using one of the two T7 promoters with lower strength (T7xx, T7xxiii), the system seemed to be balanced in a way, which is beneficial for soluble FabZ production. In case of strong T7 promoters (T7, T7xxix), on the other hand, soluble FabZ concentrations were lower than the total soluble FabZ concentration obtained with the plasmid-free reference strain.

Introduced modifications decreased the extent of interference

Exploiting the high productivity of the plasmid-free FabZ production strain B<FabZ-kanR>(02) by improving periplasmic folding capacities was the aim of this part of the study. Thus, the strong influence of the empty basic plasmid on target gene expression was highly undesired. When the plasmid-free strain was cultivated, filtration during sampling was hardly possible, sedimentation of cells yielded a viscous and not well-defined pellet, prominent product-related bands were detected by SDS PAGE analysis and a low OD₅₅₀ value of 2.0 (n = 4, SD = 0.27) was measured at the end of the experiment. The plasmid pBI4iSara.2 (P_{araBAD} instead of P_{T7} promoter module) carries neither *lacI* nor *lacO* sites nor a T7 promoter. To validate the interference hypothesis stated above, the genome-integrated production strain was transformed with this plasmid and employed in shake flask experiments described in Section 2.4.1. Note that a production temperature of 30 °C and a duration of 10 h was applied in these experiments due to optimization purposes. Results from this single experiment were compared to the results obtained with the plasmid-free strain under the same conditions. In case of the pBI4iSara.2-bearing strain, a final OD₅₅₀ value of 2.1 was measured. SDS PAGE analysis revealed prominent product-related protein bands. Filtration during sampling was impeded and sedimentation of cells yielded a viscous and not well-defined pellet. In conclusion, the FabZ-production strain bearing plasmid pBI4iSara.2 was characterized by behavior similar to the basic plasmid-free production strain. As none of the identified interfering elements (T7 promoter, *lacI*, *lacO*) was present on the plasmid, the result suggests that one or several of the elements were, indeed, involved in observed interference phenomena.

To evaluate the influence of the additional *lacI* copies, the repressor gene of plasmid pBI4iST7.2 was replaced by a non-functional spacer sequence. The resulting plasmid pBI4iST7.5 (native T7 promoter), with *lacO* sites but without the *lacI* module, was used to transform the FabZ-producing strain. Strains were employed in a total of five shake flask experiments. In two experiments, preculture duration had to be extended by approximately 9 h in order to obtain cell densities sufficient for inoculation of main culture. Main culture growth phases, *i.e.* phases before induction of recombinant protein produc-

tion, were prolonged by up to 12 h in some experiments. Prolonged growth phases were observed before (see Section 3.2.3) and revealed significant basal expression. Apparently, a high number of *lacO* sites on the co-expression plasmid caused a reduction of LacI repressor and, thus *lacO* sites of promoters of the T7 RNA polymerase and target gene expression cassettes could no longer be sufficiently blocked. Using two T7 expression systems for target and folding modulator gene seemed to require a high degree of fine-tuning to enable both efficient target protein production and improved folding capabilities. A plasmid system without the regulatory elements *lacI* and *lacO* seemed more suitable for helper factor co-synthesis.

Therefore, co-expression plasmids pBI4iST7(n).7 were generated (see Table 2.3). These bore the different T7 promoters (T7, T7xxix, T7xx, T7xxiii) and were devoid of *lacI* as well as *lacO*. Basis for the genetic construction were the plasmids of the pBI4iST7(n).5 series (devoid of the *lacI* module and bearing the T7 promoters of different strength). New T7 promoter modules (see Section 2.2.1) devoid of *lacO* sites were used to replace according *lacO*-containing T7 promoters in plasmids of the pBI4iST7(n).5 series.

Newly designed regulator-free T7 promoter plasmids were examined concerning their influence on target gene expression. The four basic plasmids were used to transform the basic production strain B<FabZ-kanR>(02). Resulting strains were employed in shake flask experiments, which were performed as described in Section 2.4.1. Concentrations of soluble (intracellular and supernatant) and total (intracellular) FabZ were determined by ELISA and SDS PAGE analysis, respectively. To estimate growth behavior and cell vitality, OD₅₅₀ values and culturable cell numbers were determined. Results are presented in Figure 3.26.

For the plasmid-free FabZ production strain, a mean total soluble product concentration ($n = 4$) of 4.0 mg l^{-1} (0.7 mg l^{-1} (intracellular, $n = 6$) + 3.44 mg l^{-1} (supernatant, $n = 4$)) was determined. When empty co-expression plasmids were present, soluble product titers of 6.5 ($1.6 + 4.9$), 12.8 ($7.9 + 4.9$), 10.2 ($9.9 + 0.3$) and 6.1 ($6.1 + 0.0$) mg l^{-1} were determined for promoters T7xxiii, T7xx, T7xxix and T7, respectively (see Figure 3.26 (A)). For the reference strain (plasmid-free), a mean OD value ($n = 6$) of 2.0 was determined at the end of cultivation. In case of empty plasmid-carrying cells, mean OD₅₅₀ values of 2.2 (promoter T7xxiii), 7.7 (T7xx), 8.7 (T7xxix) and 8.1 (T7) were determined (see Figure 3.26 (B)). Culturable cell numbers of 2.03×10^5 , 2.85×10^7 , 2.93×10^9 and 4×10^9 cfu ml⁻¹ were determined for cells harboring the empty basic plasmids pBI4iST7xxiii.7 (3 %), pBI4iST7xx.7 (20 %), pBI4iST7xxix.7 (61 %) and pBI4iST7.7 (100 %), respectively (see Figure 3.26 (C)). Product-related bands on SDS PAGE gels were most pronounced when

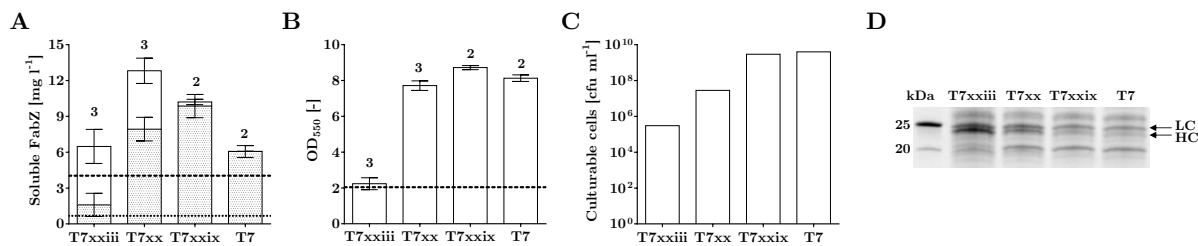


Figure 3.26: Influence of the empty co-expression plasmids with T7 promoters of different strength and without regulatory elements. Four plasmids with modified T7 promoters and devoid of *lacI* and *lacO* site were created (pBI4iST7(n).7). T7xxiii, T7xx, T7xxix and T7 indicate the strength of different T7 promoters (3, 20, 61 and 100 %, respectively). The FabZ-producing strain B<FabZ-kanR>(02) was transformed with these plasmids. Shake flask experiments were performed as described in Section 2.4.1. Different analyses were performed based on samples prepared at the end of the experiment. Multiple shake flask experiments were performed for all strains. Numbers atop bars indicate the number of performed experiments. **A:** intracellular (dotted bars) and extracellular (blank bars) soluble FabZ concentrations in dependence of the promoter used. Values were determined by ELISA analysis. Bar heights represent mean values determined for multiple shake flask experiments. Error bars indicate the standard deviations. **B:** OD_{550} values at the end of the experiment in dependence of the employed promoter. Bar height represents mean values determined from multiple measurements. Error bars represent the standard deviations. **C:** culturable cell numbers determined at the end of the experiment in single shake flask experiments. Note: logarithmic value axis. **D:** SDS PAGE analysis (reducing and denaturing conditions) gel sections of cell samples affiliated with the use of different T7 promoters. Product monomers have a molecular weight of 23.14 (LC) and 23.20 (HC) kDa, respectively.

the weakest T7xxiii promoter was applied and the protein band pattern resembled the pattern associated with the plasmid-free reference strain closely (data not shown). New promoter constructs were also tested in promoter tests using BL21(DE3) cells transformed with one of four pBI4iST7(n).7-gfp.1 plasmids (data not shown). The weakest of the T7 promoters (T7xxiii) yielded the highest fluorescence values. Furthermore, a significantly increased OD_{550} of 11.1 was determined (T7: 4.1, T7xxix: 4.5, T7xx: 5.1).

In conclusion, the performance (*e.g.* total soluble FabZ yield, product appearance in the supernatant, SDS PAGE band pattern and OD value at EoF) of the production strain transformed with the basic plasmid bearing no regulatory elements and the weakest T7xxiii promoter (pBI4iST7xxiii.7) was comparable to that of the plasmid-free strain. Positive correlation of cell count and T7 promoter strength (see Figure 3.26 (C)) indicated a tendency towards decreased cell viability with lowered influence of regulatory elements. This low viability was also expected for the plasmid-free reference strain. It also appeared that excessive GFP production under control of all but the weakest promoters resulted in metabolic burdening of the cell in promoter tests.

Note that certain combinations of regulatory elements on basic plasmids even improved soluble FabZ yields in shake flask (see Figures 3.24 (C) and 3.26 (A)). Apparently, additional regulatory elements enable improved harmonization of target gene expression rate

and capacities of cellular machineries involved in synthesis of soluble and properly folded FabZ molecules. However, the aim of this part of the study was to design a T7-based co-expression system enabling adjustment of different cellular levels of periplasmic folding modulators without interfering with target gene expression. Thus, the only appropriate of several examined basic plasmids appeared to be pBI4iST7xxiii.7, which bears a weak T7 promoter and no regulatory elements (see Figure 3.27).

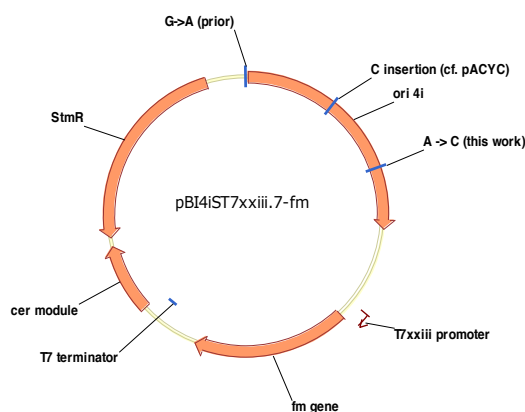


Figure 3.27: Schematic plasmid map of pBI4iST7xxiii.7 bearing a folding modulator (fm) gene. The plasmid bears the p15A origin of replication (ori 4i), a streptomycin resistance gene (StmR), the T7 terminator and the *cer* module. Mutations in the origin of replication sequence are indicated. These appeared prior to this work (“prior”), existed in the original pACYC plasmid (“cf. pACYC”) or were introduced during this study (“this work”). Transcription is regulated by the weak T7xxiii promoter. Note that the plasmid is devoid of both regulatory elements *lacI* and *lacO*, which are usually associated with the T7 promoter system. This plasmid was used for co-synthesis of different folding modulators in production experiments.

Influence of T7 RNA polymerase-based folding modulator co-synthesis on FabZ production in shake flask experiments

In the previous section, a basic plasmid with a weak T7 promoter and without regulatory elements *lacI* and *lacO* (pBI4iST7xxiii.7) was identified as appropriate for co-synthesis of folding modulators. Folding modulator co-synthesis experiments with this basic plasmid were performed in this part of the study. The influence of high-level folding modulator co-synthesis and interferences between plasmid elements and target gene expression were examined by also applying the stronger T7xx promoter variant. Co-expression plasmids were created by cloning folding modulator genes *dsbA*, *dsbC*, *fkpA*, *ppiD*, *skp* and *surA* into basic plasmids pBI4iST7xxiii.7 and pBI4iST7xx.7 individually. Plasmids were used to transform the FabZ-producing strain B<FabZ-kanR>(02) with genome-integrated target gene. Strains were employed in shake flask experiments. These were performed as described in Section 2.4.1. Multiple experiments were performed for the weak T7xxiii promoter variant, while single co-expression experiments were performed for the stronger T7xx promoter. Samples from the end of the production phase (25 °C, 1 mM IPTG, 12 h) were analyzed concerning FabZ concentration by ELISA and SDS PAGE analysis according to Sections 2.6.3 and 2.6.4, respectively. Furthermore, OD₅₅₀ values were determined at EoF. Results are presented in Figure 3.28.

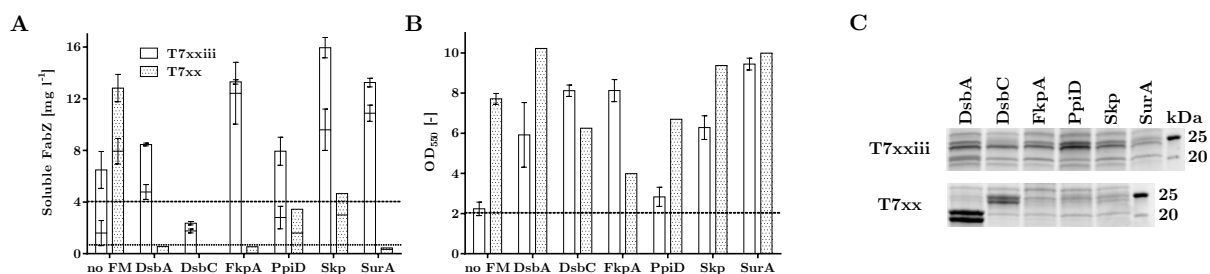


Figure 3.28: Influence of co-synthesis of periplasmic folding modulators using a newly developed T7 RNA polymerase-based system in FabZ production experiments in shake flasks. Twelve plasmids bearing one of two T7 promoters (T7xxiii and T7xx with low and moderate activity, respectively) and one of six folding modulator genes (*dsbA*, *dsbC*, *fkpA*, *ppiD*, *skp* and *surA*) were created. These were used for transformation of the FabZ-producing host strain. Shake flask experiments were performed as described in Section 2.4.1. Duplicate and single shake flask experiments were performed for folding modulator synthesis-controlling promoters T7xxiii and T7xx, respectively. **A**: Soluble FabZ concentrations inside the cells (lower bar section) and in the supernatants (upper bar sections) were determined by ELISA analysis. Blank and dotted bars for each folding modulator represent FabZ concentrations associated with co-expression promoters T7xxiii and T7xx, respectively. Dotted and dashed lines represent intracellular ($n = 6$, $SD = 0.3 \text{ mg l}^{-1}$) and total ($n = 4$, $SD = 0.7 \text{ mg l}^{-1}$) soluble FabZ concentrations determined for the plasmid-free reference strain, respectively. **B**: OD₅₅₀ values were determined at the end of the recombinant protein production phase. The dashed line represents the mean value obtained in six shake flask experiments using the plasmid-free reference strain. **C**: Cell samples were analyzed by SDS PAGE analysis (reducing and denaturing conditions). Product-associated sections (20 to 25 kDa) of the gel are presented.

As reported above, the presence of the empty basic plasmids had a slight (T7xxiii: 6.5 mg l⁻¹) or even strong (T7xx: 12.8 mg l⁻¹) positive influence on total soluble product yields compared to the plasmid-free strain (4.0 mg l⁻¹). Total soluble FabZ titers were improved between 3- and 4-fold when folding modulators FkpA (13.3 mg l⁻¹), Skp (16.0 mg l⁻¹) and SurA (13.3 mg l⁻¹) were co-synthesized using the weak T7xxiii promoter. Co-synthesis of DsbA (8.5 mg l⁻¹) and PpiD (7.9 mg l⁻¹) resulted moderately increased titers. A negative effect on total soluble product titers was observed for DsbC co-synthesis (2.4 mg l⁻¹). Different folding modulators also affected OD values and the overall target product amounts (SDS PAGE) in a different manner. Folding modulator co-synthesis under control of the stronger T7xx promoter had no or negative effects on total soluble FabZ titers. Specifically, FabZ titers were at or below the levels of strains without increased folding modulator level (depending on the reference used; empty basic plasmid-bearing with 12.8 mg l⁻¹ or plasmid-free strain with 4.0 mg l⁻¹). The maximum soluble FabZ titer of co-expression strains using this promoter was obtained when combination T7xx-*skp* was used (4.7 mg l⁻¹). Product-related bands on SDS gels were more intense in case of the weaker T7xxiii promoter (see Figure 3.28 (C)). Note that the dominant bands in the bottom SDS PAGE gel cut out (T7xx) probably derived from co-synthesized folding modulators (DsbA and DsbC). Interaction plot and ANOVA analysis (data not shown) suggest that the influence of a folding modulator on output parameters (total soluble, intracellular and extracellular soluble FabZ, OD₅₅₀) depends on the choice of the promoter (T7xx or T7xxiii) and *vice versa*. This supports the widely spread assumption, that suitability of genetic strategies is barely predictable from rational considerations (Overton, 2014). A significant positive correlation ($R^2 = 0.83$, $p = 5.67 \times 10^{-10}$) was detected between intracellular and total soluble FabZ concentrations. Linear regression analysis did not reveal any other significant correlations between the parameters folding modulator gene length, OD₅₅₀, total, intra- and extracellular FabZ concentration ($R^2 < 0.40$, $p > 0.001$). It appeared that high-level folding modulator production prevented increased titers of both total and soluble FabZ. Co-expression of selected folding modulators using the weak T7xxiii promoter, on the other hand, appeared to promote production and folding of FabZ and, thus, enable higher total (intracellular + supernatant) soluble product titers. However, it remained elusive if the influence really derived from the folding modulator or from the mere gene co-expression interacting with FabZ production.

Influence of T7 RNA polymerase-based folding modulator co-synthesis on FabZ production in fermentations

Selected strains were employed in fermentation experiments. These were performed as described in Section 2.4.2. A production phase temperature of 37°C was applied (no temperature ramp, feeding strategy 3, induction strategy 1). Similar to shake flask experiments, ELISA and SDS PAGE analyses were performed to examine product concentrations. Growth behavior in fermentation experiments was examined by assessment of dry cell weight (DCW) and OD₅₅₀ at the end of the fermentation. Values for co-expression strains (“n”) were compared with values from fermentations employing plasmid-free reference strain B<FabZ-kanR>(02) (“ref”) according to Equation 3.1. Results are presented in Figure 3.29 (p. 170).

$$Cell\ density_{EoF, n} [\%] = \frac{\frac{DCW_{EoF, n}}{DCW_{EoF, ref}} + \frac{OD_{550, EoF, n}}{OD_{550, EoF, ref}}}{2} \times 100 \quad (3.1)$$

Compared to the plasmid-free reference strain (385 mg l⁻¹, SD = 33 mg l⁻¹, n = 3), strains with empty basic plasmids bearing the T7xxiii and T7xx promoter had a moderate positive (470 mg l⁻¹) and negative (318 mg l⁻¹) influence on total soluble FabZ concentrations at EoF, respectively (see Figure 3.29 (A)). SDS PAGE analysis of samples associated with strains bearing the empty basic plasmid with the weak T7xxiii promoter revealed FabZ band intensities comparable to the plasmid-free strain (see Figure 3.29 (C), data for plasmid-free strain not shown). In case of the stronger T7xx co-expression promoter in the empty basic plasmid, product-related protein bands were less intense but unambiguously detectable. For both promoters, co-synthesis of folding modulators resulted in decreased total soluble FabZ titers. These ranged considerably from 4 (T7xxiii-*dsbC*) to 345 (T7xxiii-*skp*) mg l⁻¹. Most co-expression constructs did not severely affect bacterial growth and cell densities of ≥ 90 % compared to the cell density at EoF in fermentations of the plasmid-free strain were determined (see Figure 3.29 (B)). However, co-expression combinations T7xx-*dsbA* (75 %), T7xx-*fkpA* (54 %) and T7xxiii-*ppiD* (80 %) had a moderate to strong negative influence on growth. When the weak T7xxiii promoter was employed for co-expression of folding modulator genes, FabZ bands were detectable on SDS PAGE gels in all cases. In case of the T7xx promoter (moderate activity), no FabZ product-related bands were detected on SDS PAGE gels. For all combinations of folding modulator gene and co-expression promoter except T7xxiii-*skp*, protein bands of folding modulators were detectable on SDS PAGE gels. Band intensity from samples of experiments based on the weak T7xxiii promoter varied between different folding modula-

tors. In case of the stronger T7xx promoter (*dsbA* and *fkpA* co-expression), band intensity appeared to be (slightly) increased. In summary, an, at least, moderate improvement by folding modulator co-synthesis was expected based on results from shake flask experiments. However, in fed-batch fermentations, co-synthesis of folding modulators resulted in mostly decreased titers for all tested combinations. Strong folding modulator bands on SDS gels suggest that high-level co-expression might constitute an issue during (soluble) FabZ formation.

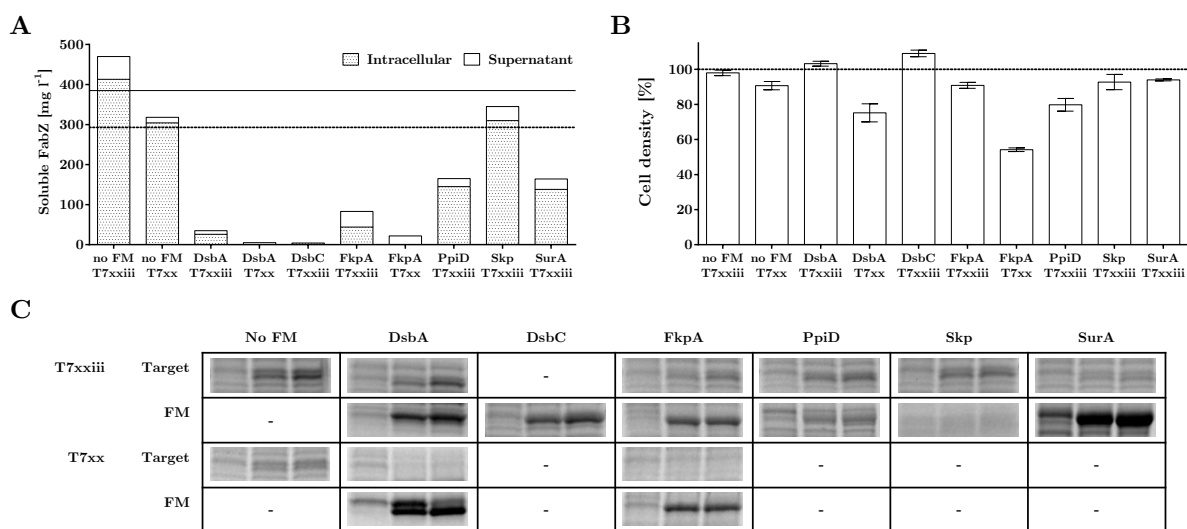


Figure 3.29: Production- and growth related results of T7 RNA polymerase-based folding modulator co-synthesis in FabZ production experiments in 5 l fed-batch fermentations. Genes of the folding modulators DsbA, DsbC, FkpA, PpiD, Skp and SurA were cloned into basic plasmids bearing one of two T7 promoters (T7xxiii with low and T7xx with moderate activity). These plasmids were used to transform the FabZ-producing strain B<FabZ-kanR>(02) with genome-integrated target gene cassette. Strains were employed in single fermentation experiments as described in Section 2.4.2 with a production phase temperature of 37°C. Names on the category axes of **A** and **B** indicate the co-produced folding modulator and T7 promoter used for control of helper gene co-expression. **A**: Samples from the end of fermentation (EoF) time point were analyzed concerning intracellular (dotted bar sections) and supernatant-borne (blank bar sections) soluble and correctly folded FabZ. Dashed and solid lines indicate intracellular and total soluble FabZ titers, which were determined in triplicate fermentation experiment employing the plasmid-free reference strain. **B**: Relative cell density values calculated with Equation 3.1 based on DCW and OD₅₅₀ values measured at EoF. The dashed line in B represents the reference cell density obtained in fermentations with the plasmid-free FabZ-producing strain B<FabZ-kanR>(02) (DCW = 74 ± 2 g l⁻¹, OD₅₅₀ = 268 ± 16; mean of triplicate fermentations ± standard deviations). Error bars represent SD values of mean calculations of the two relative cell densities (DCW and OD₅₅₀). **C**: Cell samples from three different time points (prior to and 9 and 13 h (EoF) after IPTG inducer addition) were subjected to SDS PAGE analysis under reducing and denaturing conditions. Gel sections associated with the molecular weight of the FabZ product or the folding modulator are presented. Note that DsbC and FabZ monomers have comparable molecular weight. Thus, only one cutout is presented.

Conclusion: T7 promoters for folding modulator co-synthesis

Of the systems tested so far, the T7 system with the weakest T7xxiii promoter and without regulatory elements (*lacI* and *lacO* sites) constitutes the only potentially suitable system for co-expression of folding modulator genes. The associated empty basic plasmid exerted no significant influence on growth and production behavior in neither shake flask nor fermentation experiments. The empty basic plasmid bearing the T7xx promoter in shake flask experiments resulted in higher soluble product titers in comparison to the plasmid-free production strain. On the other hand, presence of the empty basic plasmid bearing the T7xxiii promoter in cells increased soluble FabZ titers in fermentations. Ranking of achieved total soluble FabZ titers by co-expression combination revealed that results from T7xx-based co-expression in shake flasks were similar to results of T7xxiii-based co-expression in fermentations. Regression analysis supported this assumption ($R^2 = 0.83$, $p = 0.004$). On the other hand, obtained total soluble FabZ titers using the T7xxiii promoter for folding modulator co-synthesis in shake flask and fermentation approaches did not correlate ($R^2 = 0.06$, $p = 0.592$). In conclusion, it seemed that the level of folding modulator co-synthesis, which clearly influenced the FabZ productivity, differed when the same strain was employed in another cultivation set up (shake flasks vs. fermentations). It is known that significant differences between these two processes exist. For instance, the production phase temperature was increased in fermentation experiments (37 vs. 25 °C). Furthermore, in the fermentation process, recombinant protein production was induced later in the process and at higher cell densities. Apart from time- and cellular “age”-related effects, this results in different IPTG concentrations per cell. Identification of influential differences would require extensive experiments on cellular level. As establishing a practicable toolbox for efficient antibody fragment production was the main aim of this study, reasons for difference in strain performance were not further elucidated. It appears that, under the applied fermentation conditions, lower folding modulator amounts would be required in order to prevent interference with target protein production. However, no such system was available. Under different fermentation conditions and/or for plasmid-based target gene expression, the developed folding modulator co-synthesis system with the weak T7xxiii promoter might be better suitable.

3.4.3 Constitutive Promoter Systems for Co-Synthesis of Folding Modulators

Rationale: constitutive promoters

Using T7 promoter systems for both a genome-integrated target gene and the plasmid-borne folding modulator expression cassettes was found to be associated with several drawbacks (see Section 3.4.2). After system optimization, a positive influence of folding modulators was identified in shake flask but not in fermentation experiments. Interference of target and folding modulator gene expression was already observed in a master thesis prior to this study (Heistingering, 2013). Specifically, expression cassettes of target and folding modulator genes were placed on pBAD (pBR322 ori, araBAD promoter, kanR) and pACYCDuet-1 (p15A, native T7 promoter, clmR), respectively. In this set up, it was proposed that the strength of the native T7 promoter prevented a pronounced positive influence of folding modulators. Constitutive transcription steering systems represent a viable alternative to inducible systems. These systems have previously been successfully applied to improve production of soluble target proteins by co-expression of folding modulator genes (mainly by employing the pTUM4 co-expression plasmid) (Friedrich et al., 2010; O'Reilly et al., 2014; Outchkourov et al., 2008; Schlapschy et al., 2006; Schlapschy and Skerra, 2011). The pTUM4 approach relies on two constitutive *E. coli* promoters (P_{dsbA} , P_{fkpA}). Note that the activity of constitutive promoters might still be dependent on environmental conditions. Thus, strictly speaking, most constitutive promoters are at least somewhat condition-inducible. For instance, the “constitutive” P_{fkpA} promoter employed in the pTUM4 co-expression plasmid (Schlapschy et al., 2006) is σ^E (RpoE)-dependent (Dartigalongue et al., 2001). RpoE, in turn, acts as an effector molecule upon protein misfolding in the cell envelope (Missiakas and Raina, 1998), which might vary depending on many factors. However, for the sake of discussion of selected promoters, transcription is assumed to be sufficiently “constitutive”. If these promoters are employed for folding modulator co-synthesis, the modulator levels were assumed to be permanently elevated. Like increased folding modulator production using inducible systems, extensive co-synthesis under control of constitutive promoters might drain cellular resources required for target protein synthesis in production experiments. However, an interference with the regulatory elements of the T7-based target gene expression system was not expected. In conclusion, constitutive co-expression appeared to be a reasonable extension to the current soluble production toolbox, promoters of different strength were developed and tested in co-expression experiments in the final part of the study.

Establishing a constitutive promoter system

Identification of a suitable constitutive promoter system was performed in the course of a master thesis under supervision of the author of this thesis (Huber, 2015). Constitutive promoters sequences were chosen from the publically available “Anderson” (Anderson, 2006) and “Davis” (Davis et al., 2011) promoter collections. Activities of these promoters depend on the Sigma70 sigma factor, which enables transcription in a fashion very close to constitutive (Woesten, 1998). Sequences are presented in the Confidential Supplement (see Tables 5.6 and 5.7). To mitigate effects of read-through events under strong over-expression of folding modulator genes, the leaky T7 terminator (Giacomelli and Depetris, 2012; Macdonald et al., 1994; Mairhofer et al., 2014; Sousa et al., 1992; Telesnitsky and Chamberlin, 1989a,b) was replaced by the more efficient tZenit terminator (Mairhofer et al., 2014; Wittwer, 2010). Furthermore, the *cer* element, which was located directly downstream of the terminator, was removed from the plasmid backbone, resulting in basic plasmid pBI4iSC(n).3 (see Table 2.3). Cloning and promoter test results obtained with modified plasmids suggested inappropriately high folding modulator synthesis rates. The first set of promoters bore insulation sequences for independence of genetic context (Davis et al., 2011) and a strong RBS. By removal of insulation sequences and utilization of a weaker RBS, the expression rate was successfully reduced. In conclusion, three promoters with weak to moderate constitutive expression were identified as suitable for steering co-expression. Promoter strengths of 32 % (promoter module Ci, with insulation, strong RBS), 78 % (C8, without insulation, weak RBS), 100 % (C2, without insulation, weak RBS) were calculated with methods slightly different from the master thesis (see Figure 3.30).

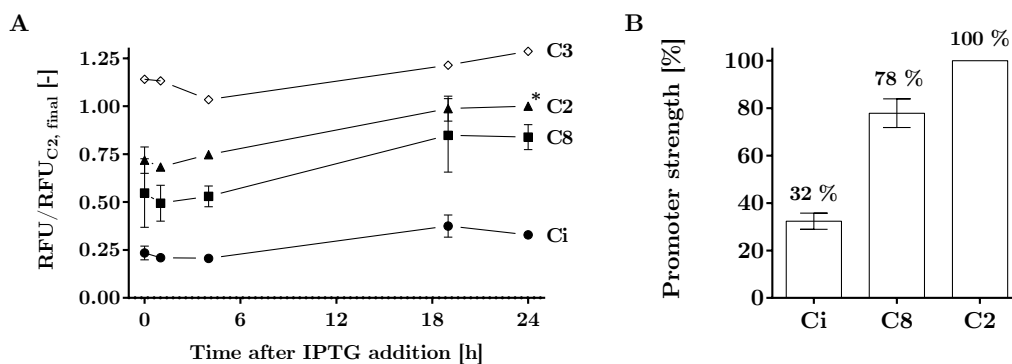


Figure 3.30: Strength of transcription steering modules bearing different constitutive promoters. pBI4iSC(n).3 plasmids bearing the *gfp.1* gene and promoter modules Ci, C8, C2 and C3 were used to transform BL21(DE3) cells. Cells transformed with basic plasmid pBI4iSC8.3 were used as a reference for non-GFP-producing strains. Resulting strains were subjected to promoter tests as described in Section 2.3. Except for promoter module C3, duplicate shake flask experiments were performed in two blocks. Samples were prepared prior to and 1, 2, 4, 19 and 24 h after addition of IPTG. Fluorescence of cells was determined for all time points. Values for the non-GFP-producing strain were subtracted from all values determined at the respective time point. For both experimental blocks, resulting RFU values of each time point were divided by the RFU value obtained for the C2 promoter at the end of the experiment ($\text{RFU}/\text{RFU}_{\text{C2, final}} = 1.00$). The asterisk indicates the reference value ($\text{RFU}/\text{RFU}_{\text{C2, final}} = 1.00$). Means and SD were calculated and are presented as points and error bars in **A**. For promoter modules Ci and C8, mean values of each time point of duplicate measurements were divided by respective mean values determined for the reference module C2. The mean and SD of the five resulting ratios (prior to and 1, 2, 4, 19 and 24 h after inducer addition) were calculated. Values are represented as bar height and error bar length, respectively, in **B**.

Effects of folding modulator co-synthesis on FabZ production in shake flasks

In a previous part of this study, plasmid-free expression strains for production of three model antibody fragments yielded periplasmic aggregates in the g l^{-1} range. To overcome the apparent folding machinery overload, a system employing three different constitutive promoter modules of appropriate strength was designed. The folding modulator genes *dsbA*, *dsbC*, *fkpA*, *ppiD*, *skp* and *surA* were cloned into each of the basid plasmids with one of three constitutive promoter modules individually.

These 18 “co-expression plasmids” were used to transform a FabZ-producing strain with genome-integrated target gene expression cassette (B<FabZ-kanR>(02), see Table 2.6). For this strain, high amounts of total product were obtained (see Sections 3.1.2 and 3.1.2), which were shown to aggregate in the periplasm (see Section 3.2.4). Thus, it appeared to be a suitable model to test co-synthesis systems with elevated periplasmic folding modulator levels. Reference strains were created by transformation of cells with the corresponding empty plasmids pBI4iSCi/8/2.3. Cryo cultures of all 21 strains were prepared as described in Section 2.2.4. Strains were employed in shake flask experiments performed as described in Section 2.4.1. Briefly, cells were incubated at 37°C. For the main cultures, 1 l un baffled Erlenmeyer shake flasks and Shake Flask Medium Version 3 with 5 g l^{-1} glycerol (see Table 5.8) were employed. Prior to addition of 1 mM IPTG at $\text{OD}_{550} = 1.0 \pm 0.3$, main culture shake flasks were transferred from 37 to 25°C prior to the production phase. OD_{550} was determined at the end of the recombinant protein production phase. Cell and supernatant samples for ELISA and SDS PAGE were prepared prior to and 12 h after IPTG addition. Soluble FabZ concentration inside of the cells and in the supernatant was determined by ELISA. Total intracellular FabZ and folding modulator concentration were estimated by SDS PAGE analysis. Selected results are presented in Figure 3.31 (p. 178).

In case of the plasmid-free reference strain a mean total soluble FabZ titer of 4.0 mg l^{-1} (0.7 (intracellular, $n = 6$) + 3.4 (supernatant, $n = 4$)) was determined. When different empty plasmids with the promoters Ci, C8 and C2 were present in the FabZ production strains, total soluble FabZ titers of 3.8 (0.5 + 3.3), 3.9 (0.7 + 3.2) and 4.0 (0.4 + 3.6) mg l^{-1} were obtained, respectively. The plasmid-free strain resulted in a mean OD_{550} of 2.0 at end of the shake flask experiment. For empty plasmid-bearing strains, an OD_{550} value of 2.0 was also determined at the end of the experiment, independent of the promoter present on the plasmid. In conclusion, none of the three empty basic plasmids had a noticeable influence on the production or growth performance of the FabZ-producing strain. Compared to previous co-expression systems, this constitutes a major improvement. As a reference for

improvement calculations, the mean ($\approx 4.0 \text{ mg l}^{-1}$) of the titers of the three basic plasmid-bearing strains and the titers obtained from shake flask experiments with plasmid-free strains was used.

FkpA co-synthesis resulted in an approximately 3-fold improved soluble FabZ titer ($4.0 \rightarrow 11.8 \text{ mg l}^{-1}$, mean value for all *fkpA* expression-controlling promoters). Note that during *fkpA* co-expression experiments performed with promoters C8 and C2, growth impairment was observed during preculture growth at 37°C . Thus, these strains were grown at reduced preculture temperature (33.5°C) in repeated experiments. As all strains were grown for at least 2.5 h at 37°C prior to induction of recombinant protein production, no influence on experimental outcome was expected. In case of Skp co-synthesis, the strongest promoter modules (C8, C2) resulted in 4.4-fold improved titers ($4.0 \rightarrow 17.5 \text{ mg l}^{-1}$, mean value), which constituted the most pronounced improvement of this experimental block. When the weak Ci promoter was used for co-expression of the *skp* gene, a moderate 1.8-fold improvement was observed ($4.0 \rightarrow 7.3 \text{ mg l}^{-1}$). Likewise, co-synthesis of SurA resulted in a moderate 2.1-fold improvement ($4.0 \rightarrow 8.5 \text{ mg l}^{-1}$, mean for all three promoters). Considering total soluble FabZ concentration, DsbA appeared to have no significant influence on FabZ productivity (Ci: 4.3, C8: 5.7, C2: 5.2 mg l^{-1}). A slight negative influence by PpiD co-synthesis was detectable (Ci: 3.3, C8: 2.8, C2: 2.8 mg l^{-1}). In case of DsbC co-synthesis, higher folding modulator levels had an increasingly negative effect on total soluble FabZ titers (Ci: 4.0, C8: 1.5, C2: 1.0 mg l^{-1}). Note that the production phase of the DsbC co-synthesis experiments using the C2 promoter was shortened from 12 to 11 h to avoid night shifts. As the production phase reduced by merely 10 %, no substantial influence on product- or growth-related results was assumed. Note that total soluble FabZ titer correlated well with intracellular soluble FabZ titers ($R^2 = 0.92$, $p = 1.81 \times 10^{-11}$) (compare ratio of lower bar segments to total bar heights in Figure 3.31 (A)).

As expected, the ratio of folding modulator to target protein determined by SDS PAGE analysis shifted in the favor of the folding modulator with increasing strength of the promoter controlling co-expression (Ci \rightarrow C8 \rightarrow C2). This suggests that stronger co-synthesis of folding modulators might have negatively influenced the total FabZ titers. However, as cell samples (one folding modulator, different promoters) were analyzed on different gels, assumptions remain tentative. Note that no ratio was examined for folding modulator DsbC, as its precursor and mature protein are of approximately the same molecular weight as the FabZ monomers.

Co-synthesis of DsbA, DsbC, FkpA, Skp and SurA resulted in 4.7-fold ($2.0 \rightarrow 9.3$, mean value) increased OD_{550} values, when promoters C8 or C2 were applied for steering co-

expression (see Figure 3.31 (C)). When promoter module Ci was used for co-expression, OD₅₅₀ values were increased maximally 2.7-fold (2.0 → 5.3 in case of Ci-*fkpA*). Values determined in case of PpiD co-synthesis remained at or below the level measured for the reference strains. To further elucidate OD₅₅₀ differences, growth experiments with the C8-controlled *skp* over-expressing strain and its empty plasmid counterpart were conducted (see Figure 3.31 (D)). OD₅₅₀ values increased similarly up to OD₅₅₀ ≈ 5 for the strain with and without *skp* co-expression at approximately 6 h after inducer addition. Afterwards, values decreased towards 2.0 at end of the experiments when Skp was not co-synthesized. Values increased towards a plateau of 8.3 (10 h of production time) when intracellular Skp levels were elevated. This illustrated that one point measurements enable only a very limited view on the growth behavior. However, in this case, Skp co-synthesis appeared to enable prolonged growth under FabZ production conditions, while “regular” cells (*i.e.* without elevated folding modulator level) stopped growing in the middle of the production phase.

In summary, data suggested that increasing the intracellular capacity of soluble product formation is key for improving overall soluble FabZ titer. Selected co-synthesis plasmids enabled significant improvements. However, if the presence of a certain folding modulator or an overall decreased target product formation and, thus, reduced secretion and folding caused by resource competition improved intracellular titers, remains elusive.

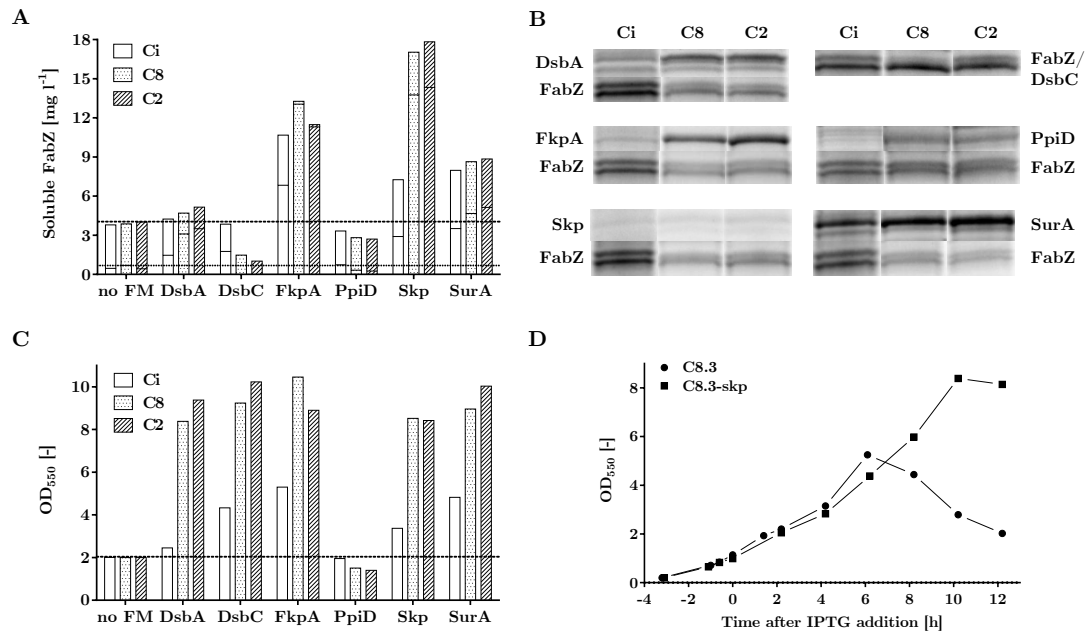


Figure 3.31: Impact of folding modulator co-synthesis using constitutive promoters in FabZ production experiments performed in shake flasks. 18 plasmids bearing one of three constitutive promoters of different strength (Ci < C8 < C2) in combination with one of six genes encoding folding modulators DsbA, DsbC, FkpA, PpiD, Skp and SurA were created. A FabZ-producing strain created by genome integration in a previous part of this study was transformed with these 18 plasmids and the three corresponding empty plasmids individually. Shake flask experiments with these 21 strains were performed as described in Section 2.4.1. **A:** intracellular (lower bar section) and supernatant (upper bar section) soluble FabZ concentrations determined by ELISA at the end of the shake flask experiment. Blank, dotted and cross-hatched bars represent the values associated with co-synthesis controlled by the promoters Ci, C8 and C2, respectively. The dotted and dashed horizontal lines represent intracellular ($n = 6$, $SD = 0.3 \text{ mg l}^{-1}$) and total ($n = 4$, $SD = 0.7 \text{ mg l}^{-1}$) soluble FabZ concentrations determined for the plasmid-free reference strain, respectively. **B:** excerpts from SDS PAGE gels of cell samples from the end of cultivation. Above the lanes, the promoter used for co-synthesis is indicated. Excerpts were chosen based on the expected molecular weight of the folding modulators and the target protein FabZ. Note that lanes associated with single folding modulators derive from different SDS gels. **C:** OD₅₅₀ was determined at the end of the recombinant protein production phase. The dashed line represents the mean value obtained in six shake flask experiments using the plasmid-free reference strain ($SD = 0.2$). **D:** progress of OD₅₅₀ values throughout the shake flask experiments. Two strains were employed in this experiment. Squares and circles represent OD₅₅₀ values obtained with strains carrying the empty and the *skp* gene-bearing pBI4iSC8.3 plasmid, respectively. Figures were adapted from a master thesis, which was supervised by the author of this work (Huber, 2015).

Effects of folding modulator co-synthesis on FabZ production in fermentations

To examine the influence of folding modulator co-synthesis on FabZ production in a different scale, selected strains were also employed in 5 l fed-batch fermentation experiments. These were performed as described in Section 2.4.2 with a production phase temperature of 37°C (no temperature ramp, feeding strategy 3, induction strategy 1). Based on results from shake flask experiments, the most promising folding modulators (FkpA, Skp and SurA) were applied and co-synthesis was evaluated with all three promoters (Ci, C8 and C2). Strains bearing empty plasmids with the two weakest promoters (Ci, C8) were employed as reference strains. Similar to shake flask experiments, ELISA and SDS PAGE analyses were performed to examine FabZ concentrations. Dry cell weight (DCW) and OD₅₅₀ at the end of fermentations (EoF) was compared with values from fermentations employing plasmid-free reference strain B<FabZ-kanR>(02) according to Equation 3.1 (p. 169). Results are presented in Figure 3.32 (p. 181). Note that SDS PAGE band intensities are one way of representing $Y_{P/X}$ values, as the same amount of cells, calculated by OD₅₅₀ values, was applied for SDS PAGE analysis. Visual evaluation of band intensities correlated well with total product per DCW ($Y_{P/X}$) values calculated with Equation 2.3 (p. 64).

When empty basic plasmids with promoters Ci (308 mg l⁻¹) and C8 (296 mg l⁻¹) were present in the cell, total soluble FabZ titers were decreased in comparison to fermentations with the plasmid-free strains (385 mg l⁻¹, n = 3, SD = 24 mg l⁻¹) (see Figure 3.32 (A)). Relative cell densities at EoF were at or above 90 % in fermentations of strains with empty basic plasmids (see Figure 3.32 (B)). FabZ-related SDS PAGE gel bands from fermentations of the empty basic plasmid-bearing strains showed comparable intensities compared to bands related to the plasmid-free strain (see Figure 3.32 (C), data for plasmid-free not shown). Empty plasmid presence resulted in slightly prolonged but comparable batch durations in comparison to the plasmid-free strain (8.5 h → 9.1 (Ci)/8.6 h (C8)). In conclusion, empty basic plasmids had a slight influence on strain performance.

FkpA co-synthesis under control of the Ci promoter resulted in a 1.5-fold increase of total soluble FabZ compared to the plasmid-free reference strain (385 → 576 mg l⁻¹). Co-synthesis of Skp and SurA under utilization of the Ci promoter resulted in slightly increased or unchanged total soluble FabZ titers in comparison to the plasmid-free reference strain (428 and 376 mg l⁻¹, respectively). Strains with Ci-controlled co-synthesis of FkpA, Skp and SurA, yielded product-related bands of similar intensity like those of the plasmid-free and basic plasmid-bearing reference strains. Folding modulator synthesis under control of Ci promoters resulted in similar batch phase durations compared to the

plasmid-free reference strain (approximately 9 h).

Co-synthesis of folding modulators FkpA, Skp and SurA using the stronger promoters C8 and C2 resulted in soluble FabZ titers of below 100 mg l^{-1} and reduced EoF cell densities compared to the plasmid-free reference strain, albeit to varying extents. Folding modulator band intensities on SDS PAGE gels varied depending on folding modulator and promoter used (see Figure 3.32 (C)). For all three promoters, SurA levels in the cells were higher than FkpA levels. Furthermore, independent of the promoter used, cellular FkpA content was always higher than Skp content. In Skp co-synthesis experiments, target product band intensity decreased ($C_i \rightarrow C8$) but then increased ($C8 \rightarrow C2$) with increasing promoter strength. Note that no Skp-related protein bands could be detected on SDS gels. In case of SurA co-synthesis, utilization of stronger co-expression promoters resulted in a decrease of overall target product. Significant glucose accumulation occurred prior to inducer addition in case of combinations *C2-fkpA* (27 g l^{-1}) and *C2-surA* (13 g l^{-1}). In these cases, the glucose feeding rate was adjusted according to the strategy described in Section 2.4.2. Glucose concentration values were slightly increased (0.55 g l^{-1}) for the strain expressing *skp* under C2 control. However, as glucose concentration decreased for the subsequent measurement time point, the feed rate was not adjusted. Batch phase durations increased uniformly with increasing co-expression promoters strength. In case of Skp co-synthesis, durations remained below 10 h. Maximum values of 13 and 12 h were observed in case of *C2-fkpA* and *C2-surA*, respectively.

In conclusion, soluble FabZ yield was improved significantly by one element of the co-expression toolbox (*Ci-fkpA*), which classifies the approach as successful. Other elements had no or a pronounced negative impact on production and growth behavior. Whether increased folding modulator co-synthesis drained cellular resources necessary for FabZ formation or impaired growth had a negative influence, remains elusive.

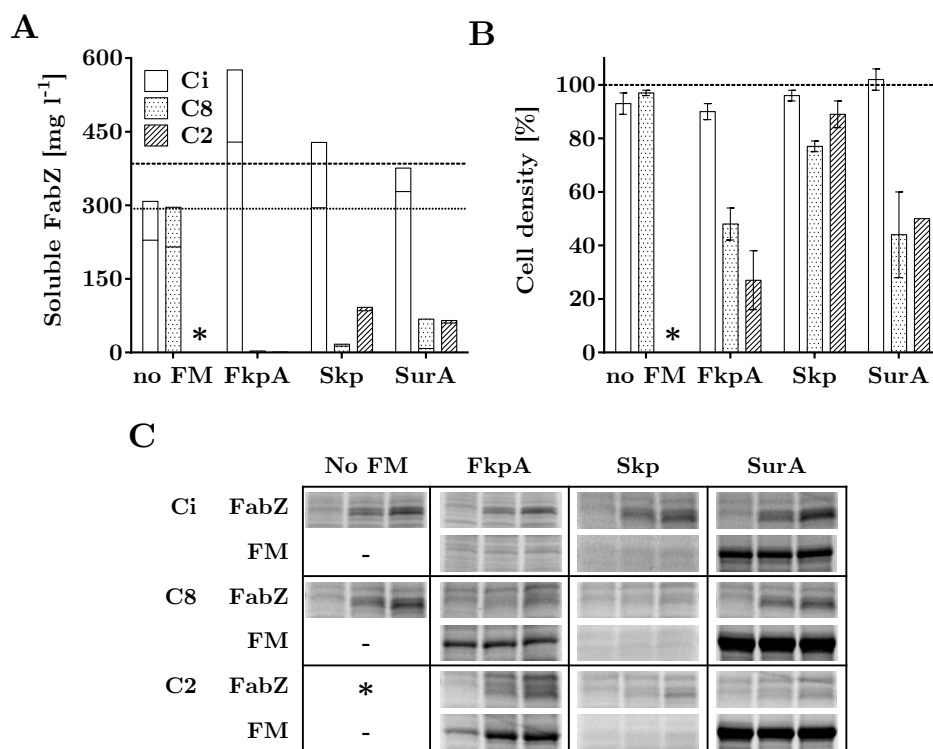


Figure 3.32: Impact of folding modulator co-synthesis controlled by constitutive promoters in FabZ production experiments performed in fed-batch fermentations. **A** FabZ-producing strain created by genome integration ($B<FabZ\text{-kanR}>(02)$) was transformed with co-expression and empty reference plasmids. Resulting strains were employed in single fermentation experiments with a production temperature of 37°C as described in Section 2.4.2. **A**: Samples from the end of fermentation (EoF) time point were analyzed concerning intracellular and supernatant-borne soluble FabZ concentration (lower and upper bar section, respectively). Note that no experiment was performed using the basic plasmid bearing the promoter C2 (asterisk). Intracellular ($SD = 9 \text{ mg l}^{-1}$) and total ($SD = 24 \text{ mg l}^{-1}$) soluble FabZ titers were determined in triplicate fermentation experiment employing the plasmid-free reference strain with equal conditions. Dotted and dashed lines indicate the respective values. **B**: Relative cell density values based on DCW and OD_{550} values measured at EoF (see Equation 3.1, p. 169, dashed line = 100 %). Names on the category axis in **A** and **B** indicate the co-produced folding modulator. Bar appearance is associated with the promoter used for co-expression of the folding modulator gene or present on the empty plasmid. **C**: Cell samples from three different time points (prior to and 6.5 and 13 h (EoF) after inducer addition; left to right) were analyzed by SDS PAGE under reducing and denaturing conditions. Gel sections associated with molecular weight of FabZ or the respective folding modulator are presented. Figures were adapted from a master thesis, which was supervised by the author of this work (Huber, 2015).

Effects of folding modulator co-synthesis on FabX production in shake flasks

As described above, successful application of folding modulator co-synthesis is, amongst other factors, dependent on the recombinant target protein (see Section 1.3.4). Thus, the newly developed co-expression system was also applied in model protein FabX production experiments. Note that the FabX-producing plasmid-free strain, which was employed in previous sections of this study, was genetically modified prior to this part of the study. However, no influence on production and growth behavior by genetic modifications was observed in comparative fermentation experiments (data not shown). Due to confidentiality reasons, further details cannot be provided. The 21 co-expression plasmids already used in FabZ experiments, were used to transform the engineered FabX production strain, which bears a reverse genome-integrated target gene expression cassette. Parent cultures of all 21 strains were prepared as described in Section 2.2.4. Subsequently, shake flask and fermentation experiments were performed. Note that strains were created within a master thesis not supervised by the author of this work (Klösch, 2016). Furthermore, fermentation experiments were performed in a collateral project.

Shake flask experiments were performed as described in Section 2.4.1 and included a shift of cultivation temperature ($37 \rightarrow 25^\circ\text{C}$) prior to induction of recombinant protein production. Note that approaches resembled shake flask experiments performed during co-expression experiments for FabZ production (see above). OD_{550} values were determined at the end of the recombinant protein production phase. Cell and supernatant samples were prepared prior to and 12 h after IPTG addition. Soluble FabX concentrations inside of the cells and in the supernatant were determined by ELISA. Total intracellular FabX and folding modulator concentrations were estimated by SDS PAGE analysis. Selected results are presented in Figure 3.33 (p. 185).

When the empty basic plasmids were present in FabX production strains, obtained total soluble FabX titers (mean: 2.3 mg l^{-1} , $\text{SD} = 0.1 \text{ mg l}^{-1}$) were comparable to the one obtained for the plasmid-free reference strain (2.4 mg l^{-1}). For empty basic plasmid-bearing strains, a mean final OD_{550} value (1.3, $\text{SD} \leq 0.1$) comparable to the one for the plasmid-free strain ($\text{OD}_{550} = 1.2$) was determined. In conclusion, empty basic plasmids exerted no significant influence on the production and growth behavior of the FabX-producing strain. Thus, following conclusions relate to both the plasmid-free and empty basic plasmid-bearing reference strains. Fold-changes were calculated in comparison to the mean values of all four reference strain cultivations (total soluble FabX: 2.4 mg l^{-1} , $\text{OD}_{550} = 1.3$).

As observed in FabZ shake flask experiments, co-expression of folding modulator genes under control of the weakest promoter Ci had no pronounced effects on examined parameters during FabX shake flask experiments. Elevated FkpA, Skp and SurA levels (controlled by promoters C8 and C2) resulted in improved soluble FabX titers. For instance, a 7-fold ($2.4 \rightarrow 16.8 \text{ mg l}^{-1}$) increase in total soluble FabX concentration was measured for co-synthesis of FkpA under control of the C8 promoter compared to the reference value. Note that during *fkpA* co-expression experiments performed with promoters C8 and C2, growth impairment was observed during preculture growth at 37°C. Thus, these strains were grown at reduced preculture temperature (33.5°C) in repeated experiments. As all strains were grown for at least 2.5 h at 37°C prior to induction of recombinant protein production, no influence on experimental outcome was expected. Co-synthesis of Skp and SurA under control of C2 yielded 4.4- and 2.5-fold improved titers, respectively (Skp: $2.4 \rightarrow 10.5 \text{ mg l}^{-1}$, SurA: $2.4 \rightarrow 6.0 \text{ mg l}^{-1}$). On the other hand, elevated DsbA, DsbC and PpiD levels (controlled by C8 or C2) had only a slight positive (DsbA: 2.6 mg l^{-1} (C8), 3.0 mg l^{-1} (C2)) or even a negative influence (DsbC: 1.1 mg l^{-1} (C8), 0.4 mg l^{-1} (C2); PpiD: 2.2 mg l^{-1} (C8), 1.6 mg l^{-1} (C2)) on soluble product titers. Note that FabX-producing cells with increased DsbC levels (C2 promoter) were characterized by slow growth. To prevent the necessity of night shifts, inducer was added without adaptation duration to the 25°C shaker incubator temperature at an OD_{550} value below the target range ($\text{OD}_{550} = 0.6$). Soluble product concentration at the end of cultivation was low and OD_{550} values were high for this genetic combination. As this was in accordance with results obtained for FabZ, the experiment was not repeated despite the deviations described.

Co-synthesis of DsbA, Skp and SurA under control of the stronger promoters (C8 and C2) resulted in moderately increased final OD_{550} values of up to 4.9. When the same promoters were used for co-expression of the *dsbC* and *fkpA* gene, significantly increased OD_{550} values of up to 9.5 were measured. No correlation between final OD_{550} values and total soluble FabX titer was found. To estimate $Y_{P/X}$ values for intracellular soluble FabX, volumetric concentrations were divided by the corresponding OD_{550} values according to Equation 2.3. Cell density-specific titers correlated with the volumetric soluble intracellular concentrations ($R^2 = 0.94$, $p = 8.65 \times 10^{-13}$). Thus, only volumetric titers were used for discussion. Furthermore, this indicated that the soluble product content per cell rather than growth behavior was improved.

SDS PAGE images revealed detectable product-related bands of similar intensity for all tested strains. OD-normalized total (soluble + insoluble) intracellular FabX ranged between 10.6 and 20.2 mg per OD unit. No significant influence of a folding modulator (p

= 0.22) or the associated promoter ($p = 0.08$) could be determined. Folding modulator content of cells was not calculated. However, in case of synthesis of DsbA, FkpA, PpiD and SurA, band intensities increased with promoter strength. Note that excerpts were derived from the same SDS PAGE gel for each folding modulator controlled by different promoters and, thus, should allow comparisons.

Overall (soluble + insoluble) intracellular product content (calculated per OD₅₅₀ unit) suggested that total productivity was not markedly affected by presence of any co-expression construct considering the inherent uncertainty of the method. Intracellular content of soluble FabX, on the other hand, was significantly increased for some co-synthesis set ups. The quotient of soluble (ELISA) and total intracellular (SDS PAGE) FabX revealed a share of 12 % soluble product in case of the best-performing co-expression strain (combination C8-*fkpA*). The four reference strains (plasmid-free and basic plasmid-bearing), on the other hand, resulted in a mean soluble product fraction of under 1 %.

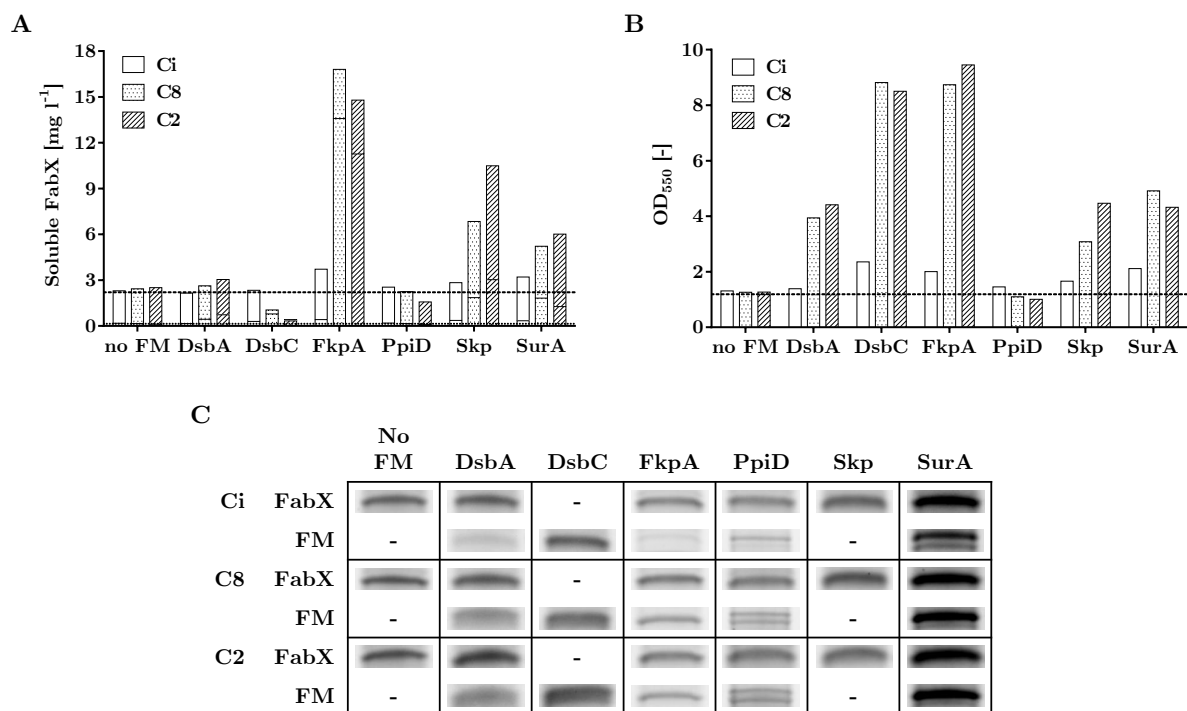


Figure 3.33: Impact of folding modulator co-synthesis using constitutive promoters during FabX production experiments in shake flasks. 18 plasmids bearing one of three constitutive promoters of different activity ($Ci < C8 < C2$) and one of six genes encoding the folding modulators DsbA, DsbC, FkpA, PpiD, Skp and SurA were created. A FabX-producing strain created by reverse genome integration was transformed with the 18 resulting plasmids and the respective empty basic plasmid devoid of folding modulator genes. Shake flask experiments were performed as described in Section 2.4.1. Note that all data originated from single shake flask experiments. **A**: soluble FabX concentrations inside the cell (lower bar section) and in the supernatant (upper bar section) were determined by ELISA at the end of the shake flask experiments. Dotted (just above category axis) and dashed lines represent intracellular and total soluble FabX concentrations determined in samples from the plasmid-free reference strain. **B**: OD_{550} values were determined at the end of the recombinant protein production phase. The dashed line represents the OD_{550} measured when the plasmid-free reference strain was employed. Blank, dotted and cross-hatched bars in **A** and **B** represent values associated with co-synthesis controlled by the promoters Ci, C8 and C2, respectively. **C**: SDS PAGE cut outs of samples from the end of the experiments. Columns and rows correspond to folding modulators and promoters used in co-expression approaches, respectively. Excerpts were chosen based on the expected molecular weight of the co-synthesized folding modulator and target protein FabX. Note also that DsbC and FabX monomers have a similar molecular weight. In this case, the image could not be interpreted. Again, Skp could not be detected on gels.

Effects of folding modulator co-synthesis on FabX production in fermentations

During fed-batch fermentation experiments with the plasmid-free FabX production strain, target product content in the supernatant increased, while cell density values decreased towards the end of fermentation (EoF). A high total intracellular product concentration was detected, which was shown to consist of periplasmic aggregates. The periplasmic folding machinery appeared to be overburdened and not sufficiently capable to properly fold and assemble the FabX product. High amounts of total intracellular target protein were also detected in shake flask FabX production experiments. Similar to fermentation experiments, soluble target product was detected in the supernatant and OD₅₅₀ values at the end of shake flask experiments were low, suggesting cell lysis. After promising results concerning improved soluble product titers were obtained in shake flask experiments, selected FabX-producing strains were also tested in fed-batch fermentation experiments. Since co-synthesis of FkpA, Skp and SurA resulted in highest titer improvements in shake flask experiments, the associated strains were selected for fermentation. In order to also examine the influence of different folding modulator levels in fed-batch processes, respective strains bearing co-expression constructs with promoters Ci, C8 and C2 were employed. Processes were performed according to Section 2.4.3. A production phase temperature of 25°C was applied. Furthermore, a hold step of several hours at elevated pH after EoF was used to promote product leakage to the supernatant. The process format applied for FabX production had been proven to be robust and reproducible in the associated project. The plasmid-free strain was employed as a reference. In fermentation and shake flask approaches, the same production phase temperature was applied. Thus, results from shake flask approaches were expected to be somewhat more transferable compared to FabZ co-expression experiments. Cell and supernatant samples were prepared throughout the process and analyzed by ELISA, 2D-HPLC and SDS PAGE analysis (see Table 2.10). Growth behavior of cells was examined by determination of OD₅₅₀ and dry cell weight (DCW) throughout the process. A relative value for EoF DCW and OD₅₅₀ in relation to the values determined for the plasmid-free strain was calculated for all strains according to Equation 3.1 (p. 169). Glucose concentration was monitored in regular intervals throughout the fermentation. Selected results are presented in Figure 3.34 and described below.

In case of the plasmid-free strain, DCW and OD₅₅₀ were determined to be 76 mg l⁻¹ and 250 at EoF, respectively. The calculated relative cell density values at EoF ranged between 83 and 113 % for the folding modulator gene co-expression strains (data not shown). DCW and OD of all co-expression strains were, at least somewhat, decreased

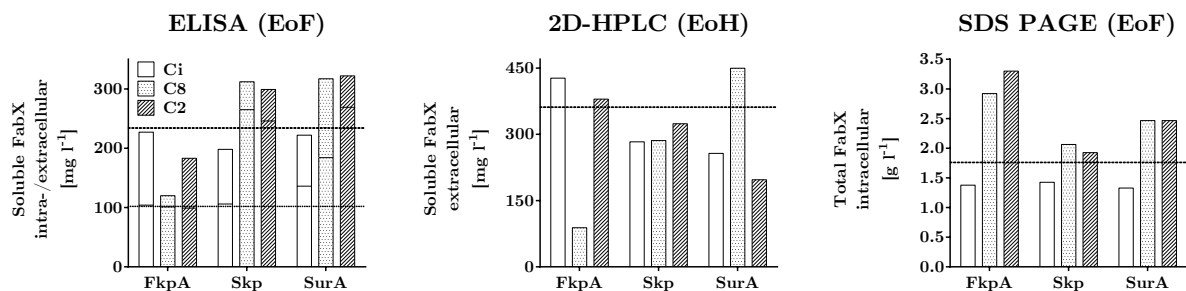


Figure 3.34: Impact of folding modulator co-synthesis using constitutive promoters on product titers in FabX production experiments performed as fed-batch fermentations. FabX-producing strains created by genome integration and transformed with folding modulator co-synthesis plasmids were employed in single fermentation experiments. These were performed as described in Section 2.4.3 with a production phase temperature of 25°C. Samples from different time points throughout the production process were analyzed for product-related values. Names on the category axis indicate the used folding modulators. In all figures, blank, dotted and cross-hatched bars represent promoters Ci, C8 and C2, which were used for co-expression, respectively. **ELISA (EoF)**: Samples from the end of fermentation (EoF) time point were examined concerning intracellular and supernatant-borne soluble FabX concentration (lower and upper bar section, respectively) using ELISA analysis. A reference fermentation employing the plasmid-free FabX-producing strain was performed. Dotted and dashed lines indicate intracellular and total soluble FabX titers determined in that experiment. **2D-HPLC (EoH)**: A hold phase at elevated pH was employed after EoF to promote cell lysis. Thereafter, the fermentation suspension was filtrated and analyzed concerning soluble and correctly folded FabX by 2D-HPLC. **SDS PAGE (EoF)**: Total FabX concentration in EoF cell samples was examined by SDS PAGE under reducing and denaturing conditions. Dashed lines in center and right graphic represent respective titer values obtained when the plasmid-free reference strain was employed.

compared to the plasmid-free reference strain.

Bands associated with folding modulators FkpA and SurA were more pronounced when the stronger promoter modules (C8, C2) were used for steering co-synthesis (no SDS PAGE images shown). A slightly more intense SurA protein band was detectable when comparing the Ci-*surA* strain with the plasmid-free FabX production strain. In all other cases (Ci-*skp/fkpA*, C8-*skp*, C2-*skp*), no bands with increased intensity were detected in the molecular weight ranges of the folding modulators. When co-expression of the three folding modulator genes *fkpA*, *skp* and *surA* was controlled by the Ci promoter, overall (soluble + insoluble) intracellular product yields were slightly decreased compared to the plasmid-free reference strain ($1.8 \rightarrow 1.4 \text{ g l}^{-1}$). Overall product titers generated by all other strains were increased compared to the plasmid-free reference strain and depended on the folding modulator (see Figure 3.34 (SDS PAGE (EoF))). Cell density-normalized yields at EoF were calculated according to Equation 2.3 (p. 64) using intracellular overall product titers and DCW values (data not shown). These correlated well with volumetric titers ($R^2 = 0.91$, $p = 1.69 \times 10^{-5}$). Total (soluble + insoluble) productivity appeared to be unaffected or even improved by elevated folding modulator levels.

FkpA co-synthesis did not affect intracellular soluble FabX concentrations ($\approx 100 \text{ mg l}^{-1}$,

see lower bar sections in Figure 3.34 (ELISA)). Why utilization of different promoters in case of FkpA co-synthesis led to different extracellular product concentrations (Ci: 123, C8: 19, C2: 84 mg l⁻¹) remains elusive. Skp co-synthesis under control of the strong promoter modules C8 and C2 increased the intracellular soluble product titer approximately 2.5-fold (102 → 256 mg l⁻¹, mean). SurA co-synthesis also led to improved intracellular soluble FabX titers (102 → 136 (Ci)/184 (C8)/269 (C2)). DCW-normalized intracellular soluble product titers at EoF were calculated (see Equation 2.3, p. 64) and correlated well with volumetric titers ($R^2 = 0.98$, $p = 7.01 \times 10^{-7}$, details not presented). It appeared that the intracellular capacity for synthesis of soluble FabX was increased by selected co-expression toolbox elements. Furthermore, high overall product concentrations in the cells suggested that competition for cellular resources was not limiting. Thus, the positive effect seemed to derive from an improved folding capacity mediated by increased folding modulator levels.

When FkpA was co-synthesized, total soluble FabX titers determined by ELISA were decreased compared to fermentations with the plasmid-free reference strain (234 → 227 (Ci)/120 (C8)/183 (C2), see entire bar height in Figure 3.34 (ELISA)). Note that FkpA co-synthesis under control of the C8 promoter resulted in accumulation of 6.8 g l⁻¹ glucose prior to induction of recombinant protein production. According to control strategy defined for the fermentation process, the feed rate was reduced to facilitate glucose consumption and a return to glucose-limited conditions. An increased glucose value (1.0 g l⁻¹) was also determined for combination C2-*skp*, albeit only at the end of the fermentation time point. If this phenomenon was an inherent characteristic of this specific combination (C8-*fkpA*) or due to an unnoticed handling mistake, remained elusive. When Skp-over-producing strains were employed, total soluble product titers were decreased (Ci: 198 mg l⁻¹) or increased (C8: 312, C2: 299 mg l⁻¹). Likewise, SurA co-synthesis resulted in decreased (Ci: 223 mg l⁻¹) or increased (C8: 317, C2: 323 mg l⁻¹) titers. In conclusion, four out of nine combinations of promoter and folding modulator gene resulted in moderately improved titers compared to the plasmid-free reference strain (C8-*skp*: + 33 %, C2-*skp*: + 28 %, C8-*surA*: + 35 %, C2-*surA*: + 38 %).

Within the associated project, results from 2D-HPLC analysis after the employed hold step (EoF) at elevated pH were used to assess overall fermentation success (see Figure 3.34 (2D-HPLC (EoH))). Also, downstream processing is performed with the culture supernatant, stressing the importance of this parameter. When FkpA-over-producing strains were employed, extracellular titers were increased (Ci: 427, C2: 380 mg l⁻¹) or decreased (C8: 89 mg l⁻¹) compared to the plasmid-free reference strain (362 mg l⁻¹). Skp co-synthesis resulted in decreased titers (Ci: 283, C8: 286, C2: 324 mg l⁻¹). When SurA

was co-synthesized, increased (C8: 450 mg l⁻¹) or decreased (Ci: 257, C2: 197 mg l⁻¹) titers were determined. Trends in values correlated well with results from ELISA analysis ($R^2 = 0.86$, $p = 1.28 \times 10^{-4}$). In conclusion, three out of nine combinations of promoter and folding modulator gene resulted in slightly to moderately improved titers compared to the plasmid-free reference strain (Ci-*fkpA*: + 18 %, C2-*fkpA*: + 5 %, C8-*surA*: + 24 %). Based on results from shake flask experiments, more pronounced positive effects were expected. However, the co-expression of folding modulator genes under control of constitutive promoters facilitated the generation of improved FabX production strains.

3.4.4 Conclusions

Folding modulator co-synthesis has frequently been presented as an advisable method for improvement of bacterial expression systems (see, *e.g.*, Kolaj et al. (2009), Baneyx and Mujacic (2004), de Marco (2009) and Section 1.3.4). For instance, the pTUM4 plasmid was reported to constitute a ready-to-use solution for improving periplasmic folding capacities (Breustedt et al., 2006; Friedrich et al., 2010; Nasreen et al., 2005; O'Reilly et al., 2014; Schlapschy et al., 2006; Wiebe et al., 2010). Based on SDS PAGE analysis, it showed high level co-synthesis of folding modulators (Breustedt et al., 2006; Schlapschy et al., 2006). In this study (see above), excessive folding modulators over-production was often associated with severe drawbacks. In accordance, pTUM4-carrying bacterial cells showed impaired growth in a previous study (Schlapschy et al., 2006). pTUM4 enables simultaneous co-synthesis of four different folding modulators (DsbA, DsbC, FkpA and SurA). Certain folding modulators might not be necessary or might even cause adverse effects. For instance, co-synthesis of Skp and FkpA was reported to have a positive effect on scFv yield, while additional DsbC co-synthesis exerted a negative influence on productivity (Sonoda et al., 2011). To assess its potential, pTUM4 was ordered and used to transform FabZ-producing strain B<FabZ-kanR>(02). This host was previously generated by genomic integration of the FabZ gene expression cassette. The resulting pTUM4-bearing strain was employed in shake flask and fermentation experiments. In shake flask experiments, total soluble FabZ titers were reduced 2.7-fold ($4.0 \rightarrow 1.5 \text{ mg l}^{-1}$), whereas final OD₅₅₀ values were increased ($2.0 \rightarrow 9.8$) in comparison to experiments with the plasmid-free strain. Folding modulator-related bands on SDS gels were clearly detectable while FabZ-associated bands were less intense compared to the plasmid-free reference strain (data not shown). In fermentation experiments, maximum total soluble FabZ titers were also drastically reduced (4.5-fold, $385 \rightarrow 86 \text{ mg l}^{-1}$). Prominent protein bands of the folding modulators were detected by SDS PAGE analysis (data not shown). In summary, application of pTUM4 did not result in the expected improvement concerning soluble FabZ concentration. On the contrary, it even reduced overall and soluble FabZ yields. It appeared that excessive synthesis of folding modulators and/or the presence of one of the four folding modulators (DsbA, DsbC, FkpA, SurA) negatively impacted FabZ product yields.

Background of this part of the work was the assumption that co-synthesis of folding modulators should be considered with caution prior to application. As described in Section 3.4.1, successful co-synthesis approaches might require different folding modulators and various levels, depending on, for instance, target protein, *E. coli* strain and produc-

tion process (Delic et al., 2014; Martinez-Alonso et al., 2010; Overton, 2014; Schaefer and Plueckthun, 2010). Several systems for adjusting co-synthesis levels of selected folding modulators were tested. Condition-inducible systems prevented a reliable adjustment of folding modulator levels due to changing environmental conditions. By application of a T7 RNA polymerase system bearing additional regulatory elements *lacI* and *lacO* sites, titers in shake flasks were improved by mere presence of the basic plasmid. Folding modulator co-synthesis based on those plasmids resulted in reduced FabZ product titers compared to the empty basic plasmid-bearing cells. However, titers were still increased compared to the plasmid-free strain. Subsequently, a T7-based co-expression system without regulatory elements (*lacI*, *lacO* sites) was designed. This empty basic plasmid had a minimal impact on recombinant target protein production. Co-synthesis of folding modulators using this system improved titers in shake flask experiments. However, product titers in fed-batch fermentations were reduced in almost all cases. To avoid interactions between folding modulator and target gene expression, a co-expression system based on constitutive promoters was designed. In shake flask experiments, empty basic plasmids with constitutive promoters did not affect host cell growth or recombinant target protein productivity. Folding modulator co-synthesis increased soluble FabZ and FabX titers in shake flask experiments. Furthermore, slight to moderate titer improvements were observed in fed-batch fermentations for FabX and FabZ production. A summary of results obtained in co-synthesis approaches is provided in Table 3.10 (p. 196).

In summary, several steps of optimization yielded the folding modulator co-synthesis system with constitutive promoters. Its application improved soluble product titers in shake flask and fed-batch fermentation experiments and for two different products (FabX and FabZ). As the co-expression plasmid pTUM4 resulted in significantly reduced FabZ product titers in both shake flask and fermentation approaches, co-synthesis of single folding modulators rather than simultaneous co-production of several modulators appeared to be the method of choice. When the T7 system was employed in co-expression experiments, the weakest promoter T7xxiii resulted in the highest soluble product titers in all cases. Possibly, this was due to the observed interference phenomenon of regulatory elements (*lacI*, *lacO*), which was least significant in this genetic set up. Results obtained using the final co-expression system (constitutive promoters) verified that genetic background (*e.g.* host strain and product) and environmental conditions (*e.g.* shake flask vs. fermentation) define the success of a co-expression approach. The constitutive promoter system created within this study offers tools to identify beneficial constellations. Selected combinations of folding modulator and promoter appeared to reduce product aggregation in the periplasm and, thus, avoid cytotoxic influences, which were previously reported (de Marco,

2013). Reduction in protein synthesis rates was previously reported to support soluble target product formation (Hoffmann and Rinas, 2004). This reduction in transcription and translation caused by co-synthesis might also be the root cause for improved Fab titers. However, as total productivity was not significantly reduced in case of FabX shake flask experiments, the presence of folding modulators appeared to evoke a positive influence. In conclusion, experiments confirmed that co-synthesis of folding modulators has to be considered thoroughly. Possibly, singly-copy genomic integration of a target gene results in different requirements concerning folding modulator co-synthesis. To date, most co-synthesis approaches have been developed to support multi-copy plasmid-based target gene expression. For plasmid-free production of recombinant proteins, increased synthesis of native *E. coli* folding modulators possibly rather disturbed cellular balance of translocation and folding than improved it. Results obtained during this study when using plasmid-free strains indicated that an overload of the periplasmic folding machinery negatively influenced the soluble product yields. However, mere limited success of co-expression experiments in fed-batch fermentations suggests that other host cell resources might also be limiting.

Despite remaining questions, it was shown that catalytic activity of some of the tested folding modulators improved soluble product titers. The effect of folding modulator co-synthesis was probably superimposed by the basic plasmid influence for the T7 co-expression system bearing regulatory elements *lacI* and *lacO*. Thus, these results were excluded from the discussion. The oxidoreductases DsbA and DsbC, the PPIases FkpA, PpiD and SurA and the general chaperone Skp were tested in this study. An overview of previous co-expression studies is presented in the introductory section (see Table 1.1, p. 26).

DsbA has previously been used to improve soluble titers of several products (Jeong and Lee, 2000; Joly et al., 1998; Lee et al., 2004; Mao et al., 2010; Wulfing and Pluckthun, 1994; Wulfing and Rappuoli, 1997; Xu et al., 2008). In this study, rather ambiguous results were obtained. Using the optimized co-expression system (constitutive promoters), DsbA co-synthesis had slight positive or negative effect on soluble FabX and FabZ titers, depending on the format of the production experiments.

Also DsbC has previously been used to improve the soluble titer of disulfide bond-bearing recombinant proteins (Hu et al., 2007; Joly et al., 1998; Kurokawa et al., 2000; Lee et al., 2013; Maskos et al., 2003; Qiu et al., 1998; Sun et al., 2014; Zhuo et al., 2014). However, it has also been reported to be expandable for folding of recombinant proteins with disulfide bonds between consecutive cysteines (Berkmen, 2012; Joly and Swartz, 1997). As FabX

and FabZ belong to this class of proteins, it was not surprising that DsbC co-synthesis did not have positive effects on the folding of these model proteins. However, results of this study suggest that increased co-synthesis exerts detrimental effects on soluble product titer without impairing total productivity. Dimerization of DsbC is hypothesized to avoid interaction with correctly folded proteins. Gratuitous DsbC over-production might overburden efficient DsbC dimerization, yielding monomers that reduce disulfide bonds of correctly folded product. According to one of two models for DsbC-mediated disulfide bond isomerization (Berkmen, 2012), DsbA is necessary to facilitate re-oxidation of the target protein previously reduced by DsbC. As DsbC was over-produced individually in this study, a lack of DsbA might have led to increased levels of reduced target protein, which, thereby, became susceptible to aggregation or proteolytic degradation. As described in Section 1.3.2, catalytic function of oxidoreductases relies on a sophisticated system. As increased levels of single components might yield an unbalanced system, special care should be taken in co-synthesis approaches.

FkpA was previously reported to not be essential within the *E. coli* cell (Horne and Young, 1995; Kleerebezem et al., 1995). However, it was successfully applied in previous co-expression experiments (Arie et al., 2001; Bothmann and Plueckthun, 2000; Gunnarsen et al., 2010; Ow et al., 2010; Padiolleau-Lefevre et al., 2006; Sonoda et al., 2011; Wu et al., 2007; Zhang et al., 2003). Positive effects on soluble product titers were observed when the co-expression system using constitutive promoters was used in FabX and FabZ production experiments (note: only slight improvement for FabX-related fermentations). Apparently and in accordance with previous reports, one of FkpA's functions (PPIase or general chaperone) supports formation of soluble Fab fragments.

As described above, PpiD was reported to not improve soluble product titers (Kolaj et al., 2009). In accordance, no positive influence of the folding modulator as such was observed after the basic plasmid influence was successfully abolished. Despite its central function amongst periplasmic folding modulators, PpiD appeared not to be a relevant candidate to support soluble antibody fragment production in *E. coli*.

SurA, which conveys PPIase and general chaperone activity, was shown to mitigate protein misfolding (Missiakas et al., 1996). It is also encoded on the pTUM4 plasmid (Schlapschy et al., 2006), even though it was previously reported to have failed to improve production of a scFv fragment (Bothmann and Plueckthun, 2000). In this study, SurA co-synthesis unambiguously increased Fab titers in shake flask experiments, independent of the target product. In fed-batch fermentations, soluble FabX titers were moderately improved, while FabZ titers remained unchanged. SurA was reported to have an Ar-X-Ar substrate selec-

tivity (“X” represents any and “Ar” any aromatic amino acid) (Bitto and McKay, 2003). FabX LC and HC chain contain two and five Ar-X-Ar motifs, respectively. LC and HC monomers of FabZ bear three and two motifs, respectively. Differences in these motif counts could explain product-dependence of SurA effects. However, differences in the fermentation process and scale (*e.g.* temperature during the production phase, induction and feeding strategy) are also likely to influence the outcome.

Skp has successfully been used in previous co-synthesis experiments (Bothmann and Pluckthun, 1998; Hayhurst et al., 2003; Hayhurst and Harris, 1999; Lin et al., 2008; Mavrangelos et al., 2001; Maynard et al., 2005; Ow et al., 2010; Sonoda et al., 2011; Wang et al., 2013). In shake flask experiments of this study, it unambiguously improved soluble FabX and FabZ titers. However, it showed no significant effects on the FabZ product concentrations in fermentation processes. During FabX production processes in fermentation scale, the observed effect depended on the time point of sampling and the method of analysis (see Figure 3.34). In summary, Skp appears to bear the potential to improve soluble antibody fragment production in *E. coli*, which would be in accordance with previous reports. Skp availability to prevent target protein aggregation might be the basis for improvements in Fab titers. Different scFv products were shown to exhibit varying Skp-sensitivity (Entzminger et al., 2012). While a successful application of Skp in co-synthesis approaches apparently depended on the target product, it appears to be a valuable toolbox element for improvement of further recombinant protein’s folding.

The initially used co-expression system (T7 promoter-based) was significantly improved in this study. The final system (constitutive promoters) enabled studying the effects of folding modulators on the target products FabX and FabZ in different process formats. In conclusion, folding modulators DsbA, DsbC and PpiD did not exert relevant effects, whereas co-synthesis of FkpA, SurA and Skp partially led to significantly increased titers. As expected, the optimal folding modulator co-synthesis level had to be identified in trial-and-error approaches. To further evaluate the usefulness of the folding modulator co-synthesis system, it should be applied for production of further recombinant products (*e.g.* scFv-Z-1), in other genetic backgrounds (*e.g.* GoI encoded on multi-copy plasmids) or in other production processes (*e.g.* during fermentation process optimization). To decrease the number of experiments, using only one of the two stronger promoters (C8 or C2) appears reasonable. To further develop the folding modulator co-synthesis system, a plethora of possibilities is available. For instance, new (*e.g.* the periplasmic chaperone Spy (Lennon et al., 2015)) or engineered (Arredondo et al., 2008; Bessette et al., 2000) folding modulators or combinations of folding modulators could be evaluated. Especially for Dsb proteins, combinations of several folding modulators have successfully been employed

(Hoshino et al., 2002; Kohda et al., 2002; Kurokawa et al., 2000, 2001; Sandee et al., 2005; Soanes et al., 2008; Sun et al., 2014). Application of different promoters (*e.g.* the titratable XylS/Pm system (Balzer et al., 2013)) might increase flexibility concerning time of synthesis and level of folding modulators. Single-copy plasmids could further reduce plasmid-related effects on the host cell. Further details on the interactions between recombinant protein product (or parts thereof) and the used folding modulators would generate a more holistic understanding of cellular processes. For instance, examination of protein interaction (*e.g.* by co-immunoprecipitation) could reveal if target product and the overproduced folding modulator really interact. Furthermore, mass spectrometry analysis of the host cell proteome, the target product and the folding modulator might reveal relevant regulatory effects of folding modulator overproduction.

Table 3.10: Summary of folding modulator co-synthesis results. The evaluation is ordered by the used promoter system (see column headers) and folding modulator (row identifiers in bold). Results presented in the Ind_{env} (**I**nduction by **e**nvironmental conditions) column were obtained using condition-inducible promoters P_{bglA} , P_{dnaK} or P_{ibpA} . **T7**, **T7 Δ lacI Δ lacO** and **constitutive** indicate that results were obtained in co-expression experiments using the regular T7 system with *lacI* and *lacO*, a T7 system devoid of regulatory elements *lacI* and *lacO* and constitutive promoters, respectively. Shake flask (“SF”) and fermentation (“F”) experiments were interpreted. The “Impact” row summarizes the results of the cultivations with production strains bearing the best combination of promoter and folding modulator compared to the reference strain (see respective chapters). Symbols +, ~ and – and identify increased, unchanged and decreased soluble product titers, respectively. Changes of more than 20 % compared to the reference value were considered as relevant. In the “P_{opt}” row, the promoter, for which the best result was determined, is presented. The very last row represents the maximum improvement (in %) of soluble product titer compared to the titer yielded by the respective plasmid-free reference strain and scale.

| | Ind_{env} ¹ | | T7 | | T7ΔlacIΔlacO | | constitutive | | | |
|------------------|---------------------------------------|-------|----------------|----------------------|---|----------------------|---------------------|-----------------|--------|----------------|
| | FabZ | | FabZ | | FabZ | | FabZ | | FabX | |
| | SF | F | SF | F | SF | F | SF | F | SF | F ⁴ |
| DsbA | | | | | | | | | | |
| Impact | + | ~ | + ² | – | + | – | + | – | + | n.a. |
| P _{opt} | bglA | bglA | T7xxiii | T7xxiii ³ | T7xxiii | T7xxiii | C2 | Ci ³ | C2 | n.a. |
| DsbC | | | | | | | | | | |
| Impact | n.a. | n.a. | – | n.a. | – | – | ~ | – | ~ | n.a. |
| P _{opt} | n.a. | n.a. | T7xxiii | n.a. | T7xxiii | T7xxiii ³ | Ci | Ci ³ | Ci | n.a. |
| FkpA | | | | | | | | | | |
| Impact | n.a. | n.a. | ~ | n.a. | + | – | + | + | + | ~ |
| P _{opt} | n.a. | n.a. | T7xxiii | n.a. | T7xxiii | T7xxiii | C8 | Ci | C8 | Ci |
| PpiD | | | | | | | | | | |
| Impact | n.a. | n.a. | ~ | – | + | – | ~ | – | ~ | n.a. |
| P _{opt} | n.a. | n.a. | T7xxiii | T7xxiii ³ | T7xxiii | T7xxiii ³ | Ci | Ci ³ | Ci | n.a. |
| Skp | | | | | | | | | | |
| Impact | n.a. | n.a. | ~ | – | + | ~ | + | ~ | + | ~ |
| P _{opt} | n.a. | n.a. | T7xxiii | T7xxiii ³ | T7xxiii | T7xxiii ³ | C2 | Ci | C2 | C2 |
| SurA | | | | | | | | | | |
| Impact | n.a. | n.a. | ~ | – | + | – | + | ~ | + | + |
| P _{opt} | n.a. | n.a. | T7xxiii | T7xxiii ³ | T7xxiii | T7xxiii ³ | C2 | Ci | C2 | C8 |
| | +89 % | -14 % | +28 % | -83 % | +295 % | -10 % | +341 % | +50 % | +661 % | +24 % |

¹Due to experimental issues, only results from the second experiment set are presented.

²+ compared to plasmid-free, – compared to empty basic plasmid-bearing reference strain.

³Only this promoter was tested.

⁴Assessment based on 2D-HPLC results after “hold step”.

Conclusion and Outlook

Overall conclusions

Genomic integration of the target gene and co-synthesis of folding modulators were applied to improve folding and secretion of recombinant proteins in *E. coli*. Plasmid-free expression was previously suggested for production of recombinant proteins in the cytoplasm (Chen et al., 2008; Mairhofer et al., 2013; Spadiut et al., 2014; Striedner et al., 2010) (also: patent application WO2008142028). In this study, the method was used for periplasmic production of model proteins under control of the T7 RNA polymerase system. Soluble product yields of three antibody fragments were significantly improved by application of genome-integrated strains (FabX: 3.3-fold, 53 → 174 mg l⁻¹; FabZ: 4.4-fold, 83 → 363 mg l⁻¹; scFv-Z-1: 8.6-fold, no absolute quantification). Thus, the suitability of plasmid-free strains for periplasmic production of antibody fragments was proven.

Plasmid-associated issues (plasmid loss, negative influence of the basic plasmid *per se*, high level of basal expression, growth cessation after inducer addition and overload of the secretion machinery), negatively influenced target product yields in plasmid-bearing strains (see Table 1.2 (p. 32) and Section 3.2.5). These issues were mitigated by application of genome-integrated strains. These findings were in accordance with results from previous studies (Mairhofer et al., 2013; Striedner et al., 2010). Despite improvements caused by genomic integration, product amounts in the g l⁻¹ range were shown to aggregate in the periplasm. High cellular content of recombinant product and the absence of significant amounts of product fragments for all target proteins indicated that proteolytic degradation in the periplasm was not a major issue.

Within the course of this project, aside from the direct comparison of plasmid-free and plasmid-bearing production strains, further observations concerning genetic improvement strategies were made. For instance, a genome-integrated LC-HC dicistronic FabZ gene expression cassette was found to be more productive than its HC-LC counterpart. An influence of the gene order was also previously reported (Humphreys et al., 2002; Jeong et al., 2011; Kirsch et al., 2005). Interestingly and contrary to previous reports (Block et al., 2012; Sabri et al., 2013), integration into the reverse strand of the *attTn7* site of *E. coli* BL21(DE3) appeared to result in higher expression rates compared to strains bearing a forward-integrated target gene. Furthermore, in plasmid-free GFP-producing *E. coli* cells, a single copy of the kanamycin resistance gene cassette did not noticeably

influence the strain's growth and production behavior, which is somewhat contrary to a previous report (Mairhofer et al., 2010).

In fermentation experiments, a lowered production temperature did not significantly improve soluble product titer but decreased the overall productivity. This result was somewhat unexpected as lowered temperatures are frequently used to improve protein folding (Berlec and Strukelj, 2013; Lebendiker and Danieli, 2014; Overton, 2014; Rodriguez-Carmona et al., 2012; Rosano and Ceccarelli, 2014; Sugiki et al., 2014). A fed-batch fermentation process with a production phase temperature of 37°C and strong induction with IPTG for screening of co-expression strains was established (see Section 3.3.1).

The observation of periplasmic aggregates observed in fermentations of plasmid-free expression systems suggested the utilization of periplasmic folding modulators to increase soluble recombinant protein yield. Thus, periplasmic folding modulators DsbA, DsbC, FkpA, PpiD, Skp and SurA were co-synthesized individually. Their co-synthesis was controlled by different promoter modules and they were located on different plasmid backbones. Both the condition-inducible (see Section 3.4.1) and the T7 RNA polymerase-based (see Section 3.4.2) systems interfered significantly with the overall production of the target protein FabZ. Those interactions were successfully prevented by application of constitutive promoters to control folding modulator gene expression. Application of this system substantially increased titers in shake flask experiments (FabZ: > 300 %, FabX: > 600 %) and moderately in fermentations (FabZ: + 50 %, FabX: + 24 %) (see Table 3.10, p. 196).

Successful application of folding modulators was reported in numerous reports (see Table 1.1). Accordingly and judging by the drastic improvement in shake flasks, similar results were expected in fermentation experiments. However, merely moderate improvements were observed. Also, the widely spread pTUM4 co-expression system (DsbA, DsbC, FkpA and SurA under transcriptional control of native *E. coli* promoters) was applied but even decreased soluble product titer dramatically (2.7-fold in shake flask experiments, 4.5-fold in fermentations).

In conclusion, the developed system of “synergistic combinations of cell engineering approaches” (Delic et al., 2014) significantly improved production of two Fab fragments in fed-batch fermentations with industrially relevant process parameters. Specifically, soluble FabX titers were improved 3.3-fold by genomic integration and, additionally, increased by 24 % by folding modulator co-synthesis, resulting in a maximum titer of 450 mg l⁻¹. Soluble FabZ titers were improved 4.3-fold by genomic integration and, additionally, increased by 50 % by folding modulator co-synthesis, resulting in a maximum titer of 576 mg l⁻¹.

Application of genetic toolbox elements

To obtain a robust and lean toolbox, single elements of the developed approach should be easily combinable. Three PoI-producing strains (plasmid-bearing, forward and reverse genome-integrated) should be the basis, as they are characterized by different expression levels. The relevance of FkpA, Skp and SurA as co-synthesis targets was previously suggested (see Table 1.1, p. 26) and supported by the results of this study. DsbC exerted no or a negative influence on target protein production. However, it might show its positive influence in case of products with non-consecutive disulfide bonds (Berkmen, 2012). In accordance with a previous report (Kolaj et al., 2009), PpiD did not have a positive influence and, thus, seemed not to be relevant as a toolbox element. Concerning promoter modules for folding modulator co-synthesis, constitutive promoters Ci and C8 cover the obviously relevant range of promoter strength. Furthermore, results from fermentation experiments suggested that lower amounts of over-produced folding modulator were more appropriate compared to very high levels. Thus, the strongest constitutive promoter C2 was removed from the standard toolbox. The remaining elements (three locations of the target gene, five folding modulators, two promoter modules) appeared relevant. Combination of this limited number of toolbox elements resulted in 33 strains (including strains without co-expression plasmids). As the success of their application can hardly be predicted, a trial-and-error screening approach is required, which is in accordance with previous reports (Berkmen, 2012; Delic et al., 2014; Fahnert, 2012; Kolaj et al., 2009; Overton, 2014).

Creation of 33 strains for productivity screening experiments seemed reasonable. Otherwise, a step-wise approach might also be pursued. An overview of such a strategy is presented in Table 4.1. Initially, a strain for plasmid-based target protein production could be created using, *e.g.* the basic plasmid pBI1KT7.1. If analyses of production experiments show plasmid-related issues in the production strain, plasmid-free expression systems might be tested. Production strains with forward or reverse genome-integrated target gene expression cassette could be created within approximately two to three weeks. Shortening the timeline for genomic integration would be a major benefit. Again, production experiments should be performed. If one of the strains forms periplasmic aggregates at a relevant level, which appears feasible, application of folding modulators would become relevant. After preparation of competent cells using each of the three basic strains (plasmid-bearing, forward and reverse genome-integrated), the 10 co-expression plasmids (pBI4iSCi/C8.3-dsbA/dsbC/fkpA/skp/surA) could be used to individually transform the production strains. The resulting 30 strains with elevated folding modulator levels have

to be screened for the most efficient producer.

Table 4.1: A possible step-wise application of the developed toolbox. Details are described in the text. Time frames indicate maximum cumulative times for strain creation procedures under consideration of the local BI RCV environment (see Materials and Methods). Production experiments were not considered in the calculation.

| Step | Molecular biology | Strain evaluation |
|--------------------------------|---|---|
| Plasmid-based expression | Cloning of GoI into the target plasmid (<i>e.g.</i> pBI1KT7.1); transformation and cryo culture preparation → 2 weeks, 1 strain | Soluble/total product titer (cytoplasm/periplasm/supernatant); plasmid loss → bottleneck analysis; process development |
| Plasmid-free expression | Genomic integration of GoI into forward and reverse strand; preparation of cryo cultures → 2 - 3 weeks, 3 strains | Soluble/total product titer (cytoplasm/periplasm/supernatant) → bottleneck analysis; process development |
| Periplasmic folding modulators | Preparation of competent cells; transformation of production strains with selected folding modulator plasmids (Ci/C8, <i>dsbA/dsbC/fkpA/skp/surA</i>); cryo culture preparation → 2 weeks, 33 strains | Soluble/total product titer (periplasm/supernatant); quantification of folding modulators → bottleneck analysis; process development |

The established genetic toolbox was supposed to enable a quick progress through the host engineering phase. The step-wise approach might help shortening the timeline of currently six to seven weeks. Note that parallelization is possible in several steps. For instance, genome-integrated strains could be prepared while plasmid-bearing production strains are examined in production experiments. However, parallelization is dependent on equipment and resources in the respective laboratory. Assessment of the resulting strains in production experiments after each step could reveal the necessity of further optimization.

Outlook

Concerning further screening experiments, establishing high-throughput capacities, analytical techniques and an enlargement of the genetic toolbox would be feasible. Initially, the separation of host engineering and process development was desired. Thus, the process parameters were locked after initiation of the co-expression work. However, results from shake flask experiments using co-expression strains were barely predictive for the results of fed-batch fermentations. Under these circumstances, a range of different standard production processes should be defined. These should be used to screen a relevant number of strains.

However, a small scale screening system which allows for either i) screening of a large number of hosts in a process predictive of the final scale or ii) simultaneous host screening and testing of process parameters would be preferred. Efforts to establish a small scale system suitable to predict the results of 5 l fed-batch approaches are currently undertaken in the course of a master thesis (Klösch, 2016). Furthermore, also a multi-fermentation system is being developed for these needs at BI RCV.

An advantage of the developed toolbox is its expandability and compatibility with further techniques in the BI RCV environment. Detailed suggestions to further develop single techniques are presented in the respective conclusion sections (see Sections 3.1.5, 3.2.5 and 3.4.4). Due to the robustness of the created strains and the employed genetic elements, the established genetic toolbox is compatible with both host and process engineering (see Sections 1.3.4 and 1.3.5). Note that folding modulator plasmids are also compatible with plasmid-based GoI expression using widely spread commercially available plasmids with compatible ori and resistance gene cassette. Down-regulation of specific targets (*e.g.* periplasmic proteases; see Section 1.3.2) was not addressed in this study. However, proteases were also reported to cause reduced yields of recombinant proteins (Carneiro et al., 2013; Duguay and Silhavy, 2004; Hunke and Betton, 2003). Thus, an sRNA-based approach to enable knock-down of *E. coli*-endogenous genes is currently evaluated within the course of a master thesis (Klösch, 2016).

In addition, more detailed examinations of strains in production experiments would facilitate purposeful application of genetic toolbox elements. Specifically, formation of target protein, up- and down-regulated targets and the entire host cell transcriptome and proteome appear relevant. Appropriate techniques (*e.g.* RT-PCR, DNA microarrays, MS/MS-based analysis) are available and have previously been used for similar approaches (Mairhofer et al., 2013; Marisch et al., 2013b).

In conclusion, a robust and lean toolbox, which allows for i) adjusting the target gene expression rate, ii) up-regulation of supportive proteins and iii) down-regulation of undesired endogenous genes seems most promising for further development. Moreover, the predictability of the small scale screening system and the analytical methods should be further improved.

Confidential Supplement

The following chapter contains internal information of Boehringer Ingelheim RCV GmbH & Co KG and cannot be disclosed to the public due to confidentiality reasons.

References

- Addgene (2014). Plasmid Modification by Annealed Oligo Cloning. URL: http://www.addgene.org/plasmid_protocols/annealed_oligo_cloning/.
- Adolf, G., Heider, K., Miglietta, J., Ostermann, E., Patzelt, E., Sproll, M., and Van, D. (2002). Antibodies specific for cd44v6. WO Patent App. PCT/EP2002/005,467.
- Ahmad, Z. A., Yeap, S. K., Ali, A. M., Ho, W. Y., Alitheen, N. B. M., and Hamid, M. (2012). scFv Antibody: Principles and Clinical Application. *Clinical and Developmental Immunology*.
- Albermann, C., Trachtmann, N., and Sprenger, G. A. (2010). A simple and reliable method to conduct and monitor expression cassette integration into the *Escherichia coli* chromosome. *Biotechnology Journal*, 5(1):32–38.
- Amann, E., Brosius, J., and Ptashne, M. (1983). Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene*, 25:167–178.
- Amann, E., Ochs, B., and Abel, K. J. (1988). Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene*, 69(2):301–315.
- Andersen, C. L., Matthey-Dupraz, A., Missiakas, D., and Raina, S. (1997). A new *Escherichia coli* gene, *dsbG*, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. *Molecular Microbiology*, 26(1):121–132.
- Anderson, C. (2006). The Anderson promoter collection. URL: <http://parts.igem.org/Promoters/Catalog/Anderson>.
- Andersson, L., Yang, S., Neubauer, P., and olof Enfors, S. (1996). Impact of plasmid presence and induction on cellular responses in fed batch cultures of *Escherichia coli*. *Journal of Biotechnology*, 46(3):255 – 263.
- Antonoaea, R., Fürst, M., Nishiyama, K. i., and Müller, M. (2008). The Periplasmic Chaperone PpiD Interacts with Secretory Proteins Exiting from the SecYEG Translocon. *Biochemistry*, 47(20):5649–5656.

- Arakawa, T. and Ejima, D. (2014). Refolding Technologies for Antibody Fragments. *Antibodies*, 3(2):232.
- Arie, J.-P., Sassoon, N., and Betton, J.-M. (2001). Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*. *Molecular Microbiology*, 39(1):199–210.
- Arpino, J. A., Hancock, E. J., Anderson, J., Barahona, M., Stan, G. B., Pappachristodoulou, A., and Polizzi, K. (2013). Tuning the dials of Synthetic Biology. *Microbiology*, 159(Pt 7):1236–1253.
- Arredondo, S. and Georgiou, G. (2011). The Problem of Expression of Multidisulfide Bonded Recombinant Proteins in *E. coli*. In Chang, R. J. Y. and Ventura, S., editors, *Folding of Disulfide Proteins*, volume 14 of *Protein Reviews*, pages 183–215. Springer New York.
- Arredondo, S., Segatori, L., Gilbert, H. F., and Georgiou, G. (2008). *De Novo* Design and Evolution of Artificial Disulfide Isomerase Enzymes Analogous to the Bacterial DsbC. *The Journal of Biological Chemistry*, 283(46):31469–31476.
- Arsene, F., Tomoyasu, T., and Bukau, B. (2000). The heat shock response of *Escherichia coli*. *International Journal of Food Microbiology*, 55(1-3):3–9.
- Aune, T. E. V. (2008). *High level recombinant protein production in Escherichia coli by engineering broad-host-range plasmid vectors containing the Pm/xylS expression cassette*. PhD thesis, Norwegian University of Science and Technology Faculty of Natural Sciences and Technology Department of Biotechnology.
- Aune, V., Bakke, I., Drablos, F., Lale, F., Brautaset, T., and Valla, S. (2010). Direct evolution of the transcription factor XylS for the development of improved expression systems. *Microbial Biotechnology*, 3:38–47.
- Baars, L. (2007). *Protein targeting, translocation and insertion in Escherichia coli - Proteomic analysis of substrate-pathway relationships*. PhD thesis, Stockholm University.
- Baars, L., Ytterberg, A., Drew, D., Wagner, S., Thilo, C., van Wijk, K., and de Gier, J. (2006). Defining the role of the *Escherichia coli* chaperone SecB using comparative proteomics. *Journal of Biological Chemistry*, 281(15):10024–10034.
- Bader, M., Muse, W., Ballou, D. P., Gassner, C., and Bardwell, J. C. (1999). Oxidative Protein Folding Is Driven by the Electron Transport System. *Cell*, 98(2):217 – 227.

- Balzer, S., Kucharova, V., Megerle, J., Lale, R., Brautaset, T., and Valla, S. (2013). A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in *Escherichia coli*. *Microbial Cell Factories*, 12:26.
- Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology*, 10(5):411–421.
- Baneyx, F. and Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*, 22(11):1399–1408.
- Bardwell, J. C., Lee, J. O., Jander, G., Martin, N., Belin, D., and Beckwith, J. (1993). A pathway for disulfide bond formation *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America*, 90(3):1038–1042.
- Bardwell, J. C., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulfide bond formation *in vivo*. *Cell*, 67(3):581–589.
- Bartlett, A. I. and Radford, S. E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nature Structural and Molecular Biology*, 16(6):582–588.
- Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., and Hwa, T. (2015). Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature*, 528(7580):99–104.
- Battesti, A., Majdalani, N., and Gottesman, S. (2011). The RpoS-Mediated General Stress Response in *Escherichia coli*. *Annual Review of Microbiology*, 65(1):189–213. PMID: 21639793.
- Battistoni, M. A., Carri, A. T., Mazzetti, P., and Rotilio, G. (1992). Temperature-dependent protein folding *in vivo* - Lower growth temperature increases yield of two genetic variants of *Xenopus laevis* Cu, Zn superoxide dismutase in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 186(3):1339 – 1344.
- Behrens, S., Maier, R., de Cock, H., Schmid, F. X., and Gross, C. A. (2001). The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *The EMBO Journal*, 20(1-2):285–294.
- Behrens-Kneip, S. (2010). The role of SurA factor in outer membrane protein transport and virulence. *International Journal of Medical Microbiology*, 300(7):421 – 428.
- Bentley, W. and Kompala, D. (1990). Plasmid instability in batch cultures of recombinant bacteria. *Chemical Engineering Education*, 24:168–172.

- Bentley, W., Mirjalili, N., Andersen, D., Davis, R., and Kompala, D. (1990). Plasmid-encoded protein: the principal factor in the “metabolic burden” associated with recombinant bacteria. *Biotechnology and Bioengineering*, 35:668–681.
- Berkmen, M. (2012). Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expression and Purification*, 82(1):240 – 251.
- Berlec, A. and Strukelj, B. (2013). Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *Journal of Industrial Microbiology & Biotechnology*, 40(3-4):257–274.
- Beshay, U., Miksch, G., Friehs, K., and Flaschel, E. (2007). Increasing the secretion ability of the *kil* gene for recombinant proteins in *Escherichia coli* by using a strong stationary-phase promoter. *Biotechnology Letters*, 29(12):1893–1901.
- Bessette, P. H., Qiu, J., Bardwell, J. C. A., Swartz, J. R., and Georgiou, G. (2000). Effect of Sequences of the Active-Site Dipeptides of DsbA and DsbC on *In Vivo* Folding of Multidisulfide Proteins in *Escherichia coli*. *Journal of Bacteriology*, 183(3):980–988.
- Better, M., Chang, C., Robinson, R., and Horwitz, A. (1988). *Escherichia coli* secretion of an active chimeric antibody fragment. *Science*, 240(4855):1041–1043.
- Bird, R., Hardman, K., Jacobson, J., Johnson, S., Kaufman, B., Lee, S., Lee, T., Pope, S., Riordan, G., and Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science*, 242(4877):423–426.
- Bird, R. E. (1981). Homology between *Escherichia coli* plasmids ColE1 and p15A. *Journal of Bacteriology*, 145(3):1305–1309.
- Birnbaum, S. and Bailey, J. (1991). Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinant *Escherichia coli*. *Biotechnology and Bioengineering*, 37(8):736–745.
- Bitto, E. and McKay, D. B. (2002). Crystallographic Structure of SurA, a Molecular Chaperone that Facilitates Folding of Outer Membrane Porins. *Structure*, 10(11):1489 – 1498.
- Bitto, E. and McKay, D. B. (2003). The Periplasmic Molecular Chaperone Protein SurA Binds a Peptide Motif That Is Characteristic of Integral Outer Membrane Proteins. *Journal of Biological Chemistry*, 278(49):49316–49322.
- Bjoerck, L. (1988). Protein L. A novel bacterial cell wall protein with affinity for Ig L chains. *The Journal of Immunology*, 140(4):1194–7.

- Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Gregor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B., and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, 277:1453–1474.
- Block, D. H. S., Hussein, R., Liang, L. W., and Lim, H. N. (2012). Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Research*, 40(18):8979–8992.
- Bonomo, J. and Gill, R. T. (2005). Amino acid content of recombinant proteins influences the metabolic burden response. *Biotechnology and Bioengineering*, 90(1):116–126.
- Borjesson, P. K. E., Postema, E. J., Roos, J. C., Colnot, D. R., Marres, H. A. M., van Schie, M. H., Stehle, G., de Bree, R., Snow, G. B., Oyen, W. J. G., and van Dongen, G. A. M. S. (2003). Phase I Therapy Study with 186Re-labeled Humanized Monoclonal Antibody BIWA 4 (Bivatuzumab) in Patients with Head and Neck Squamous Cell Carcinoma. *Clinical Cancer Research*, 9(10):3961s–3972s.
- Bothmann, H. and Pluckthun, A. (1998). Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nature Biotechnology*, 16(4):376–380.
- Bothmann, H. and Plueckthun, A. (2000). The Periplasmic *Escherichia coli* Peptidylprolyl *cis,trans*-isomerase FkpA: I. Increased Functional Expression of Antibody Fragments with and without *cis*-prolines. *Journal of Biological Chemistry*, 275(22):17100–17105.
- Braakman, I. and Bulleid, N. J. (2011). Protein Folding and Modification in the Mammalian Endoplasmic Reticulum. *Annual Review of Biochemistry*, 80(1):71–99. PMID: 21495850.
- Breustedt, D. A., Schönfeld, D. L., and Skerra, A. (2006). Comparative ligand-binding analysis of ten human lipocalins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1764(2):161 – 173.
- Brondyk, W. H. (2009). Chapter 11: Selecting an Appropriate Method for Expressing a Recombinant Protein. In Burgess, R. R. and Deutscher, M. P., editors, *Guide to Protein Purification, 2nd Edition*, volume 463 of *Methods in Enzymology*, pages 131 – 147. Academic Press.
- Brosius, J., Erfle, M., and Storella, J. (1985). Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity. *Journal of Biological Chemistry*, 260(6):3539–3541.

- Buehler, P., Wetterbauer, D., Gierschner, D., Wetterbauer, U., Elsaesser-Beile, U., and Wolf, P. (2010). Influence of Structural Variations on Biological Activity of Anti-PSMA scFv and Immunotoxins Targeting Prostate Cancer. *Anticancer Research*, 30(9):3373–3379.
- Burgess, R. R. (2009). Chapter 17: Refolding Solubilized Inclusion Body Proteins. In Burgess, R. R. and Deutscher, M. P., editors, *Guide to Protein Purification, 2nd Edition*, volume 463 of *Methods in Enzymology*, pages 259 – 282. Academic Press.
- Calloni, G., Chen, T., Schermann, S., Chang, H. c., Genevaux, P., Agostini, F., Tartaglia, G., Hayer-Hartl, M., and Hartl, F. (2012). DnaK Functions as a Central Hub in the *E. coli* Chaperone Network. *Cell Reports*, 1(3):251–264.
- Calos, M. P. (1978). DNA sequence for a low-level promoter of the *lac* repressor gene and an ‘up’ promoter mutation. *Nature*, 274(5673):762–765.
- Carneiro, S., Ferreira, E. O. C., and Rocha, I. (2013). Metabolic responses to recombinant bioprocesses in *Escherichia coli*. *Journal of Biotechnology*, 164(3):396 – 408. Highlights from the ChemPor 2011 Conference.
- Carrier, T. A. and Keasling, J. (1997). Controlling Messenger RNA Stability in Bacteria: Strategies for Engineering Gene Expression. *Biotechnology Progress*, 13(6):699–708.
- Carter, A. D., Morris, C. E., and McAllister, W. T. (1981). Revised transcription map of the late region of bacteriophage T7 DNA. *Journal of Virology*, 37(2):636–642.
- Cassland, P., Larsson, S., Nilvebrant, N.-O., and Jönsson, L. (2004). Heterologous expression of barley and wheat oxalate oxidase in an *E. coli* *trxB gor* double mutant. *Journal of Biotechnology*, 109:53 – 62. Recombinant Proteins and Host Cell Physiology.
- Cayley, S., Lewis, B. A., Guttman, H. J., and Record, M. (1991). Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity. *Journal of Molecular Biology*, 222(2):281 – 300.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W., and Prasher, D. (1994). Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805.
- Chambers, M. (2014). ChemIDplus - 214559-60-1 - Bivatuzumab [INN] - Searchable synonyms, formulas, resource links, and other chemical information.
- Chang, A. C. and Cohen, S. N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *Journal of Bacteriology*, 134(3):1141–1156.

- Chen, H., Bjercknes, M., Kumar, R., and Jay, E. (1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Research*, 22(23):4953–4957.
- Chen, H.-T., Lin, M.-S., and Hou, S.-Y. (2008). Multiple-copy-gene integration on chromosome of *Escherichia coli* for beta-galactosidase production. *Korean Journal of Chemical Engineering*, 25(5):1082–1087.
- Chen, J., Song, J.-l., Zhang, S., Wang, Y., Cui, D.-f., and Wang, C.-c. (1999). Chaperone Activity of DsbC. *Journal of Biological Chemistry*, 274(28):19601–19605.
- Chen, R. and Henning, U. (1996). A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Molecular Microbiology*, 19(6):1287–1294.
- Cho, S.-H. and Collet, J.-F. (2012). Many Roles of the Bacterial Envelope Reducing Pathways. *Antioxidants & Redox Signaling*, 18(13):1690–1698.
- Choi, J. and Lee, S. (2004). Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Applied Microbiology and Biotechnology*, 64(5):625–635.
- Choi, J. H., Keum, K. C., and Lee, S. Y. (2006). Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chemical Engineering Science*, 61(3):876 – 885. Biomolecular Engineering.
- Chou, C. (2007). Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Applied Microbiology and Biotechnology*, 76(3):521–532.
- Chuang, S. E., Burland, V., Plunkett III, G., Daniels, D. L., and Blattner, F. R. (1993). Sequence analysis of four new heat-shock genes constituting the *hslTS/ibpAB* and *hslVU* operons in *Escherichia coli*. *Gene*, 134(1):1–6.
- Clubb, R., Ferguson, S., Walsh, C. T., and Wagner, G. (1995). Three-dimensional solution structure of *Escherichia coli* periplasmic cyclophilin. *Biochemistry*, 33:2761–2772.
- Colnot, D., Roos, J., de Bree, R., Wilhelm, A., Kummer, J., Hanft, G., Heider, K.-H., Stehle, G., Snow, G., and van Dongen, G. (2003). Safety, biodistribution, pharmacokinetics, and immunogenicity of 99mTc-labeled humanized monoclonal antibody BIWA 4 (bivatuzumab) in patients with squamous cell carcinoma of the head and neck. *Cancer Immunology, Immunotherapy*, 52(9):576–582.
- Colwell, R. (2000). Viable but nonculturable bacteria: a survival strategy. *Journal of Infection and Chemotherapy*, 6(2):121–125.

- Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173(1):33 – 38.
- Cote, R. and Gherna, R. L. (2002). *Medium Formulation and Design, E. coli and Bacillus spp.* John Wiley & Sons, Inc.
- Creighton, T. E. (1986). Disulfide bonds as probes of protein folding pathways. In *Enzyme Structure Part L*, volume 131 of *Methods in Enzymology*, pages 83 – 106. Academic Press.
- Curless, C., Pope, J., and Tsai, L. (1990). Effect of Preinduction Specific Growth Rate on Recombinant Alpha Consensus Interferon Synthesis in *Escherichia coli*. *Biotechnology Progress*, 6(2):149–152.
- Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S., and Kim, J. F. (2009). Tracing Ancestors and Relatives of *Escherichia coli* B, and the Derivation of B Strains REL606 and BL21(DE3). *Journal of Molecular Biology*, 394(4):634 – 643.
- Dalal, K. and Duong, F. (2011). The SecY complex: conducting the orchestra of protein translocation. *Trends in Cell Biology*, 21(9):506 – 514.
- Dartigalongue, C., Missiakas, D., and Raina, S. (2001). Characterization of the *Escherichia coli* σ^E Regulon. *Journal of Biological Chemistry*, 276(24):20866–20875.
- Dartigalongue, C. and Raina, S. (1998). A new heat-shock gene, *ppiD*, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in *Escherichia coli*. *The EMBO Journal*, 17(14):3968–3980.
- Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- Datta, S., Costantino, N., and Court, D. (2006). A set of recombineering plasmids for gram-negative bacteria. *Gene*, 379(0):109–115.
- Datta, S., Costantino, N., Zhou, X., and Court, D. (2008). Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proceedings of the National Academy of Sciences*, 105:1626–1631.
- Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Research*, 39(3):1131–1141.

- de Boer, H., Comstock, L., and Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proceedings of the National Academy of Sciences*, 80(1):21–25.
- de Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microbial Cell Factories*, 8(1):26.
- de Marco, A. (2011). Biotechnological applications of recombinant single-domain antibody fragments. *Microbial Cell Factories*, 10(1):44.
- de Marco, A. (2012). Recent contributions in the field of the recombinant expression of disulfide bonded proteins in bacteria. *Microbial Cell Factories*, 11(1):129.
- de Marco, A. (2013). Recombinant polypeptide production in *E. coli*: towards a rational approach to improve the yields of functional proteins. *Microbial Cell Factories*, 12(1):101.
- de Marco, A. and de Marco, V. (2004). Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *Journal of Biotechnology*, 109:45–52.
- de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., and Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC Biotechnology*, 7(1):32–.
- Delic, M., Gongrich, R., Mattanovich, D., and Gasser, B. (2014). Engineering of protein folding and secretion-strategies to overcome bottlenecks for efficient production of recombinant proteins. *Antioxidants and Redox Signalling*, 21(3):414–37.
- Denoncin, K. and Collet, J. (2012). Disulfide bond formation in the bacterial periplasm: major achievements and challenges ahead. *Antioxidants and Redox Signalling*, 19(1):63–71.
- Denoncin, K., Schwalm, J., Vertommen, D., Silhavy, T. J., and Collet, J.-F. (2012). Dissecting the *Escherichia coli* periplasmic chaperone network using differential proteomics. *Proteomics*, 12(9):1391–1401.
- Depuydt, M., Messens, J., and Collet, J.-F. (2010). How Proteins Form Disulfide Bonds. *Antioxidants and Redox Signaling*, 15(1):49–66.
- Derman, A. I. and Beckwith, J. (1991). *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. *Journal of Bacteriology*, 173(23):7719–7722.

- Deuschle, U., Kammerer, W., Gentz, R., and Bujard, H. (1986). Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *The EMBO Journal*, 5:2987–2994.
- Doelle, H. W., Ewings, K. N., and Hollywood, N. W. (1982). *Microbial Reactions*, chapter Regulation of glucose metabolism in bacterial systems, pages 1–35. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Dong, H., Nilsson, L., and Kurland, C. (1995). Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *Journal of Bacteriology*, 177(6):1497–1504.
- Dragosits, M., Frascotti, G., Bernard-Granger, L., Vázquez, F., Giuliani, M., Baumann, K., Rodriguez-Carmona, E., Tokkanen, J., Parrilli, E., Wiebe, M. G., Kunert, R., Maurer, M., Gasser, B., Sauer, M., Branduardi, P., Pakula, T., Saloheimo, M., Penttilä, M., Ferrer, P., Luisa Tutino, M., Villaverde, A., Porro, D., and Mattanovich, D. (2011). Influence of growth temperature on the production of antibody Fab fragments in different microbes: A host comparative analysis. *Biotechnology Progress*, 27(1):38–46.
- Dragosits, M., Nicklas, D., and Tagkopoulos, I. (2012). A synthetic biology approach to self-regulatory recombinant protein production in *Escherichia coli*. *Journal of Biological Engineering*, 6(1):2–.
- Driessen, A. J. (2001). SecB, a molecular chaperone with two faces. *Trends in Microbiology*, 9(5):193–196.
- Driessen, A. J. and Nouwen, N. (2008). Protein Translocation Across the Bacterial Cytoplasmic Membrane. *Annu.Rev.Biochem.*, 77(1):643–667.
- du Plessis, D. J., Nouwen, N., and Driessen, A. J. (2011). The Sec translocase. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(3):851–865.
- Dubendorf, J. W. and Studier, F. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of Molecular Biology*, 219(1):45 – 59.
- Duguay, A. and Silhavy, T. J. (2004). Quality control in the bacterial periplasm. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694:121–134.
- Dumon-Seignovert, L., Cariot, G., and Vuillard, L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification*, 37(1):203 – 206.

- Duong, F., Eichler, J., Price, A., Leonard, M. R., and Wickner, W. (1997). Biogenesis of the Gram-Negative Bacterial Envelope. *Cell*, 91(5):567 – 573.
- Dürkop, M. (2014). Improving Periplasmic Fab Production in *Escherichia coli* by Folding Modulator Gene Co-Expression. Master’s thesis, University of Natural Resources and Life Sciences, Vienna.
- Edelman, G. M. (1971). Antibody Structure and Molecular Immunology. *Annals of the New York Academy of Sciences*, 190(1):5–25.
- Edman, P. (1950). Method for Determination of the Amino Acid Sequence in Peptides. *Acta Chemica Scandinavica*, 4:283–293.
- Edman, P. (1956). On the mechanism of the phenyl isothiocyanate degradation of peptides. *Acta Chemica Scandinavica*, 10:761–768.
- Eiteman, M. A. and Altman, E. (2006). Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends in Biotechnology*, 24(11):530 – 536.
- Elksne, L. E. and Rasmussen, B. A. (1996). Dsb-insensitive expression of CcrA, a metallo-beta-lactamase from *Bacteroides fragilis*, in *Escherichia coli* after amino acid substitution at two cysteine residues within CcrA. *Journal of Bacteriology*, 178(14):4306–4309.
- EMA (2007). Lucentis EPAR Scientific Discussion. Technical report, EMA.
- Entzminger, K. C., Chang, C., Myhre, R. O., McCallum, K. C., and Maynard, J. A. (2012). The Skp Chaperone Helps Fold Soluble Proteins *in Vitro* by Inhibiting Aggregation. *Biochemistry*, 51(24):4822–4834. PMID: 22650963.
- Fahnert, B. (2012). Using Folding Promoting Agents in Recombinant Protein Production: A Review. In Lorence, A., editor, *Recombinant Gene Expression*, volume 824 of *Methods in Molecular Biology*, pages 3–36. Humana Press.
- FDA (2009). Questions and answers on current good manufacturing practices, good guidance practices, level 2 guidance - buildings and facilities.
- Fernandez-Castane, A., Vine, C. E., Caminal, G., and Lopez-Santin, J. (2012). Evidencing the role of lactose permease in IPTG uptake by *Escherichia coli* in fed-batch high cell density cultures. *Journal of Biotechnology*, 157(3):391–398.
- Ferrara, N., Damico, L., Shams, N., Lowman, H., and Kim, R. (2006). Development of Ranibizumab, an Anti-Vascular Endothelial Growth Factor Antigen Binding Fragment, as Therapy for Neovascular Age-Related Macular Degeneration. *Retina*, 26(8):859–870.

- Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J., Vazquez, E., and Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, 8(1):17.
- Francis, D. M. and Page, R. (2010). Strategies to Optimize Protein Expression in *E. coli*. In *Current Protocols in Protein Science*, volume 61, chapter 5.24, pages 1–29. John Wiley & Sons, Inc.
- Fredrick, K. and Ibba, M. (2010). How the Sequence of a Gene Can Tune Its Translation. *Cell*, 141(2):227–229.
- Frenzel, A., Hust, M., and Schirrmann, T. (2013). Expression of recombinant antibodies. *Frontiers in Immunology*, 4(217).
- Freund, C., Ross, A., Guth, B., Plueckthun, A., and Holak, T. A. (1993). Characterization of the linker peptide of the single-chain Fv fragment of an antibody by NMR spectroscopy. *FEBS Letters*, 320(2):97–100.
- Friedrich, L., Stangl, S., Hahne, H., Küster, B., Köhler, P., Multhoff, G., and Skerra, A. (2010). Bacterial production and functional characterization of the Fab fragment of the murine IgG1/ λ monoclonal antibody cmHsp70.1, a reagent for tumour diagnostics. *Protein Engineering Design and Selection*, 23(4):161–168.
- Friehs, K. (2004). Plasmid Copy Number and Plasmid Stability. In Scheper, T., editor, *New Trends and Developments in Biochemical Engineering*, volume 86 of *Advances in Biochemical Engineering*, pages 47–82. Springer Berlin Heidelberg.
- Gajiwala, K. S. and Burley, S. K. (2000). HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *Journal of Molecular Biology*, 295(3):605 – 612.
- Garcia-Fruitos, E., Rodriguez-Carmona, E., Diez-Gil, C., Ferraz, R. M., Vazquez, E., Corchero, J. L., Cano-Sarabia, M., Ratera, I., Ventosa, N., Veciana, J., and Villaverde, A. (2009). Surface Cell Growth Engineering Assisted by a Novel Bacterial Nanomaterial. *Advanced Materials*, 21(42):4249–4253.
- Garcia-Fruitos, E., Vazquez, E., Diez-Gil, C., Corchero, J. L., Seras-Franzoso, J., Ratera, I., Veciana, J., and Villaverde, A. (2012). Bacterial inclusion bodies: making gold from waste. *Trends in Biotechnology*, 30(2):65 – 70.
- Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodriguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino,

- M., Mattanovich, D., and Villaverde, A. (2008). Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial Cell Factories*, 7(1):11.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In Walker, J. M., editor, *The Proteomics Protocols Handbook*, pages 571–607. Humana Press.
- GeneBridges (2013). Technical Protocol for FLPe Expression Plasmids for *E. coli*. Technical report.
- Giacomelli, G. and Depetris, A. (2012). Part:BBa_K731721. URL: http://parts.igem.org/Part:BBa_K731721?title=Part:BBa_K731721.
- Gilbert, W. and Muller-Hill, B. (1966). Isolation of the Lac Repressor. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6):1891–1898.
- Giuliodori, A. M., Pietro, F. D., Marzi, S., Masquida, B., Wagner, R., Romby, P., Gualerzi, C. O., and Pon, C. L. (2010). The *cspA* mRNA Is a Thermosensor that Modulates Translation of the Cold-Shock Protein CspA. *Molecular Cell*, 37(1):21–33.
- Glick, B. R. (1995). Metabolic load and heterologous gene expression. *Biotechnology Advances*, 13(2):247–261.
- Goemans, C., Denoncin, K., and Collet, J.-F. (2014). Folding mechanisms of periplasmic proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8):1517 – 1528. Protein trafficking and secretion in bacteria.
- Goldenberg, D., Azar, I., and Oppenheim, A. B. (1996). Differential mRNA stability of the *cspA* gene in the cold-shock response of *Escherichia coli*. *Molecular Microbiology*, 19(2):241–248.
- Goldenberg, D. M. (2002). Targeted therapy of cancer with radiolabeled antibodies. *Journal of Nuclear Medicine*, 43(5):693–713.
- Goldstein, J., Pollitt, N., and Inouye, M. (1990). Major cold shock protein of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 87(1):283–287.
- Goloubinoff, P., Gatenby, A. A., and Lorimer, G. H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*. *Nature*, 337(6202):44–47.

- Gonzalez-Montalban, N., Garcia-Fruitos, E., and Villaverde, A. (2007). Recombinant protein solubility - does more mean better? *Nature Biotechnology*, 25(7):718–720.
- Goodman, D. B., Church, G. M., and Kosuri, S. (2013). Causes and Effects of N-Terminal Codon Bias in Bacterial Genes. *Science*, 342(6157):475–479.
- Gottesman, S. (1996). Proteases and Their Targets in *Escherichia coli*. *Annual Review of Genetics*, 30(1):465–506.
- Grabherr, R., Nilsson, E., Striedner, G., and Bayer, K. (2002). Stabilizing plasmid copy number to improve recombinant protein production. *Biotechnology and Bioengineering*, 77(2):142–147.
- Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G. A., Gottesman, M. E., and Nikiforov, V. (1992). Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 89(21):10341–10344.
- Gray, C. H. and Tatum, E. L. (1944). X-Ray Induced Growth Factor Requirements in Bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 30(12):404–410.
- Green, M. and Sambrook, J., editors (2012). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Grodberg, J. and Dunn, J. J. (1988). ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *Journal of Bacteriology*, 170(3):1245–1253.
- Gronenborn, B. (1976). Overproduction of phage Lambda repressor under control of the *lac* promoter of *Escherichia coli*. *Molecular and General Genetics*, 148(3):243–250.
- Grunberg-Manago, M. (1999). Messenger RNA Stability and Its Role in Control of Gene Expression in Bacteria and Phages. *Annual Review of Genetics*, 33(1):193–227.
- Gu, W., Zhou, T., and Wilke, C. O. (2010). A Universal Trend of Reduced mRNA Stability near the Translation-Initiation Site in Prokaryotes and Eukaryotes. *PLOS Computational Biology*, 6(2):e1000664–.
- Gunnarsen, K., Lunde, E., Kristiansen, P., Bogen, B., Sandlie, I., and Loset, G. (2010). Periplasmic expression of soluble single chain T cell receptors is rescued by the chaperone FkpA. *BMC Biotechnology*, 10(1):8.

- Guruprasad, K., Reddy, B., and Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Engineering*, 4(2):155–161.
- Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology*, 22(7):346 – 353.
- Haaland, P. (1989). *Experimental design in biotechnology*. CRC Press, first edition.
- Hartl, F., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, 475(7356):324–332.
- Hasenwinkle, D., Jarvis, E., Kops, O., Liu, C., Lesnicki, G., Haynes, C. A., and Kilburn, D. G. (1997). Very high-level production and export in *Escherichia coli* of a cellulose binding domain for use in a generic secretion-affinity fusion system. *Biotechnology and Bioengineering*, 55(6):854–863.
- Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B. L., Mori, H., and Horiuchi, T. (2006). Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Molecular Systems Biology*, 2(1).
- Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B. L., and Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *Journal of Immunological Methods*, 276:185–196.
- Hayhurst, A. and Harris, W. J. (1999). *Escherichia coli* Skp Chaperone Coexpression Improves Solubility and Phage Display of Single-Chain Antibody Fragments. *Protein Expression and Purification*, 15(3):336–343.
- Haylock, A.-K., Spiegelberg, D., Nilvebrant, J., Sandstrom, K., and Nestor, M. (2014). *In vivo* characterization of the novel CD44v6-targeting Fab fragment AbD15179 for molecular imaging of squamous cell carcinoma: a dual-isotope study. *EJNMMI Research*, 4(1):11.
- Heider, K.-H., Kuthan, H., Stehle, G., and Munzert, G. (2004). CD44v6: a target for antibody-based cancer therapy. *Cancer Immunology, Immunotherapy*, 53(7):567–579.
- Heisteringer, L. (2013). Improved Secretion and Folding of Recombinantly Produced Proteins in *Escherichia coli*. Master’s thesis, University of Natural Resources and Life Sciences, Vienna, Austria.

- Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., and Behrens, S. (2005). The Periplasmic Chaperone SurA Exploits Two Features Characteristic of Integral Outer Membrane Proteins for Selective Substrate Recognition. *Journal of Biological Chemistry*, 280(25):23540–23548.
- Heras, B., Edeling, M. A., Schirra, H. J., Raina, S., and Martin, J. L. (2004). Crystal structures of the DsbG disulfide isomerase reveal an unstable disulfide. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24):8876–8881.
- Herrmann, J. and Riemer, J. (2014). Three approaches to one problem: protein folding in the periplasm, the endoplasmic reticulum, and the intermembrane space. *Antioxidants and Redox Signaling*, 21(1523-0864 (Linking)):438–456.
- Hillyar, C. R. (2012). Genetic recombination in bacteriophage lambda. *Bioscience Horizons*, 5.
- Hiniker, A. and Bardwell, J. C. (2004). *In vivo* substrate specificity of periplasmic disulfide oxidoreductases. *Journal of Biological Chemistry*, 279(13):12967–73.
- Hiniker, A. and Bardwell, J. C. A. (2003). Disulfide Bond Isomerization in Prokaryotes. *Biochemistry*, 42(5):1179–1185.
- Hiratsu, K., Amemura, M., Nashimoto, H., Shinagawa, H., and Makino, K. (1995). The *rpoE* gene of *Escherichia coli*, which encodes sigma E, is essential for bacterial growth at high temperature. *Journal of Bacteriology*, 177(10):2918–2922.
- Hirschel, B. J., Shen, V., and Schlessinger, D. (1980). Lactose operon transcription from wild-type and L8-UV5 lac promoters in *Escherichia coli* treated with chloramphenicol. *Journal of Bacteriology*, 143(3):1534–1537.
- Hoffmann, A., Bukau, B., and Kramer, G. (2010). Structure and function of the molecular chaperone Trigger Factor. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1803(6):650–661.
- Hoffmann, F. and Rinas, U. (2001). On-line estimation of the metabolic burden resulting from the synthesis of plasmid-encoded and heat-shock proteins by monitoring respiratory energy generation. *Biotechnology and Bioengineering*, 76(4):333–340.
- Hoffmann, F. and Rinas, U. (2004). Stress Induced by Recombinant Protein Production in *Escherichia coli*. In *Advances in Biochemical Engineering*, volume 89, pages 73–92. Springer Berlin Heidelberg.

- Holt, L. J., Basran, A., Jones, K., Chorlton, J., Jespers, L. S., Brewis, N. D., and Tomlinson, I. M. (2008). Anti-serum albumin domain antibodies for extending the half-lives of short lived drugs. *Protein Engineering Design and Selection*, 21(5):283–288.
- Hong, W., Jiao, W., Hu, J., Zhang, J., Liu, C., Fu, X., Shen, D., Xia, B., and Chang, Z. (2005). Periplasmic Protein HdeA Exhibits Chaperone-like Activity Exclusively within Stomach pH Range by Transforming into Disordered Conformation. *Journal of Biological Chemistry*, 280(29):27029–27034.
- Hong, W., Wu, Y. E., Fu, X., and Chang, Z. (2012). Chaperone-dependent mechanisms for acid resistance in enteric bacteria. *Trends in Microbiology*, 20(7):328 – 335.
- Horne, S. and Young, K. (1995). *Escherichia coli* and other species of the enterobacteriaceae encode a protein similar to the family of Mip-like FK506-binding proteins. *Archives of Microbiology*, 163(5):357–365.
- Hoshino, K., Eda, A., Kurokawa, Y., and Shimizu, N. (2002). Production of brain-derived neurotrophic factor in *Escherichia coli* by coexpression of Dsb proteins. *Bioscience, Biotechnology, and Biochemistry*, 66:344–50.
- Hu, X., O’Hara, L., White, S., Magner, E., Kane, M., and Gerard Wall, J. (2007). Optimisation of production of a domoic acid-binding scFv antibody fragment in *Escherichia coli* using molecular chaperones and functional immobilisation on a mesoporous silicate support. *Protein Expression and Purification*, 52(1):194–201.
- Huang, C. J., Lin, H., and Yang, X. (2012). Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journal of Industrial Microbiology and Biotechnology*, 39(3):383–399.
- Huber, D. (2015). Improving Soluble Periplasmic Antibody Fragment Production in *Escherichia coli* by Folding-Modulator Gene Co-Expression. Master’s thesis, Graz University of Technology.
- Huber-Wunderlich, M. and Glockshuber, R. (1998). A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Folding and Design*, 3(3):161–171.
- Humphreys, D. P., Carrington, B., Bowering, L. C., Ganesh, R., Sehdev, M., Smith, B. J., King, L. M., Reeks, D. G., Lawson, A., and Popplewell, A. G. (2002). A plasmid system for optimization of Fab’ production in *Escherichia coli*: importance of balance of heavy chain and light chain synthesis. *Protein Expression and Purification*, 26:309–320.

- Hunke, S. and Betton, J. M. (2003). Temperature effect on inclusion body formation and stress response in the periplasm of *Escherichia coli*. *Molecular Microbiology*, 50(5):1579–1589.
- Huston, J., Levinson, D., Mudgett-Hunter, M., Tai, M., Novotný, J., Margolies, M., Ridge, R., Brucoleri, R., Haber, E., and Crea, R. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 85(16):5879–5883.
- Ikai, A. (1980). Thermostability and Aliphatic Index of Globular Proteins. *The Journal of Biochemistry*, 88(6):1895–1898.
- Imburgio, D., Rong, M., Ma, K., and McAllister, W. T. (2000). Studies of Promoter Recognition and Start Site Selection by T7 RNA Polymerase Using a Comprehensive Collection of Promoter Variants. *Biochemistry*, 39(34):10419–10430.
- Itakura, K., Hirose, T., Crea, R., Riggs, A., Heyneker, H., Bolivar, F., and Boyer, H. (1977). Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science*, 198(4321):1056–1063. cited By 241.
- Ito, K. and Inaba, K. (2008). The disulfide bond formation (Dsb) system. *Current Opinion in Structural Biology*, 18(4):450 – 458. Membranes / Engineering and design.
- Ivancic, T., Jammik, P., and Stopar, D. (2013). Cold shock CspA and CspB protein production during periodic temperature cycling in *Escherichia coli*. *BMC Research Notes*, 6(1):248.
- Jain, A. and Jain, S. K. (2008). PEGylation: An Approach for Drug Delivery. A Review. *Critical Reviews in Therapeutic Drug Carrier Systems*, 25(5):403–447.
- Jain, R. K. (1990). Physiological barriers to delivery of monoclonal antibodies and other macromolecules in tumors. *Cancer Research*, 50(3 Supplement):814s–819s.
- Jalalirad, R. (2013). Production of antibody fragment (Fab) throughout *Escherichia coli* fed-batch fermentation process: changes in titre, location and form of product. *Electronic Journal of Biotechnology*.
- Janeway, C. A., Travers, P., and Walport, M. (2001). *Immunobiology*. Garland Publishing, 5th edition edition.

- Jarchow, S., Lueck, C., Goerg, A., and Skerra, A. (2008). Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. *Proteomics*, 8(23-24):4987–4994.
- Jeong, K. J., Jang, S. H., and Velmurugan, N. (2011). Recombinant antibodies: Engineering and production in yeast and bacterial hosts. *Biotechnology Journal*, 6(1):16–27.
- Jeong, K. J. and Lee, S. Y. (2000). Secretory production of human leptin in *Escherichia coli*. *Biotechnology and Bioengineering*, 67(4):398–407.
- Jin, D. J., Cagliero, C., and Zhou, Y. N. (2012). Growth rate regulation in *Escherichia coli*. *FEMS Microbiology Reviews*, 36(2):269–287.
- Joly, J., Leung, W., and Swartz, J. (1998). Overexpression of *Escherichia coli* oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proceedings of the National Academy of Sciences of the United States of America*, 95(6):2773–2777.
- Joly, J. C. and Swartz, J. R. (1997). *In Vitro* and *in Vivo* Redox States of the *Escherichia coli* Periplasmic Oxidoreductases DsbA and DsbC. *Biochemistry*, 36(33):10067–10072. PMID: 9254601.
- Kadokura, H. and Beckwith, J. (2002). Four cysteines of the membrane protein DsbB act in concert to oxidize its substrate DsbA. *The EMBO Journal*, 21(10):2354–2363.
- Kadokura, H. and Beckwith, J. (2009). Detecting Folding Intermediates of a Protein as It Passes through the Bacterial Translocation Channel. *Cell*, 138(6):1164–1173.
- Kadokura, H. and Beckwith, J. (2010). Mechanisms of Oxidative Protein Folding in the Bacterial Cell Envelope. *Antioxidants and Redox Signaling*, 13(8):1231–1246.
- Kadokura, H., Katzen, F., and Beckwith, J. (2003). Protein disulfide bond formation in prokaryotes. *Annual Review of Biochemistry*, 72(1):111–135.
- Kadokura, H., Tian, H., Zander, T., Bardwell, J. C. A., and Beckwith, J. (2004). Snapshots of DsbA in Action: Detection of Proteins in the Process of Oxidative Folding. *Science*, 303(5657):534–537.
- Kamionka, M. (2011). Engineering of Therapeutic Proteins Production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, 12:268–274.
- Kane, J. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology*, 6:494–500.

- Kang, T. S. and Kini, R. M. (2009). Structural determinants of protein folding. *Cellular and Molecular Life Sciences*, 66(14):2341–2361.
- Kern, R., Malki, A., Abdallah, J., Tagourt, J., and Richarme, G. (2007). *Escherichia coli* HdeB is an acid stress chaperone. *Journal of Bacteriology*, 189(2):603–10.
- Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. c., Stines, A. P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., and Hartl, F. (2005). Proteome-wide Analysis of Chaperonin-Dependent Protein Folding in *Escherichia coli*. *Cell*, 122(2):209–220.
- Kikuchi, Y., Yoda, K., Yamasaki, M., and Tamura, G. (1981). The nucleotide sequence of the promoter and the amino-terminal region of alkaline phosphatase structural gene (*phoA*) of *Escherichia coli*. *Nucleic Acids Research*, 9(21):5671–5678.
- Kim, J., Kim, Y.-G., and Lee, G. (2012). CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied Microbiology and Biotechnology*, 93(3):917–930.
- Kim, Y. H., Han, K., Lee, K., and Lee, J. (2005). Proteome response of *Escherichia coli* fed-batch culture to temperature downshift. *Applied Microbiology and Biotechnology*, 68(6):786–793.
- Kirsch, M., Zaman, M., Meier, D., Dübel, S., and Hust, M. (2005). Parameters affecting the display of antibodies on phage. *Journal of Immunological Methods*, 301:173–185.
- Kleerebezem, M., Heutink, M., and Tommassen, J. (1995). Characterization of an *Escherichia coli* *rotA* mutant, affected in periplasmic peptidyl-prolyl *cis/trans* isomerase. *Molecular Microbiology*, 18(2):313–320.
- Klösch, B. (2016). Enhancement of Soluble Periplasmic Antibody Fragment Production in *Escherichia coli* and Improvement of Upscaling Experiments. Master’s thesis, University of Natural Resources and Life Sciences, Vienna.
- Kohda, J., Endo, Y., Okumura, N., Kurokawa, Y., Nishihara, K., Yanagi, H., Yura, T., Fukuda, H., and Kondo, A. (2002). Improvement of productivity of active form of glutamate racemase in *Escherichia coli* by coexpression of folding accessory proteins. *Biochemical Engineering Journal*, 10(1):39 – 45.
- Kolaj, O., Spada, S., Robin, S., and Wall, J. (2009). Use of folding modulators to improve heterologous protein production in *Escherichia coli*. *Microbial Cell Factories*, 8(1):9.

- Kong, B. and Guo, G. L. (2014). Soluble Expression of Disulfide Bond Containing Proteins FGF15 and FGF19 in the Cytoplasm of *Escherichia coli*. *PLoS ONE*, 9(1):e85890.
- Kosuri, P., Alegre-Cebollada, J., Feng, J., Kaplan, A., Ingles-Prieto, A., Badilla, C., Stockwell, B., Sanchez-Ruiz, J., Holmgren, A., and Fernandez, J. (2012). Protein Folding Drives Disulfide Formation. *Cell*, 151(4):794–806.
- Kramer, W., Elmecker, G., Weik, R., Mattanovich, D., and Bayer, K. (1996). Kinetic studies for the optimization of recombinant protein formation. *Annals of the New York Academy of Sciences*, 782(1):323–333.
- Kuhlman, T. E. and Cox, E. C. (2010). Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Research*, 38(6):e92–e92.
- Kurokawa, Y., Yanagi, H., and Yura, T. (2000). Overexpression of Protein Disulfide Isomerase DsbC Stabilizes Multiple-Disulfide-Bonded Recombinant Protein Produced and Transported to the Periplasm in *Escherichia coli*. *Applied and Environmental Microbiology*, 66(9):3960–3965.
- Kurokawa, Y., Yanagi, H., and Yura, T. (2001). Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in *Escherichia coli*. *Journal of Biological Chemistry*.
- Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105 – 132.
- Lamiet, A. A. and Plueckthun, A. (1989). The precursor of beta-lactamase: purification, properties and folding kinetics. *The EMBO Journal*, 8(5):1469–1477.
- Lanzer, M. and Bujard, H. (1988). Promoters largely determine the efficiency of repressor action. *Proceedings of the National Academy of Sciences*, 85(23):8973–8977.
- Larson, S. M., Carrasquillo, J. A., Krohn, K. A., Brown, J. P., McGuffin, R. W., Ferens, J. M., Graham, M. M., Hill, L. D., Beaumier, P. L., and Hellstrom, K. E. (1983). Localization of ¹³¹I-labeled p97-specific Fab fragments in human melanoma as a basis for radiotherapy. *Journal of Clinical Investigation*, 72(6):2101–2114.
- Lazar, S. W. and Kolter, R. (1996). SurA assists the folding of *Escherichia coli* outer membrane proteins. *Journal of Bacteriology*, 178(6):1770–3. Lazar, S W Kolter, R Journal Article Research Support, U.S. Gov’t, Non-P.H.S. Research Support, U.S. Gov’t, P.H.S. United states J Bacteriol. 1996 Mar;178(6):1770-3.

- Lebediker, M. and Danieli, T. (2014). Production of prone-to-aggregate proteins. *FEBS Letters*, 588(2):236 – 246. Protein Engineering.
- Lee, D.-H., Kim, M.-D., Lee, W.-H., Kweon, D.-H., and Seo, J.-H. (2004). Consortium of fold-catalyzing proteins increases soluble expression of cyclohexanone monooxygenase in recombinant *Escherichia coli*. *Applied Microbiology and Biotechnology*, 63(5):549–552.
- Lee, G., Cooney, D., Middelberg, A., and Choe, W. (2006). The economics of inclusion body processing. *Bioprocess and Biosystems Engineering*, 29(2):73–90.
- Lee, J., Saraswat, V., Koh, I., Song, K.-B., Park, Y.-H., and Rhee, S.-K. (2001). Secretory production of *Arthrobacter levan* fructotransferase from recombinant *Escherichia coli*. *FEMS Microbiology Letters*, 195(2):127–132.
- Lee, J. Y. and Bang, D. (2010). Challenges in the chemical synthesis of average sized proteins: Sequential vs. convergent ligation of multiple peptide fragments. *Peptide Science*, 94(4):441–447.
- Lee, Y. J., Kim, H. S., Ryu, A. J., and Jeong, K. J. (2013). Enhanced production of full-length immunoglobulin G via the signal recognition particle (SRP)-dependent pathway in *Escherichia coli*. *Journal of Biotechnology*, 165(2):102 – 108.
- Leichert, L. I. and Jakob, U. (2004). Protein thiol modifications visualized in vivo. *PLoS Biol*, 2(11):e333.
- Lennon, C. W., Thamsen, M., Friman, E. T., Cacciaglia, A., Sachsenhauser, V., Sorgenfrei, F. A., Wasik, M. A., and Bardwell, J. C. (2015). Folding Optimization *In Vivo* Uncovers New Chaperones. *Journal of Molecular Biology*, 427(18):2983–2994.
- Lenz, G., Doron-Faigenboim, A., Ron, E. Z., Tuller, T., and Gophna, U. (2011). Sequence Features of *E. coli* mRNAs Affect Their Degradation. *PLoS ONE*, 6(12):e28544.
- Levy, R., Weiss, R., Chen, G., Iverson, B. L., and Georgiou, G. (2001). Production of Correctly Folded Fab Antibody Fragment in the Cytoplasm of *Escherichia coli* *trxB* *gor* Mutants via the Coexpression of Molecular Chaperones. *Protein Expression and Purification*, 23(2):338–347.
- Li, J. and Zhu, Z. (2010). Research and development of next generation of antibody-based therapeutics. *Acta Pharmacologica Sinica*, 31(9):1198–1207.
- Lichtenstein, C. and Brenner, S. (1982). Unique insertion site of Tn7 in the *E. coli* chromosome. *Nature*, 297(5867):601–603.

- Lin, B., Renshaw, M. W., Autote, K., Smith, L. M., Calveley, P., Bowdish, K. S., and Frederickson, S. (2008). A step-wise approach significantly enhances protein yield of a rationally-designed agonist antibody fragment in *E. coli*. *Protein Expression and Purification*, 59(1):55–63.
- Lin, W.-J., Huang, S.-W., and Chou, C. P. (2001a). DegP-coexpression minimizes inclusion-body formation upon overproduction of recombinant penicillin acylase in *Escherichia coli*. *Biotechnology and Bioengineering*, 73(6):484–492.
- Lin, Y.-H., Fang, W.-L., Lin, W.-J., Huang, S.-W., and Chou, C. (2001b). Improving production of penicillin acylase in *Escherichia coli* via efficient DegP-mediated processing of precursors in periplasm. *Process Biochemistry*, 37(1):23–30.
- Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J., and Taddei, F. (2007). Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8):3076–3081.
- Lipinska, B., Zylicz, M., and Georgopoulos, C. (1990). The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *Journal of Bacteriology*, 172(4):1791–1797.
- Liu, J. and Walsh, C. (1990). Peptidyl-prolyl cis-trans-isomerase from *Escherichia coli*: a periplasmic homolog of cyclophilin that is not inhibited by cyclosporin A. *Proceedings of the National Academy of Sciences*, 87(11):4028–4032.
- Lobstein, J., Emrich, C., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012). SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories*, 11(1):56.
- Lu, K. P., Finn, G., Lee, T. H., and Nicholson, L. K. (2007). Prolyl *cis-trans* isomerization as a molecular timer. *Nature Chemical Biology*, 3(10):619–629.
- Luirink, J., Yu, Z., Wagner, S., and de Gier, J. (2012). Biogenesis of inner membrane proteins in *Escherichia coli*. *Biochimica et Biophysica Acta*, 1817(6):965–976.
- Macdonald, L. E., Durbin, R. K., Dunn, J. J., and McAllister, W. T. (1994). Characterization of Two Types of Termination Signal for Bacteriophage T7 RNA Polymerase. *Journal of Molecular Biology*, 238(2):145 – 158.
- Madigan, M. and Martinko, J. (2009). *Brock Mikrobiologie*. Pearson Education Deutschland GmbH, 11th edition.

- Maeng, B., Nam, D., and Kim, Y. (2011). Coexpression of molecular chaperones to enhance functional expression of anti-BNP scFv in the cytoplasm of *Escherichia coli* for the detection of B-type natriuretic peptide. *World Journal of Microbiology and Biotechnology*, 27(6):1391–1398.
- Mairhofer, J., Cserjan-Puschmann, M., Striedner, G., Nöbauer, K., Razzazi-Fazeli, E., and Grabherr, R. (2010). Marker-free plasmids for gene therapeutic applications - Lack of antibiotic resistance gene substantially improves the manufacturing process. *Journal of Biotechnology*, 146(3):130–137.
- Mairhofer, J., Scharl, T., Marisch, K., Cserjan-Puschmann, M., and Striedner, G. (2013). Comparative transcription profiling and in depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. *Applied and Environmental Microbiology*, 79(12):3802–3812.
- Mairhofer, J., Wittwer, A., Cserjan, M., and Striedner, G. (2014). Preventing T7 RNA Polymerase Read-through Transcription - A Synthetic Termination Signal Capable of Improving Bioprocess Stability. *ACS Synthetic Biology*, 4(3):265–273.
- Maisonneuve, E., Ezraty, B., and Dukan, S. (2008). Protein Aggregates: an Aging Factor Involved in Cell Death. *Journal of Bacteriology*, 190(18):6070–6075.
- Mao, L., Stathopoulos, P. B., Ikura, M., and Inouye, M. (2010). Secretion of human superoxide dismutase in *Escherichia coli* using the condensed single-protein-production system. *Protein Science*, 19(12):2330–2335.
- Maresca, M., Erler, A., Fu, J., Friedrich, A., Zhang, Y., and Stewart, A. (2010). Single-stranded heteroduplex intermediates in Δ Red homologous recombination. *BMC Molecular Biology*, 11:54.
- Marisch, K., Bayer, K., Cserjan-Puschmann, M., Luchner, M., and Striedner, G. (2013a). Evaluation of three industrial *Escherichia coli* strains in fed-batch cultivations during high-level SOD protein production. *Microbial Cell Factories*, 12(1):58.
- Marisch, K., Bayer, K., Scharl, T., Mairhofer, J., Krempl, P. M., Hummel, K., Razzazi-Fazeli, E., and Striedner, G. (2013b). A Comparative Analysis of Industrial *Escherichia coli* K-12 and B Strains in High-Glucose Batch Cultivations on Process-, Transcriptome- and Proteome Level. *PLoS ONE*, 8(8):e70516.
- Martinez-Alonso, M., Garcia-Fruitos, E., Ferrer-Miralles, N., Rinas, U., and Villaverde, A. (2010). Side effects of chaperone gene co-expression in recombinant protein production. *Microbial Cell Factories*, 9(1):64.

- Martinez-Alonso, M., Gonzalez-Montalban, N., Garcia-Fruitos, E., and Villaverde, A. (2008). The Functional Quality of Soluble Recombinant Polypeptides Produced in *Escherichia coli* Is Defined by a Wide Conformational Spectrum. *Applied and Environmental Microbiology*, 74(23):7431–7433.
- Martinez-Gomez, K., Flores, N., Castaneda, H. M., Martinez-Batallar, G., Hernandez-Chavez, G., Ramirez, O. T., Gosset, G., Encarnacion, S., and Bolivar, F. (2012). New insights into *Escherichia coli* metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. *Microbial Cell Factories*, 11:46–46.
- Maskos, K., Huber-Wunderlich, M., and Glockshuber, R. (2003). DsbA and DsbC-catalyzed Oxidative Folding of Proteins with Complex Disulfide Bridge Patterns *In Vitro* and *In Vivo*. *Journal of Molecular Biology*, 325(3):495 – 513.
- Matern, Y., Barion, B., and Behrens-Kneip, S. (2010). PpiD is a player in the network of periplasmic chaperones in *Escherichia coli*. *BMC Microbiology*, 10(1):251.
- Mavrangelos, C., Thiel, M., Adamson, P. J., Millard, D. J., Nobbs, S., Zola, H., and Nicholson, I. C. (2001). Increased Yield and Activity of Soluble Single-Chain Antibody Fragments by Combining High-Level Expression and the Skp Periplasmic Chaperonin. *Protein Expression and Purification*, 23(2):289–295.
- Maynard, J., Adams, E. J., Krogsgaard, M., Petersson, K., Liu, C. W., and Garcia, K. C. (2005). High-level bacterial secretion of single-chain $\alpha\beta$ T-cell receptors. *Journal of Immunological Methods*, 306(1-2):51–67.
- McCarthy, A. A., Haebel, P. W., Torronen, A., Rybin, V., Baker, E. N., and Metcalf, P. (2000). Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nature Structural and Molecular Biology*, 7(3):196–199.
- Mecenas, J., Rouviere, P. E., Erickson, J. W., Donohue, T. J., and Gross, C. A. (1993). The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes & Development*, 7(12b):2618–2628.
- Merdanovic, M., Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011). Protein Quality Control in the Bacterial Periplasm. *Annual Review of Microbiology*, 65(1):149–168.
- Mergulhão, F., Summers, D., and Monteiro, G. (2005). Recombinant protein secretion in *Escherichia coli*. *Biotechnol Advances*, 23(3):177–202.

- Mergulhão, F. J. M., Monteiro, G. A., Larsson, G., Sandén, A. M., Farewell, A., Nystrom, T., Cabral, J. M. S., and Taipa, M. A. (2003). Medium and copy number effects on the secretion of human proinsulin in *Escherichia coli* using the universal stress promoters *uspA* and *uspB*. *Applied Microbiology and Biotechnology*, 61(5):495–501.
- Messens, J. and Collet, J.-F. (2006). Pathways of disulfide bond formation in *Escherichia coli*. *The International Journal of Biochemistry and Cell Biology*, 38:1050–1062.
- Meyer, H.-P. and Schmidhalter, D. R. (2012). Microbial expression systems and manufacturing from a market and economic perspective. In Agbo, E., editor, *Innovations in Biotechnology*, chapter 10, pages 211–250. InTech.
- Miot, M. and Betton, J.-M. (2004). Protein quality control in the bacterial periplasm. *Microbial Cell Factories*, 3:4–4.
- Miroux, B. and Walker, J. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high level. *Journal of Molecular Biology*, 260:289–298.
- Missiakas, D., Betton, J.-M., and Raina, S. (1996). New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Molecular Microbiology*, 21(4):871–884.
- Missiakas, D., Georgopoulos, C., and Raina, S. (1994). The *Escherichia coli dsbC* (*xprA*) gene encodes a periplasmic protein involved in disulfide bond formation. *The EMBO Journal*, 13(8):2013–2020.
- Missiakas, D. and Raina, S. (1998). The extracytoplasmic function sigma factors: role and regulation. *Molecular Microbiology*, 28(6):1059–1066.
- Moffatt, B. A., Dunn, J. J., and Studier, F. (1984). Nucleotide sequence of the gene for bacteriophage T7 RNA polymerase. *Journal of Molecular Biology*, 173(2):265 – 269.
- Muller-Hill, B., Crapo, L., and Gilbert, W. (1968). Mutants that make more lac repressor. *Proceedings of the National Academy of Sciences of the United States of America*, 59(4):1259–1264.
- Murphy, K., Campellone, K., and Poteete, A. (2000). PCR-mediated gene replacement in *Escherichia coli*. *Gene*, 246:321–330.

- Murphy, K. C. (2012). Chapter 8 - Phage Recombinases and Their Applications. In *Advances in Virus Research Bacteriophages, Part B*, volume 83, pages 367–414. Academic Press.
- Nakamoto, H. and Bardwell, J. C. (2004). Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694:111–119.
- Nannenga, B. and Baneyx, F. (2011). Enhanced expression of membrane proteins in *E. coli* with a P_{BAD} promoter mutant: synergies with chaperone pathway engineering strategies. *Microbial Cell Factories*, 10:105.
- Nasreen, A., Vogt, M., Kim, H. J., Eichinger, A., and Skerra, A. (2005). Solubility engineering and crystallization of human apolipoprotein D. *Protein Science*, 15(1):190–199.
- Natale, P., Brüser, T., and Driessen, A. J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane - Distinct translocases and mechanisms. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(9):1735–1756.
- Nelson, A. L. and Reichert, J. M. (2009). Development trends for therapeutic antibody fragments. *Nature Biotechnology*, 27(4):331–337.
- Nemecek, S., Marisch, K., Juric, R., and Bayer, K. (2008). Design of transcriptional fusions of stress sensitive promoters and GFP to monitor the overburden of *Escherichia coli* hosts during recombinant protein production. *Bioprocess and Biosystems Engineering*, 31(1):47–53.
- Newbury, S. F., Smith, N. H., and Higgins, C. F. (1987). Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell*, 51(6):1131 – 1143.
- Nilson, B. H., Lögdberg, L., Kastern, W., Björck, L., and Åkerström, B. (1993). Purification of antibodies using protein L-binding framework structures in the light chain variable domain. *Journal of Immunological Methods*, 164(1):33 – 40.
- Nilson, B. H., Solomon, A., Björck, L., and Åkerström, B. (1992). Protein L from *Peptostreptococcus magnus* binds to the kappa light chain variable domain. *Journal of Biological Chemistry*, 267(4):2234–2239.
- Nilvebrant, J., Kuku, G., Bjorkelund, H., and Nestor, M. (2012). Selection and *in vitro* characterization of human CD44v6-binding antibody fragments. *Biotechnology and Applied Biochemistry*, 59(5):367–380.

- Oliveira, P. H. and Mairhofer, J. (2013). Marker-free plasmids for biotechnological applications - implications and perspectives. *Trends in Biotechnology*, 31(9):539 – 547.
- Oliver, J. (2005). The viable but nonculturable state in bacteria. *The Journal of Microbiology*, 43(1225-8873 (Print)):93–100.
- O'Reilly, A., Cole, A., Lopes, J., Lampert, A., and Wallace, B. (2014). Chaperone-mediated native folding of a β -scorpion toxin in the periplasm of *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(1):10 – 15.
- Osterman, I. A., Evfratov, S. A., Sergiev, P. V., and Dontsova, O. A. (2013). Comparison of mRNA features affecting translation initiation and reinitiation. *Nucleic Acids Research*, 41(1):474–486.
- Outchkourov, N. S., Roeffen, W., Kaan, A., Jansen, J., Luty, A., Schuiffel, D., van Gemert, G. J., van de Vegte-Bolmer, M., Sauerwein, R. W., and Stunnenberg, H. G. (2008). Correctly folded Pfs48/45 protein of *Plasmodium falciparum* elicits malaria transmission-blocking immunity in mice. *Proceedings of the National Academy of Sciences*, 105(11):4301–4305.
- Overton, T. W. (2014). Recombinant protein production in bacterial hosts. *Drug Discovery Today*, 19(5):590–601.
- Ow, D., Lim, D., Morin Nissom, P., Camattari, A., and Wong, V. (2010). Co-expression of Skp and FkpA chaperones improves cell viability and alters the global expression of stress response genes during scFvD1.3 production. *Microbial Cell Factories*, 9(1):22–.
- Ow, D. S.-W., Lee, D. Y., Yap, M. G.-S., and Oh, S. K.-W. (2009). Identification of cellular objective for elucidating the physiological state of plasmid-bearing *Escherichia coli* using genome-scale in silico analysis. *Biotechnology Progress*, 25(1):61–67.
- Ow, D. S.-W., Nissom, P. M., Philp, R., Oh, S. K.-W., and Yap, M. G.-S. (2006). Global transcriptional analysis of metabolic burden due to plasmid maintenance in *Escherichia coli* DH5 α during batch fermentation. *Enzyme and Microbial Technology*, 39(3):391–398.
- Padiolleau-Lefevre, S., Debat, H., Phichith, D., Thomas, D., Friboulet, A., and Avalle, B. (2006). Expression of a functional scFv fragment of an anti-idiotypic antibody with a β -lactam hydrolytic activity. *Immunology Letters*, 103(1):39–44.
- Pan, J. L. and Bardwell, J. C. (2006). The origami of thioredoxin-like folds. *Protein Science*, 15(10):2217–2227.

- Pan, K.-L., Hsiao, H.-C., Weng, C.-L., Wu, M.-S., and Chou, C. P. (2003). Roles of DegP in Prevention of Protein Misfolding in the Periplasm upon Overexpression of Penicillin Acylase in *Escherichia coli*. *Journal of Bacteriology*, 185(10):3020–3030.
- Papanikou, E., Karamanou, S., and Economou, A. (2007). Bacterial protein secretion through the translocase nanomachine. *Nature Reviews Microbiology*, 5(11):839–851.
- Pasini, M., Fernandez-Castane, A., Jaramillo, A., de Mas, C., Caminal, G., and Ferrer, P. (2016). Using promoter libraries to reduce metabolic burden due to plasmid-encoded proteins in recombinant *Escherichia coli*. *New Biotechnology*, 33(1):78 – 90.
- Patkar, A., Vijayasankaran, N., Urry, D. W., and Srienc, F. (2002). Flow cytometry as a useful tool for process development: rapid evaluation of expression systems. *Journal of Biotechnology*, 93(3):217 – 229.
- Perez-Perez, J., Martinezcaja, C., Barbero, J., and Gutierrez, J. (1995). DnaK/DnaJ Supplementation Improves the Periplasmic Production of Human Granulocyte-Colony Stimulating Factor in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 210(2):524–529.
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10):785–786.
- Phue, J.-N., Noronha, S. B., Hattacharyya, R., Wolfe, A. J., and Shiloach, J. (2005). Glucose metabolism at high density growth of *E. coli* B and *E. coli* K: Differences in metabolic pathways are responsible for efficient glucose utilization in *E. coli* B as determined by microarrays and Northern blot analyses. *Biotechnology and Bioengineering*, 90(7):805–820.
- Pliyev, B. and Gurvits, B. (1999). Peptidyl-prolyl cis-trans isomerases: structure and functions. *Biochemistry (Moscow)*, 64:738–51.
- Plotkin, J. B. and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 12(1):32–42.
- Popov, M., Petrov, S., Nacheva, G., Ivanov, I., and Reichl, U. (2011). Effects of a recombinant gene expression on ColE1-like plasmid segregation in *Escherichia coli*. *BMC Biotechnology*, 11(1):18.
- Postema, E. J., Börjesson, P. K., Buijs, W. C., Roos, J. C., Marres, H. A., Boerman, O. C., de Bree, R., Lang, M., Munzert, G., van Dongen, G. A., and Oyen, W. J. (2003). Dosimetric Analysis of Radioimmunotherapy with ¹⁸⁶Re-Labeled Bivatuzumab

- in Patients with Head and Neck Cancer. *Journal of Nuclear Medicine*, 44(10):1690–1699.
- Prasad, I., Young, B., and Schaefer, S. (1973). Genetic Determination of the Constitutive Biosynthesis of Phospho- β -Glucosidase A in *Escherichia coli* K-12. *Journal of Bacteriology*, 114(3):909–915.
- Price, N. L. and Raivio, T. L. (2008). Characterization of the Cpx Regulon in *Escherichia coli* Strain MC4100. *Journal of Bacteriology*, 191(6):1798–1815.
- Prinz, W. A., Åslund, F., Holmgren, A., and Beckwith, J. (1997). The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the *Escherichia coli* Cytoplasm. *Journal of Biological Chemistry*, 272(25):15661–15667.
- Puigbo, P., Guzman, E., Romeu, A., and Garcia-Vallve, S. (2007). OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Research*, 35(Web Server issue):W126–W131.
- Qiu, J., Swartz, J. R., and Georgiou, G. (1998). Expression of Active Human Tissue-Type Plasminogen Activator in *Escherichia coli*. *Applied and Environmental Microbiology*, 64(12):4891–4896.
- Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K. M., Pfizenmaier, J., Shi, R., Hofmann, S., Foit, L., Ren, G., Jakob, U., Xu, Z., Cygler, M., and Bardwell, J. C. A. (2011). Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nature Structural and Molecular Biology*, 18(3):262–269.
- Raghavan, M. and Bjorkman, P. J. (1996). Fc Receptors and Their Interactions with Immunoglobulins. *Annual Review of Cell and Developmental Biology*, 12(1):181–220.
- Ramm, K. and Plückthun, A. (2000). The Periplasmic *Escherichia coli* Peptidylprolyl *cis/trans*-isomerase FkpA: II. Isomerase-independent Chaperone Activity *in Vitro*. *Journal of Biological Chemistry*, 275(22):17106–17113.
- Reichert, J. M. (2012). Marketed therapeutic antibodies compendium. *mAbs*, 4(1942-0862):413–415.
- Riechelmann, H., Sauter, A., Golze, W., Hanft, G., Schroen, C., Hoermann, K., Erhardt, T., and Gronau, S. (2008). Phase I trial with the CD44v6-targeting immunoconjugate bivatuzumab mertansine in head and neck squamous cell carcinoma. *Oral Oncology*, 44(9):823–829.

- Rietsch, A., Bessette, P., Georgiou, G., and Beckwith, J. (1997). Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *Journal of Bacteriology*, 179(21):6602–6608.
- Ritz, D. and Beckwith, J. (2001). Roles of Thiol-Redox Pathways in Bacteria. *Annual Review of Microbiology*, 55(1):21–48.
- Rizzitello, A. E., Harper, J. R., and Silhavy, T. J. (2001). Genetic Evidence for Parallel Pathways of Chaperone Activity in the Periplasm of *Escherichia coli*. *Journal of Bacteriology*, 183(23):6794–6800.
- Rodrigo, G., Gruvegard, M., and Van Alstine, J. M. (2015). Antibody Fragments and Their Purification by Protein L Affinity Chromatography. *Antibodies*, 4(3):259.
- Rodriguez-Carmona, E., Cano-Garrido, O., Dragosits, M., Maurer, M., Mader, A., Kunert, R., Mattanovich, D., Villaverde, A., and Vazquez, F. (2012). Recombinant Fab expression and secretion in *Escherichia coli* continuous culture at medium cell densities: Influence of temperature. *Process Biochemistry*, 47(3):446–452.
- Rosano, G. L. and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5(172).
- Roszak, D. B. and Colwell, R. R. (1987). Survival strategies of bacteria in the natural environment. *Microbiological Reviews*, 51(3):365–379.
- Rouviere, P. E., De Las Penas, A., Mecsas, J., Lu, C. Z., Rudd, K. E., and Gross, C. A. (1995). rpoE, the gene encoding the second heat-shock sigma factor, sigma E, in *Escherichia coli*. *The EMBO Journal*, 14(5):1032–1042.
- Rouviere, P. E. and Gross, C. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes & Development*, 10(24):3170–3182.
- Rozkov, A., Avignone-Rossa, C., Ertl, P., Jones, P., O’Kennedy, R., Smith, J., Dale, J., and Bushell, M. (2004). Characterization of the metabolic burden on *Escherichia coli* DH1 cells imposed by the presence of a plasmid containing a gene therapy sequence. *Biotechnology and Bioengineering*, 88(7):909–915.
- Ruiz, N., Kahne, D., and Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nature Reviews Microbiology*, 4(1):57–66.
- Rusch, S. L. and Kendall, D. A. (2007). Interactions That Drive Sec-Dependent Bacterial Protein Transport. *Biochemistry*, 46(34):9665–9673.

- Sabri, S., Steen, J., Bongers, M., Nielsen, L., and Vickers, C. (2013). Knock-in/Knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the *Escherichia coli* chromosome at well-characterised loci. *Microbial Cell Factories*, 12(1):60–.
- Sahdev, S., Khattar, S., and Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*, 307(1-2):249–264.
- Saida, F., Uzan, M., Odaert, B., and Bontems, F. (2006). Expression of highly toxic genes in *E. coli*: special strategies and genetic tools. *Current Protein & Peptide Science*, 7(1):47–56.
- Salinas, G., Pellizza, L., Margenat, M., Flo, M., and Fernandez, C. (2011). Tuned *Escherichia coli* as a host for the expression of disulfide-rich proteins. *Biotechnology Journal*, 6(6):686–699.
- Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, 27(10):946–950.
- Sambrook, J. and Russell, D. (2001). *Molecular Cloning: A Laboratory Manual/Third Edition*. Nature Publishing Group.
- Samuelson, J. (2011a). Recent Developments in Difficult Protein Expression: A Guide to *E. coli* Strains, Promoters, and Relevant Host Mutations. In Evans, J. and Xu, M. Q., editors, *Methods in Molecular Biology*, volume 705, pages 195–209. Humana Press.
- Samuelson, J. C. (2011b). Expression Systems. In Skaja Robinson, A., editor, *Production of Membrane Proteins: Strategies for Expression and Isolation*, number 1 in Production of Membrane Proteins: Strategies for Expression and Isolation, pages 11–36. Wiley-VCH Verlag GmbH & Co. KGaA.
- Sandee, D., Tungpradabkul, S., Kurokawa, Y., Fukui, K., and Takagi, M. (2005). Combination of Dsb coexpression and an addition of sorbitol markedly enhanced soluble expression of single-chain Fv in *Escherichia coli*. *Biotechnology and Bioengineering*, 91(4):418–424.
- Saul, F., Arié, J.-P., le Normand, B. V., Kahn, R., Betton, J.-M., and Bentley, G. (2004). Structural and Functional Studies of FkpA from *Escherichia coli*, a *cis/trans* Peptidyl-prolyl Isomerase with Chaperone Activity. *Journal of Molecular Biology*, 335(2):595 – 608.

- Sawitzke, J., Bubunenko, M., Thomason, L., Li, X., Costantino, N., and Court, D. (2013). Recombineering: A modern approach to genetic engineering. In in Chief: Stanley Maloy, E. and Kelly, H., editors, *Brenner's Encyclopedia of Genetics (Second Edition)*, pages 109–112. Academic Press, San Diego.
- Schaefer, J. V. and Plueckthun, A. (2010). Improving Expression of scFv Fragments by Co-expression of Periplasmic Chaperones. In Kontermann, R. and Dübel, S., editors, *Antibody Engineering*, pages 345–361. Springer Berlin Heidelberg.
- Schafer, U., Beck, K., and Muller, M. (1999). Skp, a Molecular Chaperone of Gram-negative Bacteria, Is Required for the Formation of Soluble Periplasmic Intermediates of Outer Membrane Proteins. *Journal of Biological Chemistry*, 274(35):24567–24574.
- Schierle, C., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. (2003). The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway. *Journal of Bacteriology*, 185(19):5706–5713.
- Schlapschy, M., Dommel, M., Hadian, K., Fogarasi, M., Korndorfer, I., and A., S. (2005). The periplasmic *E. coli* chaperone Skp is a trimer in solution: biophysical and preliminary crystallographic characterization. *Biological Chemistry*, 385:137.
- Schlapschy, M., Grimm, S., and Skerra, A. (2006). A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli*. *Protein Engineering Design and Selection*, 19(8):385–390.
- Schlapschy, M. and Skerra, A. (2011). Periplasmic Chaperones Used to Enhance Functional Secretion of Proteins in *E. coli*. In Evans, J. and Xu, M. Q., editors, *Methods in Molecular Biology*, volume 705, pages 211–224. Humana Press.
- Schlegel, S., Rujas, E., Ytterberg, A. J., Zubarev, R., Luirink, J., and de Gier, J. W. (2013). Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microbial Cell Factories*, 12(1):24–.
- Schuller, A. (2015). Improving Soluble Periplasmic Fab Production in *Escherichia coli* by Fine-tuned Folding-Modulator Co-Production. Master's thesis, University of Natural Resources and Applied Life Sciences, Vienna, Austria.
- Schwalm, J., Mahoney, T. F., Soltes, G. R., and Silhavy, T. J. (2013). Role for Skp in LptD Assembly in *Escherichia coli*. *Journal of Bacteriology*, 195(16):3734–3742.

- Selzer, G., Som, T., Itoh, T., and Tomizawa, J. (1983). The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. *Cell*, 32(1):119 – 129.
- Seo, S. W., Yang, J.-S., Kim, I., Yang, J., Min, B. E., Kim, S., and Jung, G. Y. (2013). Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metabolic Engineering*, 15(0):67 – 74.
- Sharan, S. K., Thomason, L. C., Kuznetsov, S. G., and Court, D. (2009). Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protocols*, 4(2):206–223.
- Shiloach, J. and Fass, R. (2005). Growing *E. coli* to high cell density - A historical perspective on method development. *Biotechnology Advances*, 23(5):345 – 357.
- Shiloach, J., Kaufman, J., Guillard, A. S., and Fass, R. (1996). Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by *Escherichia coli* BL21 (λ DE3) and *Escherichia coli* JM109. *Biotechnology and Bioengineering*, 49(4):421–428.
- Silva, F., Queiroz, J. A., and Domingues, F. C. (2012). Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnology Advances*, 30(3):691–708.
- Silverstone, A. E., Arditti, R. R., and Magasanik, B. (1970). Catabolite-insensitive revertants of lac promoter mutants. *Proceedings of the National Academy of Sciences of the United States of America*, 66(3):773–779.
- Simmons, L. and Yansura, D. (1996). Translational level is a critical factor for the secretion of heterologous proteins in *Escherichia coli*. *Nature Biotechnology*, 14(5):629–634.
- Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., and Panda, A. K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, 14:41.
- Singh, S. M. and Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*, 99(4):303 – 310.
- Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes and Development*, 21(19):2473–2484.

- Skorko-Glonek, J., Sobiecka-Szkatula, A., Narkiewicz, J., and Lipinska, B. (2008). The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures. *Microbiology*, 154(12):3649–3658.
- Slos, P., Speck, D., Accart, N., Kolbe, H., Schubnel, D., Bouchon, B., Bischoff, R., and Kieny, M. (1994). Recombinant Cholera Toxin B-Subunit in *Escherichia coli*: High-Level Secretion, Purification, and Characterization. *Protein Expression and Purification*, 5(5):518 – 526.
- Smith, S. (1996). Ten years of Orthoclone OKT3 (muromonab-CD3): a review. *Journal of Transplant Coordination*, 6(3):109–121.
- Snyder, L. (2013). *Molecular Genetics of Bacteria*. ASM Press, 4th edition.
- Soanes, K. H., Ewart, K. V., and Mattatall, N. R. (2008). Recombinant production and characterization of the carbohydrate recognition domain from Atlantic salmon C-type lectin receptor C (SCLRC). *Protein Expression and Purification*, 59(1):38 – 46.
- Son, Y.-J., Phue, J.-N., Trinh, L. B., Lee, S. J., and Shiloach, J. (2011). The role of Cra in regulating acetate excretion and osmotic tolerance in *E. coli* K-12 and *E. coli* B at high density growth. *Microbial Cell Factories*, 10:52–52.
- Sonoda, H., Kumada, Y., Katsuda, T., and Yamaji, H. (2010). Functional expression of single-chain Fv antibody in the cytoplasm of *Escherichia coli* by thioredoxin fusion and co-expression of molecular chaperones. *Protein Expression and Purification*, 70(2):248–253.
- Sonoda, H., Kumada, Y., Katsuda, T., and Yamaji, H. (2011). Effects of cytoplasmic and periplasmic chaperones on secretory production of single-chain Fv antibody in *Escherichia coli*. *Journal of Bioscience and Bioengineering*, 111(4):465–470.
- Sorensen, H. P. and Mortensen, K. K. (2005a). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology*, 115(2):113–128.
- Sorensen, H. P. and Mortensen, K. K. (2005b). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories*, 4(1):1–8.
- Sorensen, M. A., Kurland, C., and Pedersen, S. (1989). Codon usage determines translation rate in *Escherichia coli*. *Journal of Molecular Biology*, 207(2):365–377.
- Sousa, R., Patra, D., and Lafer, E. M. (1992). Model for the mechanism of bacteriophage T7 RNAP transcription initiation and termination. *Journal of Molecular Biology*, 224(2):319 – 334.

- Spadiut, O., Capone, S., Krainer, F., Glieder, A., and Herwig, C. (2014). Microbials for the production of monoclonal antibodies and antibody fragments. *Trends in Biotechnology*, 32(1):54–60.
- Spiess, C., Beil, A., and Ehrmann, M. (1999). A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, 97(3):339 – 347.
- Steiner, D., Forrer, P., Stumpp, M., and Pluckthun, A. (2006). Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nature Biotechnology*, 24(7):823–831.
- Stenberg, F., Chovanec, P., Maslen, S. L., Robinson, C. V., Ilag, L. L., von Heijne, G., and Daley, D. O. (2005). Protein Complexes of the *Escherichia coli* Cell Envelope. *Journal of Biological Chemistry*, 280(41):34409–34419.
- Stewart, E. J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. *The EMBO Journal*, 17(19):5543–5550.
- Stirling, C., Stewart, G., and Sherratt, D. (1988). Multicopy plasmid stability in *Escherichia coli* requires host-encoded functions that lead to plasmid site-specific recombination. *Molecular and General Genetics*, 214(1):80–84.
- Strauch, K., Johnson, K., and Beckwith, J. (1989). Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. *Journal of Bacteriology*, 171(5):2689–2696.
- Striedner, G., Huber, J., and Keller, D. (2008). Method for Producing a Recombinant Protein on a Manufacturing Scale (WO2008142028).
- Striedner, G., Pfaffenzeller, I., Markus, L., Nemecek, S., Grabherr, R., and Bayer, K. (2010). Plasmid-free T7-based *Escherichia coli* expression systems. *Biotechnol. Bioeng.*, 105(4):786–794.
- Stroomer, J. W. G., Roos, J. C., Sproll, M., Quak, J. J., Heider, K.-H., Wilhelm, B. J., Castelijns, J. A., Meyer, R., Kwakkelstein, M. O., Snow, G. B., Adolf, G. R., and van Dongen, G. A. M. S. (2000). Safety and Biodistribution of 99mTechnetium-labeled Anti-CD44v6 Monoclonal Antibody BIWA 1 in Head and Neck Cancer Patients. *Clinical Cancer Research*, 6(8):3046–3055.

- Studier, F. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1):37–44.
- Studier, F. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1):113–130.
- Studier, F., Parichehre Davanloo, B., Alan, H., Barbara, A., and John, J. (1990). Cloning and expression of the gene for bacteriophage T7 RNA polymerase (US4952496).
- Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S., and Kim, J. F. (2009). Understanding the Differences between Genome Sequences of *Escherichia coli* B Strains REL606 and BL21(DE3) and Comparison of the *E. coli* B and K-12 Genomes. *Journal of Molecular Biology*, 394(4):653 – 680.
- Stymest, K. H. and Klappa, P. (2008). The periplasmic peptidyl prolyl *cis-trans* isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS Journal*, 275(13):3470–3479.
- Sugiki, T., Fujiwara, T., and Kojima, C. (2014). Latest approaches for efficient protein production in drug discovery. *Expert Opinion on Drug Discovery*, 9(10):1189–1204.
- Summers, D. and Sherratt, D. (1984). Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell*, 36:1097–1103.
- Summers, D. and Sherratt, D. (1988). Resolution of ColE1 dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. *The EMBO Journal*, 7(3):851–858.
- Summers, D. K. (1991). The kinetics of plasmid loss. *Trends in Biotechnology*, 9(1):273–278.
- Sun, X.-W., Wang, X.-H., and Yao, Y.-B. (2014). Co-expression of Dsb proteins enables soluble expression of a single-chain variable fragment (scFv) against human type 1 insulin-like growth factor receptor (IGF-1R) in *E. coli*. *World Journal of Microbiology and Biotechnology*, 30(12):3221–3227.
- Sun, X.-x. and Wang, C.-c. (2000). The N-terminal Sequence (Residues 1 - 65) Is Essential for Dimerization, Activities, and Peptide Binding of *Escherichia coli* DsbC. *Journal of Biological Chemistry*, 275(30):22743–22749.

- Sundström, H. (2007). *Analytical Tools for Monitoring and Control of Fermentation Processes*. PhD thesis, Department of Bioprocess Technology, School of Biotechnology, Royal Institute of Technology.
- Sundström, H., Wallberg, F., Ledung, E., Norrman, B., Hewitt, C., and Enfors, S. O. (2004). Segregation to non-dividing cells in recombinant *Escherichia coli* fed-batch fermentation processes. *Biotechnology Letters*, 26(19):1533–1539.
- Swamy, K. H. and Goldberg, A. L. (1982). Subcellular distribution of various proteases in *Escherichia coli*. *Journal of Bacteriology*, 149(3):1027–1033.
- Tabor, S. and 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(4):1074–1078.
- Tajima, T., Yokota, N., Ichi Matsuyama, S., and Tokuda, H. (1998). Genetic analyses of the in vivo function of LolA, a periplasmic chaperone involved in the outer membrane localization of *Escherichia coli* lipoproteins. *FEBS Letters*, 439(1-2):51–54.
- Talmadge, K. and Gilbert, W. (1982). Cellular location affects protein stability in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 79(6):1830–1833.
- Tanabe, H., Goldstein, J., Yang, M., and Inouye, M. (1992). Identification of the promoter region of the *Escherichia coli* major cold shock gene, *cspA*. *Journal of Bacteriology*, 174(12):3867–3873.
- Tatum, E. L. (1945). X-Ray Induced Mutant Strains of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 31(8):215–219.
- Tatum, E. L. and Lederberg, J. (1947). Gene Recombination in the Bacterium *Escherichia coli*. *Journal of Bacteriology*, 53(6):673–684.
- Teich, A., Lin, H., Andersson, L., Meyer, S., and Neubauer, P. (1998). Amplification of *ColE1* related plasmids in recombinant cultures of *Escherichia coli* after IPTG induction. *Journal of Biotechnology*, 64:197 – 210.
- Telesnitsky, A. and Chamberlin, M. J. (1989a). Terminator-distal sequences determine the *in vitro* efficiency of the early terminators of bacteriophages T3 and T7. *Biochemistry*, 28(12):5210–5218.

- Telesnitsky, A. P. and Chamberlin, M. J. (1989b). Sequences linked to prokaryotic promoters can affect the efficiency of downstream termination sites. *Journal of Molecular Biology*, 205(2):315 – 330.
- Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 72(2):211–222.
- Thermo Fisher (2016). BL21-AI One Shot Chemically Competent *E. coli*. URL: <https://www.thermofisher.com/order/catalog/product/C607003>.
- Thermo Fisher Scientific Inc. (2015). Pierce[®] Recombinant Protein L, Peroxidase Conjugated. URL: <https://www.thermofisher.com/order/catalog/product/32420>.
- Thie, H., Schirrmann, T., Paschke, M., Dübel, S., and Hust, M. (2008). SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in *E. coli*. *New Biotechnology*, 25(1):49 – 54.
- Thieringer, H. A., Jones, P. G., and Inouye, M. (1998). Cold shock and adaptation. *BioEssays*, 20(1):49–57.
- Tolia, N. H. and Joshua-Tor, L. (2006). Strategies for protein coexpression in *Escherichia coli*. *Nature Methods*, 3(1):55–64.
- Tormo, A., Almirón, M., and Kolter, R. (1990). *surA*, an *Escherichia coli* gene essential for survival in stationary phase. *Journal of Bacteriology*, 172(8):4339–4347.
- Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010a). An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell*, 141(2):344–354.
- Tuller, T., Waldman, Y. Y., Kupiec, M., and Ruppín, E. (2010b). Translation efficiency is determined by both codon bias and folding energy. *Proceedings of the National Academy of Sciences*, 107(8):3645–3650.
- Turan, S., Zehe, C., Kuehle, J., Qiao, J., and Bode, J. (2013). Recombinase-mediated cassette exchange (RMCE) - A rapidly-expanding toolbox for targeted genomic modifications. *Gene*, 515(1):1–27.
- Ukkonen, K., Veijola, J., Vasala, A., and Neubauer, P. (2013). Effect of culture medium, host strain and oxygen transfer on recombinant Fab antibody fragment yield and leakage to medium in shaken *E. coli* cultures. *Microbial Cell Factories*, 12(1):73–

- Valax, P. and Georgiou, G. (1993). Molecular Characterization of β -Lactamase Inclusion Bodies Produced in *Escherichia coli*. 1. Composition. *Biotechnology Progress*, 9(5):539–547.
- Valdez-Cruz, N., Caspeta, L., Perez, N., Ramirez, O., and Trujillo-Roldan, M. (2010). Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. *Microbial Cell Factories*, 9(1):18.
- Valent, Q., Scotti, P., High, S., de Gier, J., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998). The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *The EMBO Journal*, 17(9):2504–2512.
- Vallejo, L. F. and Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microbial Cell Factories*, 3:11–11.
- Van Wielink, J. E. and Duine, J. A. (1990). How big is the periplasmic space? *Trends in Biochemical Sciences*, 15(4):136–137.
- Varma, A. and Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and Environmental Microbiology*, 60(10):3724–3731.
- Vasina, J. and Baneyx, F. (1996). Recombinant protein expression at low temperatures under the transcriptional control of the major *Escherichia coli* cold shock promoter *cspA*. *Applied and Environmental Microbiology*, 62(4):1444–1447.
- Vasina, J. A. and Baneyx, F. (1997). Expression of Aggregation-Prone Recombinant Proteins at Low Temperatures: A Comparative Study of the *Escherichia coli cspA* and *tac* Promoter Systems. *Protein Expression and Purification*, 9(2):211–218.
- Vemuri, G. N., Altman, E., Sangurdekar, D. P., Khodursky, A. B., and Eiteman, M. A. (2006). Overflow Metabolism in *Escherichia coli* during Steady-State Growth: Transcriptional Regulation and Effect of the Redox Ratio. *Applied and Environmental Microbiology*, 72(5):3653–3661.
- Verel, I., Heider, K.-H., Siegmund, M., Ostermann, E., Patzelt, E., Sproll, M., Snow, G. B., Adolf, G. R., and van Dongen, G. A. (2002). Tumor targeting properties of monoclonal antibodies with different affinity for target antigen CD44V6 in nude mice bearing head-and-neck cancer xenografts. *International Journal of Cancer*, 99(3):396–402.

- Vimberg, V., Tats, A., Remm, M., and Tenson, T. (2007). Translation initiation region sequence preferences in *Escherichia coli*. *BMC Molecular Biology*, 8(1):100.
- Wagner, S., Klepsch, M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K., Slotboom, D., and Persson, J. (2008). Tuning *Escherichia coli* for membrane protein overexpression. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38):14371–14376.
- Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, 28(9):917–924.
- Walsh, G. (2014). Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32(10):992–1000.
- Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. (2009). The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proceedings of the National Academy of Sciences*, 106(6):1772–1777.
- Walton, T. A. and Sousa, M. C. (2004). Crystal Structure of Skp, a Prefoldin-like Chaperone that Protects Soluble and Membrane Proteins from Aggregation. *Molecular Cell*, 15(3):367–374.
- Wang, R., Xiang, S., Feng, Y., Srinivas, S., Zhang, Y., Lin, M., and Wang, S. (2013). Engineering production of functional scFv antibody in *E. coli* by co-expressing the molecule chaperone Skp. *Frontiers in Cellular and Infection Microbiology*, 3(72).
- Wang, Z., Xiang, L., Shao, J., Wegrzyn, A., and Wegrzyn, G. (2006). Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism. *Microbial Cell Factories*, 5(1):34–.
- Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature*, 341(6242):544–546.
- Weiner, J. H. and Li, L. (2008). Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(9):1698–1713.
- Weininger, U., Jakob, R. P., Kovermann, M., Balbach, J., and Schmid, F. X. (2009). The prolyl isomerase domain of PpiD from *Escherichia coli* shows a parvulin fold but is devoid of catalytic activity. *Protein Science*, 19(1):6–18.

- Wiebe, J. C., Schller, C., Reiche, J. A., Kramer, K., Skerra, A., and Hock, B. (2010). An Expression System for the *E. coli* Fermentation of Recombinant Antibody F_{ab} Fragments from Mice and Rabbits. *Journal of AOAC International*, 93(1):80–88.
- Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A., and Bukau, B. (2009). Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *The EMBO Journal*, 29(5):910–923.
- Wittwer, A. (2010). A novel transcriptional termination signal capable to terminate T7 polymerase mediated transcription more efficiently. Bachelor thesis, Universität Wien, <http://othes.univie.ac.at/9213/>.
- Woesten, M. (1998). Eubacterial sigma-factors. *FEMS Microbiology Reviews*, 22(3):127–150.
- Wrobel, B. and We, G. (1998). Replication Regulation of ColE1-like Plasmids in Amino Acid-Starved *Escherichia coli*. *Plasmid*, 39(1):48–62.
- Wu, M., Pan, K., and Chou, C. (2007). Effect of heat-shock proteins for relieving physiological stress and enhancing the production of penicillin acylase in *Escherichia coli*. *Biotechnology and Bioengineering*, 96:956–66.
- Wulfig, C. and Pluckthun, A. (1994). Correctly Folded T-cell Receptor Fragments in the Periplasm of *Escherichia coli*: Influence of Folding Catalysts. *Journal of Molecular Biology*, 242(5):655–669.
- Wulfig, C. and Rappuoli, R. (1997). Efficient production of heat-labile enterotoxin mutant proteins by overexpression of dsbA in a degP-deficient *Escherichia coli* strain. *Archives of Microbiology*, 167(5):280–283.
- Wyre, C. and Overton, T. (2014). Use of a stress-minimisation paradigm in high cell density fed-batch *Escherichia coli* fermentations to optimise recombinant protein production. *Journal of Industrial Microbiology & Biotechnology*, 41(9):1391–1404.
- Xu, X., Wang, S., Hu, Y. X., and McKay, D. B. (2007). The Periplasmic Bacterial Molecular Chaperone SurA Adapts its Structure to Bind Peptides in Different Conformations to Assert a Sequence Preference for Aromatic Residues. *Journal of Molecular Biology*, 373(2):367–381.

- Xu, Y., Lewis, D., and Chou, C. P. (2008). Effect of folding factors in rescuing unstable heterologous lipase B to enhance its overexpression in the periplasm of *Escherichia coli*. *Appl Microbiol Biotechnol*, 79(6):1035–44.
- Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., Lan, W., Wang, J., Wu, H., Cao, C., and Wang, X. (2011). Highly Efficient Production of Soluble Proteins from Insoluble Inclusion Bodies by a Two-Step-Denaturing and Refolding Method. *PLoS ONE*, 6(7):e22981.
- Yavachev, L. and Ivanov, I. (1988). What does the homology between *E. coli* tRNAs and RNAs controlling ColE1 plasmid replication mean? *Journal of Theoretical Biology*, 131(2):235 – 241.
- Yokota, T., Milenic, D. E., Whitlow, M., and Schlom, J. (1992). Rapid Tumor Penetration of a Single-Chain Fv and Comparison with Other Immunoglobulin Forms. *Cancer Research*, 52(12):3402–3408.
- Yu, D., Ellis, H. M., Lee, E.-C., Jenkins, N. A., Copeland, N. G., and Court, D. (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 97(11):5978–5983.
- Zhang, T., Bertelsen, E., and Alber, T. (1994). Entropic effects of disulphide bonds on protein stability. *Nature Structural and Molecular Biology*, 1(7):434–438.
- Zhang, Z., Song, L., Fang, M., Wang, F., He, D., Zhao, R., Liu, J., Zhou, Z., Yin, C.-c., Lin, Q., and Huang, H. (2003). Production of soluble and functional engineered antibodies in *Escherichia coli* improved by FkpA. *Biotechniques*, 35(5):1032–1038.
- Zhuo, X.-F., Zhang, Y.-Y., Guan, Y.-X., and Yao, S.-J. (2014). Co-expression of disulfide oxidoreductases DsbA/DsbC markedly enhanced soluble and functional expression of reteplase in *Escherichia coli*. *Journal of Biotechnology*, 192, Part A:197 – 203.
- Zimmerman, S. B. and Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *Journal of Molecular Biology*, 222(3):599 – 620.
- Zucca, S., Pasotti, L., Politi, N., Cusella De Angelis, M., and Magni, P. (2013). A standard vector for the chromosomal integration and characterization of BioBrick™ parts in *Escherichia coli*. *Journal of Biological Engineering*, 7(1):12–.