Studies on bioprocesses for protein secretion with *Corynebacterium glutamicum*

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften genehmigte Dissertation

vorgelegt von

Dipl.-Ing. (FH) Johannes Hemmerich

aus Neuwied am Rhein

Berichter: Univ.-Prof. Dr. rer. nat. Marco Oldiges
Univ.-Prof. Dr. rer. nat. Ulrich Schwaneberg
Univ.-Prof. Dr. rer. nat. Wolfgang Wiechert

Tag der mündlichen Prüfung: 17. Juli 2018

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

I hereby affirm that I independently authored this thesis and did not use any source or appliance other than those cited in this study. This thesis has not been submitted in either whole or part for a degree at this university or any other institution.

Aachen, July 2018

(Johannes Hemmerich)

"Merde!"

Captain Jean-Luc Picard
Star Trek: The Next Generation, Season 2, Episode 3.

Contents

Acknowledgements

First of all, I owe Marco Oldiges a deep thank, who supervised my work. He gave me the opportunity to become a member of his research group and I do believe that the mutual acceptance and esteem significantly contributed to the success of my work. Furthermore, he was always keen in my personal and professional progress and was always a sounding board for my weired ideas. I also would like to thank Wolfgang Wiechert for the opportunity to become a member of his institute and giving valuable feedback for my work when needed. Also, his continous efforts in providing us with the excellent working conditions here at the institute is greatly acknowledged. Both Marco Oldiges and Wolfgang Wiechert are also acknowledged for examination of this thesis. A big thank goes to Uli Schwaneberg and Lars M. Blank for being members of my examination committee, as well as for their time spent in the examination of this thesis. Jan Schirawski is acknowledged for serving as chairman of my examination committee.

Conducting a PhD project is not possible without fruitful cooperations. Therefore, I would like to thank several people: The close and uncomplicated daily cooperation with Sarah Jurischka and Roland Freudl is greatly acknowledged. Sarah helped a lot with her patient assistance in plasmid transformation, and the provision of several cutinase secretion strains. Roland was always willing to critically discuss results, which was extremely valuable feedback for me, especially during preparation of manuscripts. Also, all partners within the BioExpresSPro project are acknowledged for discussions during our project meetings. Besides, there have been more successful cooperations: The collaboration with Eric von Lieres, Lars "Kaffee-Buddy" Freier and Katja Schöler yielded highly interesting results. The help and cooperation of Tino Polen and Ulli Degner with conducting, analyzing and interpretation of microarray experiments is also gratefully acknowledged.

Furthermore, the whole working group with its varying members is acknowledged. Those people made the time just fly away, I always felt very comfortable being a member of the group. In particular, I would like to thank Matthias Moch for his assistance with lots of bioreactor cultivations, Bianca Klein for conducting many proteomic analyses, Jannick Kappelmann and Stephan Noack for discussions and help with manuscript preparation, as well as Hogler Morschett and Niklas Tenhaef for critical proofreading. Finally, I thank the beloved current and former members of R136 (Bowser, Donkey Kong,

Wario and Yoshi), especially for being real competitors in Mario Kart 64.

There is another group of persons that significantly contributed to the outcome of this thesis: the student workers Sebastian J. Reich, Marc Weiske, Carmen Steffens, Tim Langhorst and Mohamed Labib (in order of appearance). You did not only show your braveness and thirst for adventure by choosing my as supervisor, but all of you did a fantastic job. It was my privilege and pleasure to work with you. I really appreciate your hard work, thank you very much!

In the end, I would like to express my deepest gratitude to all my friends and family that helped me to make my way. In particular, I want to deeply thank my mother, who always believed in me. Mom, this is for you!

Finally, I dedicate this work to my little family Sarah and Carlo. Your love is the most precious thing I could ever imagine and without you, nothing would make sense any more. Everything I do, I do for you. I love you!

Published work

This work has been conducted at the Institute of Bio- and Geosciences – Biotechnology (IBG-1) at Forschungszentrum Jülich from October 2014 until February 2018 as part of the *BioExpresSPro* project which is financially supported by the Ministry of Innovation, Science and Resesarch within the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-002 13). Parts of this work have already been published in peer-reviewed journals, presented on conferences and in supervised students' theses. Any thought, methodology, result and conclusion that are part of an existing own publication is not explicitly referred to, but considered as properly cited by the following listing:

Peer-reviewed publications

Publication I: Freier L, Hemmerich J, Schöler K, Wiechert W, Oldiges M, von Lieres E: Framework for Kriging-based iterative experimental analysis and design: Optimization of secretory protein production in *Corynebacterium glutamicum*. ***Engineering in Life Sciences*** 2016 **16**:538-549

Publication II: Hemmerich J, Rohe P, Kleine B, Jurischka S, Wiechert W, Freudl R, Oldiges M: Use of a Sec signal peptide library from *Bacillus subtilis* for the optimization of cutinase secretion in *Corynebacterium glutamicum*. ***Microbial Cell Factories*** 2016 **15**:208

Publication III: Hemmerich J, Wiechert W, Oldiges M: Automated growth rate determination in high-throughput microbioreactor systems. ***BMC Research Notes*** 2017 **10**:617.

Publication IV: Hemmerich J, Freier L, Wiechert W, von Lieres E, Oldiges M: Generic protocol for optimization of heterologous protein production using automated microbioreactor technology and Kriging-based experimental design. ***Journal of Visualized Experiments*** 2017 **130**:e56234.

Publication V: Hemmerich J, Wiechert W, Noack S, Oldiges M: Microbioreactor systems for accelerated bioprocess development. ***Biotechnology Journal*** 2018 **13**:1700141.

Manuscripts

Manuscript I: Hemmerich J, Moch M, Jurischka S, Wiechert W, Freudl R, Oldiges M: Combinatorial impact of Sec signal peptides from *Bacillus subtilis* and bioprocess conditions on heterologous cutinase secretion by *Corynebacterium glutamicum*. *Manuscript submitted*.

Manuscript II: Hemmerich J, Tenhaef N, Steffens C, Kappelmann J, Weiske M, Reich SJ, Wiechert W, Oldiges M, Noack S: Increasing process information from high-throughput microbioreactor cultivations by repeated low-volume sampling from microtiter plates. *Manuscript submitted*.

Manuscript III: Hemmerich J, Labib M, Steffens C, Weiske M, Reich SJ, Jurischka S, Wiechert W, Freudl R, Polen T, Oldiges M: Comprehensive evaluation of a library of genome-reduced *Corynebacterium glutamicum* strains for heterologous cutinase secretion – An explorative study comprising analyses of transcriptome, proteome as well as extracellular yields and rates from high-throughput microbioreactor experimentation. *Manuscript in preparation*.

Conference talks

- Hemmerich J, Freier L, Jurischka S, Schöler K, Wiechert W, Freudl R, von Lieres E, Oldiges M: Application of *Corynebacterium glutamicum* beyond amino acid production: Adjustment of medium composition unravels potential for secretory protein production. **ProcessNet 2016**, Aachen/DE. Day of presentation: 14.09.2016
- Hemmerich J, Freier L, Steffens C, Wiechert W, von Lieres E, Oldiges M: Hyphenating Mini-Pilot-Plant and Design-of-Experiments: Detailed bioprocess profiling for different optimization objectives. **Himmelfahrtstagung 2017**, Neu-Ulm/DE. Day of presentation: 24.05.2017

Poster presentations

- Jurischka S, Hemmerich J, Oldiges M, Freudl R: Optimization of Sec-dependent heterologous protein secretion in *Corynebacterium glutamicum*. **VAAM Annual Conference**, Jena/DE, 13. - 16.03.2016
- Hemmerich J, Rohe P, Kleine B, Jurischka S, Wiechert W, Freudl R, Oldiges M.: Proper selection of signal peptides for heterologous expression and secretion: Challenging, but worthwile. **PhD Day NRW "Future Bioeconomy"**, Düsseldorf/DE, 22.09.2016

- Hemmerich J, Steffens C, Freier L, Rohe P, Kleine B, Jurischka S, Freudl R, Wiechert W, von Lieres E, Oldiges M: Mini Pilot Plant methods for quantitative phenotyping of recombinant protein secretion with *Corynebacterium glutamicum*. **Bioprocessing Days**, Recklinghausen/DE, 20. - 22.02.2017
- Hemmerich J, Moch M, Rohe P, Kleine B, Jurischka S, Wiechert W, Freudl R, Oldiges M: Secretory production of cutinase from *Fusarium solani pisi* with *Corynebacterium glutamicum* and *Bacillus subtilis*: Impact of Sec signal peptide, host organism and bioprocess condition on secretion performance. **9th International Conference on Recombinant Protein Production (RPP9)**, Dubrovnik/HR, 23. - 25.04.2017

Supervised students' projects

- Reich SJ: **Impact of the genomic background on recombinant cutinase secretion in *Corynebacterium glutamicum*.** Master Thesis 2016, RWTH Aachen, Aachen/DE
- Weiske M: **Einfluss des Hintergundgenoms und des Signalpeptides auf die Cutinasesekretion durch *Corynebacterium glutamicum* in Mikrobioreaktorsystemen.** Bachelor Thesis 2016, HS Niederrhein, Krefeld/DE
- Steffens C: **Phänotypisierung von *Corynebacterium glutamicum* Genomreduktionsmutanten für die Cutinasesekretion: Methodenentwicklung und erste Anwendungsbeispiele.** Bachelor Thesis 2016, FH Aachen Abt. Jülich, Jülich/DE
- Langhorst T: **Stammspezifische Sensitivitätsuntersuchungen von Nährmedienkomponenten für *C. glutamicum* Genomreduktionsstämme (running title)**, Research Internship Project 2017, RWTH Aachen, Aachen/DE
- Labib M: **Influence of *rrnX* deletions on the intra- and extracellular phenotype of *Corynebacterium glutamicum* cutinase secretion strains.** Master Thesis 2017, Universität Bonn, Bonn/DE

Abstract

After being discovered 60 years ago, *Corynebacterium glutamicum* is a major industrial workhorse for the production of amino acids like L-glutamate and L-lysine at several million tons per year. Intense and still ongoing basic and applied research fueled this great biotechnological success story. Currently, *C. glutamicum* is increasingly getting into focus as production host for heterologous proteins of both technical and clinical interest. Consequently, research is shifting towards the use of *C. glutamicum* microbial cell factory for protein production.

To facilitate rapid and reliable characterization of newly constructed strain variants from metabolic engineering, microbioreactor (MBR) systems emerged as versatile tools. Such systems provide an increased experimental throughput with the ability to control environmental cultivation parameters, as well as monitoring of culture dynamics like biomass formation. In this study, methods for MBR systems are developed to suite the demands of a specific screening objective, namely the quantitative phenotyping of *C. glutamicum* strains secreting green fluorescent protein (GFP) and cutinase as heterologous model proteins. This involves the careful selection of standard operating conditions with respect to physiological demands like oxygen-unlimited metabolization of glucose as main carbon source, as well as determination of accuracy and imprecision of online biomass monitoring. Furthermore, to obtain time-resolved data on secretory protein formation and substrate consumption, an improved method is presented and validated that does not accompany a loss in MBR cultivation throughput.

The need for (semi-)automated data processing from high-throughput MBR cultivations is also discussed on the example of derived performance indicators (PIs) that represent condensed evaluation metrics for rapid evaluation of whole cultivation experiments. As application example, a method for automated growth rate determination is presented in detail, since this PI is probably the most often applied characteristic in biological fitness testing of mutant strain libraries. In addition, on the example of maximizing secreted GFP titer, it is shown how MBR systems, integration of laboratory automation and Kriging-based Design of Experiments (DoE) complement each other in a synergistic way. As a result, an iterative workflow is presented that serves as blueprint for development of further biotechnological applications. Unexpectedly, secreted GFP titer could be doubled, showing

that routinely applied nutrition media designed for amino acid production with *C. glutamicum* need to be carefully adapted and optimized with changing screening objectives, that is here secretion of heterologous proteins.

To complement the current knowledge on how to select the optimal signal peptide (SP) for different expression hosts and different target proteins of choice, the interrelation of bioprocess control strategy and choice of SP to optimize cutinase secretion with *C. glutamicum* is investigated in detail. Since the envisaged degree of process control could not be realized with the available MBR systems, a consistent data set was generated relying on more than 150 bench scale bioreactor runs. Furthermore, the results are discussed and interpreted in the light of changing bioprocess optimization objectives, which again highlights the need for careful definition of optimization objectives.

Representing a typical application example of MBR systems, the quantitative microbial phenotyping of a library of genome reduced *C. glutamicum* strains for heterologous cutinase secretion was conducted. The collected data comprise growth rates and cutinase yields as extracellular phenotypes, as well as detailed analysis at the transcriptome and proteome level for a small subset of strains. Next to surprising phenotypes due to specific genomic deletions, as well as differential analysis of phenotypes from strains with overlapping genomic deletions, attempts were made to explain the metabolic perturbations from observed significantly differential regulation at the protein level. Also, by incorporating all data on extracellular phenotypes, a data-driven, phenomenological multiple regression approach was used to identify the minimum set of genomic deletions needed in terms of improved cutinase secretion.

Finally, a few future aspects outreaching the scope of this work are presented as outlook. These aspects concern the application of *C. glutamicum* as potential alternative host for heterologous protein production, the demand for further development of microbioreactor systems and the need for smart solutions for warehousing, (re-)processing and interpretation of heterogeneous data sets to cope with the foreseeable increase of information output generated from high-throughput experimentation in combination with powerful analytical methods.

Zusammenfassung

Nach seiner Entdeckung vor mehr als 60 Jahren ist heutzutage *Corynebacterium glutamicum* der etablierte Produktionsorganismus für die industrielle Herstellung von Aminosäuren wie L-Glutamat und L-Lysin im Millionen-Tonnen Maßstab. Intensive grundlegende und angewandte Forschung, die immer noch andauert, war und ist eine treibende Kraft für diese Erfolgsgeschichte der Biotechnologie. Aktuell rückt *C. glutamicum* in den Fokus als Produktionsorganismus für die Herstellung von heterologen Proteinen, die für technische und klinische Anwendungen interessant sind. Als Konsequenz richtet sich die Forschung hin zur Anwendung von *C. glutamicum* als mikrobielle Zellfabrik für die Proteinproduktion.

Mikrobioreaktor (MBR) Systeme sind wertvolle Werkzeuge für die schnelle und zuverlässige Charakterisierung von neuen, mittels *metabolic engineering* konstruierten Stammvarianten. MBR Systeme bieten einen erhöhten experimentellen Durchsatz bei gleichzeitiger Kontrolle von Kultivierungsparametern, sowie die Möglichkeit der Aufzeichnung von Kultivierungsdynamiken. In dieser Arbeit wurden verbesserte Anwendungsmethoden für MBR Systeme entwickelt, zugeschnitten auf die Anforderungen, die sich ergeben bei der quantitativen Phänotypisierung von *C. glutamicum* Stämmen, welche das *green fluorescent protein* (GFP) und Cutinase als Modelprotein sekretieren. Die Arbeiten erstrecken sich dabei über die sorgfältige Auswahl von Standardarbeitsbedingungen in Bezug auf physiologische Anforderungen wie die Sauerstoff-unlimitierte Metabolisierung von Glukose, sowie die Bestimmung von Genauigkeit und Präzision der online Biomassemessung. Darüber hinaus wird eine Methode präsentiert, welche es erlaubt, zeitlich aufgelöste Konzentrationsänderungen von sekretiertem Protein und Substrat zu bestimmen, ohne dass dies mit einem Verlust an experimentellem Durchsatz einhergeht.

Die Notwendigkeit von (halb-)automatischer Datenprozessierung aus Hochdurchsatzkultivierungen in MBR Systemen wird diskutiert anhand abgeleiteter *Performance*-Indikatoren (PI). Diese stellen kondensierte Maßzahlen für die schnelle Evaluierung der gesamten Kultivierung dar. Als konkretes Anwendungsbeispiel wird eine Methode für die automatisierte Bestimmung der spezifischen Wachstumsrate detailliert beschrieben, denn dieser PI ist mutmaßlich die am häufigsten herangezogene Charakteristik bei der Testung von Mutantenbibliotheken hinsichtlich biologischer Fitness. Des Weit-

eren wird anhand des Beispiels der Maximierung des extrazellulären GFP Titers gezeigt wie sich MBR Systeme, die Integration von Laborrobotern und Kriging-basierte Planung und Auswertung von Experimenten in synergistischer Weise ergänzen. Als Ergebnis wird ein iterativer Arbeitsablauf präsentiert, der als Blaupause für die Entwicklung von weiteren biotechnologischen Anwendungen dienen kann. Unerwarteterweise konnte der sekretierte GFP Titer verdoppelt werden, was zeigt, dass routinemäßig angewandte Nährmedien, welche für die Aminosäureproduktion mit *C. glutamicum* entwickelt wurde, für andere Optimierungsziele sorgfältig angepasst werden müssen.

Ein weiterer Teil dieser Arbeit untersucht den Einfluss des bisher nicht untersuchten Zusammenhangs zwischen Bioprozesskontrolle und Wahl des Signalpeptides auf die Cutinasesekretionsleistung mit *C. glutamicum*. Da der gewünschte Umfang an Bioprozesskontrolle im verwendeten MBR System nicht zur Verfügung steht, wurde ein konsistenter Datensatz aus über 150 Bioreaktorkultivierungen im Labormaßstab generiert. Die Diskussion und Interpretation der gewonnen Daten erfolgt unter dem Gesichtspunkt verschiedener Optimierungsziele.

Als ein typisches Anwendungsbeispiel für MBR Systeme ist die quantitative Phänotypisierung einer Bibliothek von Genom-reduzierten *C. glutamicum* Stämmen für die sekretorische Produktion von Cutinase gezeigt. Es wurden vielschichtige Datensätze erhoben, welche Wachstumsraten und Cutinaseausbeuten als extrazelluläre Phänotypen umfassen, sowie die detaillierte Analyse von Transkriptom und Proteom als intrazelluläre Phänotypen für ausgesuchte Stammvarianten. Neben überraschenden phänotypischen Ausprägungen für verschiedene genomische Deletionen, sowie differentiellen Analysen von Genotypen mit überlappenden Deletionen zwischen verschiedenen Stämmen, werden Erklärungsversuche geliefert, welche mögliche metabolische Zustände als Folge von als signifikant erkannten Änderungen im Proteom beschreiben. Des Weiteren wurden die gesammelten Daten zu extrazellulären Phänotypen in einem Daten-getriebenen, rein phänomenologischen Regressionsansatz mit Dummyvariablen genutzt zur Ermittlung eines minimal reduzierten Genoms, welches mit einer erhöhten Cutinaseausbeute einhergehen sollte.

Schließlich werden einige Aspekte über die Zielsetzung dieser Arbeit hinaus als Ausblick gegeben. Dies umfasst die Anwendung von *C. glutamicum* als potentiellen alternativen Organismus für die Produktion heterologer Proteine und zukünftige Erfordernisse in der Weiterentwicklung von MBR Systemen. Die Notwendigkeit von Lösungsansätzen für die Verwaltung, (Re-)Prozessierung und Interpretation von kontinuierlich wachsenden heterogenen Datenströmen wird angesprochen, da es abzusehen ist, dass sich der Output von Daten aus der Kombination von Hochdurchsatzkultivierung und leistungsstarken analytischen Methoden stark erhöhen wird und diesem begegnet werden muss.

Nomenclature

δ	Measurement error
ΔA_{410}	Linear increase in absorption of diluted sample at 410 nm over time
μ	Specific growth rate
μ_{Exp}	Observed specific growth rate
μ_{Set}	Adjusted specific growth rate as parameter for exponentially increasing feeding profile
A_{340}	Absorption reading of diluted sample at 340 nm
A_{595}	Absorption reading of sample at 595 nm
$a_{Bradford}$	Slope for protein calibration of Bradford assay
a_{CDW}	Slope for BS / cell dry weight calibration
a_{Glc}	Slope for glucose calibration
a_{pNP}	Slope for pNP calibration
$b_{Bradford}$	Offset for protein calibration of Bradford assay
b_{CDW}	Offset for BS / cell dry weight calibration
b_{Glc}	Slope for glucose calibration
c_i	Concentration of component i
c_P	Product concentration (i.e., cutinase activity)
$c_{P,Sup}$	Product concentration (i.e., cutinase activity) in supernatant
c_S	Substrate (i.e., glucose) concentration
$c_{S,0}$	Initial substrate (i.e., glucose) concentration
c_{SF}	Substrate (i.e., glucose) concentration of feed
c_X	Biomass concentration
EA	Enzymatic activity
f_D	Sample dilution factor

I_{BS}	Backscatter signal
$F\,(t)$	Feed rate
$k_{c_X \rightarrow c_{Biovol}}$	Empirical biomass / biovolume conversion factor
m_P	Product mass (i.e., total cutinase activity)
m_S	Maintenance coefficient for substrate consumption at zero growth as parameter for exponentially increasing feeding profile
m_X	Biomass
q_P	Biomass specific product (i.e., cutinase) formation rate
q_S	Biomass specific substrate consumption rate
r_P	Volumetric product (i.e., cutinase) formation rate
r_S	Volumetric substrate consumption rate
r_X	Volumetric biomass growth rate
t	Process runtime
t_{Feed}	Feeding process runtime
V_{Base}	Volume added base
$V_{Diluent1}$	Volume of first diluent
$V_{Diluent2}$	Volume of second diluent
$V_{DilutedSample}$	Volume of diluted sample
V_R	Cultivation volume
$V_{R,0}$	Initial cultivation volume
V_{Sample}	Sample volume
$Y_{P/X}$	Biomass specific product (i.e., cutinase) yield
$Y^*_{P/S}$	Non-observable substrate specific product (i.e., cutinase) yield
$Y_{X/S}$	Substrate specific biomass yield
$Y^*_{X/S}$	Non-observable substrate specific biomass yield

Abbreviations

BHI brain heart infusion

BS backscatter

CI confidence interval

DO dissolved oxygen

DoE Design of Experiments

FDR false discovery rate

GFP green fluorescent protein

GRS genome reduced strain

LHS liquid handling system

LOD limit of detection

LOQ limitation of quantification

LR linear regression

MBR microbioreactor

MCB master cell bank

MCF microbial cell factory

MRD maximum relative deviation

MTP microtiter plate

NLR non-linear regression

NTU nephelometric turbidity unit

OLS ordinary least squares

PBS phosphate buffered saline

PI performance indicator

pNP *p*-nitrophenol

pNPP *p*-nitrophenylpalmitate

SP signal peptide

SRP signal recognition particle

TLS total least squares

WCB working cell bank

WLR weighted linear regression

1. Introduction and background

This chapter provides the background for this work. A short introduction is given for *Corynebacterium glutamicum* and its biotechnologically relevant protein secretion pathways. In the light of industrial biotechnology, the current status on the application of signal peptide libraries to optimize heterologous protein production is described. Following this, process development for protein production is introduced and how current efforts in genome streamlining can contribute to the development of improved bioprocesses. Finally, black box models are presented that allow the characterization of such bioprocesses from a biochemical engineering point of view. Parts of this chapter are based on *Publication II*, *Publication V* and *Manuscript I*.

Author contributions:
Johannes Hemmerich wrote the chapter, and prepared the non-reprinted figures. Roland Freudl and Marco Oldiges helped with writing sections 1.2 and 1.4. Stephan Noack and Wolfgang Wiechert helped with writing section 1.4.

1.1. *Corynebacterium glutamicum*

Corynebacterium glutamicum was discovered in 1957 during a screening project for the isolation of glutamate producing microorganisms [1]. It is a Gram-positive, aerobically growing, non-pathogenic member of the Corynebacteriaceae family. *C. glutamicum* belongs to the phylum Actinobacteria, showing a high GC-content with mycolic acid layers attached to the cell wall. Its club-shape and the snapping cell division is typical (Figure 1.1). *C. glutamicum* possess a single circular chromosome with approximately 3.3 mio bp and about 3000 protein-coding regions [2].

As major producer for food and feed amino acids at industrial scale for decades [4, 5], extensive bioprocess knowledge with *C. glutamicum* and methods for genetic manipulation are available. Furthermore, *C. glutamicum* is classified as GRAS ("generally regarded as safe"), can be cultivated to high cell densities cultivations [6–8] and was shown to be very robust towards process inhomogeneities encountered in large scale cultivations [9–11]. Besides, like other mostly monoderm Gram-positive

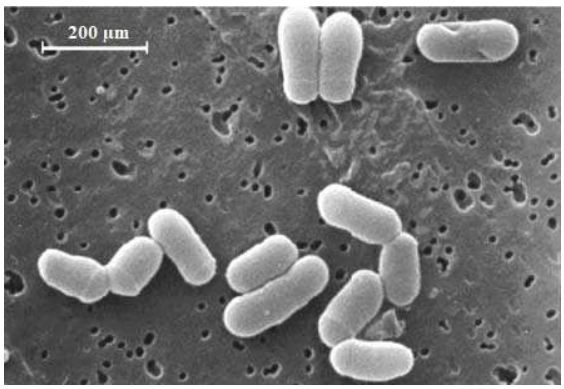


Figure 1.1.: Scanning electron microscope picture of *C. glutamicum*. Reprinted from [3] with permission.

expression hosts (e.g., *Bacillus* sp.), *C. glutamicum* exhibits further favorable features with respect to secretory enzyme production, namely low nutritional demand, low amount of endogenously secreted proteins and proteolytic activity, as well as the ability to secrete heterologous proteins into the extracellular medium in the g L^{-1} range [8, 12–15].

1.2. The general secretory and twin-arginine translocation pathway, with special emphasis on *C. glutamicum*

To enable extracellular protein production with microbial cells, target proteins of choice must be directed to secretion pathways that enable the translocation of newly synthesized protein across the cytoplasma membrane as well as other outer compartments, depending on the nature of the expression host used (e.g., Gram-positive or Gram-negative). Mainly, there are two biotechnologically relevant protein secretion pathways for industrial microbials, the general secretory (Sec) pathway and the twin-arginine translocation (Tat) pathway. In short, proteins secreted via Sec are translocated across the cytoplasma membrane in an unfolded state, either co- or posttranslationally, while via Tat proteins are translocated that have gained their functional conformation already in the cytoplasm. To mediate the secretion of newly synthesized protein by either pathway, specific small peptides are needed fused at the N-terminus of the nascent polypeptide of the protein of interest. These signal peptides (SPs) exhibit distinct features allowing the cell to discriminate how to further process newly synthesized proteins. For example, SPs directing proteins into the Sec machinery are characterized by a highly conserved tripartite structure (cf. section 1.2.1) while SPs specifically addressing the Tat pathway show a conserved motif of two arginine residues [16]. A schematic draw-

ing describing the corresponding sequence of events during protein secretion via either Sec or Tat pathway in *C. glutamicum* is shown in Figure 1.2 [17].

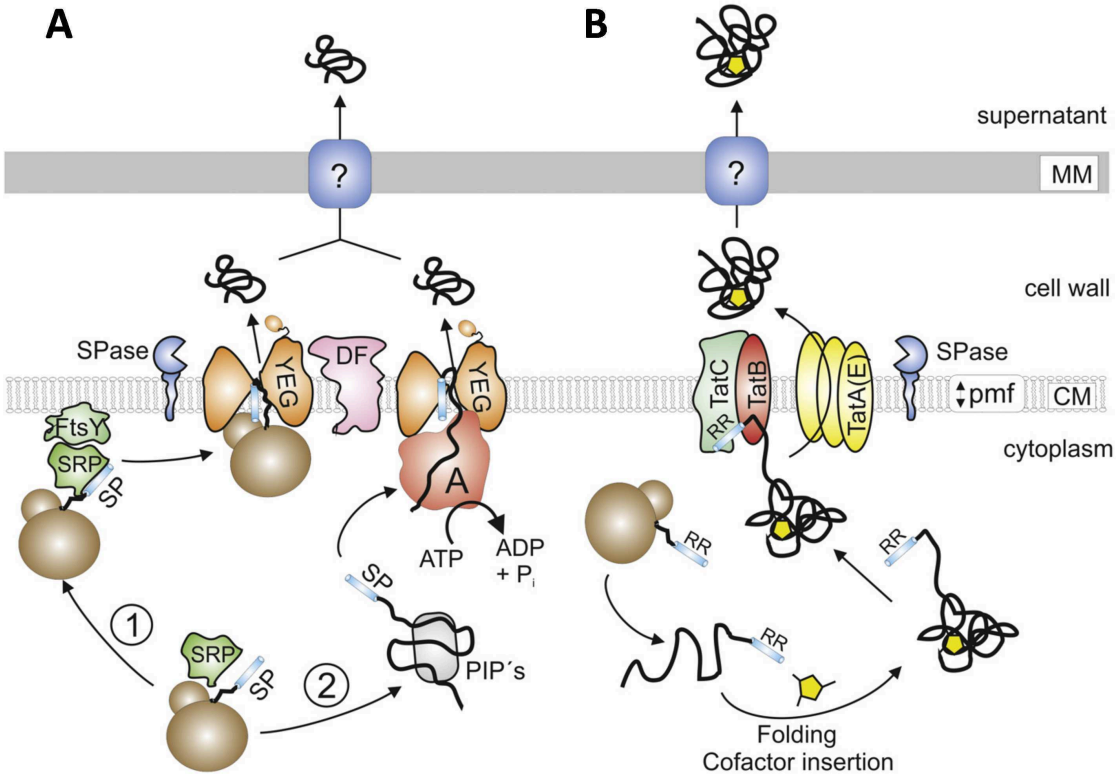


Figure 1.2.: Overview of Sec (A, co-translational branch (1), post-translational branch (2)) and Tat (B) protein secretion pathways in *C. glutamicum*. Figure reprinted from from [17] with permission.

During Tat-mediated protein secretion, the nascent polypeptide is completely translated before it folds into its functional conformation in the cytoplasm, which may include the incorporation of cofactors. Afterwards, the SP-decorated protein is recognized by the TatC and TatB complex and after recruitment of TatA(E), the substrate (i.e., protein) is translocated across the cytoplasma membrane, fueled by the proton-motive-force. At the trans-side of the membrane, the protein is released, and the SP is cleaved by signal peptidase (SPase).

For Sec-mediated translocation, the target protein is fused to a Sec specific SP at the N-terminus and upon emerging the ribosome, the SP is recognized by the signal recognition particle (SRP) in the co-translational mode, forming the ribosome-nascent chain (RNC)-SRP complex, which is directed to the membrane-bound receptor FtsY and the RNC is handed over from SRP over FtsY to the SecYEG

translocon spanning the cytoplasma membrane. Then, the continuing translation pushes the nascent polypeptide through the translocon and at the trans-side of the membrane, the SP is cleaved off and the protein starts to fold into its functional conformation.

Delivery of the target protein to the SecYEG translocon can also occur in a post-translational manner, where the newly synthesized protein is kept in a translocation competent (i.e., unfolded) state by posttranslationally interacting proteins (PIPs). Then, the SP is bound by the SecA protein, which directs the protein to the translocon and also pushes it through the translocation pore in an ATP-dependent manner. Again, the protein starts to fold at the trans-side of the membrane and the SP is cleaved.

The protein translocation across the membrane is assisted by SecDF for both the co- and post-translational branch, which converge at the SecYEG translocon at the cis-side of the membrane. It is assumed that the decision for either co- or post-translational secretion is based on length of the hydrophobic domain of the Sec SP, at least this is the case in *E. coli* [18]. In contrast to other Gram-positive bacteria like *B. subtilis*, *C. glutamicum* possesses an additional outer mycomembrane and although proteins can obviously pass this membrane, it is not yet known whether any dedicated transport systems exist [17].

Most knowledge on protein secretion pathways was gained from studies using *E. coli* as model organism for the Sec pathway [19, 20] as well as for Gram-negatives, while *B. subtilis* [21, 22] serves as model organism for Gram-positives [16]. Since the Sec pathway is found to be highly conserved among all kingdoms of life [23], functionality of Sec pathway in other bacteria is inferred from studies on *E. coli* as model organism. However, some organism-specific questions remain, for example the fact that there is no homologue of the *E. coli* SecB chaperone found in Gram-positives that keeps proteins in a secretion competent state and targets them to the translocon during post-translational protein translocation. However, since nearly all Sec components discovered in *E. coli* were also found in Gram-positives, it is often assumed that pathways do function in a similar manner and thus, the existence of PIPs fulfilling the function of SecB is postulated [17]. General chaperones (GroEL-GroES / DnaK-DnaJ-GrpE / trigger factor) or the soluble form of SecA are suggested to act as such PIPs [17]. Also, there is indication that SecA may play an important role during both co- and post-translational Sec-dependent translocation [24, 25]. Another difference in components of the protein secretion pathways is observed for the structure of SRP, which lacks the Alu domain in *E. coli* but not in *B. subtilis* [26] In eukaryotic cells, the Alu domain of SRP mediates arrest of translation until the translocation across the cytoplasma membrane occurs, as the energy needed for translocation is provided by the elongation of the polypeptide [27]. The Alu domain of *B. subtilis* SRP was also suggested

to function likewise [28], meaning that translation arrest in Gram-positives during co-translational Sec-mediated protein secretion could occur.

In conclusion, there is still basic research needed to increase understanding of the complexity of protein secretion pathways. However, with respect to biotechnological application, research could be translated towards powerful strategies like SP library screening (cf. section 1.2.2) that allow fine-tuning and optimizing heterologous protein secretion bioprocesses with various expression hosts.

1.2.1. Structure and function of Sec signal peptides

SPs are small sequences that are fused in front of newly synthesized proteins in order to mark a protein as a one having a different destination than the cytoplasm. Sec SPs show a highly conserved three-domain structure (Figure 1.3A): a positively charged N-region, with a high preference for lysine residues at P2 and P3 [21, 29], the longest region, termed H-region, consists of hydrophobic amino acids, and a C-domain, which contains a cleavage site (typical motif: A-X-A) between the SP and target (pre-)protein to be recognized by signal peptidase (SPase). In contrast to the highly conserved three-domain structure of SPs, a high diversity in aa composition within these domains is observed.

Depending on the absence or presence of a Sec SP in combination with further specific features, newly synthesized proteins are recognized by the Sec machinery of a bacterial cell and transported to different possible destinations. For example, in case the protein possesses trans-membrane domains, it is positioned into the membrane while the presence of a cell wall binding motif causes the retention of the protein in the cell wall after translocation over the membrane (Figure 1.3B). The transportation out of the microbial cell into the extracellular space is mostly of biotechnological interest.

1.2.2. Application of signal peptide libraries for the optimization of recombinant protein secretion

The use of SPs for secretory production of proteins is common technology. In several studies, the SP has been shown to be a critical major factor in optimizing recombinant protein secretion.

From the genome of *B. subtilis*, 173 Sec SPs have been identified based on the highly conserved three-domain structure [21], which have been compiled by Brockmeier et al. to a genetic library to optimize recombinant protein secretion in *B. subtilis* [30]. The study revealed that systematic screening of all natural SPs in *B. subtilis* is a powerful strategy to improve extracellular enzymatic activity using cutinase from *Fusarium solani pisi* as model enzyme. The SP was found to affect the processing kinetics of cutinase preprotein during translocation, but a fast or low cutinase preprotein processing

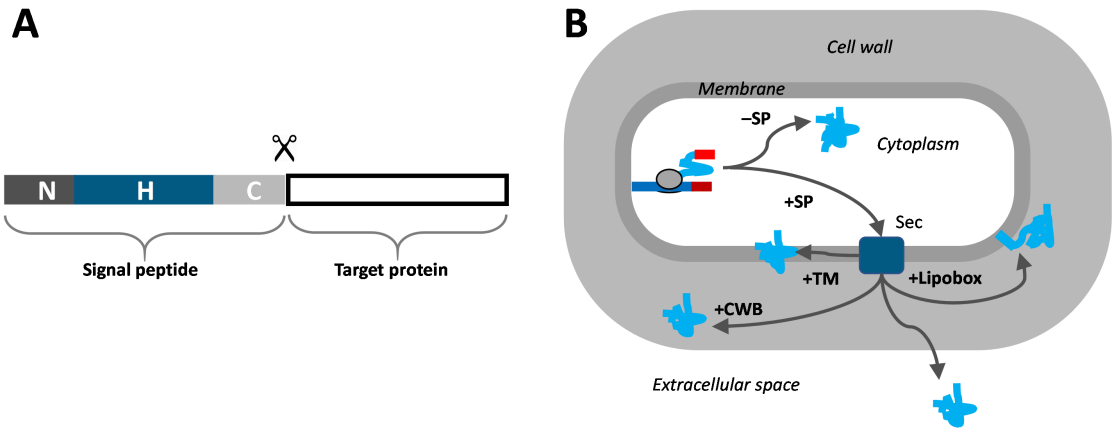


Figure 1.3.: Typical structure of Sec signal peptides and possible destinations of newly synthesized proteins directed to the Sec machinery of Gram-positive microorganisms. (A): Sec SPs possess a highly conserved tri-partite structure, the positively charged N-domain, the longest H-domain built from hydrophobic amino acids and the C-domain containing a cleavage site for release of the target protein. (B): Figure part B modified from [22]. *TM*, transmembrane domain. *CWB*, cell wall binding motif.

could not be correlated to high or low extracellular cutinase activities, respectively. Furthermore, the SP screening results for cutinase could not be transferred to a second model enzyme, an esterase from metagenomics origin.

Essentially, the same observation was made in a screening study regarding a saturation mutagenesis of the N-domain from the *B. subtilis* AmyE SP. Using again cutinase as model enzyme, extracellular protein production was varied between 20 to 400% compared to the use of the wild type AmyE SP. Again different obtained mutants resulted in different processing kinetics of cutinase precursor, but these could not be correlated to the amount of extracellular cutinase activity [31].

In a recent study employing the *B. subtilis* SP library described by Brockmeier et al. [30], it was found that SPs show a high correlation in secretion performance of an alkaline xylanase for the comparison of two promotors of different strength (P43 and P*glvm*) [32]. This indicates that transcript level of SP-xylanase fusion is of minor importance to the secretion efficiency. Clearly, the SP was the major factor for xylanase secretion.

For another Gram-positive expression host, *Lactobacillus planatarum*, screening of homologous SPs was applied to optimize secretory production using staphylococcal nuclease (NucA) and lactobacillal amylase (AmyA) as model enzymes. As seen in other studies, no correlation of SP performance between different target protein, i.e. NucA and AmyA in this case, could be observed [33].

Browsing the genome of *C. glutamicum R* identified predicted SPs, which were then screened for

secretion of α-amylase from *Geobacillus stearothermophilus* [34]. Several of these SPs were shown to outperform the well-known corynebacterial PS2 SP with respect to extracellular amylase activity, highlighting again that for each target protein, the optimal SP has to be identified to achieve maximal secretion efficiency.

In another study, a library comprising Sec SPs from *B. subtilis* and *B. licheniformis* (approx. 400 SPs in total) was evaluated for secretion of subtilisin BPN' from *B. amyloliquefaciens* as model enzyme, using *B. subtilis* as expression host [35]. Besides the finding that heterologous SPs function in general, it was seen that the highest extracellular subtilisin BPN' activity in *B. subtilis* was achieved with a heterologous SP from *B. licheniformis*. The comparison of the eight best performing SPs yielded considerably higher extracellular subtilisin BPN' activities (approximately 750% to 350%) compared to the native SP of subtilisin BPN' used as reference. Furthermore, these eight SPs were applied in two *B. licheniformis* strains for subtilisin BPN' secretion and it was found that the results were comparable between *B. subtilis* and the two *B. licheniformis* strains, and highly comparable between the two *B. licheniformis* strains.

As seen in previous studies, Sec SPs from low-GC firmicutes *B. subtilis* are functional in high-GC actinobacterium *C. glutamicum* [36] and therefore, the assessment of the Sec SP library introduced by Brockmeier et al. [30] was a promising approach to optimize recombinant protein secretion also in *C. glutamicum*. In contrast to the results obtained by Degering et al. [35], it was found that specific SPs cause different secretion efficiencies for the same model enzyme (cutinase) when using *B. subtilis* [30] or *C. glutamicum* [37] as host. More specifically, some SPs showed diametrical performance in both hosts, whereas others did not, and some SPs showed comparable performance in one host but not in the other host. This indicates that the phylogenetic distance of expression hosts correlates with these observations and that the "expression environment" of the host cells is most likely the more similar, the closer hosts are related.

To date, it is not possible to predict an optimal combination of SP, target protein and expression strain. Bioinformatics tools (e.g., SignalP [38]) have been proven useful for the prediction of SPs from sequenced genomes and thus, the prediction of secreted proteins [21, 34] but attempts to correlate secretion efficiency of both homologous and heterologous proteins with scores calculated from SP sequences (e.g., hydrophobicity, net charge, D-score) have failed so far [30, 35]. Consequently, SP library screening to optimize protein secretion is an important step in expression engineering and thus, recognized to be crucial for the development of a biotechnological production process for technical bulk enzymes. Anyhow, with increasing size of SP libraries, accordingly increasing screening effort has to be considered for practical reasons.

1.3. Reduced genomes for the construction of platform organisms

When talking about genome reduction, several expressions are frequently found in literature which are defined based on the outcome/aims of such projects, for example the generation of a chassis, minimal or streamlined cell (cf. [39, 40]). In this sense, minimal cells may be considered as of major interest for fundamental research, while a chassis or streamlined cell is more interesting for applied research since such cells may serve as platform technology:

First, in basic research, projects aim to contribute to the fundamental understanding on how a cell works since a "minimal" gene set of a cell is considered to reduce the complexity of a resulting phenotype [40, 41]. In the end, it is postulated that the knowledge generated will facilitate the construction of functional cells from scratch for different purposes [42], true to the motto "What I cannot create, I do not understand" [1]. To this end, a bottom-up approach for the construction of a minimal genome seems to be the preferred strategy. Consequently, attempts in genome reduction and the characterization of corresponding phenotypes are of fundamental interest in the field of (synthetic) biology.

Second, in applied research, contrary to the bottom-up approach typically seen in basic research, a top-down strategy attempts to reduce the genome of a given organism to the only necessary encoded functionality with respect to given constraints (e.g., growth on a certain carbon source). This approach is considered as useful for the construction of platform strains, which are well characterized and predictable in behavior due to a less complex genome. This would serve as starting point for the development of specific strains, for example by the introduction of foreign or synthetically designed metabolic pathways. Furthermore, removing non-essential genes is deemed to dispense unwanted sources of energy and metabolic building blocks consumption in production strains, so that those resources are preserved for improved product biosynthesis yield [43]. Therefore, the construction of streamlined production chassis strains as configurable modules for industrial biotechnology development projects can be considered as driving interest for applied research on minimal/reduced genomes [44, 45]. Some results from genome reduction projects concerning industrially important microorganisms are collected in Table 1.1.

1.3.1. Genome reduction of *C. glutamicum*

The complete sequencing of its genome [2] paved the way towards the construction of genome streamlined *C. glutamicum* strains for application in industrial biotechnology. With respect to het-

[1] Once noted on a blackboard by Richard P. Feynman.

Table 1.1.: Results from selected genome reduction reports in literature.

Organism	Results	Reference
B. subtilis	Construction of strains with genome reductions of up to 36%. Comparable growth rates to the wild type strain in complex medium are reported. Multi-omics analyses of genome reduced strains revealed, e.g., differential regulation of proteins involved in protein biosynthesis and assembly of Fe-S clusters. Moreover, results led to a refined *i*Bsu1103 genome-scale metabolic model with improved predicition of wild type phenotypes.	[46, 47]
C. glutamicum	Strain C1* has a 14.3% reduced genome compared to the wild type strain ATCC13032. Strain C1* is robust towards different stresses including oxygen limitation and long-term stability of growth in mineral and complex medium.	[48]
E. coli	Several strains, with up 15.3% reduced genome (strain MDS43) compared to *E. coli* MG1655, a close relative to K-12 were generated. No impairment in growth fitness, a higher transformation efficiency and lower mutation rate were observed.	[49]
P. putida	Prophage curation enhanced tolerance against environmental stresses. Strain EM383 with 4.3% reduced genome compared to parental strain KT2440 shows higher biomass yield and lower maintenance. Growth rate, biomass specific product yield and product synthesis rate using GFP as heterologous model protein are increased.	[50, 51]
S. avermitis	A series of deletions mutants was generated with approx. up to 19% genome reduction. Expression of heterologous gene clusters for streptomycin from *S. griseus* and for cephamycin from *S. clavuligerus* resulted in higher titers compared to the native producers.	[52]

erologous protein production, the curation of the *C. glutamicum* genome from prophage regions [53] and insertional sequence (IS) elements [54] was reported to enhance heterologous protein production and transformation efficiency. Starting from the prophage-cured strain MB001 [53], the construction of a *C. glutamicum* chassis strain was started in a top-down approach by identification of potentially irrelevant genes [48, 55]. The irrelevance or essentiality of a gene was defined with respect to the desired phenotype of non-impaired submerged growth in comparison to the wild type strain ATCC13032 in the widely used defined CgXII medium [56] with glucose as main carbon source [48, 55]. Based on this definition, phylogenetic conservation, gene expression levels as well as literature knock-out data, 41 gene clusters were classified as not necessary. From these, 36 clusters could be successfully deleted (i.e., resulted in non-lethal phenotypes), meaning that the remaining ones do provide unknown essential function to the cell or may encode for unknown regulatory mechanisms. Out of these, the individual deletion of 26 clusters resulted in strains with a non-reduced biological fitness compared to the wild type strain. Surprisingly, several combinatorial deletions comprising

two of these 26 clusters generated growth-impaired strains, meaning that some (unknown) redundant essential function is encoded within those clusters. Finally, by deleting non-critical combinations of gene clusters, the *C. glutamicum* chassis strain C1* was constructed, showing a 13.4% genome reduction, and was successfully tested for biological fitness against several stressors and for long-term growth stability [48]. In addition to the construction of strain C1*, the project provided a large strain library and a revised genome annotation list, which are now available for further research.

1.4. Industrial biotechnology

Industrial biotechnology deals with the biological production of bulk, fine or specialty chemicals, pharmaceuticals and precursors as well as enzymes for food, feed or technical applications [57–61]. These value-added molecules are produced by the conversion of raw materials using the metabolic capacity of microbial cell factories (MCFs). Advances in synthetic biology made it possible to specifically engineer MCFs to produce a broad range of natural and non-natural products [62]. Nowadays, genetic engineering of MCFs is possible in short times [63], owed to the availability of genetic strain libraries, the standardization of genetic elements (e.g., "BioBricks" and variants thereof) [64], non-targeted high-throughput strain engineering using biosensor technology [65] as well as the development of advanced of gene editing technologies, like the CRISPR system [66–68].

1.4.1. Development of protein production processes

With an estimated market of 1 billion US-$ in 2010 [58], the production of technical bulk enzymes makes up a major part of industrial biotechnology. To supply this market with sufficient quantities and in an economically feasible manner, platform technologies are developed continuously. With extracellular, secretory enzyme formation, product recovery is facilitated because time- and energy-consuming cell disruption is avoided. This is reflected by the fact that dominating expression hosts for the production of commercial enzymes are secretory hosts, like *Aspergillus* sp. and *Bacillus* sp. with a share of about 27% and 17%, respectively [69], representing almost half of total commercial enzyme production (Figure 1.4).

In industrial biotechnology the development of protein production processes often follow a framework that can be described in a simplified form as a four-stage linear workflow [70, 71]. First, an expression system is constructed enabling a microbial host to produce the recombinant protein of interest in principle. Second, strains are engineered to increase product titer and yield. This can be done, for example, by testing different promotors [7, 72] and/or ribosome binding sites [73] to

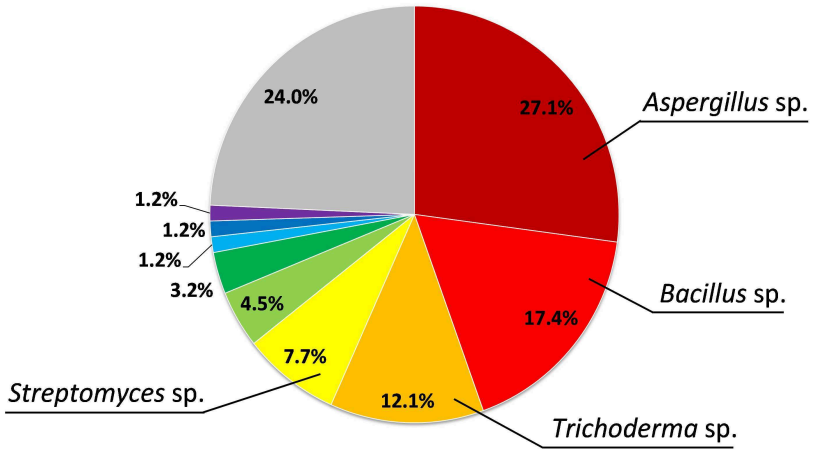


Figure 1.4.: Host organisms used for the production of enzymes, that are commercialized by AMFEP [69], based on number of products. *Aspergillus* sp. summarizes *A. melleus*, *A. niger* and *A. oryzae*. *Bacillus* sp. summarizes *B. amyloliquefaciens*, *B. subtilis*, *B. licheniformis*, *B. stearothermophilus* and others. *Trichoderma* sp. summarizes mainly *T. reesei* and *T. longibrachiatum*. *Streptomyces* sp. summarizes mainly *S. violaceoruber*.

increase expression of the target gene, as well as screening of secretion signals to increase target protein export [30, 32, 35, 37]. At this stage with focus on metabolic engineering, methods for growth experiments with subsequent target protein quantification are needed that are characterized by a high degree of standardization, parallelization and throughput to handle the resulting amount of constructed strain variants [74, 75]. Typically, these screening experiments are conducted using microtiter plates (MTPs) or shake flasks. Third, a few most promising strains are cultivated in bench-scale bioreactors, where environmental conditions like pH or temperature are tightly controlled. Also, the search for fed-batch control strategies that improve the formation of the target protein [76, 77] is currently performed at this stage. At bioreactor scale, cultivation conditions matching those in large scale bioreactors (e.g., oxygen and substrate gradients) can also be evaluated for their impact on the performance of producer strain candidates using scale-down simulators [78, 79]. Fourth and finally, after identification of the apparent best combination of producer strain and bioprocess control strategy, the transfer to the production scale follows. Consequently, successful biotechnological protein production requires the evaluation of many biological and bioprocess variables, which cannot be experimentally investigated in all combinations due to the huge design space. Therefore, during development stages, decisions must be made to subsequently reduce the number of variables, with the ultimate goal to setup an economically feasible production of the target protein of choice. An

overview on commercialized technical enzymes and their area of application is given in Table 1.2.

Table 1.2.: Selected examples of industrial application of enzymes.

Application / industry	Enzyme(s)	Purpose	Reference
Animal nutrition	Phytase	Increases phosphate availability from feed for monogastric animals	[80]
	Cellulase	Facilitated digestion of silage	[81]
Bulk chemicals	Nitrile hydratase	Acrylamide production	[82]
	Lipase	Chiral amines, polyester synthesis	[83]
Clothing	Cellulase	Smoothening of cotton fabric	[81]
	Laccase	Bleaching ("stone-washed") effect of denim	[84]
DNA amplification and synthesis	Taq polymerase	Heat stable DNA polymerase for PCR reaction with changing temperatures	[85]
Food industry	Amylase, glucoamylase, glucose isomerase	Conversion of starch into high fructose corn syrup	[86]
	Pectinase	Clarification and viscosity reduction of natural juices	[87]
Laundry detergents	Lipase, Protease	Washing at low temperatures (hydrolysis of fatty and proteinaceous dirt)	[88, 89]
Pharmaceutics	*o*-phtalyl amidase	Synthesis of intermediates in production of antibiotic loracarbef	[90]
	Tyrosine phenollyase	Production of L-3,4-dihydroxyphenylalanine (L-DOPA)	[91]
	Ketoreductase	Production of Montelukast, active against asthma and allergies	[92]
Pulp and paper industry	Hemicellulases (Xylanase, Mannanase)	Bleaching of pulp during Kraft process for paper production	[93]

1.5. Bioprocess kinetics and operation of bioreactors

The cultivation of microorganisms in bioreactors is often described by rather simple models. Such models have proven over decades to be sufficient in order to foresee cultivation dynamics in most cases. Typically, in bioprocess engineering the stirred tank reactor (STR) is the vessel of choice to carry out the (bio-)reaction, which is basically the growth of microorganisms and associated reactions like heterologous protein synthesis or by-product excretion. Equation (1.1) describes the general mass balance for component i in the reactor, with the boundary set around the reactor. Here, $\frac{dm_i}{dt}$ denotes the accumulation term, which is positive for an inflow or formation of component i and is negative for a drain of consumption of component i. $F_{in} \cdot c_{i,in}$ denotes the inflow term for component i and $F_{out} \cdot c_{i,out}$ denotes the drain term. Finally, $r_i \cdot V_R$ denotes the reaction term, again positive for formation of component i and negative for consumption of component i.

$$\frac{dm_i}{dt} = \frac{d\left(c_i \cdot V_R\right)}{dt} = F_{in} \cdot c_{i,in} - F_{out} \cdot c_{i,out} + r_i \cdot V_R \tag{1.1}$$

1.5.1. Description of the batch process

In batch cultivations, all nutrients for the microbial culture in the cultivation medium are added at the beginning of the whole process [2]. After inoculation, the growth of the culture starts and ends upon depletion of the limiting substrate, which is mostly the carbon source. After completion of growth, the whole reactor content is harvested for product recovery. Changes in the cultivation volume V_R due to pH adjustment, sampling, evaporation, etc. are considered to be negligible (i.e., $V_R = const.$) since those represent typically only up to 1 - 2% of the total cultivation volume [3]. This also means that necessarily both the inflow and outflow terms in the above depicted general mass balance become zero. Therefore, the mass balance can be written as concentration balance for the component i being biomass (c_X), substrate (c_S) and product (c_P) which contain only the reaction terms, as depicted in (1.2), (1.4) and (1.3), respectively.

$$\frac{dc_X}{dt} = r_X = \mu \cdot c_X \tag{1.2}$$

[2] In aerobic bioprocesses, oxygen needs to be continuously delivered to the culture due to its very low solubility in aqueous solutions, meaning that in case of insufficient aeration, oxygen becomes the limiting nutrient within seconds. Therefore, oxygen is gassed into the bioreactor constantly ensuring a non-limiting concentration in the cell suspension and hence, does not need to be considered in the mass balance.

[3] Based on observations in this work.

$$\frac{dc_S}{dt} = r_S = -\frac{1}{Y_{X/S}} \cdot \frac{dc_X}{dt} \tag{1.3}$$

$$\frac{dc_P}{dt} = r_P = Y_{P/X} \cdot \frac{dc_X}{dt} \tag{1.4}$$

The reaction term regarding biomass (index X), $r_X = \mu \cdot c_X$, cf. (1.2), describes growth as an autocatalytic reaction with biomass as catalyst and μ as proportionality factor, which is assumed to be constant during the resulting exponential growth phase. A decline in growth rate to model the transition from exponential growth phase into the stationary phase due to depletion of the substrate can be described using different well accepted empirical models, for example Monod's equation [94]. However, in this study only the exponentially growing culture is of interest, and thus, investigations on and modeling of culture behavior during other growth phases are out of scope.

Product formation is considered to be directly growth-coupled which is a reasonable assumption for a protein production, cf. (1.3). For the sake of simplicity, a constant relation between biomass growth and product formation is assumed. This is expressed by a yield coefficient relating product formation with biomass growth, $Y_{P/X}$, which corresponds to the well-known Luedeking-Piret model of product formation without a non-growth associated production term [95]. Indeed, experimental data obtained in this study supports these assumptions (e.g., cf. Chapter 3 and Chapter 5).

Consumption of substrate is typically modeled to be proportional to the formation of biomass and product, expressed by the proportional factors (i.e., yield coefficients) $Y^*_{X/S}$ and $Y^*_{P/S}$, respectively, cf. (1.5) to (1.7), where the * indicates the non-observability of these yield coefficients, cf. below. In this study, products are heterologous proteins (i.e., green fluorescent protein (GFP) and cutinase) which are synthesized by the cell, meaning that the substrate needed for heterologous protein synthesis is consumed and metabolized by the biomass first and then the cell distributes the gained energy and building blocks into growth and product formation. This means the yield coefficients $Y^*_{X/S}$ and $Y^*_{P/S}$ are not separately observable without high experimental effort. Following this, substrate consumption is modeled to be caused by biomass growth only, expressed by the yield coefficient $Y_{X/S}$ in (1.3), which is experimentally easily accessible. Intuitively, the amount of substrate accounting for the total "cost" of heterologous protein production is then reflected by a lower value of $Y_{X/S}$ obtained for the protein production strain compared to the non-producing strain. Such cost would be, for example, upregulation of chaperones, lower abundance of host proteins due to the competition with the heterologous protein, an increased energy demand due to biosynthesis of target protein mRNA or production of the plasmid-encoded selection marker. This total cost is often summarized as "metabolic burden" [96, 97]. Inserting (1.2) and (1.4) into (1.5) yields (1.6), and with a little bit of algebra it is seen how the yield coefficients are related with each other, cf. (1.7).

Thereby, the term $Y_{P/X}/Y^*_{P/S}$ has the dimension mass of substrate per mass of biomass (here, with the unit $g_S\ g_X^{-1}$), describing which amount of substrate is needed by one unit of biomass for product formation (i.e., the metabolic burden to the cells). In case the cell does not synthesize a product, $Y_{P/X}$ becomes zero and thereby, $Y^*_{X/S}$ equals $Y_{X/S}$.

$$\frac{dc_S}{dt} = r_S = -\frac{1}{Y^*_{X/S}} \cdot \frac{dc_X}{dt} - \frac{1}{Y^*_{P/S}} \cdot \frac{dc_P}{dt} \tag{1.5}$$

$$r_S = -\frac{1}{Y^*_{X/S}} \cdot \mu \cdot c_X - \frac{1}{Y^*_{P/S}} \cdot Y_{P/X} \cdot \mu \cdot c_X = -\mu \cdot c_X \cdot \left(\frac{1}{Y^*_{X/S}} + \frac{Y_{P/X}}{Y^*_{P/S}} \right) \tag{1.6}$$

$$\frac{1}{Y_{X/S}} = \frac{1}{Y^*_{X/S}} + \frac{Y_{P/X}}{Y^*_{P/S}} \tag{1.7}$$

Conclusively, the batch bioprocess can be characterized by three to five parameters (μ, q_S, q_P, $Y_{X/S}$ and $Y_{P/X}$), of which q_S and q_P contain the information on μ as well as $Y_{X/S}$ and $Y_{P/X}$, respectively. Consequently, the batch bioprocess can be expressed in the summarized form depicted in (1.8). The corresponding parameters are versatile performance indicators (PIs) of the bioprocess. The relevance and interpretation of PIs is discussed later in the text (e.g., cf. section 2.7).

$$\frac{d}{dt} \begin{pmatrix} c_X \\ c_P \\ c_S \end{pmatrix} = \begin{pmatrix} r_X \\ r_P \\ r_S \end{pmatrix} = \begin{pmatrix} \mu \\ q_P \\ q_S \end{pmatrix} \cdot c_X = \begin{pmatrix} 1 \\ Y_{P/X} \\ -Y^{-1}_{X/S} \end{pmatrix} \cdot \mu \cdot c_X \tag{1.8}$$

1.5.2. Description of the fed-batch process

The fed-batch bioprocess in characterized by the addition of concentrated nutrient solution, which is in most cases the carbon source (glucose in this study). Furthermore, the feeding solution is added in a growth-limiting manner, meaning that the glucose concentration in the culture is nearly zero since the added glucose is immediately consumed by the culture mass. In turn, the biomass formation is restricted by the applied feed rate, which can be constant, linearly, exponentially increasing or pulsed [98]. This mode of operation is termed open-loop control because the feed rate follows a pre-programmed function. In contrast, in case the feed rate is controlled (adjusted) by considering some output signal from the bioreactor, the term closed-loop control is used [99]. An example for a closed-loop control is the continuous feed adjustment based on the online monitored respiration quotient (RQ) in production of baker's yeast. This strategy allows to control the amount of carbon sources in the bioreactor to reduce the "crabtree" effect [98]. Another typical example is the pH

adjustment in a bioreactor where the addition of base or acid depends on the measured pH value in the bioreactor.

The application of an exponentially increasing feed rate allows to adjust a constant specific growth rate of the cultivated cells. Depending on the desired growth rate of the culture, the feed rate is adjusted, as described in (8.1) in section 8.7.3. To derive culture characteristics analogously to the batch processes, the changing culture volume has to be considered, meaning that instead of biomass concentration (c_X) and product concentration (c_P) the masses of biomass (m_X), cf. (1.9) and product (m_P), cf. (1.10) have to be considered. The changing volume is attributed to the inflow of feeding solution which is determined by $F(t)$, the volume of pH adjusting agents $V_{Acid/Base}(t)$ added to maintain the pH setpoint and the removal of samples $V_{Sample}(t)$, cf. (1.11).

$$\frac{dm_X}{dt} = \frac{d\left(c_X \cdot V_R\right)}{dt} = \mu \cdot m_X \tag{1.9}$$

$$\frac{dm_P}{dt} = \frac{d\left(c_P \cdot V_R\right)}{dt} = q_P \cdot m_X \tag{1.10}$$

$$\frac{dV_R}{dt} = F\left(t\right) + V_{Acid/Base}\left(t\right) - V_{Sample}\left(t\right) \tag{1.11}$$

Since in this study only data from the exponentially fed growth phases are used to derive culture parameters μ, $Y_{P/X}$ and q_P as performance indicators, the general considerations of the previous section on batch bioprocesses can be applied. The fed-batch bioreactor cultivations are used to evaluate the impact of different growth rates on heterologous cutinase secretion efficiency in terms of cutinase yield $Y_{P/X}$ and productivity q_P. More specifically, the fed-batch cultivations serve to determine the empirical relations $Y_{P/X} = f(\mu)$ and $q_P = f(\mu)$ in Chapter 5, based on experimental data from the exponential fed-batch phase according to (1.12), (1.13) and equation (1.14). In practice, growth rate μ and cutinase yield $Y_{P/X}$ are often estimated from the experimental data using regression methods [100], as indicated in (1.12) and equation (1.13).

$$\mu = \frac{1}{m_X} \cdot \frac{dm_X}{dt} \approx \frac{ln\left(\Delta \boldsymbol{m_X}\right)}{\Delta \boldsymbol{t}} \tag{1.12}$$

$$Y_{P/X} = \frac{dm_X}{dm_P} \approx \frac{\Delta \boldsymbol{m_P}}{\Delta \boldsymbol{m_X}} \tag{1.13}$$

$$q_P = Y_{P/X} \cdot \mu \tag{1.14}$$

The obtained empirical relations can be used to determine a growth rate μ for which either cutinase yield $Y_{P/X}$ or cutinase productivity q_P is maximal. Once these empirical relations are known

for a specific production strain, a closed-loop control strategy can be applied to ensure that a bioproduction process is operated at the optimal value for growth rate μ [101]. However, such strategy requires signals from the bioreactor to be fed into the controller, and the required signals are ideally recorded in a high temporal solution and do not need manual experimental effort to be read out. Furthermore, implemented software controllers needed for closed-loop control can be arbitrarily complex [102] like, for example, sophisticated software controllers (also known as soft-sensors) that process a multitude of online monitored signals to control a non-directly experimentally observable parameter [103].

2. Microbioreactor systems for accelerated bioprocess development and quantitative microbial phenotyping

This chapter introduces microbioreactor systems as versatile tools for biotechnological applications and discusses current progress as well as future needs and prospectives. The need for automated data processing and the corresponding derivation of performance indicators is considered as major issue in projects of quantitative phenotyping. To reflect this point, the chapter describes and discusses in detail a procedure for automated determination of growth rate, one of the most important characteristics of a microbial culture. This chapter is based on *Publication III* and *Publication V*.

Author contributions:

Johannes Hemmerich prepared the figures, developed the MATLAB code, analyzed the data, interpreted results, and wrote the chapter with the help of Marco Oldiges and Wolfgang Wiechert. Stephan Noack helped with writing sections 2.1 to 2.7.

2.1. The need for accelerated bioprocess development in industrial biotechnology

As outlined in the previous chapter (e.g., section 1.2.2, section 1.3, section 1.4), current advances in molecular biology and metabolic engineering allow to quickly generate large strain libraries for a great diversity of biotechnological applications, including bioprocess development. These strain libraries need to be characterized for a set of well-defined and controllable cultivation conditions, which is the purpose of quantitative microbial phenotyping. However, the experimental throughput of corresponding methods that allow for such quantitative phenotyping is rather limited in comparison to the size of typical strain libraries. Therefore, novel technologies and approaches are urgently needed to close the gap between the availability of genetically engineered strains and their detailed,

quantitative characterization under process-relevant conditions, to finally enable the joint development, characterization and optimization of MCFs and bioprocesses.

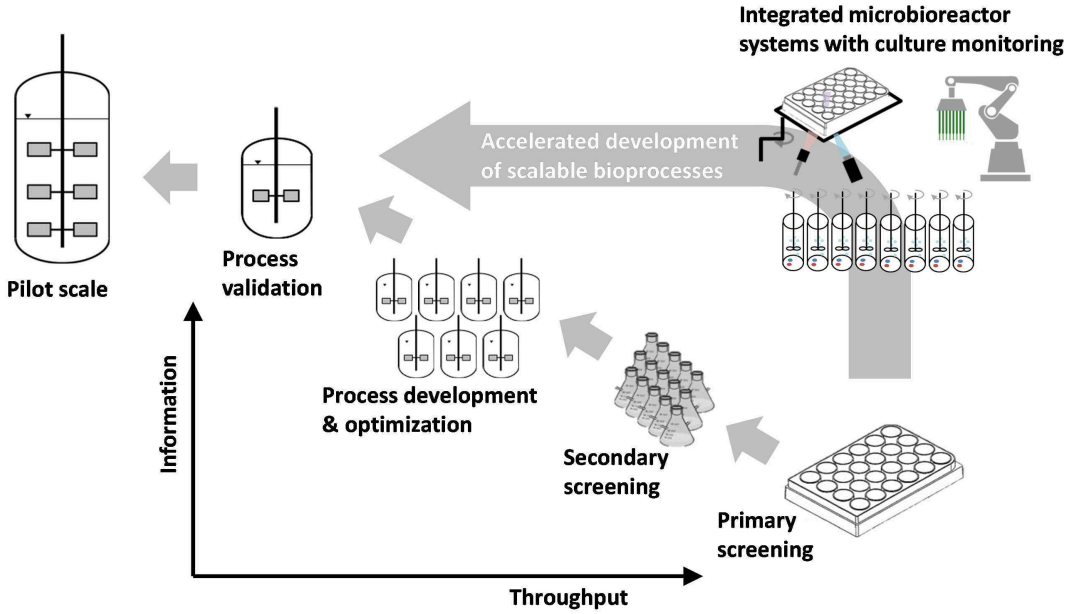


Figure 2.1.: Traditional bioprocess development versus microbioreactor systems. While microtiter plates and shake flasks provide high cultivation throughput, the gained information on strain and process properties is limited. Bench-scale bioreactors allow precise environmental control and can be equipped with different sensing probes, but cultivation throughput is limited. microbioreactor (MBR) systems can offer a shortcut to enable both, that is a fast and reliable bioprocess development. Figure reprinted from [104] with permisson, originally modified from [105].

The need for increased experimental throughput in quantitative microbial phenotyping under well controlled conditions gave rise to the development of MBR systems, which ultimately aim to replace shake flasks and lab-scale bioreactor systems for early and mid-stages of bioprocess development, as schematically depicted in Figure 2.1. Following this aim, several necessary specifications and desired properties of MBR systems can be formulated:

(*i*) miniaturized volume per single cultivation
(*ii*) high parallelization of single cultivations to realize necessary throughput
(*iii*) minimal setup-time between experiments
(*iv*) user-friendly and fail-safe operation
(*v*) automated and unsupervised operation
(*vi*) monitoring of bioprocess variables at high temporal resolution

(*vii*) controllable cultivation environment regarding temperature, pH and feeding profiles
(*viii*) accessibility of cultures for dosing/sampling and further analytics
(*ix*) batch, fed-batch and continuous mode of cultivation
(*x*) option for hardware and software integration with lab robotics and additional lab equipment
(*xi*) scalability of results with respect to standard lab-scale bioreactor operation

In recent years, enormous progress has been made in the field of MBR systems. However, not all necessary specifications and desired properties mentioned above have been realized and great efforts in fundamental and applied research are still ongoing. This is reflected by the existence of a variety of different MBR systems covering a broad technological range and a wide field of application (Table 2.1).

2.2. Microbioreactor systems based on stirred cultivation units

An intuitive approach to increase the cultivation throughput of stirred tank bioreactors is the miniaturization of this bioreactor type. The rationale behind this downsizing approach is that the geometric similar reactor setup at lab-scale and micro-scale can facilitate the scale-up of bioprocesses developed in the MBR systems. Another advantage of downsizing existing stirred tank bioreactors is the applicability of already established methods for bioprocess monitoring by using adequately sensitive sensors. For example, small-scale sensors enabling low range torque measurements allowed the experimental determination of detailed power consumption characteristics [120]. In combination with computational fluid dynamic simulations further understanding of hydrodynamics with different stirrers was gained [123]. Clearly, the high degree of parallelization is advantageous for the high resolution of operating conditions enabling detailed process insight. Since these MBR systems rely on arrays of individual down-sized stirred bioreactors, they can be operated with a varying number of individual vessels and cultivation conditions, fitting to specific demands of throughput. However, with increasing flexibility of individual reactor control, the technical setup of equipment becomes more complex, which then may counteract ease of application.

Up to now, two systems with typical cultivation volumes in the range of 10 mL are commercially available, namely "ambr 15" from TAP Biosystems and "bioREACTOR" from 2mag (see Table 2.1 for more details). The ambr 15 system was originally introduced for application in cell culture, but has been reported very recently for the evaluation of recombinant *E. coli* strains [110]. The bioREACTOR system is mainly used for microbial application and different impeller types can be used, for example to cultivate shear-sensitive mycelium forming microorganisms [120] or to conduct enzymatic

Table 2.1.: Comparison of commercially available MBR systems with respect to specifications and applications.

MBR	Specifications and capabilities	Application examples	Literature
ambr 15 *tapbiosystems.com*	Single-use stirred mini-reactors, 24 or 48 in parallel. Culture volume from 10 – 15 mL. Designed for cell culture applications. Temperature control of 12 mini-reactors at once. Quasi-continuous monitoring of pH and DO via optodes. Integrated into liquid handling system for automated feed addition, sampling and recalibration of pH optodes. Individual submerged feeding of gas mixtures for DO control.	Cell line development; medium, feed and growth parameter optimization	[106–110]
BioLector *m2p-labs.com*	Single-use MTPs with transparent bottom, 48 well Flowerplates are mostly applied. Culture volume from 0.8 – 2.4 mL. Quasi-continuous monitoring of biomass formation via backscatter, fluorescence and via optodes pH and DO. Integration of liquid handlers for sampling, feeding and pH adjustment. Environmental control of temperature, humidity and gas atmosphere.	Strain screening; phenotyping; medium, feed and growth parameter optimization; toxicity screening; OTR screening	[55, 111–119]
bioREACTOR *2mag.de*	Single-use stirred mini-reactors, 8 or 48 in parallel. Quasi-continuous monitoring of pH and DO via optodes. Culture volume from 8 – 15 mL. Specialized stirrers for mycelium-forming organisms. Integration into liquid handlers for sampling, feeding and pH adjustment.	Strain screening; medium, feed and growth parameter optimization; biotransformation; continuous cultivation	[120–126]
BioScreen C *bioscreen.fi*	Two parallel 100 well "honeycomb" MTPs. Environmental control of temperature. Culture volume up to 0.4 mL. Quasi-continuous monitoring of biomass formation via optical density.	Strain screening; phenotyping; toxicity screening; biotransformation	[127–132]
Growth profiler *enzyscreen.com*	Up to 10 parallel MTPs with transparent bottom. Quasi-continuous monitoring of biomass formation via image scanning. Environmental control of temperature.	Strain screening; toxicity screening	[113, 119, 133–136]
Micro-24 *pall.com*	Shaken single-use cassettes with 24 bubble columns. Culture volume from 3 – 7 mL. Quasi-continuous monitoring of pH and DO via optodes. Regulation of pH via gassing of NH_3 and CO_2. Individual temperature control.	Cell line and strain screening; medium, feed and growth parameter optimization	[137–140]
micro-Matrix *applikonbio.com*	Single-use 24 square well MTP with quasi-continuous monitoring of pH and DO via optodes. Culture volume from 1 - 7 mL. Individual temperature control. Control for up to 4 gasses and liquid addition.	Feed and growth parameter optimization	
SensorDishReader *presens.de*	Incubator shaker mounting system for MTPs (6, 24 or 96 well format). Culture volume from 0.2 - 10 mL with different MTPs. Quasi-continuous monitoring of pH or DO via optodes.	Cell line, strain and biocatalyst screening; medium optimization; toxicity screening;	[141–144]

hydrolysis of plant material [123]. In addition, this system was successfully applied for parallel continuous cultivations [124, 126] and a first prototype of an add-on system consisting of micropumps was reported for feeding and pH correction using small volumes at high frequency [145]. However, the application of non-invasive optical biomass measurements based on turbidity or light scattering is difficult in these MBR systems, since the continuously agitated impellers and resulting submerse gas bubbles interfere with such optical measurements. Therefore, the two systems currently rely on sampling for biomass determination, and here robotic integration seems to be mandatory to avoid shifting the bottleneck from cultivation to analysis.

2.3. Microbioreactor systems based on shaken cultivation units

The recognition of MTPs as cost-effective, easy-to-handle and parallelized cultivation units stimulated the thorough characterization of operation conditions from a biochemical engineering perspective. Consequently, the understanding of shaken small- and micro-scale culture systems was greatly improved. For instance, the discovery of the "out-of-phase phenomenon", which now can be described by the non-dimensional Phase number (Ph) [146, 147], greatly advanced the understanding of fluid dynamics in such devices. In addition, methods were developed to characterize hydrodynamics, mass transfer and power input in MTPs [147–152]. Results from fundamental research were transferred into application leading, for example, to the development of MTPs specifically designed for the cultivation of MCFs with high oxygen demand [111, 153]. Culture aeration in MTPs was proven to be affected by a variety of operating conditions, for example cultivation volume, shaking frequency, shaking diameter and geometry of MTP cavity [111, 147, 148, 150, 151, 153–156]. A comprehensive overview about oxygen transfer characteristics in miniaturized cultivation systems is given elsewhere [157].

An inherent drawback of using MTPs as cultivation devices is that individual culture well manipulations require a stop of the shaking movement, which can negatively impact the oxygen transfer and mixing of all cultivation wells [115, 139]. This issue becomes more critical when the number of necessary interactions increases and then also requires additional hardware that is capable of dosing/aspirating liquids to the individual cultivation wells under constant shaking of the microplate [112, 158]. Moreover, depending on the technical specifications of the shaken MBR system, some culture parameters can only be controlled globally for all cultivation wells in one MTP, e.g., the head space gas mixture or temperature [118, 159]. Consequently, detailed investigations on the impact of these parameters on process performance can only be carried out in separate MBR cultivation exper-

iments, substantially lowering the throughput during bioprocess development [118, 159]. Noteworthy, the Society for Laboratory Automation and Screening (SLAS) has developed industry standards [1] for MTP-based devices, which could facilitate the integration of such MTP-based MBR systems into robotic environments. Currently, several MBR systems based on MTPs have become commercially available commercially available, namely "Bioscreen C" from Oy Growth Curves, "BioLector" from m2p-labs, "Growth profiler" from Enzyscreen, "Micro-24" from Pall and "micro-Matrix" from Applikon (see Table 2.1 for more details).

2.4. Integration of microbioreactor systems into lab robotics

The integration of MBR systems into robotic environments greatly enhances their capabilities towards an improved quantitative microbial phenotyping and bioprocess development [36, 37, 112, 121, 124, 160–162]. In particular, liquid handling robotics allow for automated manipulation of growing cultures, for example, by adding inducers, taking samples, adding nutrients and adjusting pH. All of these events can be executed based on pre-programmed timetables or scripted protocols using online (e.g., biomass, pH) or atline (e.g., substrate, product concentration) measurements. Scripted protocols allow for a flexible and highly standardized approach to handle and treat individual cultivations. As a result, more precise and reliable biological results can be obtained. While the MBR cultivation is running, samples can be processed in parallel to enable their quick and safe storage, for example by cell separation [115, 160, 161] or to analyze them directly, for example by spectrophotometric assays [115, 124, 163]. Clearly, the interplay and orchestration of many lab devices including the MBR systems within the lab robotic environment requires easy to access and compatible software and hardware interfaces, as well as scheduling software to supervise the overall process in its whole complexity [161, 163]. Up to now, several studies have been reported that hyphenate various high-throughput lab automation devices and MBR systems to obtain such data densities that enable almost holistic strain and bioprocess development [113, 115, 164–167].

In the past, the integration of robotic systems for downstream process development led to significant increases in the throughput of several standard operation procedures. For example, cell lysis has been optimized for improved release of intracellular products [168, 169], the screening of chromatographic material and binding/elution conditions was facilitated [170], and the identification of optimal compositions of aqueous two-phase systems [171] or precipitation conditions [172] has become experimentally feasible. So far only a few but highly interesting studies have been reported

[1] *ANSI/SLAS 1-2004* to *ANSI/SLAS 4-2004*, *ANSI/SLAS 1-2012*, available at www.slas.org

that combine high-throughput upstream and downstream bioprocess development [165, 166, 173]. This is certainly due to the much higher complexity of such wholistic approaches. For further information on recent developments in high-throughput downstream processing the reader is referred to the following reviews [174, 175].

2.5. Scalability of microbioreactor bioprocesses

The purpose of MBR systems is to facilitate the development of bioprocesses that are scalable to lab-scale and, in an optimal case, up to pilot-scale. For typical robust microbial production strains, both the maximum oxygen transfer rate (OTR_{max}) [36, 176] and the k_La-value [177] are often reported to be suitable scale-up criteria. Depending on the methods used for OTR_{max} determination at MBR and bioreactor scale, correction factors can be applied [178]. Maintaining a dissolved oxygen level >20% as scale-up criterion resulted in highly comparable growth and heterologous cutinase secretion profiles of *C. glutamicum* strains in MBR (1 mL) and bioreactor (1 and 20 L) cultivations. Moreover, the biomass specific product yield $Y_{P/X}$ was comparable along the different cultivation scales and strains tested [36].

In case shear-sensitive microorganisms are cultivated, power input was proven to be another suitable scale-up criterion. By adjusting the mean power input to the same value for 10 mL MBR and 2 L bioreactor cultivations of *Streptomyces tendae*, highly comparable culture dynamics including biomass and product formation were achieved [120]. In another study, an MTP-based MBR system was validated for bioprocess development with *Streptomyces lividans* [179]. Operating conditions for the MBR and lab-scale bioreactor were chosen to match the same culture morphology, which was used as biological scale-up criterion.

As a more complex example, the simultaneous consideration of two scale-up criteria was reported to be crucial for the comparability of biomass specific yields of a soluble ribonuclease inhibitor fusion protein among cultivations in shake flasks and bioreactors [180]. The first (technical) scale-up criterion was to maintain a growth-limiting fed-batch cultivation mode, while the second (biological) scale-up criterion was to induce the culture at a certain growth rate.

A more detailed review on scale-up from MBR systems can be found elsewhere [181].

2.6. Need for automated data processing

Apart from untargeted screening approaches for the characterization of whole strain libraries, the targeted optimization of bioprocess conditions (e.g., media optimization) requires further sophisticated experimental procedures. In general, to maximize the experimental power of high-throughput MBR cultivations, the focus should rely on those experiments only from which the highest information gain can be expected. Preferably, this is done in iterative procedures consisting of dry-lab experimental planning, wet-lab MBR experimentation and dry-lab data evaluation. Here, smart Design of Experiments (DoE) is suited to accomplish this task [167, 182, 183]. Intentionally, DoE was developed to gain maximal information from a limited number of experiments and mainly uses black-box models to describe functional relationships between input factors (e.g., strains and cultivation conditions) and output factors (e.g., PIs) [184]. The importance of PI determination for quantitative phenotyping projects by automated data processing is discussed later (see sections 2.7 and 3.4).

However, with the option for high-throughput experimentation more input factors can be taken into account, increasing the predictive power of DoE approaches. Noteworthy, optimization by DoE provides information about improved conditions, but lacks mechanistic understanding unless mechanistic instead of black-box models are used [185]. The perhaps most valuable feature of DoE applications is the systematic investigation of non-intuitive input factors that are not considered to be important when relying on biased expert knowledge or educated guesses alone. In turn, DoE rapidly helps to identify bioprocess conditions (e.g., medium composition, temperature) that enable improved phenotypes and process performance indicators. An important aspect for data processing is the use of mathematical correction methods, which has been successfully applied to correct optical measurements suffering from interference [186]. Since optical measurements are widely applied in MBR systems due to fast and non-invasive data acquisition, it is important to know the experimental conditions that may cause problems with this measurement technique [186–188].

2.7. Determination of culture performance indicators by automated data processing

Through parallelized and automated MBR cultivations a variety of raw data sets are easily generated in short time. The evaluation of such amounts of data is not possible with manual procedures anymore. Therefore, the current challenge with high-throughput MBR cultivations is to extract the relevant biological and process information in a non-delayed and efficient manner. Resulting raw

data streams need to be (semi-)automatically processed and visualized in a meaningful way for interpretation of results and decision-making, compare also Figure 3.8C.

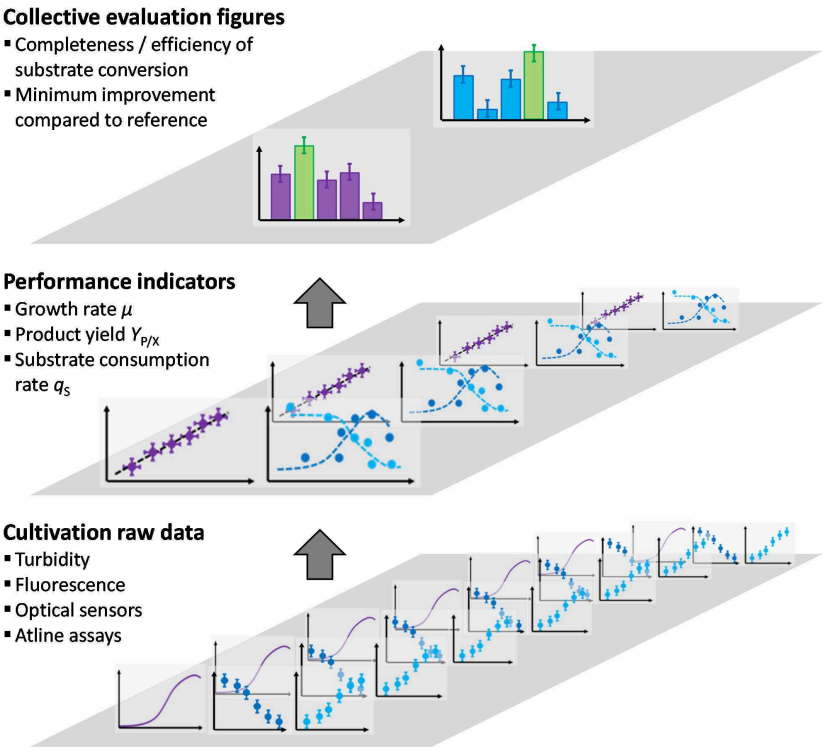


Figure 2.2.: Multilayer data processing workflow to handle the complex output from automated MBR cultivations. Raw data obtained from parallelized microbioreactor cultivations from the bottom layer, including data from online monitoring (e.g., turbidity, fluorescence) and data from atline assays obtained from transient sampling. These datasets are further processed using model-based approaches to derive important performance indicators (PIs). Finally, these PIs are evaluated in comparison to reference values or defined threshold values as minimal optimization targets. Figure reprinted from [104] with permisson.

Typically, relevant PIs are determined, including biomass and product titers (c_X, c_P), specific rates for biomass growth (μ), substrate consumption (q_S) and product formation (q_P) as well as specific yields relating substrate conversion and product formation to biomass growth ($Y_{X/S}$, $Y_{P/X}$). These PIs are usually obtained by fitting experimental data to single mathematical expressions or small process models [117, 189–194] , preferably in an automated way [163, 193]. By comparing such PIs with reference values or defined benchmark settings, good performing strain variants and/or cultivation conditions can be found (see Figure 2.2). Noteworthy, automated data processing for PI determination is not restricted to MBR cultivations but is helpful for any kind of bioprocesses conducted at

different scales. However, automated data processing is of particular importance for the experimental high-throughput nature of MBR systems.

To illustrate a specific example, the following section deals with the determination of the growth rate μ from high-throughput MBR cultivations, as growth rate is probably the most important and most often used characteristic PI in biological fitness testing.

2.8. Automated growth rate determination from high-throughput microbioreactor cultivations

The growth rate μ is an important PI, because it is a suitable metric to assess biological fitness of microbial mutant strains [195]. Typically, parallel high-throughput growth experiments in MBR systems involve strain mutants which show deviating growth behavior like different lag-phases, final biomass concentrations, maximum growth rates or even biphasic growth patterns. Often such phenotypes cannot be predicted but need to be revealed in high-throughput MBR experiments. Furthermore, strain-specific growth patterns may change completely and unpredictably in different nutrition media with, e.g., different carbon sources. Therefore, to fully characterize mutant strain libraries, many high-throughput MBR growth experiments are needed, resulting in big data sets that have to be evaluated accordingly.

To calculate the growth rate from an individual cultivation, the corresponding data subset covering only the exponential growth phase must be determined. With typically short microbial cultivation times of one day, the generation frequency of such data sets to be evaluated is very high. Most importantly, the task of biological fitness testing of microbial mutant strain libraries should follow a standardized protocol covering both wet-lab experiments and data processing to enable fair comparison. Apparently deviating results, for example a significantly higher growth rate of a mutant strain compared to the wild type strain, are ideally discovered in an automated procedure, and then followed by further in-depth manual evaluation including confirmatory growth experiments.

Efficient handling of the data load obtained by MBR systems is enabled by scripting languages which are ideally suited for rapid development of automated data processing routines, yet a very few with specific features are reported and freely available [196–198]. Here, a MATLAB code is presented that facilitates calculation of growth rates in an automated way based on biomass readings at high temporal resolution from individual cultivations conducted in parallel. This calculation is easily parallelized, as it is demonstrated for all 48 cultivations taking place simultaneously in one cultivation plate in the BioLector [199] MBR device. Growth rate calculation is based on the exponential

growth model, $\frac{dc_X}{dt} = \mu \cdot c_X$, considering the growth rate μ to be constant during the exponential growth phase. Therefore, cultivation conditions have to be defined in a meaningful way to avoid limited growth originating from insufficient maximum oxygen transfer rates or insufficient pH buffering capacity.

2.8.1. Specification of MATLAB code

The presented MATLAB code requires five input arguments:

(1) a vector of time stamps
(2) a vector of corresponding BS readings (= biomass data) blanked by the initial value
(3) a vector of corresponding BS measurement errors
(4) a BS value as user-defined limitation of quantification (LOQ)
(5) the adjusted R^2 that has to be reached for the regression-based growth rate calculation from the data set

The code, depicted in Figure A.1, is designed to detect the exponential growth phase from a given data set by iteratively calculating a growth rate. From that calculation, several stopping criteria serving as metric for recognition of the exponential growth phase have to be fulfilled. During the first iteration, a time series containing all BS signals between the first measurement where BS exceeds the LOQ and the final measurement is evaluated. Typically, the final measurement is taken during the stationary phase, that is after completion of growth. If the stopping criteria (cf. below) are not met, the final measurement is removed for the next iteration, that is the penultimate measurement is set as new final measurement.

Three conditions are defined as stopping criteria: First, the adjusted R^2 from the regression must reach a certain threshold, a value > 0.99 was found to be suitable. This criterion alone is not sufficient, since the data used for regression show a high temporal resolution, so that adjusted R^2 is still satisfied if several non-wanted data points from the stationary phase are included. Thus, the second stopping criterion is that the increase in biomass in the last measurement cycle needs to be higher than in the previous one. Finally, the third criterion is that these two biomass increases must not be negative, which is sometimes observed as technical measurement artifact during transition from exponential to stationary phase.

The blanked BS readings are transformed by natural logarithm, $\overline{c_X} = ln\,(c_X)$, to linearize the growth data for calculating the growth rate according to $\mu = \frac{1}{c_X} \cdot \frac{dc_X}{dt} \approx \frac{\Delta ln(c_X)}{\Delta t}$. Although nonlinear regression (NLR) is considered as "gold-standard", linear regression (LR) after data transformation results in highly comparable growth rates as discussed below. NLR requires an initial guess

since it is an iterative procedure, and this initial guess is reasonably calculated by LR from transformed data. Because each BS signal c_X is connected with a corresponding measurement error δ_{c_X}, the growth rate calculation is performed as weighted linear regression (WLR). Therefore, errors are transformed accordingly by $\delta_{\overline{c_X}} = ln\,(c_X)' \cdot \delta_{c_X} = \frac{1}{c_X} \cdot \delta_{c_X}$ and the inverse squared transformed errors, $\delta_{\overline{c_X}}^{-2}$, are used as weights.

2.8.2. Detailed depiction of data processing

To obtain valid growth rates, cultivation conditions must be applied ensuring that growth is limited by internal factors of the cell only and not by external factors. This is seen for *C. glutamicum* and *E. coli* in the left panels of Figure 2.3A and Figure 2.3B, respectively.

For both growth experiments, conducted in the BioLector MBR, conditions were chosen that result in an exponential growth phase of the cultures. *C. glutamicum* was grown in defined CgXII medium [56] with 10 g L^{-1} glucose as carbon source (left panel of Figure 2.3A). For the chosen batch mode of operation with 1000 µL cultivation volume and the microplate shaken at a frequency of 1000 rpm, the BS signal increases exponentially from approx. 15 h to 24 h. The DO signal drops accordingly, showing a sharp rise back to 100% at approximately 24 h, which confirms glucose depletion at this point. Thus, the cultivation conditions are determined to be suitable for *C. glutamicum* strain screening and calculation of growth rate. In the case of *E. coli*, two growth phases can be derived from the online monitored biomass and DO signals (left panel of Figure 2.3B). Here, the defined medium M9 [201] with 20 g L^{-1} glucose as carbon source was used, with a filling volume of 1000 µL and a shaking frequency of 1400 rpm. An exponential increase of biomass signal is seen until approx. 7 h, with the DO signal dropping accordingly until a sharp rise at the same time point. Afterwards, a second, retarded growth phase is visible until approximately 15 h, accompanied by a slowly increasing DO signal. Presumably, excreted acetate and other overflow metabolites from the first exponential growth phase are consumed now, which is a known phenomenon for *E. coli* [202].

After applying the MATLAB code on the data shown in the left panels of Figure 2.3, the resulting processed data is depicted in the corresponding right panels. The single exponential growth phase of the *C. glutamicum* culture is detected precisely, also for a second replicate cultivation. For *E. coli*, exhibiting a first exponential and a second non-exponential growth phase, the MATLAB code is able to detect the first phase safely for all three replicate cultivations, of which one is shown. The growth rate for *C. glutamicum* and *E. coli* is calculated to 0.46 h^{-1} and 0.61 h^{-1} on average, respectively. In case NLR instead of WLR is used, the same growth rates are determined (cf. Table 2.2), indicating that WLR of transformed data is a suitable and reliable method for the determination of growth rates.

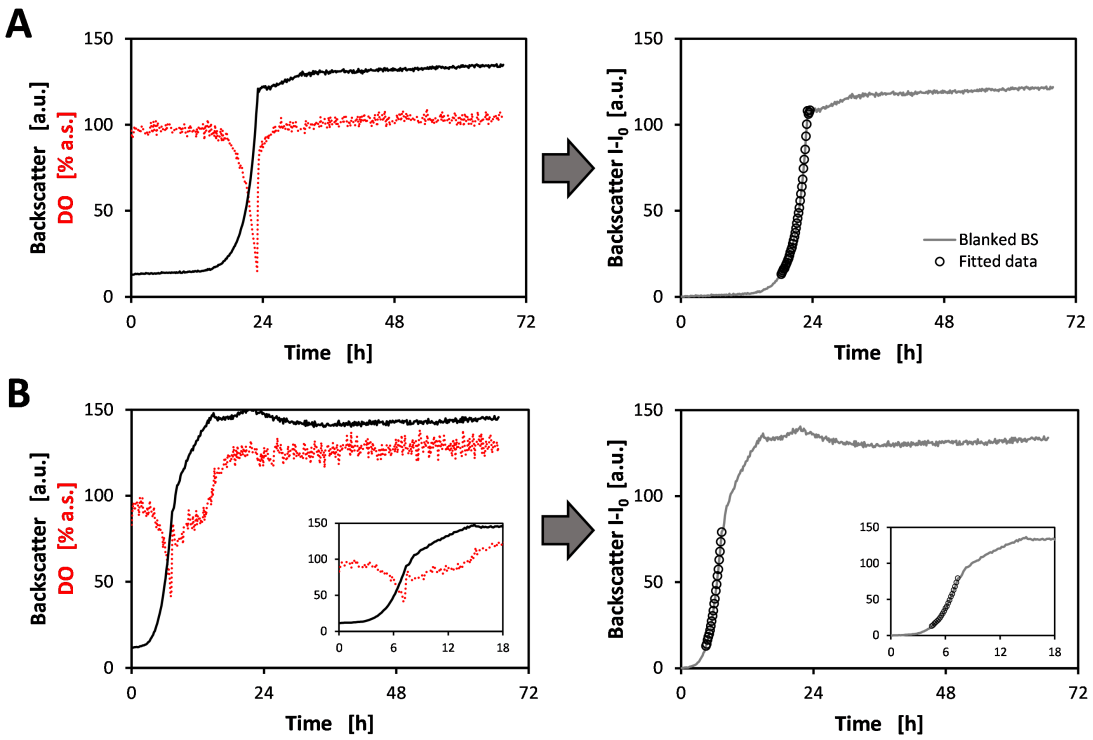


Figure 2.3.: Growth kinetics of *C. glutamicum* and *E. coli* from BioLector cultivations and depiction of data processing for automated growth rate calculation. (A): Online monitored backscatter (BS) and dissolved oxygen (DO) signal for C. glutamicum in CgXII medium with 10 g L^{-1} glucose, a filling volume of 1000 µL and a shaking frequency of 1000 rpm. The panel on the right shows the processed biomass data, i.e. the blanked BS signal ($I–I_0$) and data points determined automatically for calculation of growth rate. Propagated measurement error for blanked BS signal was calculated to 0.39 a.u. (B): Online monitored signals like in part A, but for *E. coli* in M9 medium with 20 g L^{-1} glucose, a filling volume of 1000 µL and a shaking frequency of 1400 rpm. Propagated measurement error for blanked BS signal was calculated to 0.34 a.u. Right panel analogous to the one in part A. Insets in part B magnify the first 18 h of cultivation. Measurement cycle time for recording BS and DO signals was set to 9 min for both *C. glutamicum* and *E. coli* cultivations. Figure reprinted from [200].

A further application example of the presented code on the re-evaluation of previously published growth rates is found in Chapter A.

2.8.3. Further considerations

Biomass monitoring in the BioLector MBR device relies on BS measurements, which is the key feature to temporal high-density data acquisition in MTP cultivations [199, 203]. In contrast to the determination of biomass concentration by cell dry weight, optical measurement for biomass determination

Table 2.2.: Comparison of regression methods for automated growth rate calculation.

Cultivated organism		Growth rate μ [h^{-1}] with 95% CI from regression method	
		WLR after ln-transformation	**NLR**
C. glutamicum	Replicate #1	0.455 (0.448 – 0.462)	0.456 (0.448 – 0.463)
	Replicate #2	0.456 (0.448 – 0.464)	0.457 (0.449 – 0.465)
E. coli	Replicate #1	0.601 (0.580 – 0.622)	0.603 (0.582 – 0.624)
	Replicate #2	0.615 (0.598 – 0.632)	0.616 (0.599 – 0.633)
	Replicate #3	0.617 (0.599 – 0.635)	0.618 (0.600 – 0.636)

Growth rates with corresponding 95% confidence interval (CI) are shown for each replicate cultivation of *C. glutamicum* (n = 2) and *E. coli* (n = 3). *WLR* weighted linear regression. *NLR* non-linear regression. Growth data for *C. glutamicum* and *E. coli* were kindly provided by Simon Unthan and Viola Pooth (both IBG-1, Forschungszentrum Jülich), respectively.

via BS was found to be organism specific [178]. Hence, strain specific correlations for BS and optical density measurements need to be determined and have been reported [204]. Issues like optical crosstalk or cell morphology have to be considered for biomass determination by optical measurements in general [186, 187, 205]. Consequently, for meaningful interpretation of screening results, biomass calibration should be an integral part of strain screening campaigns with the BioLector device and other MBRs. Furthermore, cultivation conditions (media composition, filling volume, shaking frequency) should be verified to fulfill the underlying assumptions of the exponential growth model since the presented MATLAB code calculates growth rates based on this model. Therefore, conditions causing diauxic growth behavior, for example oxygen limitation or the use of complex media, may result in erroneous growth rate calculation although the presented code was shown to reproducibly detect the exponential phase from multiphasic growth patterns (cf. Figure 2.3B). In such cases, a differential method that calculates a dynamic growth rate over time [206] may be more suitable for data interpretation for changing growth regimes. This holds especially true for cultivations of microorganisms that do not show an exponential growth [179] or exhibit complex morphology [207]. However, it is reasonable to assume that the presented code can be also applied to different growth conditions or changed media compositions, if the obtained data still match to the exponential growth model.

The use of MBR systems in high-throughput mutant strain screening campaigns and accelerated microbial bioprocess development easily produces a high amount of data. Consequently, resulting data needs to be processed in an automated, standardized and efficient way. Using standardized data output from MBR in spreadsheet formats, MATLAB routines can be implemented which facilitate elevated experimental throughput. Most importantly, the application of standardized growth rate calculation methods enables a fair comparison of screened strains regarding biological fitness.

3. Robotics-integrated microbioreactor methods for phenotyping of cutinase secretion by different *C. glutamicum* strains

This chapter starts with basic characterization of standard operating conditions. Special adjustments are made with respect to cutinase secretion performance as screening objective. In order to derive versatile performance indicators to characterize mutant strains, the well-established method of harvesting replicate cultivations is compared to an improved method that involves repetitive small-volume sampling from any individual culture of a high-throughput MBR cultivation experiment, including the propagation of random errors covering all experimental steps. Finally, the new method is applied for the characterization of a set of genome-reduced *C. glutamicum* strains with respect to cutinase secretion. Since this chapter deals with the characterization of MBR workflows, only swift interpretation and discussion of biological results is made. A part of this chapter is based on *Manuscript II*.

Author contributions:
Johannes Hemmerich wrote the chapter, designed and performed experiments, analyzed the data, interpreted the results, and prepared the figures. Marco Oldiges, Stephan Noack and Wolfgang Wiechert helped with writing section 3.4. Sebastian J. Reich planned and performed experiments, and analyzed the data in sections 3.2 and 3.3. Carmen Steffens planned and performed experiments, and analyzed the data in section 3.4.

3.1. Characterization of standard operating conditions

For aerobic bioprocesses conducted in MTP-based shaken MBR systems, aeration and the avoidance of oxygen limitation are an important issue because the latter induces, for example, by-product formation via overflow metabolism. Moreover, non-exponential growth phases due to consumption of these by-products may occur, which in turn makes the standard exponential growth model inappli-

cable to derive a growth rate as characteristic (i.e, PI), as discussed in section 2.7. Sufficient aeration of the culture is achieved by a combination of bioprocess parameters, of which the filling volume and shaking frequency for the used type of MTP (here, Flowerplate) are most easy to adjust. The maximal oxygen transfer rate OTR_{max} of the applied system is determined by these parameters mainly (cf. also section 2.3). To ensure that the maximal oxygen uptake rate OUR_{max} reached by the culture does not exceed the OTR_{max}, the amount of main carbon source (here, glucose) needs to be adjusted appropriately.

In this study, a filling volume of 800 µL and a shaking frequency of 1300 rpm were chosen as standard conditions for the cultivation of *C. glutamicum* in CgXII defined medium containing 20 g L^{-1} glucose as main carbon source. This was verified against another filling volume (1000 µL) and other glucose concentrations (10, 15 and 40 g L^{-1}), and resulting culture dynamics are depicted in Figure 3.1. For the filling volume of 800 µL (Figure 3.1A and C), superimposed biomass time courses for 10, 15 and 20 g L^{-1} glucose are seen, which is also the case for the corresponding DO signals.

Regarding sufficient aeration, the time course of dissolved oxygen (DO) signal is most valuable for investigation. All aforementioned DO signals show a continuously descending signal, ending with a sharp increase back to the initial signal intensity. The corresponding biomass signals increase accordingly until they reach a plateau at the time the DO signals rise back to their initial values, which indicates the depletion of glucose (Figure 3.1A and C). In contrast to that, the DO signal from cultivations with 40 g L^{-1} glucose indicates a limitation in oxygen transfer until approx. 15 h, coinciding with a slowdown in biomass accumulation (i.e., linear increase), beginning at the same time when the cultures growing on 20 g L^{-1} glucose reach stationary phase. Regarding a filling volume of 1000 µL (Figure 3.1B and D), a slight DO limitation is seen for cultivations growing on 20 g L^{-1} glucose, while the other cultivations with 10 and 15 g L^{-1} glucose behave similar to their counterparts with a filling volume of 800 µL. The metabolization of 20 g L^{-1} glucose as main carbon source by *C. glutamicum* without running into oxygen limitation is considered a major criterion for the selection of standard operation conditions. In these terms, the chosen operation conditions of 1300 rpm and 800 µL as shaking frequency and filling volume, respectively, are suitable.

Next, the chosen standard operating conditions for MBR cultivations were characterized with respect to the quasi-continuous online monitoring of the BS signal, which serves for real-time biomass quantitation. In order to exclude cell specific effects during characterization experiments, a turbidity standard ("AmcoClear", GFS Chemicals, Columbus, OH/USA) is used to provoke stable BS signal that are not subject to any biological impact. Briefly, a Flowerplate was filled with different volumes of turbidity standard of 4000 nephelometric turbidity units (NTU). This Flowerplate was incubated in a

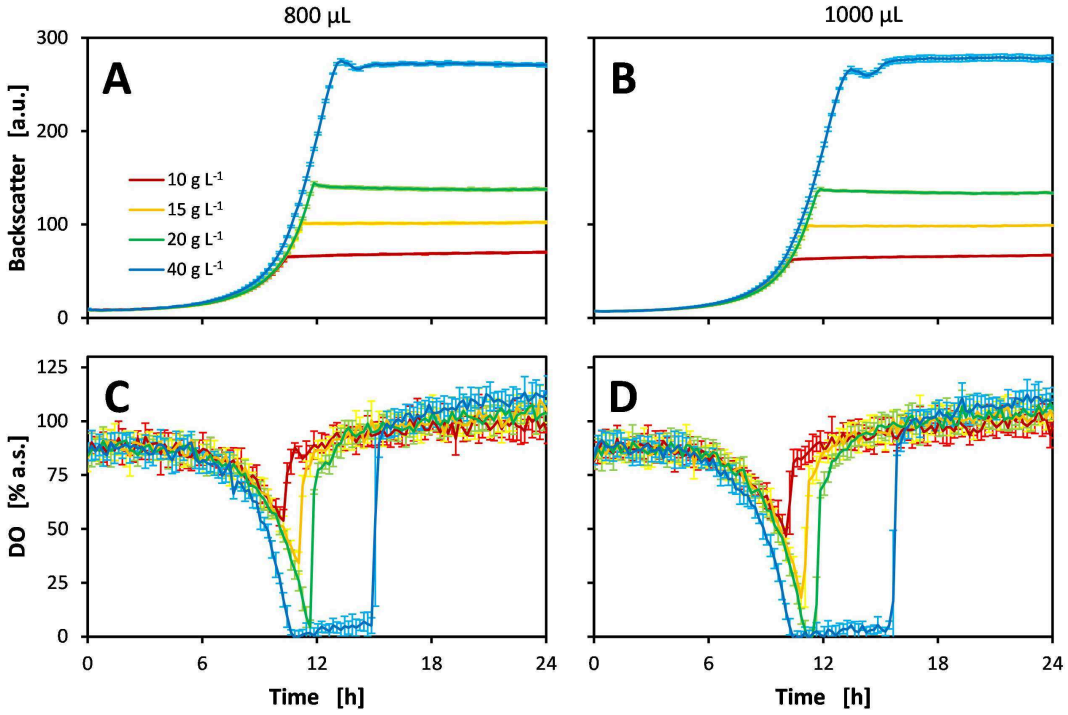


Figure 3.1.: Evaluation of operating conditions for cultivation of *C. glutamicum* in CgXII medium with different glucose concentrations. The chosen standard conditions (20 g L^{-1} glucose, 800 µL, 1300 rpm) were compared against a higher filling volume (1000 µL) and other glucose concentrations (10, 15 and 40 g L^{-1}). Depicted are biomass formations for the different glucose concentrations in a filling volume of 800 and 1000 µL (panels A and B, respectively). Panels C and D show the time courses of corresponding DO signals. Data for each condition originate from six culture replicates and are shown as mean with standard deviation as error bars. *DO* dissolved oxygen.

BioLector device using the standard cultivation protocol, but with an increasing shaking frequency from 500 to 1500 rpm (100 rpm increases after 30 min of BS signal recording, resulting in eight successive BS readings). This procedure, including the preparation of a new Flowerplate, was repeated three times. The behavior of the BS signal for the turbidity standard in dependence of filling volume and shaking frequency is depicted in Figure 3.2.

For the whole range of shaking frequencies tested (500 to 1500 rpm), filling volumes of 900 to 1200 µL result in hardly detectable differences in the BS signals. This is inferred from the corresponding data points in Figure 3.2A that do not deviate from each other and follow a straight horizontal line. When comparing the maximum relative deviations (MRDs) of the BS signals, $MRD\,[\%] = \frac{BS_{max}-BS_{min}}{BS_{Average}} \cdot 100$, of each filling volume data series, MRDs of 0.3%, 0.5%, 1.0% and 2.0% are found

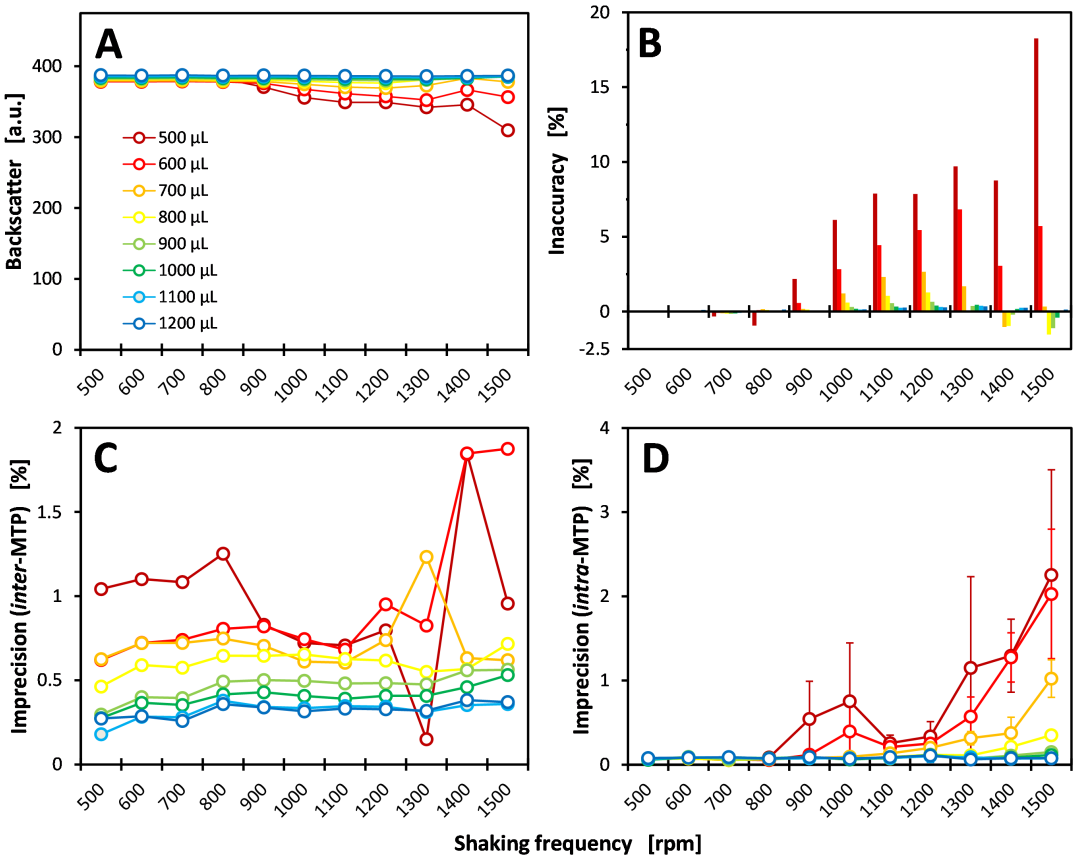


Figure 3.2.: Behavior of BS signal of a turbidity standard (4000 NTU) in the BioLector MBR system for different filling volumes (500 to 1200 µL) and shaking frequencies (500 to 1500 rpm) in 48-well Flowerplates. (A): BS signal with increasing shaking frequency. After 30 min of data acquisition, the shaking frequency was incremented by 100 rpm. Data is shown as mean value from three independently prepared and measured Flowerplates. (B): Inaccuracy of BS signals for different filling volumes and shaking frequencies. Inaccuracy refers to as the relative deviation of the BS signal at a certain shaking frequency to the BS signal of each filling volume at 500 rpm. (C): Imprecision of BS measurements for independently repeated measurements (*inter*-MTP imprecision), given as relative standard deviation from three independent measurements for each condition. (D): Imprecision values for repeated BS measurements (*intra*-MTP imprecision) for a given filling volume and shaking frequency, depicted as mean values of *intra*-MTP imprecision with standard deviation from three independently measured Flowerplates. Colors in each panel encode the different filling volumes, according to the legend shown in panel A. *NTU* nephelometric turbidity units.

for the filling volumes of 1200, 1100, 1000 and 900 µL, respectively. A slightly higher BS signal deviation is observed for filling volumes of 800 and 700 µL in the range of 1000 to 1300 rpm, as indicated by the slightly lowered data points for these shaking frequencies in comparison to their respective

values at 500 rpm and the BS signals of the filling volumes 1200 to 900 µL measured in the same rpm-range (Figure 3.2A). Corresponding MRDs are calculated to 3.0% and 3.7%, respectively.

In case filling volumes of 600 and 500 µL are applied, a much higher deviation from the corresponding starting value at 500 rpm is observed for shaking frequencies above 900 rpm. These data points are located below all other data points for all shaking frequencies above 900 rpm, which means the BS signal from turbidity standard is lower for these conditions. The highest deviation is observed for 500 µL at 1500 rpm. This is also reflected by MRDs of 7.2% and 21.6% for data series on 600 and 500 µL.

To further evaluate the impact of shaking frequency and filling volume on the BS signal, inaccuracy of BS measurement was calculated and is shown in Figure 3.2B, corresponding to the data in shown in Figure 3.2A. Inaccuracy of the BS signal was defined for each filling volume as the relative deviation of the BS signal obtained at a certain shaking frequency in comparison to the respective BS signal at 500 rpm, $Inaccuracy\,[\%] = \left|1 - \frac{BS_{xrpm}}{BS_{500rpm}}\right|$. As can be seen in Figure 3.2B, when using a filling volume of 500 µL, shaking frequencies of 500 to 800 rpm are not critical in terms of inaccuracy since from 500 to 700 rpm inaccuracies are hardly detectable, and at 800 rpm inaccuracy is as low as -1.04%. Beginning with 900 rpm, inaccuracy increases from 2.79% to a value of 19.4% at 1500 rpm.

The same trend is observed for a filling volume of 600 µL, although the considerable range of shaking frequencies subjected to detectable inaccuracy is from 1000 rpm (2.74% inaccuracy) to 1500 rpm (4.71%). The remaining filling volumes are inconspicuous in terms of inaccuracy for all shaking frequencies tested, except for 700 µL where only minor inaccuracies are observed from 1100 rpm (2.32%) to 1400 rpm (-1.12%). For the filling volume of 800 µL, noticeable yet very small inaccuracies are observed for a shaking frequency of 1200 rpm (1.2%) and 1500 rpm (-1.81%). Most importantly, for a filling volume of 800 µL operated at 1300 rpm, which is the operating condition determined to be ideally suited in terms of oxygen transfer capacity (cf. above), inaccuracy is as little as -0.09%, meaning that this condition is also suitable in terms of inaccuracy of the BS signal.

Finally, the BS signal was characterized with respect to imprecision in two senses: the *inter*-MTP imprecision, depicted in Figure 3.2C, is defined as the relative standard deviation calculated from the three independently measured BS signals, which are the mean values for eight successive BS readings obtained three times. Therefore, the *inter*-MTP imprecision characterizes the reproducibility of BS reading between independently run BioLector experiments.

In addition, the *intra*-MTP imprecision is defined as relative standard deviation of the eight successive BS readings for each condition (i.e., filling volume and shaking frequency). The *intra*-MTP imprecision was determined for each of the three independent experiments. Therefore, the *intra*-

MTP imprecision characterizes the reproducibility of the BS measurement itself in dependence of filling volume and shaking frequency. Results are shown in Figure 3.2D as mean values of *intra*-MTP imprecision with corresponding standard deviations as error bars (n = 3).

The *inter*-MTP values of BS signals in dependence of filling volume and shaking frequency are depicted in Figure 3.2C. In general, it is seen that with increasing filling volume *inter*-MTP precision is lowered, with apparent minimal *inter*-MTP imprecisions reached with a filling volume of 1100 µL, as the data points for 1200 µL do not deviated from those for 1100 µL. For the shaking frequencies tested, values of *inter*-MTP imprecision with 800 µL are typically at about 0.5%. For a filling volume of 700 µL, *inter*-MTP imprecision values show a tendency to higher values compare to 800 µL, with a distinct increase to about 1.2% at 1300 rpm for unknown reasons. When applying filling volumes of 500 and 600 µL, *inter*-MTP imprecisions show a noisy behavior which also cannot be explained here. However, this noisy behavior in terms of *inter*-MTP imprecision for 500 and 600 µL coincides with a distinct increase in inaccuracy (cf. Figure 3.2B), indicating that these operating conditions should be avoided. However, with *inter*-MTP imprecision values below 1% for most of operation conditions and below 0.5% for filling volumes of 800 µL and higher for the tested range of shaking frequencies, the reproducibility of BS measurements between independent BioLector experiments can be considered as being high.

The *intra*-MTP values of BS signals in dependence of filling volume and shaking frequency are depicted in Figure 3.2D. In general, *intra*-MTP imprecision of the BS signals is found to be rather small (max. 2.25 ± 1.25% for 500 µL at 1500 rpm) in comparison the maximum relative standard deviation of 5% for the optical measurement in the BioLector device, according to the manufacturer. However, similar to the trends observed for inaccuracy (cf. Figure 3.2B), some patterns can be observed for the determined *intra*-MTP imprecision values. Filling volumes of 1200 to 800 µL result in hardly detectable *intra*-MTP imprecisions (0.06% to 0.35%) for all shaking frequencies tested. Considerably higher *intra*-MTP imprecisions are found for the remaining filling volumes of 700, 600 and 500 µL for shaking frequencies starting with 1300, 900 and 800 rpm, respectively, as seen in Figure 3.2D. In particular, for the envisaged standard operating conditions of 800 µL and 1300 rpm an *intra*-MTP imprecision of as low as 0.11 ± 0.03% is determined.

Conclusively, the contribution of both *inter*-MTP imprecision (0.55%) and *intra*-MTP imprecision (0.11%) of the BS signal to the overall precision of culture derived PIs like growth rate determined from BioLector cultivations is considered negligible. Furthermore, analysis of a NTU dilution series in order to check linearity of the BS signal revealed essentially the same values for slopes and offsets within the observed error ranges (cf. Figure B.1). Consequently, the chosen standard operating conditions

are also suitable in terms of BS signal imprecisions.

The deviating behavior of the BS signal (expressed as inaccuracy and imprecision, cf. paragraphs above) in dependence of the shaking frequency and filling volume can be attributed to the corresponding changes in power input and in the amount of bulk liquid accessible by the measuring optics, respectively, at least with attempts of explain in a qualitative manner. Figure 3.3 shows a schematic on the hydrodynamical behavior of the bulk liquid in MTPs in dependence of shaking frequency and filling volume, as well as the impact on optical BS measurement in the BioLector MBR device.

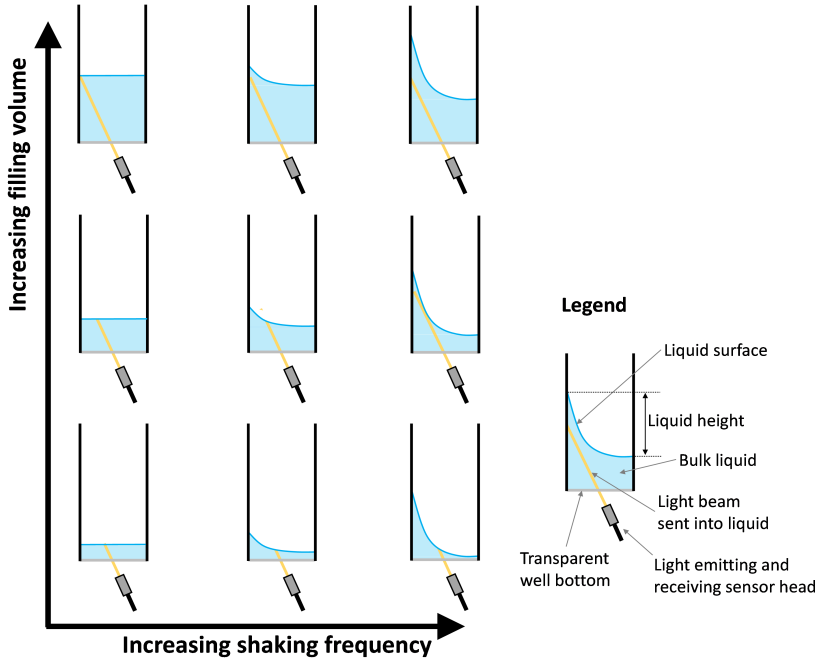


Figure 3.3.: Schematic drawing on impact of fluid dynamics in dependence of shaking frequency and filling volume on BS measurement in the BioLector MBR device. Depicted is a cross-sectional side-view of a MTP well with a transparent bottom. Beneath, the optical sensing head of the BioLector is shown. With increasing filling volume and shaking frequency, the surface of the agitated bulk liquid forms a sickle to different extends. Drawing is based on published photographic analyses on this topic [148, 156], as well as slow-motion videos available at enzyscreen.com.

With increasing shaking frequency, the power input into the bulk liquid also increases, which finally enables the rotational movement of the bulk liquid [152]. The higher the shaking frequency, the more pronounced the sickle formed by the surface of the rotating bulk liquid, meaning that the liquid height above the well center decreases and increases with increasing distance from the well center

to the outer perimeter of the well [148, 156], cf. Figure 3.3. This could provide an explanation why for low filling volumes at high shaking frequencies the BS signal is lower compared to higher filling volumes when operated at the same shaking frequency. The reason would be the lowered "liquid depth" (caused by a lowered liquid height) that can be penetrated by the light sent into the liquid, meaning that in this case there is a substantial lower number of suspended particles to scatter back the incoming light. In Figure 3.3, such reduced light depth is schematically shown as yellow light beam of varying length in dependence of shaking frequency and filling volume.

Depending on the well geometry (e.g., round, square or flower-shaped), the resulting liquid distribution of the rotating bulk liquid differs for the same shaking frequency, as inferred from photographic analyses [156] and different liquid heights measured above the well center for a variety of well geometries [111]. The rotating liquid in the commercially available flower-shaped MTPs (i.e., Flowerplates) takes a U-form, which is more pronounced with increasing shaking frequency [155]. Furthermore, it can be assumed that the position of the sensor head and its angle towards the transparent well bottom is of substantial influence on the BS signal. However, in this work a commercial BioLector device is used where these properties are fixed. Thus, further investigations of such on the behavior of light scattering for different filling volumes and shaking frequencies were not conducted.

The relevance of the collected data series shown in Figure 3.2 is that these allow the determination of operating conditions (i.e., filling volume and shaking frequency) that enable a stable backscatter measurement signal with high accuracy and low imprecision. Therefore, the chosen standard operating conditions of 800 µL and 1300 rpm are now considered to be sufficiently validated.

3.2. Adaption of microbioreactor workflows with respect to protein secretion by *C. glutamicum* as screening objective

A major part of this work is the characterization of recombinant *C. glutamicum* strains with respect to heterologous protein secretion. To achieve a sufficiently high throughput at manageable effort, MBR based cultivation protocols are versatile tools. However, such protocols have been established earlier for other screening objectives than specified in this study [160, 208] and thus, these protocols need to be revised with respect to the changed objectives concerning heterologous cutinase secretion with *C. glutamicum*.

3.2.1. Identification of proteolytic degradation of cutinase activity in *C. glutamicum* cultivation supernatants

During first MBR growth experiments, an increase in extracellular cutinase activity accompanying biomass formation is observed, as expected [36], but with inacceptable uncertainty from the cutinase assay carried out in analytical triplicates (Figure 3.4A). This observation was made for different process conditions and strains tested, indicating an insufficient MBR screening procedure. A more detailed look at the results obtained from the analytical procedure revealed that with each assay conducted from the same sample, cutinase activity decays (insets in Figure 3.4A), which is again a consistent finding among a variety of samples. Since cutinase is reported to be a very robust enzyme in cultivation broth towards presence of cells and temperature [209], it was assumed that the presence of some proteolytic activity was responsible for the high imprecision of cutinase activity determination. This hypothesis was tested by splitting a sample of culture supernatant into four parts, to which either one volume of buffer (serving as control), buffer containing 1 mM EDTA, buffer containing a protease inhibitor cocktail (PIC, "Complete", Roche Diagnostics), or buffer containing both EDTA and PIC were added. Resulting samples were incubated at room temperature for approximately 100 min and analyzed for cutinase activity over that period. This was conducted twice. Resulting time profiles of cutinase activity are shown in Figure 3.4B.

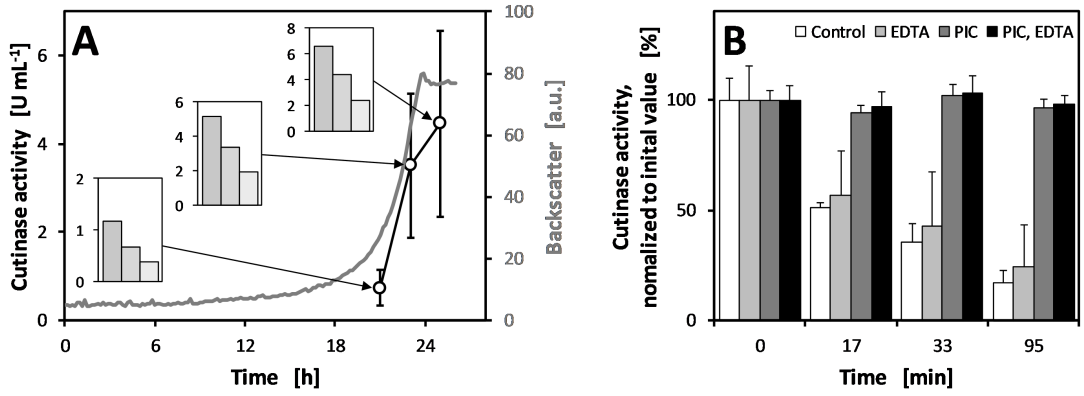


Figure 3.4.: Investigation on observed high measurement error for cutinase activity from MBR cultivations. (A): The observed high error of cutinase activity determination from MBR cultivations is based on a decay of cutinase activity from the first to the last analytical replicate (insets in panel A), which is seen in all culture samples. (B): The assumption of protease activity in *C. glutamicum* culture supernatants is verified by repeated cutinase activity determination from culture supernatant samples, which are split into four aliquots (control: sample plus buffer; EDTA: sample plus buffer containing 1 mM EDTA; PIC: sample plus buffer containing protease inhibitor cocktail; PIC, EDTA: sample plus buffer containing protease inhibitor cocktail and 1 mM EDTA). Shown data represent mean with min/max values as error bars from two replicate measurement series.

Both control and EDTA treated sample aliquots exhibit a decay of cutinase activity over time, which is not the case for the sample aliquots that are treated with the protease inhibitor cocktail. From these findings, two conclusions can be drawn: First, cultivation supernatants of *C. glutamicum* secreting cutinase contain proteolytic activity. This is rather surprising since *C. glutamicum* is described not to secrete proteases degrading heterologous proteins [17, 210], like is the case for, e.g., *B. subtilis* [211]. Second, the discovered protease is presumably no metallo-protease since EDTA has no positive effect on stability of cutinase activity. As a next step, further investigations envisage the impact of pre-culturing (i.e., the inoculation from a pre-culture or from a frozen cryo aliquot) and the presence of urea in CgXII medium. The first aspect would cover possible protease contamination that may originate from lysed cells during the freezing process of cryo-aliqouts, while the second aspect investigates the pH-stabilizing effect of urea in CgXII medium. Corresponding results are shown in Figure 3.5.

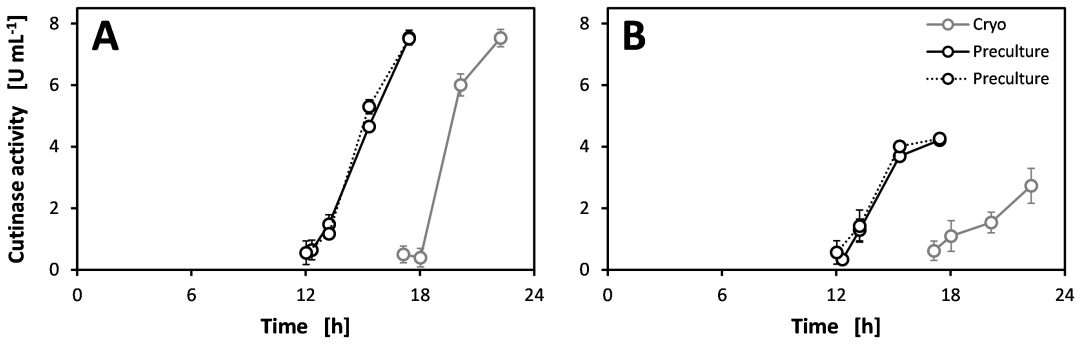


Figure 3.5.: Cutinase activity from sequentially harvested cultivations of *C. glutamicum* strain WT-NprE in CgXII medium with 5 g L^{-1} urea (panel A) or without urea included (panel B). Cultivations were inoculated using a thawn cryo vial or a pre-culture, as indicated. At a backscatter signal of 15 a.u., cultures were automatically induced with 100 µM IPTG. For clarity reasons, only cutinase time profiles are shown, which follow biomass growth as cutinase production is growth coupled.

When comparing this data from cultures that are either inoculated from a cryo culture or from an intermediate pre culturing step, it is seen that the first strategy introduces a lag-phase of approximately 6 h until growth coupled cutinase secretion starts, both for culture with urea added or not. Furthermore, it is seen that the time interval of cutinase activity accumulation remains nearly the same, which is about 5 h with urea added (Figure 3.5A) and about 6 h when urea is omitted (Figure 3.5B). On the other hand, it is seen that the presence of urea is beneficial for cutinase secretion since the obtained maximal cutinase activities are approximately doubled when urea is added to the medium compared to the cultivation without urea (ca. 7.5 U mL^{-1} vs. ca. 3 – 4 U mL^{-1}, cf. Figure 3.5A

vs. Figure 3.5B). This is rather surprising since the omission of urea, whose absence causes a drop of pH over time, does not affect the growth of *C. glutamicum* cultures (see Figure B.2). Hence, pH value of the cultivation medium seems critical for optimal cutinase secretion but not for growth of *C. glutamicum* (compare Figure 3.5A and B, as well as Figure B.2). This assumption is supported by the observation that in bioreactor cultivation samples no decay in cutinase activity from analytical triplicates is found (data not shown). In bioreactor cultivations, pH is controlled at a value of 7 which was not possible in MBR cultivations. Furthermore, when urea is not present as pH stabilizing agent, inoculation from a cryo-culture negatively affects apparent cutinase secretion efficiency, since in this case only 2.7 ± 0.6 U mL^{-1} are achieved, compared to 4.1 ± 0.2 U mL^{-1} in the cultures that are inoculated from a pre-culture (Figure 3.5B). In contrast, when urea is added as pH-stabilizing agent during growth, the inoculation strategy has no impact on the final extracellular cutinase activity, which is 7.5 ± 0.3 U mL^{-1} or 7.5 ± 0.0 U mL^{-1} for cultivations started from a cryo aliquot or from a pre-culture, respectively, as depicted in Figure 3.5A.

In summary, the observed proteolytic decay in cutinase activities is assumed to originate from intracellular proteases of *C. glutamicum* released from lysed cells due to the freezing process during the preparation of cryo strain aliquots. Furthermore, main cultivations started from freshly grown pre-cultures show a reduced culture time. Therefore, pre-culturing steps with subsequent washing before inoculating the MBR main cultivation are considered to be necessary for strain phenotyping with respect to cutinase secretion.

3.2.2. Induction profiling

The time point of induction and the induction strength (i.e., inducer concentration) have been shown often to be important bioprocess parameters, which both need to be optimized, preferably in a combinatorial approach. Using MBR systems, such optimization can be conducted efficiently and is often referred to as "induction profiling" [36, 159, 162, 212]. In order to determine the optimal induction strategy for several *C. glutamicum* strains, induction profiling was done for six strains in total: WT, Δ*cglMRR*, MB001, W65, W127 and zwf^{FBR}. These strains were tested for cutinase secretion (SP: NrpE, henceforth denoted as *strain*-NprE), in dependence of induction strength (i.e., IPTG concentrations of 50, 100, 200 and 500 µM) and biomass concentration at the time point of induction (i.e., at BS signals of 15, 30 and 45 a.u., corresponding to an early, mid-time and late stage of the growth phase). A MBR protocol (cf. section 8.7.2) was used that automatically adds IPTG to the desired final concentration upon reaching the pre-defined BS threshold. The protocol employs a liquid handling system (LHS) integrated with the BioLector MBR system. Per condition (i.e., a specific combination of IPTG concen-

tration and biomass signal to be reached for induction) four replicate cultures were grown, which were harvested after the addition of IPTG in a pre-defined sequential manner to obtain a temporal resolution of extracellular cutinase activity. This procedure is common technology to obtain time resolved concentrations when using small scale cultivation systems [116, 160]. Due to the technical specification of the actually used LHS, for addition of IPTG and harvest of the culture volume the shaking movement of the incubated Flowerplate needs to be paused.

Out of the six strains tested, for three strains the resulting time courses of extracellular cutinase activity are shown in Figure 3.6, representing a small selection of strains that include strain WT, strain W65 with genomic deletions previously classified as essential [55] and strain W127 with genomic deletions previously classified as non-essential [48]. Panels A, B and C correspond to the three cultured *C. glutamicum* strains (WT-NprE, W65-NprE and W127-NprE, respectively) while panels 1, 2 and 3 correspond to the BS signals at which the IPTG addition was conducted (15, 30 and 45 a.u., respectively). These BS signals represent an early, mid-time or late stage of the growth phase. Thereby, in each panel the time courses of extracellular cutinase activities are shown for a specific combination of strain and stage of the growth phase when induction was done. For each of the resulting nine specific combinations, results are shown for four tested IPTG concentrations. For clarity reasons, only the extracellular cutinase activities are depicted.

For the parental strain WT-NprE, induction at lower biomass concentrations, corresponding to the early exponential growth phase at a BS signal of 15 a.u. (Figure 3.6A1), is preferred since this results in higher cutinase activities compared to the other later IPTG induction at higher BS signals (Figure 3.6A1 to A3). With continuing growth, reflected by increasing biomass concentrations at time points of induction, final cutinase activities become lower (cf. Figure 3.6A2 and A3, induced at 30 and 45 a.u., respectively). This trend can be observed for all the IPTG concentrations tested. However, when inducing in the beginning of the growth phase, which results in the highest observed extracellular cutinase activities, there is hardly any impact of IPTG concentration on cutinase activity time profiles seen (Figure 3.6A1). Since all cutinase activity time profiles show saturation kinetics characterized by a steep increase for the first measurements followed by a decrease, it can be assumed that the sampling time points cover the complete growth phase.

This behavior is also found for strain W65-NprE (Figure 3.6B1 to B3). When induced at BS signals of 15 and 30 a.u. (Figure 3.6B1 and B2, respectively) and independent of the applied IPTG concentration, a constant increase in cutinase activity is seen. Only when induced at a BS signal of 45 a.u. (Figure 3.6B3), a decreasing cutinase activity accumulation similar to the results obtained with strain WT is observed. The similarity of cutinase activity time profiles for induction at BS signals of 15 and

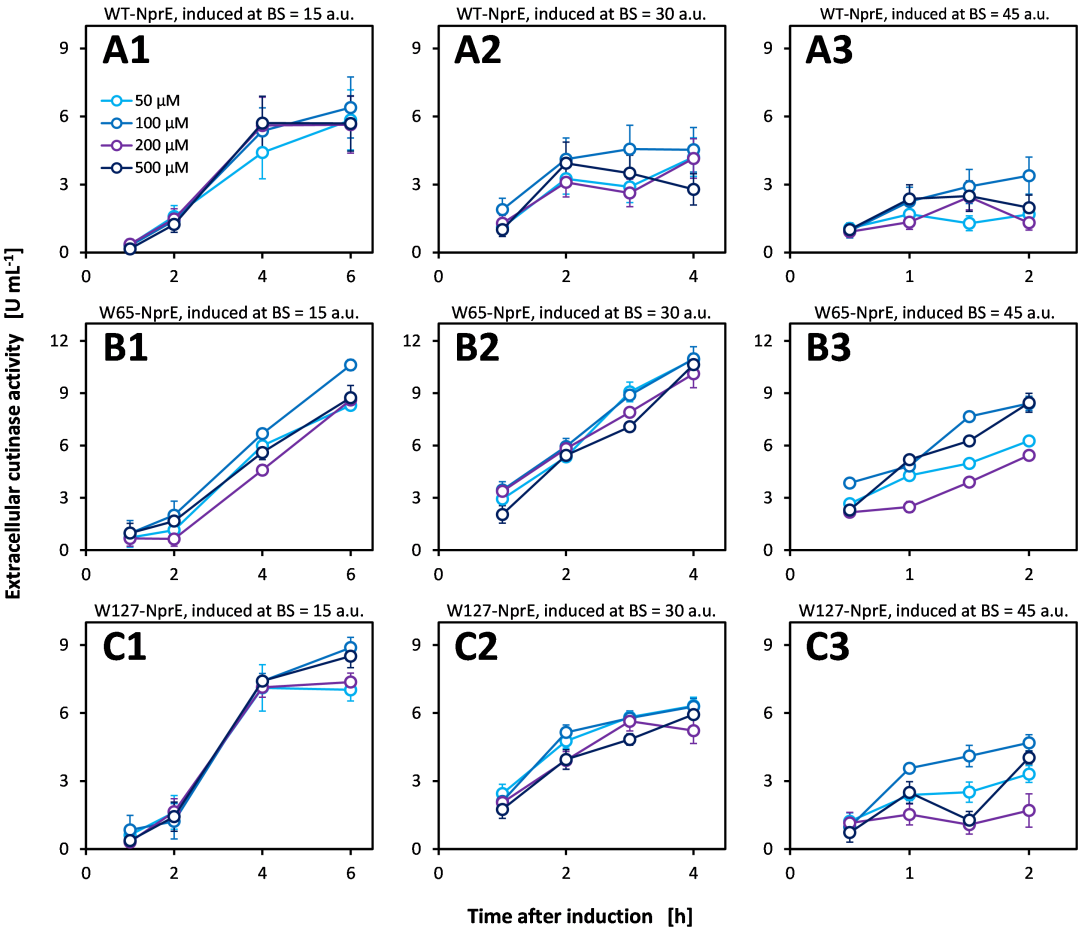


Figure 3.6.: Comparative induction profiling for *C. glutamicum* cutinase secretion strains WT-NprE, W65-NprE and W127-NprE. Each panel shows the extracellular cutinase activity over time after induction, determined from BioLector MBR cultivations. Induction profiling conditions comprise three strains, four IPTG concentrations and three different growth stages at induction. The latter parameter is represented by a BS signal of 15 a.u. (early growth phase), 30 a.u. (mid-time growth phase), and 45 a.u. (late growth phase), which serves as trigger signal for IPTG addition. Tested IPTG concentrations are 50, 100, 200 and 500 µM. Panels A, B, and C show time profiles of cutinase activities obtained with strains WT-NprE, W65-NprE, and W127-NprE, respectively, while panels 1, 2, and, 3 show results for these strains induced at early, mid-time, and late growth phase, respectively. In each panel (i.e., for a specific combination of strain and growth stage where IPTG was added, as indicated at the top of each panel), results for the four tested IPTG concentrations are shown. The corresponding color coding is given in panel A1, which applies to all panels. For clarity reasons, biomass signals are not shown. Error bars indicate standard deviations from replicates (n = 3) of cutinase activity determination from each sample.

30 a.u. can be attributed to the fact that strain W65-NprE grows significantly slower than strain WT-NprE [55]. In consequence, the predefined harvesting time points depicted in Figure 3.6B1 and B2 do not cover the transition from the exponential growth phase to the stationary phase, like for strain WT-NprE. For induction at 45 a.u., this is assumed to be the case (Figure 3.6B3), since here the decay in cutinase accumulation is observed. Like concluded for strain WT-NprE, induction is not preferred to be done in the late exponential growth phase with strain W65-NprE. The transition to stationary phase was not monitored for strain W65-NprE with induction at early and mid exponential growth phase and thus, it is not known for which of both conditions the highest cutinase activity is reached (Figure 3.6B1 and B2). It can be only speculated that early induction is preferred. However, it is reasonable to assume that also in strain W65-NprE early induction is beneficial in terms of high secreted cutinase activity.

Finally, observations on secreted cutinase activity for the third strain shown, W127-NprE (Figure 3.6C1 to C3), are highly comparable to strain WT-NprE, meaning that, the later during growth induction takes place, the lower the extracellular cutinase activity that can be reached. Strain W127-NprE was described to shown the same biological fitness as strain WT-NprE [48]. Like observed for strain WT, impact of IPTG concentration can be neglected for induction at early and mid exponential growth (induction at 15 a.u. and 30 a.u., Figure 3.6C1 and C2, respectively). Different levels of cutinase activity for the different IPTG concentrations tested are seen when inducing in the late exponential growth phase (Figure 3.6C3). However, similar to induction at BS signals of 45 a.u. with strains WT-NprE and W65-NprE (cf. Figure 3.6A3 and B3, respectively), obtained final cutinase activities are much higher with early induction at a BS signal of 15 a.u. for strain W127-NprE, irrespective of the IPTG concentration applied (ca. 8 to 9 U mL^{-1} vs. ca. 2 to 4 U mL^{-1}, compare Figure 3.6C1 vs. C3, respectively).

Results from induction profiling can be summarized for all six strains as follows (i.e., including strains zwf^{FBR}-NprE, Δ*cglMRR*-NprE, and MB001-NprE that are shown in (Figure B.3): highest extracellular cutinase activities are achieved with induction in the beginning of the exponential growth phase. In addition, IPTG concentration is of minor importance when inducing at early growth, at least for the IPTG concentrations tested. However, in most cases an IPTG concentration of 100 µM results in slightly higher final cutinase activities compared to the other concentrations tested (see corresponding data points in Figure 3.6A1, B1 and C1) and therefore, 100 µM IPTG will be used as standard induction strength in further MBR growth experiments.

In a previous study with *C. glutamicum* WT secreting cutinase using the NprE SP, an early induction with comparably low IPTG concentration was determined to be the preferred induction condition, although 200 to 300 µM IPTG were reported as optimal on the evaluation from samples taken after the

culture has entered stationary phase approximately 2 h earlier [36]. When only considering strain WT-NprE for cutinase secretion, this is in line with the observations made in this study, which refer to the data depicted in Figure 3.6A1. However, the choice of 100 µM IPTG as optimal inducer concentration is justified since this represents an optimum for strain WT-NprE, as well as five additional strains. Because these observations are made for six different *C. glutamicum* strains with different genotypes and growth phenotypes, the optimized induction strategy (i.e., 100 µM IPTG added early in growth) is also deduced for further genome reduced strains (GRS). Furthermore, for the optimal induction condition determined, the six strains provide a preview on the optimization potential of cutinase secretion efficiency, as different maximal cutinase activities are observed, namely ca. 6 U mL^{-1} (WT-NprE), > 11 U mL^{-1} (W65-NprE, with maximum cutinase activity expected to be found in stationary phase that was not recorded) and ca. 9 U mL^{-1} (W127-NprE), cf. Figure 3.6A1, Figure 3.6A2 and Figure 3.6A3, respectively, as well as ca. 7 U mL^{-1}, ca. 9 U mL^{-1} and ca. 9 U mL^{-1} (strains zwf^{FBR}-NprE, $\triangle$*cglMRR*-NprE, and MB001-NprE, respectively, shown in Figure B.3).

3.3. Increasing process information from microbioreactor cultivations by harvesting replicate cultivations at a high frequency

Next, for five of the six example strains from the GRS library, detailed bioprocess kinetics were acquired, using the optimal induction strategy as determined in the previous section. The MBR protocol of sequential harvest applied for induction profiling (see previous section) was modified in such way that at 16 timepoints three replicate cultures were harvested. Thereby, a high temporal solution of culture dynamics for biomass, cutinase secretion and glucose consumption is achieved. Also, the exponential growth phase is covered with a sufficiently high number of data points to verify that cutinase secretion is growth coupled. Up to now, this was shown only for strain WT-NprE [36], but needs to be verified for further *C. glutamicum* strain variants tested in this study.

Since in total 48 replicate cultivations are needed with this modified MBR protocol, only one strain can be cultivated during a single MBR growth experiment, which reduces the cultivation throughput of the MBR system. Like in the previous section, only results for the three strains WT-NprE, W65-NprE and W127-NprE are shown exemplarily in Figure 3.7, time courses for strains MB001-NprE and zwf^{FBR}-NprE are be found in Figure B.4.

In Figure 3.7, panels A, B and C refer to strains WT-NprE, W65-NprE and W127-NprE, respectively, and panels 1, 2 and 3 refer to the time courses of biomass, glucose concentration and extracellular cutinase activity for these strains, respectively. The data depicted in Figure 3.7 covers the complete

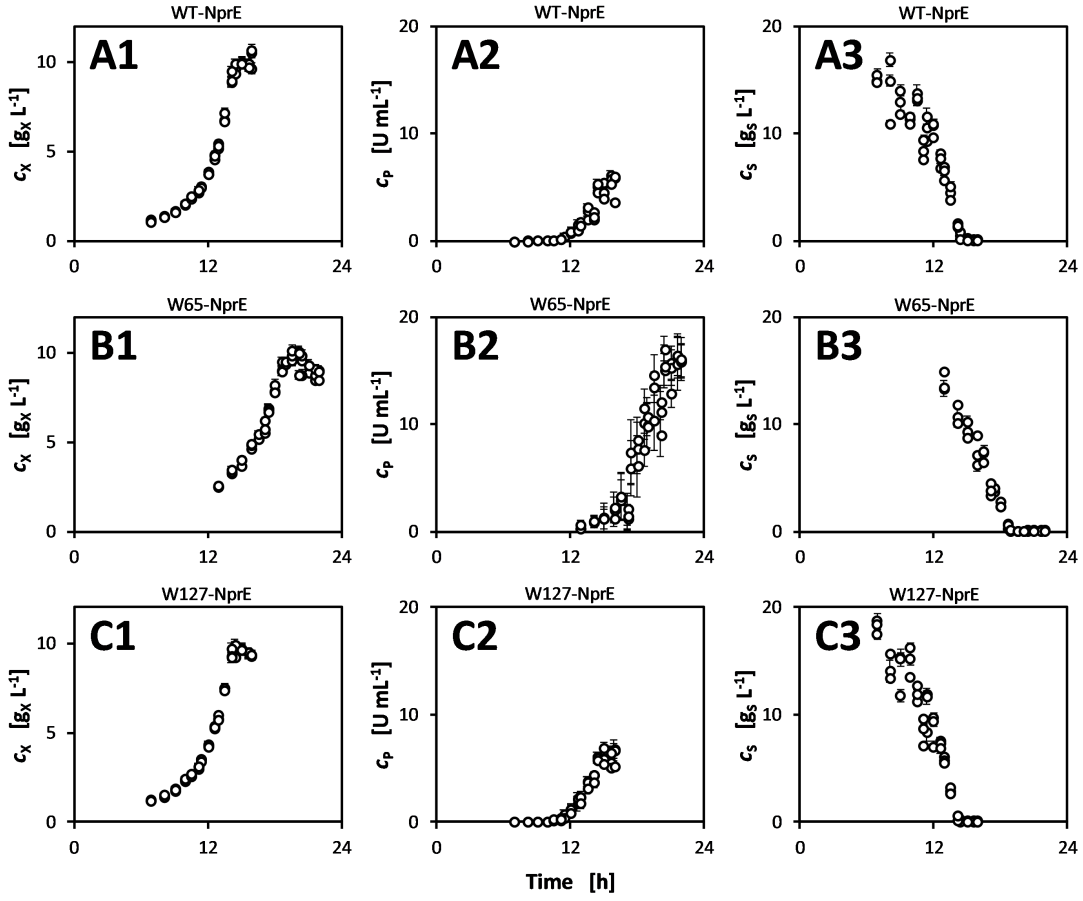


Figure 3.7.: High-resolution time courses for cell dry weight, glucose and extracellular cutinase activity (panels 1, 2 and 3, respectively) for strains WT-NprE, W65-NprE and W127-NprE (panels A, B and C, respectively). Each row of panels shows data for one of the three strains, which are indicated at the top of each the panels. Each column of panels shows one kind of data for the three strains. For induction of each culture replicate, 100 µM IPTG was added upon reaching a BS signal of 15 a.u., and each strain was cultivated in 48 replicates, which were sequentially harvested in triplicates. Cell dry weight is calculated from BS signal at time point of harvest via linear calibration. c_X cell dry weight. c_P cutinase activity. c_S glucose concentration.

exponential growth phase for the three strains shown. Biomass for all strains grows exponentially until entering the stationary phase, which is reached at about 15 h, 29 h and 15 h for strains WT-NprE (Figure 3.7A1), W65-NprE (Figure 3.7B1) and W127-NprE (Figure 3.7C1), respectively. Glucose consumption follows inversely for all strains (Figure 3.7A2 to C2), as expected. It can be seen that, by comparing the corresponding panels showing time courses of biomass and glucose for each strain, depletion of glucose as main carbon source results in stagnating cell dry weight. Upon addition of IPTG, cutinase

activity also increases accompanying the increasing cell dry weight (Figure 3.7A3 to C3). As speculated in section 3.2.2, strain W65-NprE reaches considerably higher final cutinase activities in comparison to strains WT-NprE and W127-NprE. Surprisingly, strain W65-NprE reaches almost doubled cutinase activities, although it is not known which genomic deletions are causative. At the same time, these genomic deletions in strain W65-NprE result in an impaired growth rate in comparison to strain WT-NprE, which was also reported earlier for the plasmid-free strain variants [55]. As also previously observed, strains WT-NprE and W127-NprE show comparable culture kinetics with respect to biomass formation, glucose consumption and cutinase secretion, meaning that the deleted genes in strain W127-NprE can be classified as non-essential with respect to cutinase secretion bioprocess performance.

By choosing only data points that belong to the exponential growth phase after induction, culture PIs (growth rate μ, substrate specific biomass yield $Y_{X/S}$ and biomass specific cutinase yield $Y_{P/X}$) are calculated using linear regressions methods. As outlined previously (cf. section 2.7), PIs represent a condensation of culture kinetics that allow a rapid evaluation of both strain- and cultivation-specific conditions. PI calculation was done using ordinary least squares after ln-transformation for determination of growth rate μ, as well as orthogonal least squares [213] for biomass yield $Y_{X/S}$ and cutinase yield $Y_{P/X}$. Calculations use mean values from samples harvested in triplicates and resulting standard deviation is taken into account by the application of a parametric bootstrap procedure, that is regressions were calculated repeatedly (5000 times). All measurement data points are perturbed independently according to their assumed Gaussian distribution, parametrized by the respective mean value and standard deviation. In each repetition, a value for each measuring point is sampled from its distribution. In turn, 5000 values for each μ, $Y_{X/S}$ and $Y_{P/X}$ are generated and from these, mean values and standard deviations are reported in Table 3.1.

Table 3.1.: Determined PIs from MBR method of harvesting sequential replicate cultivations.

Strain	**Growth rate μ [h^{-1}]**	**Biomass yield $Y_{X/S}$ [$g_X\ g_S^{-1}$]**	**Cutinase yield $Y_{P/X}$ [$kU\ g_X^{-1}$]**
WT-NprE	0.36 ± 0.00	0.58 ± 0.04	0.68 ± 0.06
MB001-NprE	0.32 ± 0.00	0.58 ± 0.05	0.65 ± 0.05
W65-NprE	0.22 ± 0.00	0.49 ± 0.03	1.64 ± 0.25
W127-NprE	0.34 ± 0.00	0.50 ± 0.05	0.64 ± 0.05
zwf^{FBR}-NprE	0.34 ± 0.00	0.53 ± 0.04	0.78 ± 0.07

Only data from the exponential growth phase after induction was used to derive the PIs. The bootstrapped standard deviations represent a metric for the uncertainty of determination of the corresponding PIs.

With respect to the **(!**s ((**!**s)PI) calculated (Table 3.1), literature data is available for strain WT-NprE

only [36]. Obtained values for growth rate μ (Lit: 0.41 ± 0.02 h^{-1}), biomass yield $Y_{X/S}$ (Lit.: 0.60 ± 0.05 g_X g_S^{-1}) and $Y_{P/X}$ (Lit.: 1.05 – 0.88 kU g_X^{-1}) are comparable to the values reported in literature, although a different induction strategy was applied than in this study.

Strain MB001 was only characterized in terms of intracellular production of a heterologous model protein (YFP) [53] and it was reported that a higher plasmid stability in MB001 accounted for higher YFP titer in comparison to the wild type strain. However, as seen here, MB001-NprE does not outperform its parental strain WT-NprE in terms of secretory cutinase production (0.65 vs. 0.68 kU g_X^{-1}, cf. Table 3.1).

The point mutation characterizing strain zwf^{FBR}-NprE was described to be beneficial with respect to L-lysine production in *C. glutamicum* since this mutation causes an increased NADPH generation, which is needed in the anabolic amino acid pathways [214]. However, in this study a distinct effect on cutinase yield could not be observed (0.78 vs. 0.68 kU g_X^{-1}, cf. Table 3.1).

The conclusions drawn for strains W65-NprE and W127-NprE from the comparison of culture kinetics (cf. above) are confirmed by the derived PIs depicted in Table 3.1: The genomic deletions in strain W127-NprE have no apparent negative impact on cutinase secretion performance (0.64 vs. 0.68 kU g_X^{-1}) while strain W65-NprE reaches a roughly doubled cutinase yield in comparison to strain WT-NprE (1.64 vs. 0.68 kU g_X^{-1}).

Moreover, the generated data show clearly that cutinase secretion in *C. glutamicum* strains is growth-coupled in general, which justifies the assumption of a constant yet strain-specific yield coefficient relating cutinase secretion with biomass formation (i.e., $Y_{P/X}$). Since this is observed for strain WT-NprE and further four strains, the assumption is extended to the whole GRS library.

The calculated PIs depicted in Table 3.1 represent a single snapshot since the value estimations are based on experimental data stemming from one MBR growth experiment. In general, to increase confidence, it is advised to use replicate data from multiple MBR growth experiments [115], meaning that all experimental steps including uncertainties from "day-to-day variations" are taken into account, that can neither be identified nor be quantified in most cases. Consequently, single cultivations (i.e., one cultured well per MBR growth experiment) should be considered as insufficient, multiple culture replicates per MBR growth experiment should be preferred, and ideally these should be repeated independently at least once. This was done in Chapter 6 and therefore, a detailed discussion about PIs as phenotypes determined for the different GRS in combination with further analytical results is postponed to Chapter 6.

3.4. Increasing process information from microbioreactor cultivations by repeated low-volume sampling

Clearly, for the investigation of process kinetics the method of harvesting replicate cultivations is a drawback compared to classical lab-scale bioreactor setups, since the originally intended throughput by miniaturization gets lost. Thus, to preserve the advantage of culture kinetics obtained from sequential harvest of replicate cultivations but without the drawback of a loss in throughput from MBR cultivations, a further improvement of robot-assisted sampling and analyte detection from parallel MBR cultivations is presented in the following section.

The improved workflow involves repeated sampling of small volumes (6 x 20 µL) from MTP cultivations (800 µL) to obtain time-resolved concentration data without loss of throughput. These datasets are then fitted to standard regression models to estimate specific yields and rates for each replicate experiment. Following a detailed characterization of each step in the method workflow, both technical and biological uncertainties for PI estimation are quantified using non-linear random error propagation. Due to the novelty of the improved workflow, a detailed description is given, followed by an in-depth characterization of altered process steps. The suitability of the workflow is demonstrated with a high number of cultivation replicates with *C. glutamicum* wild type, which allows performing a random error propagation covering all experimental steps involved in determination of culture PIs. Finally, the workflow is applied to a selection of cutinase secreting *C. glutamicum* GRS.

3.4.1. Description of the improved workflow

For the quantitative characterization of strain libraries, specific yields and rates are most commonly applied as PIs, as discussed earlier. Typically, these PIs are estimated by utilizing (non)-linear regression methods and in order to provide meaningful results, a minimum number of informative data points from the concentration changes in the culture's broth is required.

For example, in the case of 48-well Flowerplates [111, 153], 48 parallel cultivations can be conducted at once. By common experience a minimum of six informative data points are needed for proper regression-based PI estimation in two independent replicates. Consequently, the effective capacity for strain characterization is greatly reduced by a factor of twelve, that is only four strains can be compared in one run.

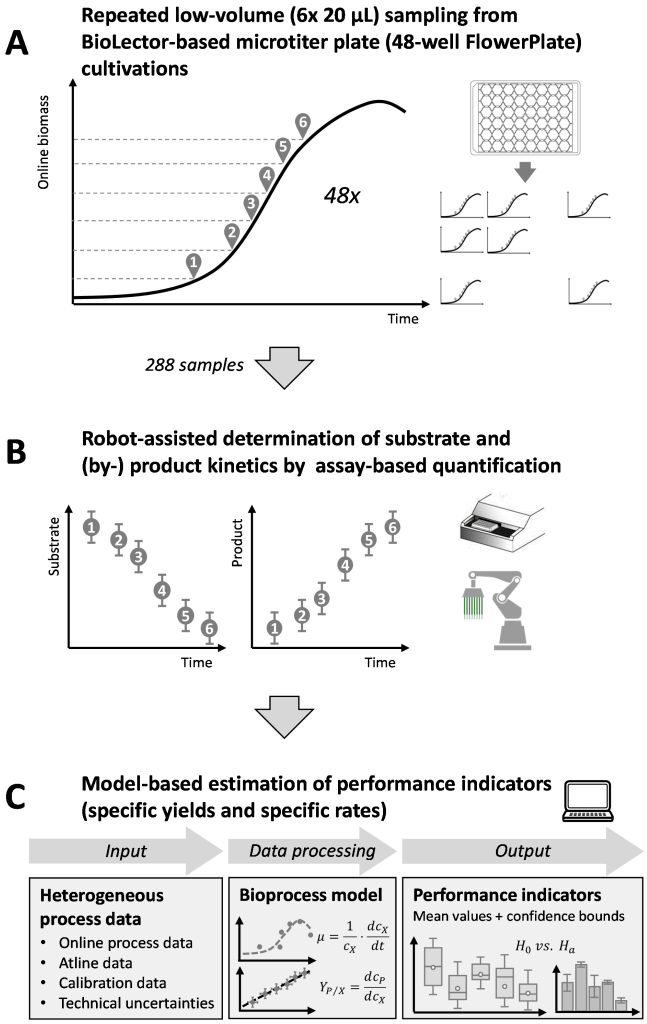


Figure 3.8.: Workflow for the estimation of PIs from individual MTP cultures by repeated low-volume sampling. (A): In the established procedure six samples of 20 µL are taken upon reaching pre-defined biomass signals to cover the exponential growth phase. (B): The samples are then stored until analysis of the primary substrate (e.g., glucose) and product (e.g., cutinase activity) after cell separation by centrifugation. (C): In a final step, the collected heterogeneous data are evaluated with the help of a process model to estimate PIs (e.g., product yield, substrate consumption rate) that are usable for trustworthy strain and process comparisons.

The novel workflow for repeated low-volume sampling of BioLector cultivations (Figure 3.8) was established by building upon the recently introduced Mini Pilot Plant technology that enables automated and pre-defined manipulation of individual MTP cultures [36, 160, 206]. For cultivation, 48-well Flowerplates were used with an initial filling volume of 800 µL. The repeated sampling (6 x 20 µL) of

the 48 individual cultivations was carried out automatically by the integrated liquid handling system and executed upon reaching pre-defined online monitored BS (i.e., biomass) signals. The latter ensured that both fast and slow growing strains were sufficiently sampled from the exponential growth phase. From each Flowerplate, in total 288 cell samples were taken, processed by centrifugation and the resulting cell-free supernatants were analyzed for concentration changes in the primary substrate and product by utilizing automated and parallelized spectrophotometric assays. Finally, the resulting data points online and atline measurements were fitted to an unstructured bioprocess model to derive estimates for the desired PIs based on mean values and corresponding confidence bounds.

3.4.2. Impact of repeated low-volume sampling on culture dynamics

The introduced workflow is only valid if the applied repeated sampling procedure has no effect on cell population growth, that is the resulting growth dynamics are comparable to non-sampled cultures. This aspect was investigated by monitoring parallel cultivations of *C. glutamicum* WT on defined CgXII medium (Figure 3.9). Pre-inoculated medium was distributed into six wells of a Flowerplate and one set of three replicates was repeatedly sampled while the other set remained untouched. As a result, the sampled culture showed a classical batch profile with balanced (exponential) growth (Figure 3.9A). Moreover, all measured backscatter values were highly correlated with those of the non-sampled culture (Figure 3.9B), proving that the implemented low-volume sampling procedure is compatible with the applied MBR cultivation approach. Indeed, the resulting gradual reduction of the culture volume from 800 µL to 680 µL does not induce an increased inaccuracy of the BS signal, nor a noticeable higher *intra*-MTP imprecision, only a slightly higher *inter*-MTP imprecision can be assumed, as seen above (cf. section 3.1, Figure 3.2).

3.4.3. Impact of shaking on microbial growth during repeated low-volume sampling

The employed MBR system is based on orbital shaken MTPs. Therefore, it is important to know if an interruption in MTP shaking during the sampling process has any influence on the batch culture. Depending on the technical specifications of a certain combination of MBR and integrated LHS devices, addition and removal of liquid to and from the culture requires a stopping of the culture agitation. In shaken cultivation systems oxygen is exclusively introduced into the culture broth via the interfacial area between the gas and liquid phase. Therefore, the growing culture may react sensitive towards a stopping of the shaking movement since the gas-liquid interfacial area is high when the liquid is rotating and low when the liquid is standing still [148]. Furthermore, without shaking there

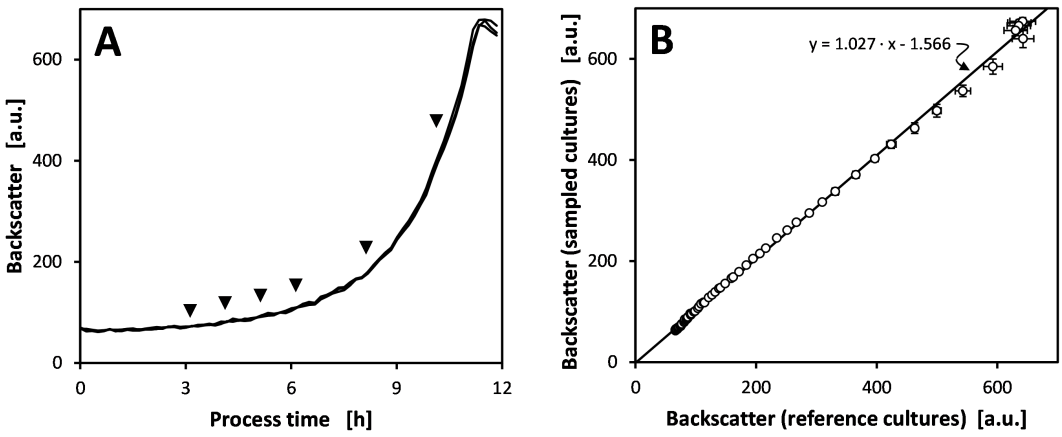


Figure 3.9.: Impact of repeated low-volume sampling on growth of *C. glutamicum* wild type. (A): Growth curves of triplicate cultures that are repeatedly sampled during the exponential growth phase as indicated by the black triangles. (B): Comparison of BS signal from repeatedly sampled cultures and non-sampled reference cultures. Data points represent mean BS values from three replicate cultivations with standard deviation as error bars. Solid line represents result from orthogonal least squares regression analysis [213]. Each cultivation was started with 800 µL from a freshly prepared CgXII medium with 20 g L^{-1} glucose as carbon and energy source. Initial OD_{600} was approx. 0.1.

is a lack of turbulence due to missing power input [152], oxygen spreads in the liquid phase only by diffusion. In turn, interruption of shaking immediately causes a drastic reduction in the OTR [215]. Because the MTP is shaken as a whole, all individual cultures might undergo short oxygen limitation if the shaking is stopped for the sampling of one specific well. A complete sampling cycle, durating approximately 3 min on average, comprises opening of the incubation chamber, aspiration of cell suspension, dispensing of the cell suspension into a sample receiving MTP, washing of the robotic tips and closing of the incubation chamber which takes approximately 3 min altogether.

To study the impact of shaking during repeated sampling two MBR cultivation experiments were compared, using the same operating conditions, the same medium batch and the same pre-cultures. In one case, shaking was stopped during sampling while in the other case sampling was performed without any interruption of shaking. Altogether, six different *C. glutamicum* strains were tested (each in eight replicates, shown in Figure B.5) and Figure 3.10 exemplarily shows the resulting growth patterns of the wild-type strain.

In the MBR sampling experiment with interrupted shaking the BS signal points at a two-phasic growth behavior (Figure 3.10A). While in the first eight hours of cultivation an exponential growth was observed, the second phase was characterized by a linear increase in the backscatter only. Dur-

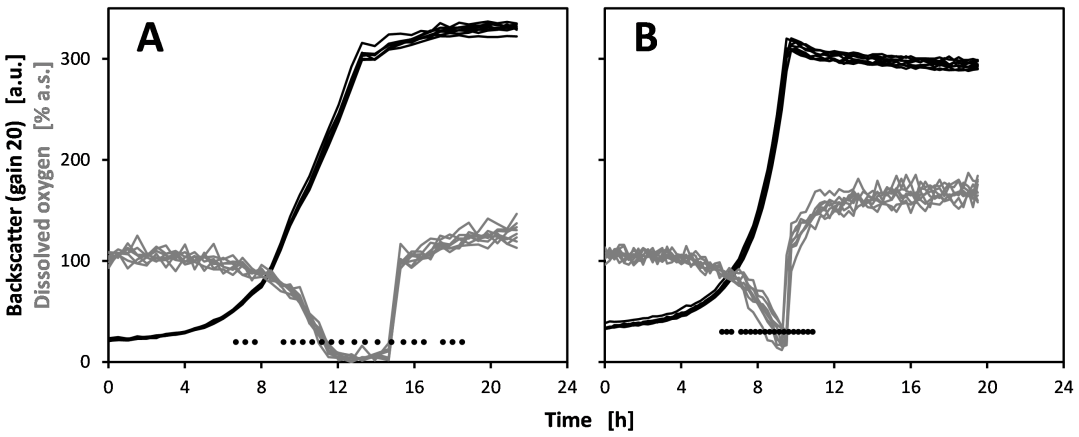


Figure 3.10.: Impact of interrupted shaking on growth of *C. glutamicum* WT during repeated low-volume sampling. (A): Cultivations with stopped shaking for sample taking. (B): Cultivations where sampling was conducted without interruption of shaking. Black dots indicate time points of sampling. Shown are eight cultivation replicates.

ing this phase the cultures were greatly affected by the frequent stops in shaking that resulted in repeated oxygen limitations (visible in the near zero DO signal). In contrast, no oxygen limitation occurred in case of repeated low-volume sampling without interruption of shaking (Figure 3.10B).

Consequently, if a higher sampling load is required in MBR experiments, that is for deriving meaningful PIs from single cultures using regression methods, then sampling needs to be performed under continuous shaking in order to maintain a sufficiently high OTR for undisturbed cell population growth.

3.4.4. Determination of performance indicators with random error propagation

In the final step of the workflow the obtained process data are evaluated with the help of an unstructured bioprocess model (Figure 3.8C). For the case of simple growth phenotyping, it has the general form:

$$\mu = \frac{1}{c_X} \cdot \frac{dc_X}{dt} \approx \frac{ln\left(\Delta \boldsymbol{c_X}\right)}{\Delta \boldsymbol{t}} \tag{3.1}$$

$$Y_{X/S} = \frac{dc_X}{dc_S} \approx \frac{\Delta \boldsymbol{c_X}}{\Delta \boldsymbol{c_S}} \tag{3.2}$$

$$q_S = -\frac{1}{Y_{X/S}} \cdot \mu \tag{3.3}$$

$$c_X = I_{BS} \cdot a_{CDW} + b_{CDW} \tag{3.4}$$

$$c_S = \frac{A_{340} - b_{Glc}}{a_{Glc}} \cdot f_{D1} \cdot f_{D2} \tag{3.5}$$

$$f_{D1} = \frac{V_{Sample} + V_{Diluent1}}{V_{Sample}} \tag{3.6}$$

$$f_{D2} = \frac{V_{DilutedSample} + V_{Diluent2}}{V_{DilutedSample}} \tag{3.7}$$

For the time-dependent changes in biomass and substrate the assumption that biomass growth is balanced holds throughout the batch experiment (μ, $Y_{X/S}$, q_S = const.), cf. (3.1) to (3.3). The corresponding raw data from optical BS measurements [1] and spectrophotometric assay are converted by two calibration models (3.4) and (3.5). The latter model also takes into account the necessary sample dilution steps (3.6) and (3.7) in order to match the linear dynamic range of the assay (for details regarding the applied calibration models see Chapter 8 and Table 3.2).

There are two sources of uncertainty, which finally affect the precision of PI determination. First, technical uncertainty arises from all random errors introduced during the execution of cultivation experiments. Here parametric bootstrapping is frequently used for error propagation analysis [216, 217], provided that random errors for all technical variables are known. Second, biological uncertainty results from the intrinsic biological noise of a microbial culture. This type of uncertainty can only be made visible when conducting the identical cultivation experiment multiple times.

As proof of concept for the novel workflow, *C. glutamicum* wild type was cultivated in 24 independent cultures in defined CgXII medium with glucose as carbon and energy source. All cultures were automatically sampled six times for determination of glucose consumption. Random errors (coefficients of variation) for the technical variables of the bioprocess model were determined in well-defined separate experiments and are listed in Table 3.2. Parametric bootstrapping was carried out in 5,000 repetitions for each culture, resulting in 24 distributions for μ, $Y_{X/S}$ and q_S that in total cover 120,000 single estimates for each PI.

The resulting boxplots from PI determination with random error propagation are shown in Figure 3.11. The corresponding absolute values are determined as μ = 0.43 ± 0.01 h^{-1}, $Y_{X/S}$ = 0.61 ± 0.08 g_X g_X^{-1} and q_S = 0.72 ± 0.10 g_X g_X^{-1} h^{-1} (Figure 3.11A, B and C, respectively), given as mean ± standard deviation from the merged distribution of each PI. Considering both technical and biological noise, these numbers translate relative uncertainties of 2.7 %, 13.4 % and 14.0 %, respectively.

[1] BS calibration is specific for each BioLector device and operating conditions, and was determined in this study as $c_X = 0.0462 \cdot I_{BS} - 1.059, R^2 > 0.99$

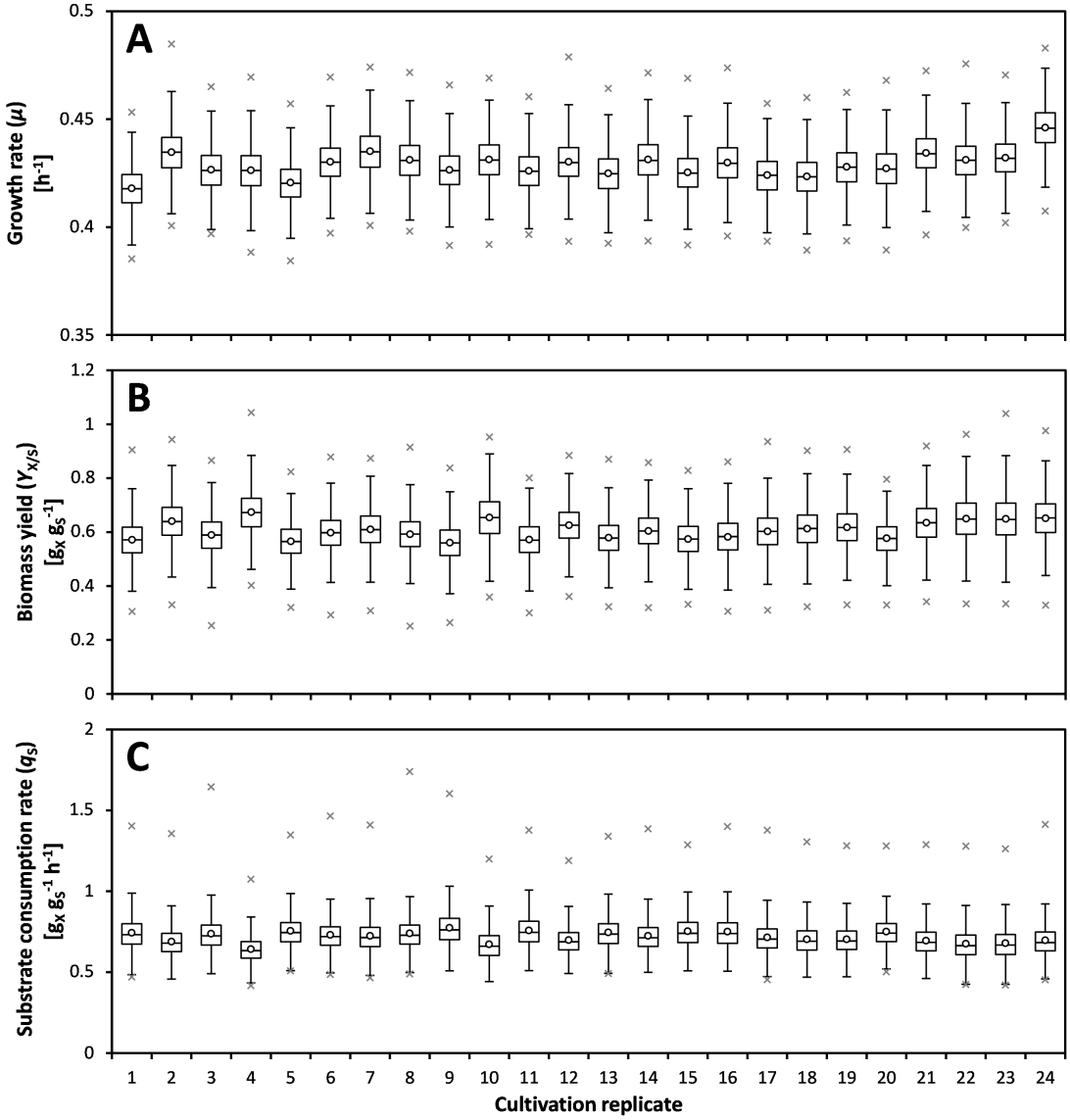


Figure 3.11.: Determination of PIs specific growth rate μ (A), substrate specific biomass yield $Y_{X/S}$ (B) and biomass specific substrate consumption rate q_S (C) for each of n_{biol} = 24 replicate cultivations of *C. glutamicum*. For each cultivation condition, a parametric bootstrap was carried out with $n_{bootstrap}$ = 5000 replicates. Resulting distributions of values are shown as box plots. Boxes indicate interquartile range with median and mean denoted as line and circle, respectively. Whiskers indicate 1.5-fold the interquartile range and grey crosses indicate maximum and minimum outliers.

Table 3.2.: Technical variables in the bioprocess model and corresponding coefficients of variation (CV).

Technical variable	CV [%]	Description of CV
BS measurement, I_{BS}	0.9	Average CV from analytical triplicates, determined from backscatter / cell dry weight calibration.
Slope of backscatter / cell dry weight calibration, a_{CDW}	1.7	Relative slope error.
Offset of backscatter / cell dry weight calibration, b_{CDW}	10.5	Relative offset error.
Sample volume, V_{Sample}	11.7	CV for volumes of < 100 µL removed from shaking BioLector device, determined for liquid handling system after installation (n = 20).
Volume added to sampled cell suspension, $V_{Diluent1}$	0.4	CV for volumes > 100 µL, determined for liquid handling system after installation (n = 96).
Diluted sample volume for second dilution step, $V_{DilutedSample}$	1.8	CV for volumes < 100 µL, determined for liquid handling system after installation (n = 96).
Volume added to diluted sample $V_{Diluent2}$	0.4	CV for volumes > 100 µµL, determined for liquid handling system after installation (n = 96).
Absorption measured at 340 nm for glucose assay, A_{340}	2.6	Average CV, determined from glucose standards (n = 128).
Slope of glucose assay calibration, a_{Glc}	0.8 – 8.2	Relative slope error in each assay.
Offset of glucose assay calibration, b_{Glc}	1.6 – 20.3	Relative offset error in each assay.
Linear increase of absorption at 410 nm of cutinase assay, ΔA_{410}	2.1	Average relative error from 288 duplicate measurements.
Slope of pNP calibration in cutinase assay, a_{pNP}	0.5 – 0.9	Relative slope error in each assay.

Determined values for μ, $Y_{X/S}$ and q_S are in good accordance with numerous reports from literature [48, 55, 218–220], proving that the presented workflow is suitable to replace laborious shake flask or bioreactor experiments for standard growth phenotyping.

3.4.5. Application on cutinase secreting *C. glutamicum* genome reduced strains

The novel workflow was applied for the characterization of a set of recently introduced GRS of *C. glutamicum* [48, 55] that were further engineered towards heterologous cutinase secretion. The GRS carry deletions of all three prophages, two insertion elements and different combinations of larger gene clusters (Table 8.1). All GRS were transformed with the same expression plasmid containing the NprE Sec SP from *Bacillus subtilis* to enable cutinase secretion [30, 36]. Each of five GRS and the corresponding wild-type strain were cultivated in 8 biological replicates, resulting in 48 paral-

lel cultivations from which 288 samples were analyzed for both glucose concentration and cutinase activity.

For advanced phenotyping of extracellular cutinase activity, the bioprocess model was extended by two additional PIs related to specific cutinase yield ($Y_{P/X}$) and secretion rate (q_P) [39]. For details regarding the applied calibration model (3.10) and determined random errors of new technical variables, see Chapter 8 and Table 3.2. Results are shown in Table 3.3.

$$Y_{P/X} = \frac{dc_P}{dc_X} \approx \frac{\Delta \boldsymbol{c_P}}{\Delta \boldsymbol{c_X}} \tag{3.8}$$

$$q_P = Y_{P/X} \cdot \mu \tag{3.9}$$

$$c_P = \Delta A_{410} \cdot a_{pNP} \cdot f_{D1} \cdot f_{D2} \tag{3.10}$$

Table 3.3.: Comparison of extracellular phenotypes and genotypes for recombinant cutinase secreting *C. glutamicum* strains with differently reduced genomes, cultivated in eight replicates each.

Characteristic	**WT-NprE**	**W31-NprE**	**W115-NprE**	**W116-NprE**	**W127-NprE**	**W130-NprE**
Genotype						
ΔCGP123		x	x	x	x	x
Δ*ISCg12*		x	x	x	x	x
Δ2312-2322					x	x
Δ2621-2643					x	x
Δ2663-2686					x	x
Δ2755-2760					x	x
Δ3102-3111			x	x	x	x
Δ3263-3301			x	x		
Δ3324-3345		x		x		x
Phenotype						
μ [h^{-1}]	0.38 ± 0.01	0.34 ± 0.01	0.22 ± 0.00	0.18 ± 0.00	0.31 ± 0.01	0.33 ± 0.01
$Y_{X/S}$ [$g_X\ g_S^{-1}$]	0.58 ± 0.08	0.50 ± 0.09	0.55 ± 0.07	0.47 ± 0.07	0.55 ± 0.08	0.58 ± 0.11
q_S [$g_X\ g_S^{-1}\ h^{-1}$]	0.66 ± 0.09	0.72 ± 0.18	0.40 ± 0.05	0.38 ± 0.07	0.57 ± 0.08	0.59 ± 0.14
$Y_{P/X}$ [$kU\ g_X^{-1}$]	0.74 ± 0.10	0.64 ± 0.09	0.37 ± 0.06	1.42 ± 0.20	0.78 ± 0.14	0.78 ± 0.12
q_P [$kU\ g_X^{-1}\ h^{-1}$]	0.28 ± 0.04	0.22 ± 0.03	0.08 ± 0.01	0.25 ± 0.03	0.24 ± 0.04	0.26 ± 0.04

Values represent mean and standard deviation calculated by a parametric bootstrap procedure (5000 bootstrap repetitions for each cultivation replicate). μ specific growth rate. $Y_{X/S}$ substrate specific biomass yield. q_S biomass specific substrate consumption rate. $Y_{P/X}$ biomass specific cutinase yield. q_P biomass specific cutinase formation rate.

Strains WT-NprE and W127-NprE have been cultivated previously using the established MBR method of harvesting replicate cultivations (cf. section 3.3). The determined PIs using the improved MBR workflow are highly comparable to these obtained with the established method, as the comparison of values shows for strains WT-NprE (μ = 0.38 vs. 0.36 h^{-1}, $Y_{X/S}$ = 0.58 vs. 0.58 g_X g_S^{-1}, $Y_{P/X}$ = 0.74 vs. 0.68 kU g_X^{-1}) and W127-NprE (μ = 0.31 vs. 0.34 h^{-1}, $Y_{X/S}$ = 0.55 vs. 0.50 g_X g_S^{-1}, $Y_{P/X}$ = 0.78 vs. 0.64 kU g_X^{-1}), compare also corresponding entries in Table 3.3 vs. Table 3.1.

The comparison shows that the improved workflow is an attractive alternative to increase experimental throughput and that reliable time-resolved culture dynamics and PIs derived thereof are obtained. However, the method of harvesting replicate cultivations is needed when low sample volumes are not sufficient due to constraints of analytical methods, like it is the case for the quantification of time-resolved proteomic changes during MBR cultivation [221].

The results on determined growth rates identify strains W115-NprE and W116-NprE as slowly growing (μ = 0.22 ± 0.00 and 0.18 ± 0.00 h^{-1}, respectively), which was previously reported for these strains without expression plasmids [222]. Surprisingly, these two strains are also remarkable in terms of cutinase yields $Y_{P/X}$: Strain W115-NprE shows a distinctly lowered cutinase yield of $Y_{P/X}$ = 0.37 ± 0.06 kU g_X^{-1}, while strain W116-NprE shows a distinctly increased cutinase yield of $Y_{P/X}$ = 1.42 ± 0.20 kU g_X^{-1} compared to the remaining strains (Table 3.3).

From the data presentation in Table 3.3, it is seen that the combinatorial deletion of cg3263-3301 and cg3324-3345 is apparently responsible for the remarkably increased cutinase yield with strain W116-NprE, while the deletion of cg3263-3301 is presumably resulting in the lowered growth rate of strains W115-NprE and W116-NprE. In section 3.3, a high cutinase yield ($Y_{P/X}$ = 1.64 ± 0.25 kU g_X^{-1}) for strain W65-NprE was also observed, which is comparable to strain W116-NprE (Table 3.3). Interestingly, these two strains are also comparable with respect to their growth phenotype (μ^{W116} = 0.18 h^{-1} vs. μ^{W65} = 0.22 h^{-1}, cf. Table 3.3 vs. Table 3.1), although the strains share no genomic deletions except ΔCGP123 and ΔIS*Cg*12, which is the case for nearly all GRS (i.e., not for MB001).

Possibly, a reduced growth rate is a prerequisite for high cutinase yields, which is discussed in detail later in section 6.5 and section 6.6 of Chapter 6, together with further results.

4. Hyphenating automated microbioreactor systems and Kriging-based Design of Experiments for bioprocess development

This chapter describes how Kriging-based Design of Experiments methodology and automated microbioreactor technology complement one another in an ideal manner. A general blueprint as well as a detailed example workflow is described, that allow easy adaption for other biotechnological optimization objectives. Performance of the jointly applied technologies is demonstrated on the maximization of GFP production with *C. glutamicum* by nutrition medium optimization. Next to results interpretation, modifications to the method are also critically discussed. This chapter is based on *Publication I* and *Publication IV*.

Author contributions:

Johannes Hemmerich and Lars Freier wrote the chapter with the help of Eric von Lieres, Marco Oldiges and Wolfgang Wiechert. Johannes Hemmerich and Lars Freier planned experiments, analyzed the data and interpreted results. Johannes Hemmerich performed experiments and prepared non-reprinted figures. Lars Freier performed Kriging analyses.

4.1. Blueprint of a general iterative optimization workflow for biotechnological applications based on Kriging-assisted Design of Experiments methodology

"Classic" DoE is a powerful tool for efficiently estimating single and combinatorial effects of input variables on noisy system outputs. It has a long tradition and was introduced by Fisher in 1935 [223]. DoE is established in several biotechnological fields and has been applied for numerous applications as summarized in [184]. Although DoE is popular for identifying significant input variables, the methodology usually lacks in approximating highly nonlinear functional relationships. For this

purpose, the Kriging approach is often more appropriate. Kriging provides a data driven, unbiased linear estimator with minimal mean square prediction error. Very recently, Kriging methodology was extended to incorporate also non-linear trend functions [224]. For further details on Kriging and its application in biotechnology, the reader is referred to the literature [167, 224–226].

A core business in industrial biotechnology using MCFs is the iterative process of strain engineering and optimization of bioprocess conditions. One important aspect is the improvement of cultivation medium to provide an optimal environment for microbial formation of the product of interest. It is well accepted that the media composition can dramatically influence overall bioprocess performance to improve recombinant protein production with microbial systems [227–229] and consequently, the adjustment of medium composition is a rewarding step in bioprocess development with respect to optimal productivity [76, 230–232]. In general, optimization studies benefit from increased experimental throughput, which becomes even more powerful in combination with DoE methodologies, such as to assess interactions between design variables or reduce high-dimensional search spaces.

On the example of medium optimization in order to maximize secretory GFP production with *C. glutamicum*, a protocol for media optimization is presented combining state-of-the-art laboratory automation, MBR technology with online process monitoring, and Kriging-based data analysis/experimental design. The Kriging methodology is implemented in a MATLAB toolbox ("KriKit") which can be downloaded and used free of charge [233].

The presented framework is divided into four steps, which are illustrated in Figure 4.1. The steps are indicated by box frames and correspond to sections of the detailed protocol found in the appendix (section C.1). The first step (Figure 4.1, first box) is to define the project goals and to determine the required methods. The combination of DoE methodologies, MBR technology, and laboratory automation allows an increased experimental throughput that demands powerful data processing. The second step (Figure 4.1, second box) aims to detect sensitive design variables (i.e., medium components) with high influence on the optimization objective. This leads to a reduced number of design variables of interest. The third step (Figure 4.1, third box) comprises an iterative optimization for a more detailed investigation of the functional relationship between the remaining design variables and the objective of interest. Using the successively extended data set, the Kriging approach is applied for predicting the experimental outcome at unmeasured locations. The iterative cycle stops as soon as the Kriging model predicts an optimum or plateau with sufficient accuracy. The results are verified in the fourth step (Figure 4.1, fourth box), beginning with a further sensitivity analysis around the identified optimum. If initially, insensitive components are found to be insensitive also in the optimal region, it is reasonable to assume that this holds true during the iterative optimization

procedure in the third step. Afterwards, it is advised to verify optimization results by application of orthogonal methods, like an activity assay or SDS-Page for the presented case.

The generic nature of the presented protocol allows for easy adaption to different laboratory equipment, expression hosts, and target proteins of choice, as well as bioprocess variables like pH value or cultivation temperature. Furthermore, optimization objectives like protein production rate, specific yield, or product quality can be chosen to fit the scope of other optimization studies.

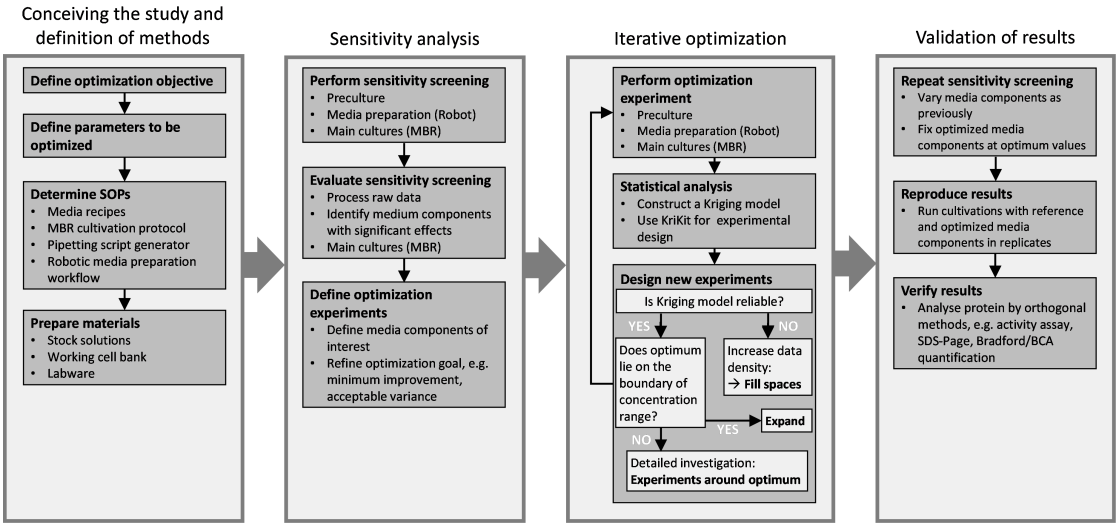


Figure 4.1.: Workflow depiction for Kriging-assisted Design of Experiments (DoE) methodology. Figure modified from [234].

4.2. Application example on media optimization for Tat-mediated green fluorescent protein secretion with *C. glutamicum*

As application example, maximization of secretory GFP production with *C. glutamicum* is shown by optimizing the composition of CgXII minimal medium. GFP titer was chosen as the optimization objective as it can be quantified easily and it is widely applied as model protein for studies on MBR systems [178, 235, 236]. To ensure reproducibility and comparability of results, conduction of reference cultivations is indispensable. Corresponding GFP signals serve as normalization standard for the GFP signal from cultivations with different medium compositions under investigation. The specific step-by-step workflow carried out for this example optimization is found in section C.1.

4.2.1. Selection of media components for investigation

A pre-selection of media components was performed on a rational basis, which served to narrow the window of design variables: glucose as the main carbon source fixed to 10 g L^{-1}, non-metabolizable 3-(N-morpholino)-propanesulfonic acid (MOPS) to provide buffering capacity, KH_2PO_4 and K_2HPO_4 providing also buffering capacity and serving as phosphate source, urea as basal nitrogen source and pH-stabilizing agent, biotin for complementation of biotin auxotrophy, protocatechuic acid (PCA) as iron chelating agent. Thus, there are still nine medium components left for investigation, namely $(NH_4)_2SO_4$, $FeSO_4 \cdot H_2O$, $ZnSO_4 \cdot 7\,H_2O$, $CuSO_4 \cdot 5\,H_2O$, $NiCl_2 \cdot 6\,H_2O$, $CaCl_2 \cdot 2\,H_2O$, $MgSO_4 \cdot 7\,H_2O$, and $CoCl_2 \cdot 6\,H_2O$. Additionally, $Na_2MoO_4 \cdot 2\,H_2O$ and H_3BO_3 were included as design variables, as those were described as additives in another CgXII formulation [237]. The medium components of interest (i.e., design variables) can be divided into three groups:

First, $(NH_4)_2SO_4$ as standard nitrogen source is a major nutrient. As mentioned before, the basal nitrogen source urea was not altered, thus the nitrogen supply was expected not to become growth limiting. In another study, modulating the nitrogen source was determined as promising target, although the overall optimization goal was to maximize biomass-specific GFP signal [36].

Second, $FeSO_4 \cdot 7\,H_2O$, $MnSO_4 \cdot H_2O$, $ZnSO_4\,7 \cdot H_2O$, $CuSO_4 \cdot 5\,H_2O$, $NiCl_2 \cdot 6\,H_2O$, and $CoCl_2 \cdot 6\,H_2O$ represent the group of trace elements, which seem to be inherited from the first publication of the CgXII medium [56]. To add more variations of this theme, $Na_2MoO_4 \cdot 2\,H_2O$ and H_3BO_3 were also investigated as mentioned before.

The third group is constituted by $MgSO_4 \cdot 7\,H_2O$ and $CaCl_2 \cdot 2\,H_2O$, as those exceed clearly the concentration range of trace elements and thus, their necessary presence to promote growth is most likely from other nature as for trace elements. Teramoto et al. reported that increased Ca^{2+} concentration correlates with increased GFP and Amylase secretion using the Tat pathway [227]. It was speculated that varying concentrations of Mg^{2+} and Ca^{2+} show significant effects on secretion of GFP, despite different conditions were used than by Teramoto et al. who applied *C. glutamicum* R as background strain, medium containing yeast extract and casamino acids, and CgR0949 as signal peptide. In this study, *C. glutamicum* ATCC13032, defined medium, and PhoD as signal peptide was employed.

4.2.2. Screening analysis

The MTP-based MBR system allows 48 experiments to be performed in parallel. Taking into account the maximal possible number of parallel experiments on one MTP (48) and the total number of media components (11) makes the 2_{IV}^{11-6} fractional design an appropriate choice. This experimental design comprises 32 experiments and allows the estimation of the main effect for each of the investigated

media components. The remaining cultivation wells (16) were used for multiple replicates of experiments with the reference medium to assess reproducibility and positional effects. That is, each experiment is conducted once (no replicates), except fpr the reference experiment (five replicates). Corresponding concentration ranges of investigated components are given in section C.1.

The statistics of all 16 normalized GFP signal curves are depicted Figure 4.2A. The relative standard deviation of the curves increases with time but never exceeds 1.7%. In the considered concentration range, varying the majority of the media components did not show a noticeable effect on the objective (Figure 4.2B). Component NH_4^+ shows a strong negative effect, while Ca^{2+} and Mg^{2+} show the strongest positive tendency. The effect of Mg^{2+} is not significant for the current concentration range but might be for a broader concentration range. Consequently, it was decided to omit NH_4^+ from the medium and to investigate the effect of Ca^{2+} and Mg^{2+} in further experiments. Nitrogen limitation is prevent by urea still present in the medium.

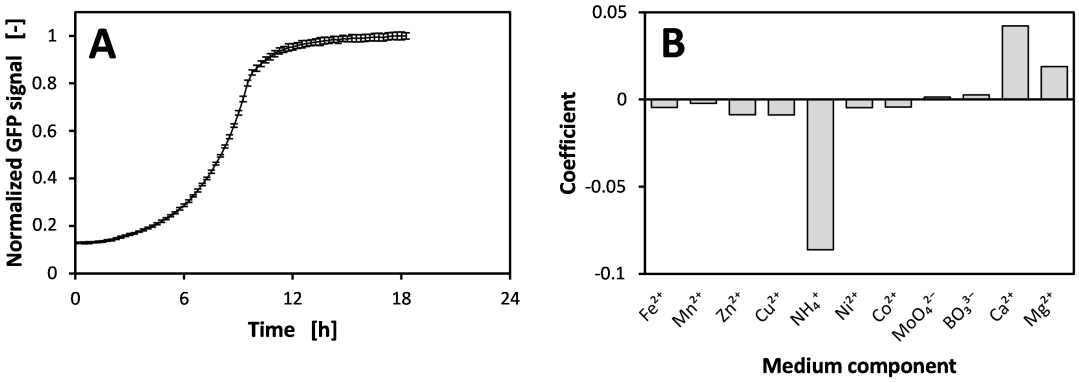


Figure 4.2.: Depiction of results from screening analysis. (A): Time course of normalized GFP signal during growth of reference cultivation experiments in one microbioreactor run (n = 16 biological replicates, depicted as mean value ± standard deviation). (B): Effect of investigated medium components on GFP signal. Figure modified from [167].

4.2.3. Iterative optimization

The experiments of the first iteration were planned such as to extend the knowledge obtained by the initial screening by investigating also potential interactions between the previously identified medium components of influence. For sake of clarity, all following concentrations refer to the concentrations used in the original reference medium (CgXII), indicated by x Ref. Furthermore, since the screening revealed that NH_4^+ has a significant negative effect on the GFP-Signal, it was also of interest if NH_4^+ could be omitted. A full factorial experimental design with eight experiments was

applied. Concentration ranges for Ca^{2+} and Mg^{2+} are chosen to be the same as in the screening analysis, that is minimal concentration was 0.4x Ref and maximal concentration was 2x Ref. For NH_4^+, the concentrations were set in the range of 0x Ref to 2x Ref. The GFP signal between the sample points was estimated using Universal Kriging [167], see Figure 4.3. As predicted by the screening analysis, Ca^{2+} and Mg^{2+} have a positive influence on the GFP signal, with the effect of Ca^{2+} being more pronounced. Furthermore, the negative effect of NH_4^+ was confirmed. Consequently, by omitting NH_4^+, the maximal GFP signal could be increased by ca. 30% (from 1.03 to 1.36, Figure 4.3A and B). The Kriging interpolations in Figure 4.3A indicate that high concentrations of Ca^{2+} and Mg^{2+} lead to high GFP signals. In order to validate this hypothesis, the maximal concentrations of Ca^{2+} and Mg^{2+} were doubled in the next experimental iteration.

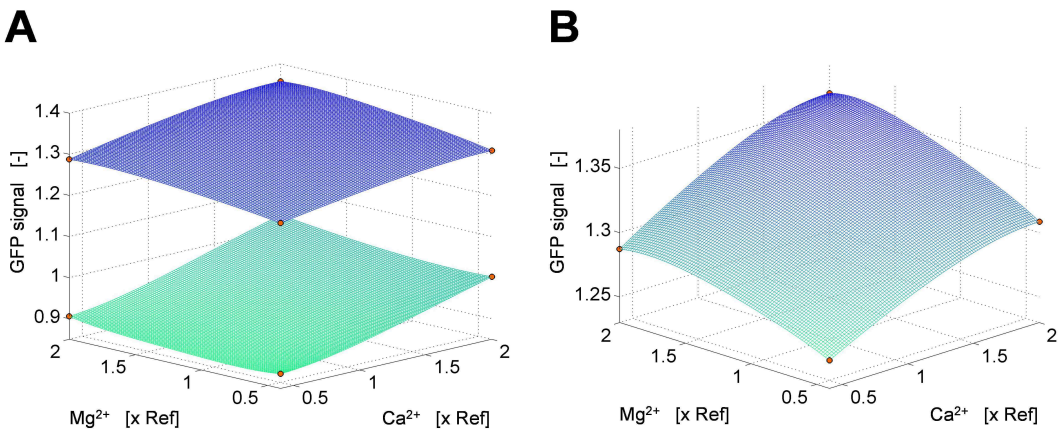


Figure 4.3.: Detailed view about impact of NH_4^+ on GFP signal. (A): Kriging interpolations of GFP signal when maximal amount of NH_4^+ (lower plane) or no NH_4^+ (upper plane) is added to medium. Sample points are indicated by red dots. X and Y axes denote relative concentration of ions used for medium preparation compared to CgXII-Medium. (B): Detailed view of Kriging interpolation using medium without NH_4^+. Figure reprinted from [167] with permission.

Two criteria were formulated that satisfy a good experimental design for Ordinary Kriging [238]: First, the design should be space filling for exploring the system and second, the design should comprise some points in close proximity for studying the dependency of the covariance between these neighboring sample points. Following these two criteria, the new experiments for the 2nd iteration were designed using a full factorial in the upper right quarter of the extended parameter space with one point located near to the maximal concentration level of Ca^{2+} and Mg^{2+} in iteration 1, see Figure 4.4A. Experimental results and the Kriging interpolation of iteration 2 confirmed the positive correlation of the GFP signal with Ca^{2+} and Mg^{2+}. Therefore, further enhancement of GFP signal was

expected for increased concentrations of both Ca^{2+} and Mg^{2+}.

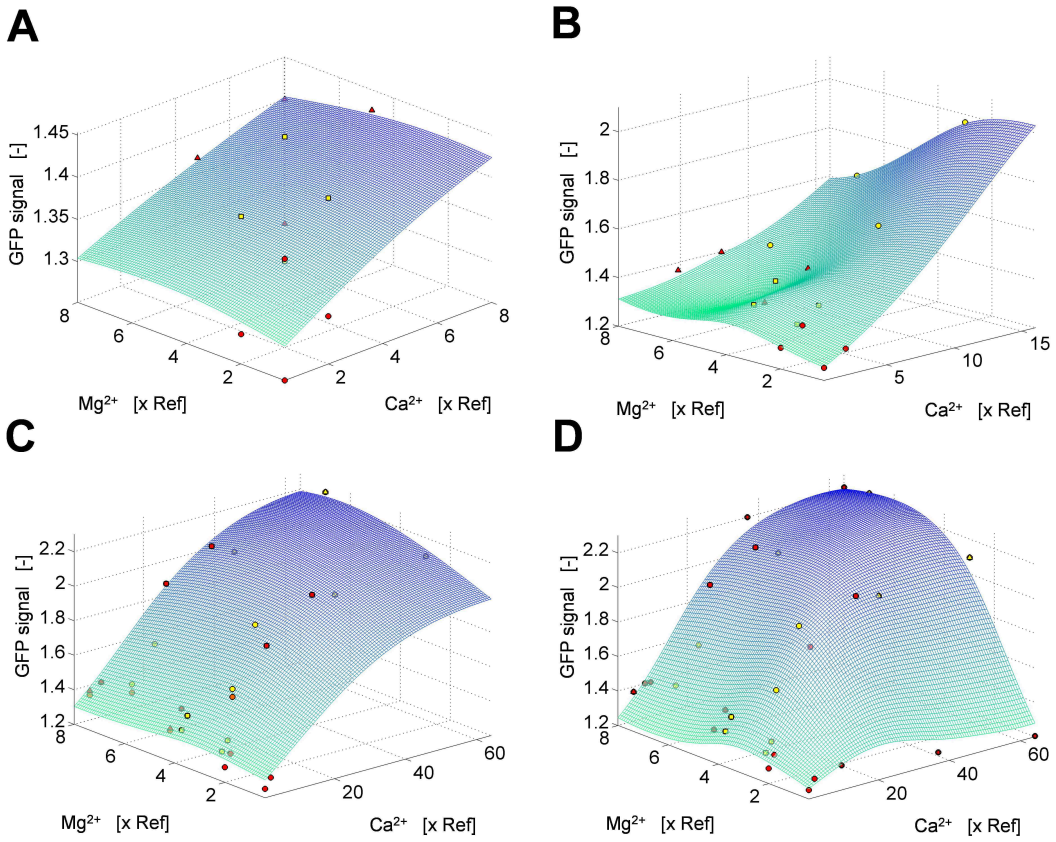


Figure 4.4.: Development of GFP signal prediction during iterative optimization. Results of iteration 3, 4, 6 and 7 are shown in panels A, B, C and D, respectively. Experimental sample points are indicated by different symbols, according to the corresponding iterations. Figure parts reprinted from [167] with permission.

The experiments of the 3rd iteration were planned in a similar manner as in the 2nd iteration: The maximal concentration of Ca^{2+} and Mg^{2+} was again doubled and a full factorial was placed in the upper right part of the new parameter space. Experimental data of the iterations 1–3 and the associated Kriging interpolation are presented in Figure 4.3A and B. Apparently, samples with a Mg^{2+} concentration of 8x Ref do not lead to higher GFP signals compared to samples with a Mg^{2+} concentration of 4x Ref, indicating that an optimal concentration range of Mg^{2+} has potentially been found. However, still no indication for a saturation of the response of GFP for increasing Ca^{2+} concentrations could be found in the third iteration. Consequently, the experimental design for iteration 4 was placed around the potential optimal Mg^{2+} concentration (4x Ref) and an increase in the Ca^{2+} concentration range

from 9.6x Ref to 16x Ref.

The results of iteration 4 are visualized in Figure 4.4B and show that the GFP signal could be doubled compared to the reference cultivations when applying a Ca^{2+} concentration of 16x Ref and a Mg^{2+} concentration of 3x Ref. Kriging interpolation predicts a steep increase in the GFP signal in the Ca^{2+} concentration range from 8x Ref to 16x Ref. Although a clear explanation for this phenomenon cannot be given, increased Ca^{2+} concentrations have been described to enhance recombinant protein secretion before [227]. Furthermore, the optimum with respect to Mg^{2+} found in the previous iteration could be confirmed.

Since an optimal Ca^{2+} concentration was not detected yet, the experimental design from iteration 4 was again shifted towards a higher Ca^{2+} concentration range in iteration five. That is, Ca^{2+} concentration was varied in a range from 19.2x Ref to 32x Ref, and range of Mg^{2+} was not changed due to the confirmed optimum. Experiments of iteration 5 and associated Kriging interpolation gave first indications for a saturation of the positive effect of Ca^{2+} on GFP signal (Figure 4.4C).

In order to verify this finding, the experimental plan for iteration 6 was constructed likewise as for iterations 4 and 5 by doubling the concentration range for Ca^{2+}, see Figure 4.4C. Experimental data from iterations 1–6 and the associated Kriging interpolation are depicted in Figure 4.4C. The Kriging interpolation clearly shows a flattening of the GFP signal at high Ca^{2+} concentrations. As will be discussed in more detail in section 4.2.6, the plateau is most likely caused by precipitation of solid Ca-complexes or other Ca-compounds in combination with other medium components, which leads to limited accessibility of soluble Ca^{2+} ions to the cells. Hence, no further increase in GFP signal is expected for higher Ca^{2+} concentrations due to the formation of solid non-bioavailable Ca-complexes or other Ca-compounds and not due to saturated uptake of Ca^{2+} by the biological system. Moreover, it appears that the negative effect of high Mg^{2+} concentrations can be neutralized by high Ca^{2+} concentrations.

The final iteration 7 was planned with the intention to explore the boundaries of the parameter space, that is for the applied minimal and maximal Mg^{2+} concentrations. Results are shown in Figure 4.4D. The Kriging interpolation reveals that very low Mg^{2+} concentrations lead to a significant decrease in the GFP signal, especially in the case of high Ca^{2+} concentrations. Consequently, the data of Figure 4.4D justifies the assumption that both cations Ca^{2+} and Mg^{2+} cannot replace each other.

4.2.4. Identifying optimal parameter regions

The results show that an optimal region with respect to Ca^{2+} and Mg^{2+} concentration exists. Identifying the boundaries of this region is a nontrivial task since noisy measurements and low number of

data points may lead to inaccuracies in the Kriging prediction model, but Kriging provides an estimation of the prediction error. This allows to make a statement about the significance of differences in predicted output signals. The KriKit toolbox allows to perform a *z*-test, based on the Kriging interpolation as well as its estimated prediction error [167]. Figure 4.5 visualizes the results of the *z*-test: The blue region indicates medium compositions that lead to significantly lower GFP signals than the best prediction. Parameter values in the red region will most likely lead to maximal GFP signals. The identified optimal region is located in the upper right corner of the investigated design space, that is high Mg^{2+} and Ca^{2+} concentrations. It is not expected that higher concentrations of Mg^{2+} and Ca^{2+} ions will lead to further improvements of GFP secretion. Considering Mg^{2+}, the results of iteration 3 showed that for a low Ca^{2+} concentration, high amounts of Mg^{2+} lead eventually to decreasing values in the GFP signal. It seems likely that this also holds for high Ca^{2+} concentration, although the positive effect is believed to be limited by precipitation of Ca^{2+} salts.

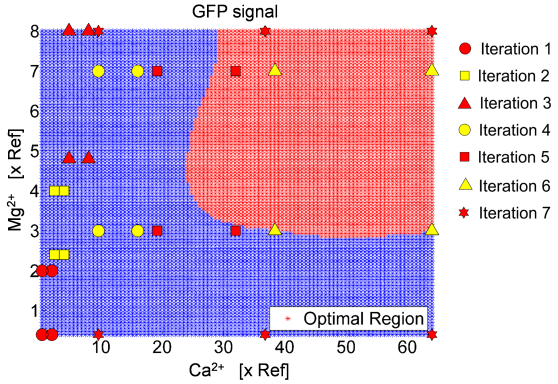


Figure 4.5.: Identification of optimal parameter region based on *z*-test and Kriging interpolation. Figure reprinted from [167] with permission.

4.2.5. Validation of results

The identified optimal concentrations for Ca^{2+} and Mg^{2+} cause a doubled GFP signal compared to the reference composition of CgXII medium. To confirm the optimized medium composition finally, a validation screening was conducted to show that the results of the screening analysis (section 4.2.2) are also valid with optimized concentrations of Ca^{2+} and Mg^{2+}. Consequently, all initial medium components of interest were varied except for Ca^{2+} and Mg^{2+}. These were fixed at values of ${Ca^{2+}}_{opt}$ = 32x Ref and ${Mg^{2+}}_{opt}$ = 6.8x Ref that yield optimal GFP signals and result in feasible pipetting volumes for the applied LHS. This sample point is part of the predicted optimal region, near the left border

(see Figure 4.5), as higher Ca^{2+} concentrations are not expected to cause a significant increase in GFP signal. The applied fractional factorial is similar to the experimental design of the initial screening with a resolution of IV and allows an analysis of main factor effects without confounding with pairwise interaction. Additionally, eleven samples were placed at the defined optimal point for assessing biological reproducibility.

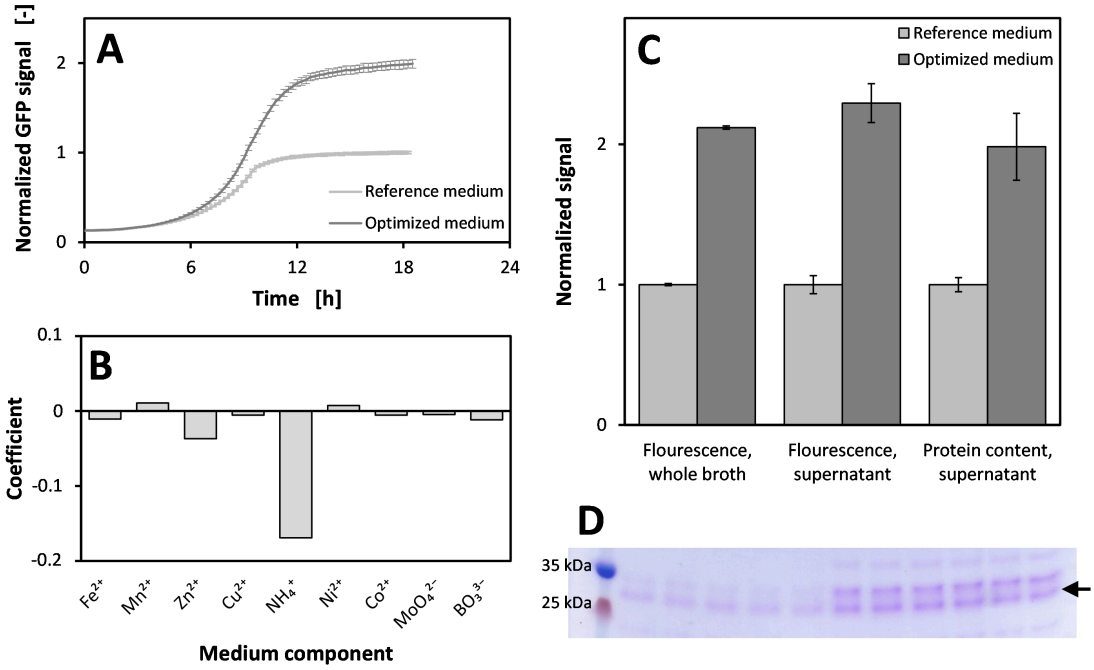


Figure 4.6.: Summary of results from validation screening. (A): Comparison of normalized GFP signal over time for reference cultivations (n = 16 replicates) and cultivations with optimized medium composition as determined after iterative optimization supported by Kriging interpolation (n = 10 replicates). Mean values ± standard deviation are shown. (B): Impact of medium components on GFP signal at optimized medium composition. (C): Comparison of GFP fluorescence from whole cultivation broth and supernatant after cultivation and protein content in supernatant, made from reference (n = 6) and optimized medium (n = 6). (D): SDS-page analysis of cultivation supernatants in reference medium (n = 5) and in optimized medium (n = 6), GFP band (26 kDa) is indicated by an arrow. Figure modified as well as reprinted in parts from [167] with permission.

Figure 4.6A depicts the statistics of ten normalized GFP signal curves at the optimal point. The remaining eleventh curve was excluded as an experimental outlier as it differed significantly from the other curves. It can be seen that the mean normalized GFP value from the cultivations with optimized Ca^{2+} and Mg^{2+} concentration is doubled compared to the reference cultivations. The mean value of the curves converges to 2x Ref. The relative standard deviation of the GFP signal increases with time but never exceeds 3.7%, which is below the maximum standard deviation (5%) of the BioLector

device measurements, according to manufacturer data. The screening analysis results are visualized in Figure 4.6B and indicate that NH_4^+ still has a strong negative effect on the GFP signal. Remaining components seem to have a very low or no effect. It can consequently be concluded that changes in Ca^{2+} and Mg^{2+} concentrations do not affect the impact of the other medium components on the GFP signal. For the used expression strain, it was reported that the majority of the mature GFP is located extracellularly [239]. Consequently, the GFP fluorescence of the fermentation suspension can be considered to be appropriate for capturing the amount of secreted GFP. However, additional analytics have been conducted from a separate cultivation run with six biological replicates for both reference and optimized medium composition. As depicted in Figure 4.6C, both fluorescence and protein content of cultivation supernatants from optimized cultivations have doubled compared to those from reference medium. This is in agreement with GFP signals from the whole fermentation suspension. Indeed, analysis by SDS-page verifies that the amount of secreted GFP is higher for the optimized medium (Figure 4.6D).

4.2.6. Interpretation of results

Response data of the optimized medium composition was re-evaluated to give indication of the previously speculated underlying reason of the limited effect of increased Ca^{2+} and Mg^{2+} addition. In cultivation wells with high $CaCl_2$ concentrations (approximately 0.1 to 5.5 mM, i.e. approx. 10 to 50x Ref), white turbidity of the medium was observed by optical inspection. Therefore, it can be assumed that Ca^{2+} forms compounds with other medium components like PO_4^{3-} and SO_4^{2-}, which results in precipitation as soon as Ca^{2+} concentration reaches a certain threshold in CgXII medium. Consequently, associated calcium ions are temporarily not accessible to the microorganism, but may re-dissolve during growth when soluble Ca^{2+} is incorporated to increasing biomass. However, associated reduction in turbidity will be covered by increasing turbidity due to biomass formation.

For the optimized medium, only concentrations of Ca^{2+} and Mg^{2+} have been increased, thus cultivation data for all iterations regarding the initial BS signal have been related to the applied concentration of those cations. Kriging was used for interpolating the functional relationship between the concentration of Ca^{2+} and Mg^{2+} and BS signal. As illustrated in Figure 4.7, the BS signal shows a positive correlation with increasing Ca^{2+} concentration, which is most likely caused by precipitation. Mg^{2+} shows only a positive influence for high Ca^{2+} concentration. The combination of qualitative optical properties showing turbidity of medium, experimental output in terms of limited effect of increased Ca^{2+} addition on GFP secretion and statistical evaluation of presumed correlation between initial BS signal and Ca^{2+} concentration suggests that the saturating effect of Ca_2^+ on GFP signal is due to the

physical constraint of Ca_2^+ solubility.

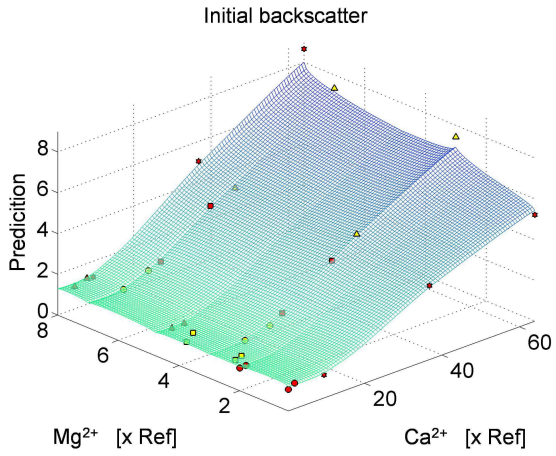


Figure 4.7.: Kriging interpolation of functional relationship between medium component Ca^{2+} and Mg^{2+} and initial backscatter signal as indicator for precipitation. Figure reprinted from [167] with permission.

Despite the verified finding that increased concentrations of Ca^{2+} promote enhanced secretion of recombinant GFP in *C. glutamicum*, the underlying reason is not clear. In the following, three hypotheses are presented:

In some cases, an increase of Ca^{2+} ions in the medium has been reported to increase the amount of active (i.e., correctly folded) secreted protein. This phenomenon was explained by the need for divalent cations to support correct folding of the target protein which incorporates those cations as cofactor. In case of GFP, such an effect can be ruled out here, as GFP does not need divalent cations as folding cofactors [240].

Another hypothesis considers the cell wall as several layers of a molecular sieve acting as a depth filter, whose pores are permeable for endogenous extracellular proteins due to evolutionary adaption. This may not be the case for secretory heterologous proteins during their passage to the outside of the cell, but the incorporation of Ca^{2+} may enlarge the pore size of this assumed depth filter around the cell which could facilitate the release of recombinant GFP.

Furthermore, Ca^{2+} ions might neutralize negative charges of the "continuum of anionic charge" formed by lipoteichonic acids and wall teichonic acids found in the cell wall of Gram-positives [241]. In combination with indications that GFP presents positive charges on the outer side [242], reduction of electrostatic interaction could facilitate the passage of GFP to the extracellular medium.

For *Bacillus brevis* it was shown that with increasing concentrations of $MgSO_4$, $MgCl_2$ or $CaCl_2$ (up to 5 mM) the biosynthesis of certain cell wall proteins is remarkably decreased [243]. In this study,

optimal concentrations of approximately 7 mM (6.8x Ref) for $MgSO_4 \cdot 7\ H_2O$ and approximately 2.9 mM (32x Ref) of $CaCl_2 \cdot 2\ H_2O$ were identified, which are in the same order of magnitude. Maybe a similar phenomenon as described for *B. brevis* occurs also in *C. glutamicum* as it is also a Gram-positive organism. Thus, it is speculated that a high load of Ca^{2+} causes down-regulation of cell wall protein synthesis and induces a "leaky" cell wall.

However, the last two assumptions imply that fluorescence of cell wall retained GFP is shadowed compared to GFP located in the medium and clearly, the presented hypotheses remain speculative.

4.3. Modifications to the method

Next to multi-purpose and expandable robotic liquid handling systems like the one used in this study, it should be mentioned that there are several smaller liquid handling systems commercially available. These systems are capable to perform this task and can be placed inside of laminar flow work benches. If no automated pipetting system is available, different media compositions according to the DoE plan can also be realized by manual pipetting using single and/or multi-channel pipettes. Since the manual preparation is more error-prone and will require highly focused work for quite a long time, it is recommended to prepare a lower number of different media compositions.

Depending on the capabilities of the employed MBR system, the corresponding cultivation protocol will vary. For instance, if no online measurement of biomass formation is available, it may be sufficient to measure biomass concentration after completion of the growth experiment. In combination with online monitoring of the pH and DO, which is implemented in several MBR systems, the growth saturation can be determined safely.

In principle, the growth experiments can be conducted in MTPs alone placed inside shaking incubators, without the use of a MBR system. In this case, proper cultivation conditions have to be ensured, which is addressed by three aspects:

(1) Oxygen-limited cultivations can be avoided by using MTPs with suitable geometries, in combination with proper shaking frequencies and shaking diameters, for example, square 96 or 24 deep well plates operated at 1000 rpm at 3 mm throw or at 250 rpm at 25 mm throw, respectively. Importantly, the lower the achievable maximal oxygen transfer rates, the lower the main carbon source should be concentrated. As mentioned above, for this study, the use of 10 g L^{-1} glucose was suitable to prevent oxygen limitation for the employed cultivation conditions.

(2) Sampling of the MTP cultures for biomass and product quantification should be reduced to a minimum. Each time the MTP is removed from the shaking incubator, oxygen transfer will

immediately break-down which may result in unfavorable cultivation conditions.

(3) In the opinion of the author, the use of MTP readers as cultivation devices is not recommended as these devices were not developed for this purpose. For example, shaking mechanics were built for occasional mixing of microplates after the reagent addition and thus, often lack robustness for long runs of continuous shaking lasting for days. Moreover, sufficient power input needed for microbial cultivations cannot be realized in these readers. The integration of optical density readings in short time intervals requires stopping of the shaking motion, resulting in repeated periods of oxygen limitation. Furthermore, evaporation in such systems over long cultivation periods will distort results severely for several reasons.

4.4. Further considerations

To speed-up iterative optimization steps, it is advised to carefully select the analytical method for product quantification. Fast and simple methods should be preferred at the cost of precision and accuracy, as the iterative experimental design strategy tolerates experimental inaccuracy. However, the final results must be verified against sufficiently precise and accurate product quantification methods that might be more complicated. In general, careful evaluation and decision making about the study procedures require effort in the beginning of the study, but pay out in the long run, after routine methods have been established.

It is highly recommended to define a reference experiment that is compared to all experiments during the optimization. That is, the applied medium component concentrations as well as measured output are normalized via dividing by reference values. This way, each applied and measured value can be interpreted as the x-fold of the reference value. To take into account variations between the plates, five reference experiments were performed on each plate in this study. The mean value of the measured outcome is used for the normalization.

It can generally not be guaranteed that the developed medium is also optimal for other strains. However, the improved medium will most likely also be appropriate for cultivating expression strains with small genetic differences, for example, when producing enzyme variants with single amino acid substitutions obtained from mutagenesis studies, although even single point mutations have been described to effect cellular metabolism and heterologous expression performance [244, 245]. In this case, the presented protocol can be a first step, followed by protocols for high-throughput expression screenings [246]. If the protocol is used for medium development with subsequent scale-up to fed-batch cultivations, the optimized medium should be verified for the corresponding bioprocess

conditions, as clone screening campaigns at the microscale identified different top performers for different feeding strategies and cultivation media [112, 247]. Furthermore, the introduced KriKit [233] can generally contribute to improved holistic bioprocess optimization. Very recently, the tool's abilities were extended to also support multi-objective optimization [248], which can be important for optimizing both upstream and downstream processes [165, 173].

5. Combinatorial impact of Sec signal peptides from *Bacillus subtilis* and bioprocess conditions on heterologous cutinase secretion by *Corynebacterium glutamicum*

This chapter leaves the field of MBR cultivation and takes a step forward into bioprocess development using lab scale bioreactors. To complement the current knowledge on how to select the optimal SP for different expression hosts and different target proteins of choice, the interrelation of bioprocess control strategy and choice of SP to optimize cutinase secretion with *C. glutamicum* is investigated in detail. Furthermore, it is discussed how the results are interpreted with changing bioprocess optimization objectives. This chapter is based on *Manuscript I*.

Author contributions:
Johannes Hemmerich designed and performed experiments, analyzed the data, interpreted the results, prepared the figures, and wrote the chapter with the help of Roland Freudl, Marco Oldiges and Wolfgang Wiechert. Matthias Moch performed experiments. Sarah Jurischka contributed strains.

5.1. Selection of Sec signal peptides for heterologous cutinase secretion

In this study, the combinatorial effect of technical and biological variables was systemically investigated for their impact on heterologous protein secretion performance with *C. glutamicum*. Until now, the impact of such variables is typically investigated in sequential order, for example starting with molecular strain optimization followed by bioprocess engineering for the apparent best strain only. Here, about 150 laboratory scale bioreactor runs are investigated systematically to cover the envisaged design space. This design space comprises feeding profile, Sec SP and IPTG concentration as well as reproducibility of results by conducting each bioreactor cultivation in biological duplicates.

Cutinase from *Fusarium solani pisi* [249] was used as model protein of eukaryotic origin for this study on secretory production of heterologous proteins with *C. glutamicum*. Application examples of cutinase comprise laundry and dishwashing detergents, food industry and textile processing [250]. Sec SPs from *B. subtilis* were shown to enable heterologous cutinase secretion in *C. glutamicum* [36], yet results from SP library screening in *B. subtilis*, to identify SPs causing high cutinase secretion efficiency in this organism, cannot be transferred to *C. glutamicum* [37]. Because it is not known which SP is best suited to secrete cutinase with *C. glutamicum* for various bioprocess conditions a priori, five Sec SPs from *B. subtilis* were chosen, associated with a high (Epr), average (YwmC) and low (YpjP) cutinase secretion performance in *B. subtilis* [30]. Additionally, two SPs (AmyE, NprE) causing nearly the same below-average secretion performance in *B. subtilis* were chosen (cf. Figure 5.1A).

The resulting recombinant *C. glutamicum* secretion strains, denoted by the respective applied Sec SP, were evaluated in different bioprocess regimes (batch, fed-batch with varying adjusted growth rates, and induction strength). Comparative evaluation of the different bioprocess control strategies is based on observed specific growth rate (μ_{Exp}) and biomass specific cutinase yield ($Y_{P/X}$) which are calculated from the experimental data obtained during the fed-batch phase of the cultivation. The application of exponential glucose feeding profiles to control the specific growth rate of the culture resulted in highly reproducible biomass formation and cutinase secretion, as seen for the time courses of cell dry weight and extracellular cutinase activity from two representative replicate cultivations (Figure 5.1B and C).

5.2. Impact of heterologous cutinase secretion on growth

Among glucose-limited fed-batch cultivations, the plasmid-free wild type strain (denoted as WT) shows a perfect match of the observed growth rates μ_{Exp} with the adjusted growth rates μ_{Set}. In contrast, the five cutinase secretion strains often show a reduction in observed growth rates compared to strain WT. For instance, observed growth rates of μ_{Exp} = 0.07 h h^{-1} are found for strains Epr and NprE with 50 µM IPTG for the fed-batch cultivation with μ_{Set} = 0.10 h^{-1} with 50 µM IPTG, corresponding to a reduction of 0.03 h^{-1} on an absolute basis compared to the value for μ_{Set} as well as the observed growth rate of μ_{Exp} = 0.10 h^{-1} obtained with strain WT. For the fed-batch bioprocesses operated at feeding rates corresponding to values of μ_{Set} = 0.15, 0.20 and 0.25 h^{-1}, maximum growth rate reductions of 0.02, 0.02 and 0.03 h^{-1}, respectively, are observed among the cutinase secretion strains. For the batch cultivations, where the glucose uptake by the culture is not restricted externally, observed growth rate for strain WT was μ_{Exp} = 0.38 ± 0.01 h^{-1}, which is in good accordance with

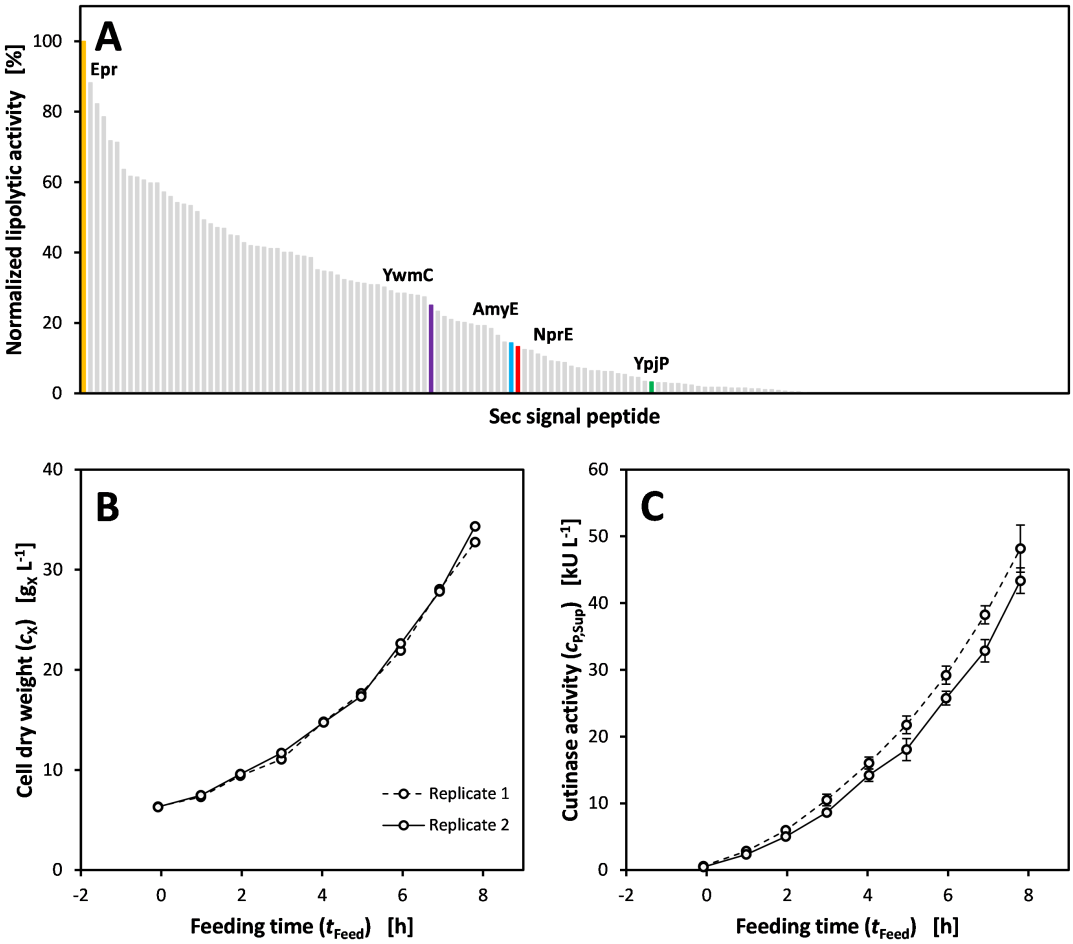


Figure 5.1.: Selection of Sec signal peptides (SPs) for cutinase secretion with *C. glutamicum* and representative process performance of secretory cutinase production. (A): Sec SPs from *Bacillus subtilis* ranked with respect to extracellular cutinase activity (expressed as normalized lipolytic activity) using *B. subtilis* as expression host as reported [30]. The five Sec SPs in this study used for cutinase secretion by *C. glutamicum* are highlighted, representing top (Epr), medium (YwmC) and poor (YpjP) performers. Additionally, two neighboring lower-mid range ranked SPs (AmyE, NprE) were chosen. Graph is reproduced from published data [30]. Color coding for Sec SPs is used throughout the whole text. (B): Time profiles of cell dry weight (c_X) for *C. glutamicum* pEKEx2-NprE-cutinase. (C): Time profiles of cutinase activity in cultivation supernatant ($c_{P,Sup}$) with error bars representing standard deviations from cutinase activity measurements carried out in analytical triplicates for *C. glutamicum* pEKEx2-NprE-cutinase. Cultures were induced with 50 µM IPTG during an exponential feed profile for a calculated growth rate of μ_{Set} = 0.25 h^{-1}.

literature [55, 219, 251]. For the cutinase secretion strains, values of observed growth rates ranged from μ_{Exp} = 0.32 ± 0.01 to 0.34 ± 0.02 h^{-1}, corresponding to a maximum reduction of 0.06 h^{-1} with strain NprE. A detailed list of all observed growth rates is given in Table D.1.

These reduced growth rates observed for the plasmid containing cutinase secretion strains could be interpreted as metabolic burden due to higher demand for energy equivalents and building blocks required for plasmid replication and heterologous gene expression [96, 252, 253]. Such phenomenon was described in literature where, for example, the impact of enhanced green fluorescent protein (EGFP) production with *C. glutamicum* was investigated [254]. The recombinant EGFP production strain showed retarded growth compared to the corresponding wild-type strain in bioreactor cultivation processes, which were designed to ensure glucose presence at non-growth limiting concentrations. Also, this was observed for secretory production of a β-peptidyl-aminopeptidase with the eukaryotic microorganism *Pichia pastoris* [255]. Specifically, in batch cultivations the recombinant strain showed 46% reduction in growth rate compared to the wild type strain, which was not seen in glucose-limited exponential fed-batch processes.

5.3. Relative cutinase yield is negatively correlated with growth rate, independently of the applied Sec signal peptide

Biomass specific product yield $Y_{P/X}$ describes the efficiency of the biomass to synthesize the desired product from consumed substrate in bioprocesses with growth-coupled product formation [100, 256]. A high or low value of $Y_{P/X}$ indicates a high or low usage of substrate for product synthesis in competition to biomass formation. Thus, $Y_{P/X}$ is considered as a strain-specific, biological parameter of the bioprocess.

The cutinase yields from bioprocesses concerning a certain Sec SP were normalized to the maximum cutinase yield obtained with that SP, to evaluate the impact of different bioprocess control strategies independently from effects of the chosen Sec SP. Hence, bioprocesses can be evaluated with respect to cutinase yield for the different applied Sec SPs on a relative basis. The corresponding data are shown in Figure 5.2. Five groups of relative cutinase yields are depicted in Figure 5.2, with each group representing the data obtained using the respective Sec SP (AmyE, Epr, NprE, YpjP and YwmC, as indicated) enabling secretion of cutinase.

Evaluation of the values for relative $Y_{P/X}$ for each SP reveals that the highest and lowest cutinase yield for each SP is always obtained for fed-batch cultivations with the lowest and highest adjusted growth rate, respectively. For example, increasing the adjusted growth rate from μ_{Set} = 0.10 to 0.15 $^{-1}$

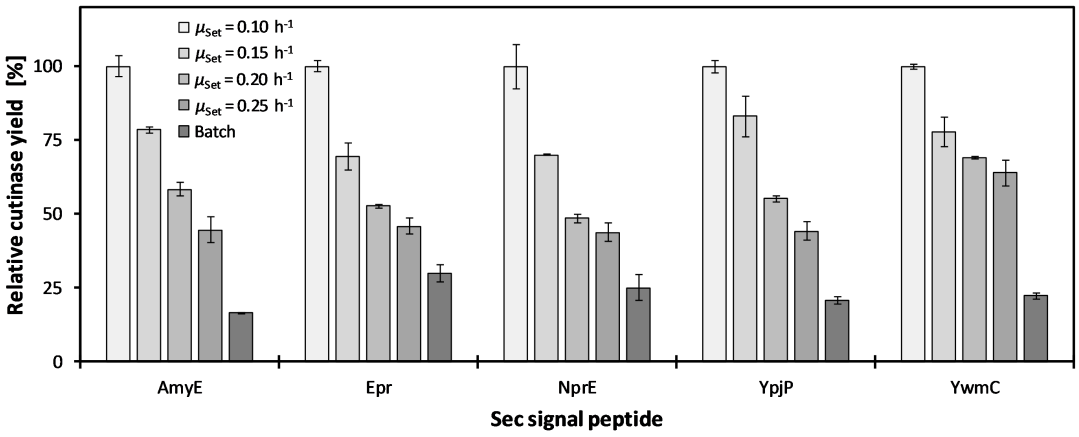


Figure 5.2.: Comparison of relative biomass specific cutinase yields for different Sec SPs and bioprocess conditions (i.e., fed batch processes with different adjusted growth rates μ_{Set} and batch process), grouped by SPs. Data series of cutinase yields for each SP were normalized to their respective absolute maximum value (AmyE: 2006 ± 100 U g_X^{-1}, Epr: 3056 ± 79 U g_X^{-1}, NprE: 3423 ± 361 U g_X^{-1}, YpjP: 2680 ± 80 U g_X^{-1}, YwmC: 1652 ± 20 U g_X^{-1}). Cutinase secretion was induced with 1000 µM IPTG. All data are mean values with min/max values indicated by error bars from two biological replicates.

is accompanied by a decrease in relative cutinase yield from 100% to roughly 70 - 80% for all SPs. Relative cutinase yields from the remaining cultivations fit nicely into this picture: With decreasing growth rate the cutinase yield increases.

Such an inverse correlation was also reported in various other studies, for example on the production of a nuclease with *E. coli* [257], a β-aminopeptidase with *E. coli* [258], GFP with *B. subtilis* [259], GFP with *E. coli* (Bienick et al. 2014), or LacZ with *E. coli* [260]. In these studies, the observed $Y_{P/X} = f(\mu)$ relations were explained by redirection of metabolic flux to meet the increased energy demand from heterologous protein production [258] and modeling of growth-rate dependent ribosomal capacity [259–261]. However, these reports concerned intracellular protein production in batch cultivations, meaning that the observed reduced growth rates were a result of heterologous protein biosynthesis. In contrast, this study deals with secretory heterologous protein production during fed-batch cultivations with adjusted growth rates, meaning that observed growth rates are not caused by endogenous factors, but rather by the applied bioprocess control strategy.

As seen in Figure 5.2, it can be assumed that the chosen Sec SP is the main determinant of cutinase secretion efficiency and with the adjustment of bioprocess conditions, this pre-determined cutinase secretion efficiency is maximal or reduced.

5.4. The choice of Sec signal peptide in combination with bioprocess conditions maximizes absolute cutinase yield

Next, a comparative overview of absolute cutinase yields observed for the different bioprocess conditions tested is given in Figure 5.3. It shows absolute cutinase yields $Y_{P/X}$ as function of observed growth rates μ_{Exp}. Data points belonging to each of the five cutinase secretion strains are color-encoded and connected by a line. For each IPTG concentration used (50, 250, and 1000 µM), one panel is shown in Figure 5.3 (A, B, and C). In addition to the data visualization, the collected data on cutinase yield $Y_{P/X}$ was subjected to a two-step multiple regression analysis. This model-based approach allows reconciling all data at once. First, the data was fitted to a model comprising an intercept, linear terms, and interaction terms between the different predictor variables. Second, the model was reduced by removing parameters with p-values above 0.05. The values of the removed parameters were zero or close to zero, meaning that the corresponding variables have no effect on cutinase yield, according to the regression analysis. Then, the data was fitted to the reduced model, and results are shown in Table 5.1.

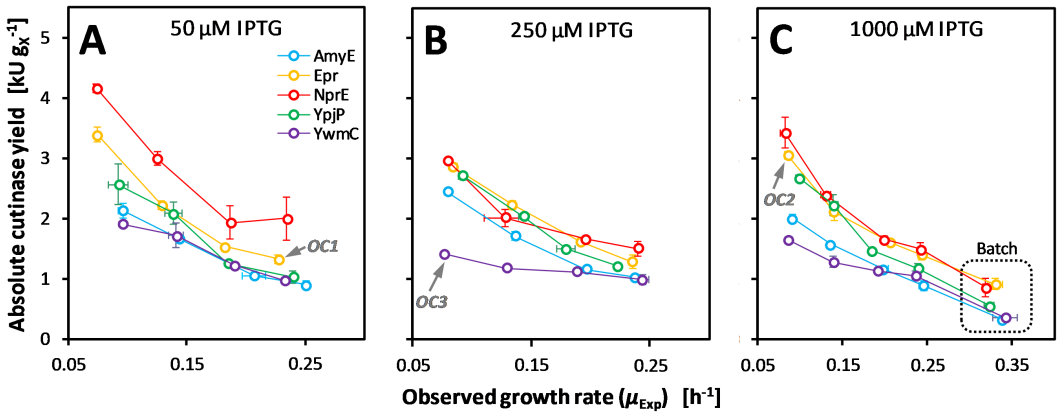


Figure 5.3.: Absolute cutinase yield ($Y_{P/X}$) as function of observed specific growth rate (μ_{Exp}). The five Sec SPs are encoded by colors (AmyE, blue; Epr, orange; NprE, red; YpjP, green; YwmC, purple) used for cutinase secretion. IPTG was added to 50 µM (B), 250 µM (C), or 1000 µM (D), respectively. Data from batch cultivations in (D) are marked by a dotted frame. Three operating conditions ("*OC1*" to "*OC3*") are indicated which are referred to later in the text. All data are mean values with min/max values indicated by error bars from two biological replicates. Detailed numerical values for the data points of $Y_{P/X}$ and μ_{Exp} shown in this figure are listed in Table D.2 and Table D.1, respectively.

Table 5.1.: Results from multiple regression analysis.

Description	Model parameter	Estimate (SE)
Linear terms		
AmyE	β_1	2.985 (0.177)
Epr	β_2	4.120 (0.177)
NprE	β_3	4.541 (0.177)
YpjP	β_4	3.666 (0.177)
YwmC	β_5	2.073 (0.177)
Interaction terms		
AmyE * μ_{Set}	$\beta_{1,6}$	-8.590 (0.962)
Epr * μ_{Set}	$\beta_{2,6}$	-11.783 (0.962)
NprE * μ_{Set}	$\beta_{3,6}$	-12.497 (0.962)
YpjP * μ_{Set}	$\beta_{4,6}$	-10.463 (0.962)
YwmC * μ_{Set}	$\beta_{5,6}$	-4.367 (0.962)
Regression diagnostics		
R^2 (adjusted R^2)	0.906 (0.899)	
RMSE	0.263	

Data were fitted against the reduced model $Y_{P/X} = \sum_{i=1}^{5} \beta_i x_i + \sum_{i=1}^{5} \beta_{i,6} x_i x_6$ with the categorical predictor variables $x_1 ... x_5$ encoding the used SP (i.e., AmyE, Epr, NprE, YpjP, and YwmC) and the predictor variable x_6 encoding the adjusted growth rate. Data from batch cultivations were not incorporated since for this kind of bioprocess no growth rate can be adjusted. Calculations were made using MATLAB's `fitlm` function. Results from the regression analysis of the full model are given in Table D.4. *SE* standard error. *RMSE* root mean squared error.

In most cases, the same rank order for the SPs with respect to absolute cutinase yield is observed (Figure 5.3). For example, the order of SPs for the lowest values of μ_{Exp} seen with 50 µM and 1000 µM IPTG (in Figure 5.3A and C, respectively) is clearly NprE, Epr, YpjP, AmyE and YwmC, from best to worst. In case an induction strength of 250 µM IPTG is used (Figure 5.3B), the NprE, Epr and YpjP results in nearly the same cutinase yield ($Y_{P/X}$ = 2.96, 2.87 and 2.72 kU g_X^{-1}, respectively) while the use of AmyE SP causes a slightly lower value of $Y_{P/X}$ = 2.46 kU g_X^{-1} and the use of the YmwC SP results in a distinct lower value of $Y_{P/X}$ = 1.42 kU g_X^{-1}. However, the SP rank order remains.

The above described particular SP rank order of bioprocesses with lowest observed growth rates and an induction strength of 50 µM IPTG becomes less clear for the fed-batch cultivations with the

highest adjusted growth rate at the same induction strength (Figure 5.3A). For this bioprocess operation, the use of the NprE SP is still the best choice ($Y_{P/X}$ = 2.00 kU g_X^{-1}) while the application of Epr results in a distinct lower cutinase yield ($Y_{P/X}$ = 1.34 kU g_X^{-1}) which is comparably low as found for the remaining three SPs (AmyE: 0.90 kU g_X^{-1}, YpjP: 1.04 kU g_X^{-1}, YwmC: 0.98 kU g_X^{-1}). The observed SP rank order is also suggested by the results of the regression analysis, reflected by the estimated parameter values of the linear terms: The higher the value, the more pronounced the (positive) effect on cutinase yield (NprE: 4.541, Epr: 4.120, YpjP: 3.666, AmyE: 2.985, YwmC: 2.073, see Table 5.1).

Values for $Y_{P/X}$ obtained with the five Sec SPs are more similar at high growth rates than at low growth rates. For example, in bioprocesses with an induction strength of 50 µM IPTG, values of $Y_{P/X}^{NprE}$ = 4.16 kU g_X^{-1} and $Y_{P/X}^{YwmC}$ = 1.91 kU g_X^{-1} are found for the lowest growth rates. For the highest growth rates, values of $Y_{P/X}^{NprE}$ = 2.00 kU g_X^{-1} and $Y_{P/X}^{YwmC}$ = 0.98 kU g_X^{-1} are seen accordingly (cf. Figure 5.3A). This scheme is also seen for the other inducer concentrations tested (i.e., 250 µM and 1000 µM IPTG in Figure 5.3B and C, respectively), indicating that a SP specific effect on overall secretion efficiency is more pronounced at lower growth rates. Consequently, for the regression analysis, interactions between different predictor variables were taken into account and the interactions between each SP and the adjusted growth rates were found to be significant ($p < 0.05$). The inverse correlation between cutinase yield and adjusted growth rate is expressed by the negative sign of the corresponding parameter values (Table 5.1). Furthermore, the absolute numerical values of the parameter estimates indicate how strong the adjusted growth rate affects cutinase yield for the specific SP. According to this, the above ranking of SPs is again observed for their interaction with the adjusted growth rate (see Table 5.1).

Although the results of the regression analysis suggest that the effect of IPTG concentration on cutinase yield is negligible, the data presentation in Figure 5.3 points to an optimal IPTG concentration for certain conditions: For the combination "NprE, μ_{Set} = 0.10 h^{-1}" cutinase yields of $Y_{P/X}$ = 4.16, 2.96 and 3.42 kU g_X^{-1} are observed with IPTG added to concentrations of 50, 250 and 1000 µM, respectively. Likewise, for the combination "Epr, μ_{Set} = 0.10 h^{-1}" cutinase yield of $Y_{P/X}$ = 3.39, 2.87 and 3.06 kU g_X^{-1} are observed (again referring to IPTG concentrations of 50, 250 and 1000 µM, respectively). On the other hand, when using the YpjP SP in fed-batch cultivations operated at μ_{Set} = 0.10 h^{-1}, highly comparable cutinase yields of $Y_{P/X}$ = 2.56, 2.72 and 2.68 kU g_X^{-1} are found for IPTG concentrations of 50, 250 and 1000 µM, respectively, indicating that IPTG concentration has almost no effect for the combination "YpjP, μ_{Set} = 0.10 h^{-1}". Hence, the corresponding result of the regression analysis is explained by this observed inconclusive effect of IPTG concentration. Therefore, the combinatorial evaluation of the data visualization (Figure 5.3) and regression analysis (Table 5.1), sug-

gests to investigate the optimal induction strength with the combination "NprE, μ_{Set} = 0.10 h^{-1}". It is well-known that for each bioprocess an optimized induction can be determined, which is typically carried out using micro- to small-scale cultivation systems. This approach is often referred to as "induction profiling" [36, 162, 262].

In general, it is seen from Figure 5.3 that the choice of bioprocess conditions can turn a Sec SP attributed with "good secretion performance" into a "bad" one and vice versa. For instance, the operating condition 1 (*OC1*), indicated in Figure 5.3A, results in a rather low cutinase secretion performance of $Y_{P/X}$ = 1.34 kU g_X^{-1}. As soon as bioprocess conditions are changed, the use of Epr SP results in a much higher cutinase secretion performance of $Y_{P/X}$ = 3.06 kU g_X^{-1}, as indicated by *OC2* in Figure 5.3C. Furthermore, *OC3* (cf. Figure 5.3B) operates at a different combination of conditions in comparison to *OC1*, but results in nearly the same cutinase secretion efficiency ($Y_{P/X}^{OC3}$ = 1.42 kU g_X^{-1} vs. $Y_{P/X}^{OC1}$ = 1.34 kU g_X^{-1}). As can be inferred from the data presentation in Figure 5.3, any *OC* that is found above or below an arbitrarily chosen *OC* (e.g., *OC1*) will result in improved or decreased cutinase secretion efficiency, respectively. Thus, this kind of data visualization as diagram of characteristic curves allows to quickly determine preferred bioprocess conditions with respect to the optimization objective (i.e., $Y_{P/X}$).

5.5. Cutinase yield and productivity are conflicting optimization objectives requiring different bioprocess conditions for optimal results

For biotechnological applications, productivity (q_P) is a measure of economic feasibility and thus, it is of high importance and desired to be as high as possible. For growth coupled product formation, productivity is the product of growth rate and product yield [100], and strictly speaking, productivity considers both the efficiency of substrate conversion into product (i.e., $Y_{P/X}$) and the time needed for formation of the product synthesizing biocatalyst (i.e., μ). Hence, productivity q_P appears as another possible optimization objective in bioprocess development, next to product yield $Y_{P/X}$. The decision which of the different objectives is to be maximized is preferred may differ case-by-case. Typically, yield ($Y_{P/X}$) and productivity (q_P) are conflicting optimization objectives [263, 264] which means maximal yield $Y_{P/X}^{max}$ and maximal productivity q_P^{max} cannot be achieved for the same combination of strain and bioprocess.

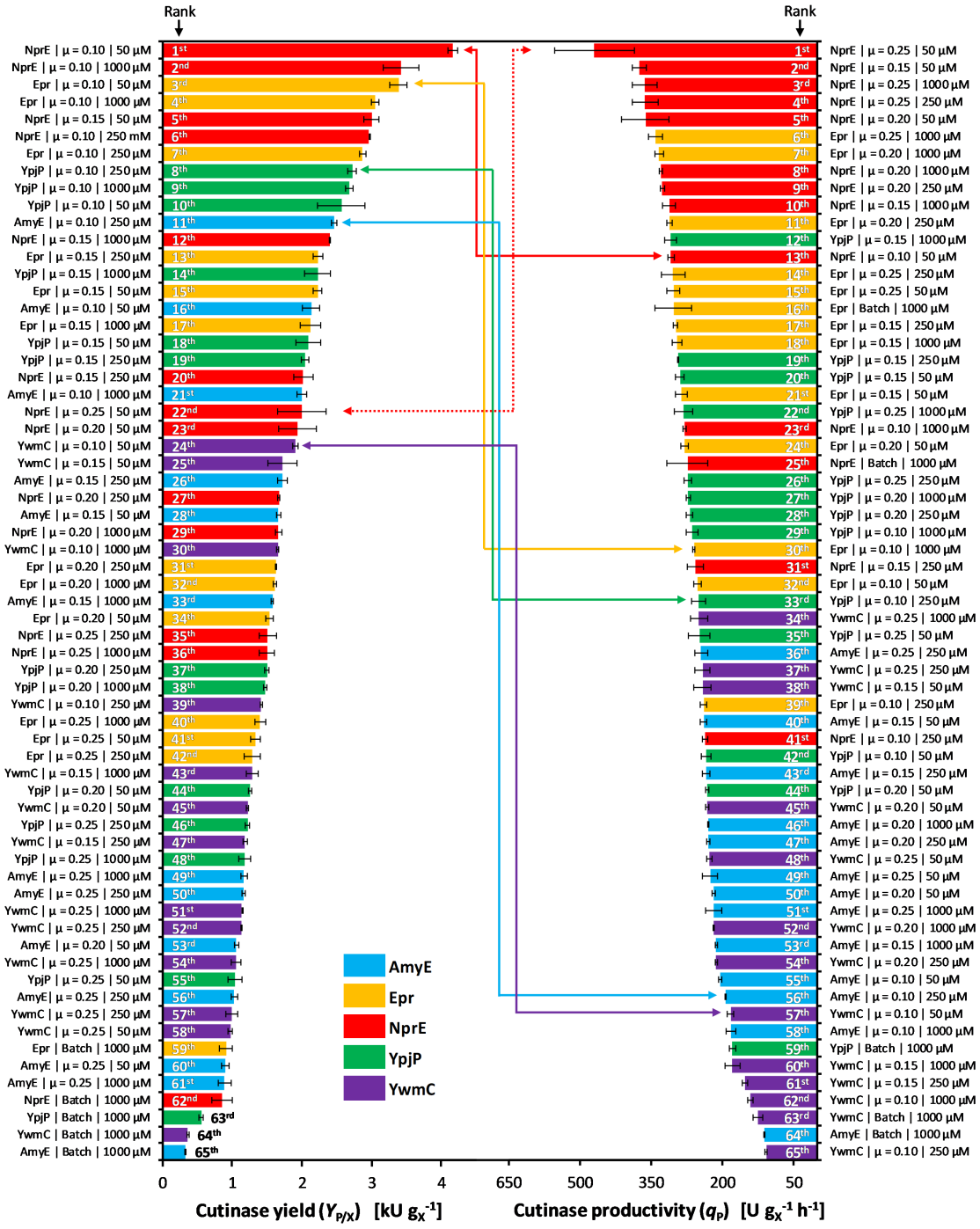


Figure 5.4.: See figure description on next page.

Figure 5.4.: (See figure on previous page) **Comparative ranking of bioprocess conditions with respect to different cutinase secretion performance optimization objectives.** Data is ordered by biomass specific cutinase yields ($Y_{P/X}$, left part) and ordered by biomass specific cutinase productivities (q_P, right part). Bar colors encode the applied Sec SP used for cutinase secretion. Description of bioprocess conditions depicts applied Sec SP (AmyE, Epr, NprE, YpjP and YwmC), adjusted growth rate (μ_{Set} = 0.10, 0.15, 0.20 and 0.25 h^{-1}) and induction strength (50, 250 or 1000 µM IPTG). Conditions for SP specific highest values of $Y_{P/X}$ are connected to the corresponding values of q_P by colored arrows. Cultivations with *C. glutamicum* WT and cultivations without IPTG added are not shown. All data are mean values with min/max values indicated by error bars from two biological replicates. Detailed numerical values for $Y_{P/X}$ and q_P shown in this figure are listed in Table D.2 and Table D.3, respectively.

As was made for cutinase yields $Y_{P/X}$, values for q_P were calculated from all bioprocesses comprising different feeding rates, induction strengths and Sec SPs, so that cutinase secretion performance for each bioprocess can be evaluated based on either $Y_{P/X}$ or q_P. The resulting data sets are depicted in Figure 5.4 in a comparative bar plot, corresponding numerical values are found in Table D.2 and Table D.3. Bioprocesses were rank-ordered from high to low values of $Y_{P/X}$ and q_P in the left and right part of Figure 5.4, respectively. The applied Sec SPs in each bioprocess are indicated by the color-coding of the bars while further details of the respective bioprocess are shown next to the plots. Since in total 65 different cutinase secretion bioprocesses are shown, those are ranked from 1st to 65th with respect to either $Y_{P/X}$ or q_P (left or right part of Figure 5.4, respectively).

A first visual evaluation of Figure 5.4 indicates that Sec SPs tend to spread among the bioprocess conditions rank-ordered with respect to $Y_{P/X}$ (left part of Figure 5.4), while SPs tend to group among bioprocess conditions rank-ordered with respect to q_P (right part of Figure 5.4). An interpretation for this would be that the impact of the applied SP is more pronounced with respect to q_P than for $Y_{P/X}$. Strikingly, it seems that to reach high values of $Y_{P/X}$ as well as q_P the use of NprE or Epr SP is needed since those SPs are found four and three times, respectively, among the top 10 ranked bioprocesses with respect to $Y_{P/X}$. Both SPs are solely found among the top 10 ranked bioprocesses with respect to q_P. Likewise, YpjP and YwmC/AmyE can be classified as mostly medium and poor performers, respectively, with respect to both optimization objectives $Y_{P/X}$ and q_P.

Regarding the application of a batch or fed-batch process, it is seen that the batch processes do reach only low cutinase secretion performance in terms of $Y_{P/X}$ irrespective of the employed Sec SP, which is reflected by corresponding low rankings: 59th, 62nd, 63rd, 64th and 65th out of 65 using the Epr, NprE, YpjP, YwmC and AmyE Sec SP, respectively (cf. left part of Figure 5.4). In contrast, this is different with respect to q_P as optimization objective where two batch processes are found among the best 25 processes, that is 16th (Epr) and 25th (NprE). The other batch processes are poorly ranked (YpjP: 59th, YwmC: 63rd, and AmyE: 64th, cf. right part of Figure 5.4). When considering the batch bioprocesses only, again Epr and NprE Sec SPs show superior cutinase secretion performance (both for $Y_{P/X}$ and

q_P) compared to the other SPs. Therefore, when aiming to reduce effort for SP selection in fed-batch cultivations, the Epr and NprE could be considered as promising choices for fed-batch cultivations, based on results obtained in batch cultivations. However, as seen in this study, optimization of further bioprocess conditions is still needed achieve the highest yield or productivity.

The findings described in section 5.4 on the example of cutinase yield $Y_{P/X}$ are presented in a very condensed way in the left part of Figure 5.4, which facilitates the direct comparison to the data on productivity q_P. By using this way of visualization the data generated from a high number of bioreactor cultivations, the interrelation of Sec SPs as biological variables and bioprocess condition as technical variables becomes easily accessible: The overall maximal cutinase yield of $Y_{P/X}^{max}$ = 4.16 kU g_X^{-1} is obtained for the condition "NprE, μ_{Set} = 0.10 $^{-1}$, 50 µM IPTG". By choosing inappropriate bioprocess conditions for the NprE-cutinase secretion strain, the overall maximum cutinase yield achievable with this strain (in combination with optimal bioprocess conditions) is lowered to a value of 0.86 kU g_X^{-1} (batch, 1000 µM IPTG), which corresponds to a fundamental reduction of 79% from $Y_{P/X}^{max}$. On the other hand, with the AmyE SP and optimized bioprocess conditions, a value of $Y_{P/X}$ = 2.46 kU g_X^{-1} (μ_{Set} = 0.10 h^{-1}, 250 µM IPTG) can be reached, which corresponds to a reduction of 41% only with respect to $Y_{P/X}^{max}$ (cf. left part of Figure 5.4). Likewise, these conclusions can be inferred from the right part of Figure 5.4, where cutinase productivities q_P for the different bioprocesses are shown: The NprE SP is needed to achieve the overall maximum productivity of q_P^{max} = 470 U g_X^{-1} h^{-1} in combination with 50 µM IPTG and a feeding profile corresponding to μ_{Set} = 0.25 h^{-1}. Similar to the observations made for $Y_{P/X}$, when choosing suboptimal bioprocess conditions (i.e., 250 µM IPTG, μ_{Set} = 0.10 h^{-1}), the productivity with NprE-strain is lowered to q_P = 195 U g_X^{-1} h^{-1}, which corresponds to a substantial reduction of 59% from q_P^{max}. In comparison to that, the application of Sec SPs that were attributed as "medium and poor performing" can result in distinct higher productivities due to carefully chosen bioprocess conditions, namely q_P = 309 U g_X^{-1} h^{-1} (YpjP, 1000 µM IPTG, μ_{Set} = 0.15 h^{-1}) and q_P = 245 U g_X^{-1} h^{-1} (AmyE, 250 µM IPTG, μ_{Set} = 0.25 h^{-1}) which represent a reduction of only 34% and 48% from q_P^{max}, respectively. This confirms that the choice of the "best" Sec SP is a prerequisite for maximizing the optimization objective, but on top of that, the accompanying optimal bioprocess conditions need to be determined to fully access this potential.

The conflict between optimizing cutinase yield $Y_{P/X}$ and cutinase productivity q_P by adjustment of bioprocess conditions and the choice of Sec SP is indicated by the colored arrows in Figure 5.4. The arrows visualize that a certain combination of bioprocess conditions and chosen Sec SP results in substantially different rankings with respect to either optimization objective. More specifically, the bioprocess ranked 1st with respect to cutinase yield comprises the NprE SP, but with respect

to cutinase productivity this specific combination is ranked only 13th, as indicated by the bold red arrow in Figure 5.4. Likewise, with respect to cutinase yield $Y_{P/X}$ the optimal bioprocesses for the Epr, YpjP, AmyE and YwmC SPs are ranked 3rd, 8th, 11th and 24th, respectively, which correspond to the 30th, 33rd, 56th and 57th ranked bioprocesses in terms of cutinase productivity (indicated by the orange, green, light blue and purple solid arrows, respectively, in Figure 5.4). In terms of cutinase productivity, the bioprocess that is ranked 1st corresponds to the bioprocess ranked 22nd in terms of cutinase yield $Y_{P/X}$, as indicated by the red dotted arrow in Figure 5.4. These observations finally raise the question how to choose a specific combination of bioprocess and Sec SP that results in satisfactory cutinase secretion performance with respect to $Y_{P/X}$ as well as q_P.

One approach would be to choose conditions that maximize $Y_{P/X}$ ($Y_{P/X}^{max}$ = 4.16 kU g_X^{-1} for conditions "NprE, 50 µM IPTG, μ_{Set} = 0.10 h^{-1}") and with this, the corresponding value for cutinase productivity is set to q_P = 308 U g_X^{-1} h^{-1}. Likewise, the condition that maximizes q_P (q_P^{max} = 470 U g_X^{-1} h^{-1} for "NprE, 50 µM IPTG, μ_{Set} = 0.25 h^{-1}") determines also the corresponding value of $Y_{P/X}$ = 2.00 kU g_X^{-1}. These interrelations between $Y_{P/X}$ and q_P are visualized by a red bold arrow (indicating $Y_{P/X}^{max}$ with corresponding q_P) and a red dotted arrow (indicating q_P^{max} with corresponding $Y_{P/X}$) in Figure 5.4. A decision for a specific set of conditions could be made by the apparent ranking distances with respect to either $Y_{P/X}$ or q_P. For example, the conditions that maximize $Y_{P/X}$ (i.e., 1st rank) result in the 13th rank with respect to q_P (cf. red bold arrow in Figure 5.4). On the other hand, the conditions that maximize q_P, are ranked 22nd with respect to $Y_{P/X}$ (cf. red dotted arrow in Figure 5.4). Following this, the decision for a certain combination of Sec SP and bioprocess conditions would be in favor for those that result in the 1st rank with respect to $Y_{P/X}$ and 13th rank with respect to q_P. However, multi-objective optimization in biotechnological process development is still an active area of research, implying that this task is not a trivial one and literature suggests several approaches to tackle this problem encountered in several fields of applied biotechnology [248, 265–268].

5.6. Concluding remarks

This study comprehensively investigated the combinatorial impact of Sec SPs and bioprocess conditions on cutinase secretion performance with *C. glutamicum*, which is of high relevance for applications in industrial biotechnology, specifically for the development of biotechnological protein production processes. The results of this study add important knowledge to the question which SP is best suited for heterologous protein secretion and how to choose such optimal SP. In addition to previous findings that for each target protein of choice the optimal but unpredictable combination

of SP and expression host has to be determined from scratch, it is seen in this study that the bioprocess conditions do also need to be carefully adjusted to maximize heterologous protein secretion efficiency. Most likely, it is not possible to predict in-silico the optimal combination of target protein, SP and expression host. However, general trends could be derived from the comprehensive data set presented in this study:

(1) A metabolic burden due to heterologous protein secretion, visible as reduction in growth rate, is more pronounced in unrestricted batch cultivation in comparison to growth-controlled fed-batch cultivations.
(2) Secreted protein yield is high when the growth rate of the culture is low and vice versa.
(3) Promising Sec SPs for further development of fed-batch cultivations could be inferred from technically less complex batch cultivations.
(4) Inducer concentration needs to be optimized in order to further maximize the given optimization objective.

Since process optimization is typically conducted in iterative rounds comprising experiment and evaluation, the overall optimization objective needs to be determined in the beginning, as optimal bioprocess conditions vary with different optimization objectives. The above mentioned general trends, based on empirical relations, can be used to increase efficiency in bioprocess development and thus, are of great importance in this field of applied biotechnology.

In this study only a few biological and technical variables (i.e., Sec SPs and IPTG concentration as well as feeding profile) were investigated for their impact on cutinase secretion efficiency in terms yield and productivity. Nevertheless, the experimental design space can be easily expanded, leading to a combinatorial explosion that cannot be handled anymore, simply for practical reasons. Especially combinatorial effects on given optimization objectives need a certain amount of data points to satisfy requirements of confidence evaluation. It is foreseeable that the challenge to acquire sufficiently large data sets to determine suitable design parameters will gain more importance in the future, which could be tackled by the application of multiplexed and down-scaled microbial cultivation systems, as well as statistics-supported experimental planning, analysis and interpretation. Moreover, the multi-dimensional design space and corresponding complexity in results clearly asks for data processing frameworks that facilitate interpretation of results for decision making, which finally is and will remain the human responsibility.

6. Evaluation of a genome reduced *C. glutamicum* strain library for heterologous cutinase secretion

This chapter is concerned with the quantitative microbial phenotyping of a library of genome reduced *C. glutamicum* strains for heterologous cutinase secretion, which represents another typical application example of MBR systems. The collected data comprise growth rates and cutinase yields as extracellular phenotypes, as well as detailed analysis at the transcriptome and proteome level for a small subset of strains. Next to surprising phenotypes due to specific genomic deletions, as well as differential analysis of strains with overlapping genomic deletions, attempts were made to explain the metabolic perturbations from observed significant differential regulation at the protein level. Also, by incorporating all data on extracellular phenotypes, a data-driven, phenomenological multiple regression approach was used to identify the minimum set of genomic deletions needed in terms of improved cutinase secretion. Parts of this chapter are based on *Manuscript II* and *Manuscript III*.

Author contributions:

Johannes Hemmerich wrote the chapter, designed and performed experiments, analyzed the data, interpreted the results, and prepared the figures. Mohamed Labib, Sebastian J. Reich, Carmen Steffens and Marc Weiske designed and performed experiments, and helped with analyzing the data in sections 6.1 to 6.6. Mohamed Labib performed microarray experiments, and helped with analyzing the data in section 6.6.

6.1. Overview of observed extracellular phenotypes: Combinatorial impact of Sec signal peptide and genomic deletions

Quantitative phenotyping of a strain library represents a typical application example for MBR systems (another typical application example of screening of cultivation conditions and media compositions is presented in Chapter 4). In particular, automated sampling and harvest procedures enable the acquisition of data that can be compared even for differently growing strain mutants, as observed

in this study, where a library of GRS of *C. glutamicum* [48, 55, 222] was screened for the ability to secrete heterologous cutinase from *F. solani pisi*. To enable cutinase secretion via Sec pathway, the plasmid-encoded cutinase gene needs to be fused to as Sec-specific SP. Since the optimal SP for a certain target protein to be secreted cannot be predicted in silico [30, 32–34] (cf. also section 1.2.2) and because it was shown that a combination of SP and target protein found to be optimal varies between expression hosts [37], cutinase secretion with *C. glutamicum* GRS was tested for five Sec SPs detached from a library comprising all natural Sec SPs from *B. subtilis* [30].

The constructed strain library comprises 13 strain variants with five Sec SPs and further 14 strains with two out of the five Sec SPs, resulting in 93 strains in total to be screened for cutinase secretion performance. The GRS are characterized by deletion of single and multiple genomic clusters. All cutinase secretion strains were assessed by MBR growth experiments in several replicates and resulting extracellular phenotypes were determined as observed growth rate μ and biomass specific cutinase yield $Y_{P/X}$.

Figure 6.1A depicts cutinase yield as function of growth rate for the tested *C. glutamicum* GRS (each data point representing one GRS) and the five different Sec SPs from *B. subtilis* (AmyE, Epr, NprE, YpjP, and YwmC, encoded by colors as already introduced in Chapter 5). In general, it is seen that a high diversity in cutinase yield is found for Epr and NprE Sec SPs. When applying the YpjP Sec SP, diversity in cutinase yield is only slightly higher than for AmyE and YwmC SPs. High values for $Y_{P/X}$ are found when using Epr and NprE SP. More specifically, obtained cutinase yields range from 0.00 – 0.28 kU g_X^{-1}, 0.15 - 1.10 kU g_X^{-1}, 0.14 - 1.15 kU g_X^{-1}, 0.13 - 0.46 kU g_X^{-1} and 0.16 - 0.47 kU g_X^{-1} for the strains with the AmyE, Epr, NprE, YpjP or YwmC SP, respectively. Based these observations, the five Sec SPs can be ranked with respect to cutinase yield as follows: Epr, NprE > YwmC, YpjP > AmyE. Since the high values for $Y_{P/X}$ are only found at low growth rates, this seems to be a prerequisite to achieve a high cutinase yield, but only when using Epr or NprE SP. All data are derived from MBR batch cultivations allowing the GRS to grow under non growth limiting glucose concentrations. Thus, the observed reduced growth rates are due to genomic deletions classified previously as "non-essential" [48, 55, 160]. However, these classifications were made for *C. glutamicum* GRS that were not imposed to a heterologous protein secretion burden [48, 55, 160].

In order to collect more data on the interrelation between genome reduction and extracellular phenotypes (i.e., μ and $Y_{P/X}$), it was decided to focus only on AmyE and NprE Sec SPs but with a higher number of different *C. glutamicum* GRS. The updated range (i.e., including more GRS tested) of obtained cutinase yields for the AmyE and NprE SP is 0.01 - 0.45 kU g_X^{-1} and 0.37 - 1.62 kU g_X^{-1}, respectively, confirming the above classification of Sec SPs in term of cutinase secretion. Corresponding

results from MBR phenotyping are shown in Figure 6.1B. The range of observed growth rates is approximately 0.15 - 0.40 h^{-1} and 0.30 - 0.45 h^{-1} for the NprE and AmyE SP, respectively, with one GRS showing an exceptionally low growth rate of approximately 0.20 h^{-1} among all GRS with the AmyE SP, coinciding with a hardly detectable cutinase yield.

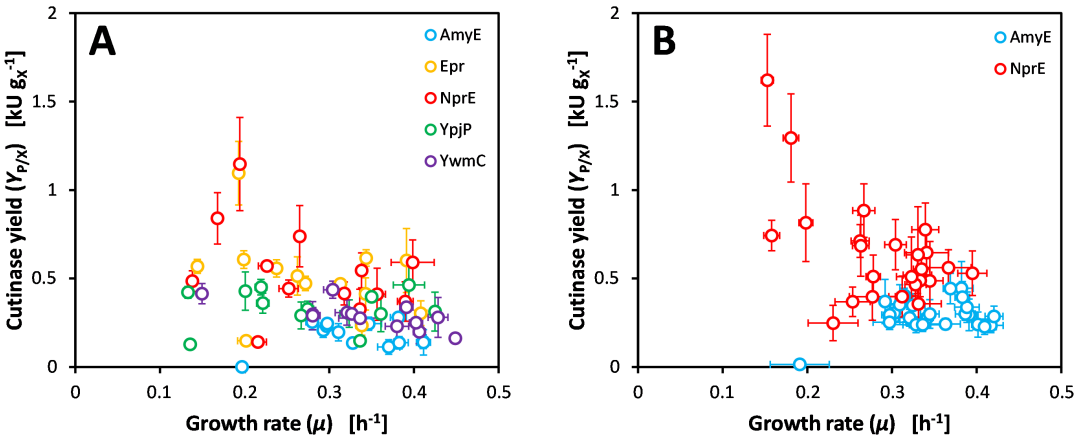


Figure 6.1.: Overview of *C. glutamicum* GRS cutinase secretion performance, depicted by determined biomass specific cutinase yields $Y_{P/X}$ as function of observed growth rate μ from *C. glutamicum* GRS library screening. (A): Results from first screenings comprising five Sec SPs, encoded by colors. (B): Results from initial and additional screening comprising a higher number of GRS while tested Sec SPs were reduced to two (AmyE and NprE). In both panels, each data point represents one GRS and the used Sec SP is indicated by different colors. Error bars indicate standard deviation from growth experiments conducted in two to three replicates (panel A) or eleven to 46 (panel B, as indicated in Table E.1).

As mentioned above, the number of Sec SPs tested was reduced from five to two, which results in a wide and narrow range of cutinase yields, that is the NprE and AmyE SP, respectively, as seen in Figure 6.2B. To enable a comparison of the impact of these two SPs for specific genomic deletions (i.e., for specific GRS), the corresponding observed growth rates and cutinase yields are presented as in comparison to the respective strains WT-NprE and WT-AmyE, as seen in Figure 6.2A (relative growth rates) and Figure 6.2B (relative cutinase yields). In addition, all observed growth rates and cutinase yields as absolute values are summarized in Table E.1.

From the GRS specific comparison of growth phenotypes due to the use of the AmyE or NprE SP (classified as poor or good performing, respectively, in terms of cutinase secretion efficiency, cf. above), it becomes obvious that for many GRS the choice of Sec SP seems not to impact growth considerably. It can be rather concluded that the genomic deletions of the GRS impose a growth burden mostly independent from the applied SP (cf. Figure 6.2A). For example, strain W25 shows the

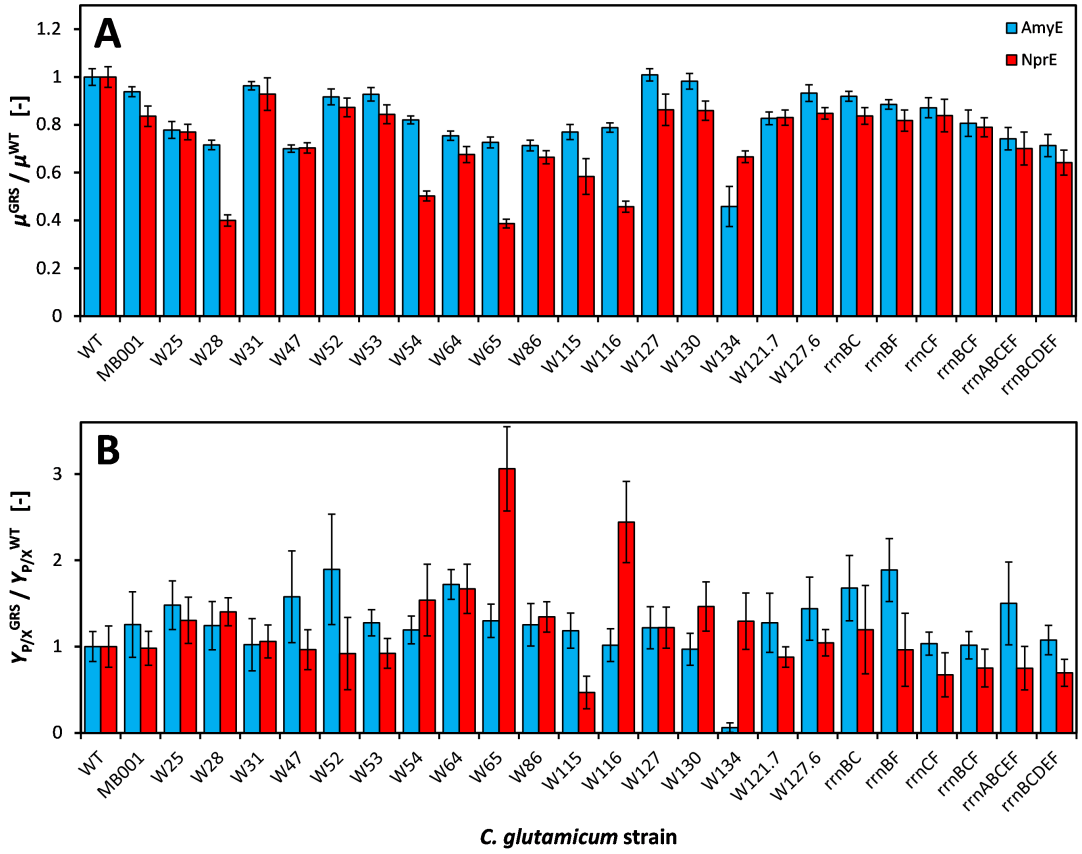


Figure 6.2.: Observed growth rates and cutinase yields from *C. glutamicum* GRS library screening in relation to the respective wild type strains containing either the AmyE or NprE SP. (A): Relative GRS specific growth rates μ^{GRS}/μ^{WT}. (B): Relative GRS specific cutinase yields $Y_{P/X}^{GRS}/(Y_{P/X}^{WT}$. Corresponding numerical data, errors and number of replicates are found in Table E.1.

same growth rate for using the AmyE or NprE SP, and this same growth rate is about 80% the growth rate obtained with the corresponding strains WT strains. The same observation is also made for strain W121.7 as well as the strains rrnBC to rrnBCDEF, which show growth rate reductions of about 10 to 30%, irrespective of the SP applied.

Interestingly, a few strains show a distinct growth phenotype when comparing both SPs, as well as in comparison to the use of the respective SPs in the WT analogue. More specifically, strain W28 AmyE shows a growth rate reduction of approximately 30%, while the same strain containing the NprE SP (i.e., strain W28-NprE) shows a growth rate reduction of approximately 60%. Similar patterns are also observed for strains W54-NprE, W65-NprE and W116-NprE. In general, in case a SP specific growth

rate reduction is observed for a GRS, the presence of the NprE SP impairs growth stronger than the presence of the AmyE SP, with the only exception for strain W134. Here, the AmyE SP imposes a higher growth burden than the NprE SP. Thus, strain W134 shows a unique and remarkable growth phenotype among all GRS tested. Some examples of strains with interesting phenotypes are discussed in more detail in later sections.

In order to enable a GRS specific evaluation of cutinase secretion for both the AmyE and NprE SP, data on cutinase yield was normalized to the respective strains WT-AmyE and WT-NprE in Figure 6.2B, as done for growth rates accordingly. Many of the strains show a cutinase yield that is comparable or slightly different to the level of the corresponding strains WT-AmyE and WT-NprE.

For example, many strains (e.g., MB001-AmyE, W31-AmyE, W116-AmyE, W121.7-AmyE or W127.6-AmyE) show cutinase levels comparable to WT-AmyE within the observed error ranges. Some of the AmyE-GRS (e.g., W25-AmyE, W64-AmyE) show a slightly higher cutinase yield in comparison to strain WT-AmyE with non-overlapping error ranges. Also, as previously seen for growth rates, few strains with the AmyE SP show a prominent difference in cutinase yield, that are W52-AmyE and rrnBF-AmyE reaching almost 2-fold the cutinase yield of WT-AmyE. Again, strain W134-AmyE stands out of the crowd since nearly no cutinase yield is detectable. However, because absolute cutinase yields are rather low when using the AmyE SP (cf. above), results on cutinase yield obtained with the NprE SP are considered as much more interesting.

In general, similar patterns for cutinase yield are observed for the GRS-NprE as for the GRS-AmyE (cf. Figure 6.2B): While most of the strains show cutinase yields that are comparable to strains WT-NprE (e.g., MB001-NprE, W52-NprE, W127-NprE, W127.6-NprE), some remarkable phenotypes in terms of cutinase secretion can be observed for a few strains. A drastic reduction is found for strain W115-NprE, which is the lowest cutinase yield among all GRS-NprE. On the other hand, for two strains a remarkably high increase in cutinase yield is found, namely W65-NprE (about 3-fold the level of WT-NprE) and W116-NprE (about 2.5-fold the level of WT-NprE). Besides the improvement in cutinase yield, these two strains are highly interesting since they share no genomic deletions except $\triangle$CGP123 and $\triangle$IS*Cg*12, which all other GRS do except MB001 ($\triangle$CGP123 only). Thus, at least two different genomic deletions can be assumed to be independently responsible for the observed increased cutinase secretion with strains W65-NprE and W116-NprE, which is discussed in more detail later in the text.

When comparing the GRS specific data shown in Figure 6.2A and Figure 6.2B, a correlation between growth phenotype and cutinase secretion phenotype can be inferred specifically for each depicted GRS. For example, the stronger growth reduction of strain W28-NprE is not resembled in a higher

cutinase yield. Like that, a higher increase is seen for strain W52-AmyE than for W52-NprE, while almost no growth rate reduction is seen for both W52-AmyE and W52-NprE. On the other hand, the remarkable increase in cutinase yield for strains W65-NprE and W116-NprE is accompanied by a strong decrease in according growth rates at the same time. More specifically, the observed growth rates for strains W65-NprE and W116-NprE are roughly 60% less compared to strain WT-NprE, while strains W65-AmyE and W116-AmyE show a growth rate reduction of about 20% only in comparison to strain WT-AmyE. Therefore, the strongly pronounced phenotypes for strains W65 and W116 can be clearly attributed to the NprE SP.

Conclusively, screening of the *C. glutamicum* GRS library under the objective of heterologous cutinase secretion revealed noticeable strain phenotypes in terms of growth and cutinase secretion that could not have been concluded from previously screening for growth rate only [48, 55, 222]. As both phenotypes (i.e., growth rate μ and cutinase yield $Y_{P/X}$) are easily accessible and could be determined repeatedly in a higher throughput by the application of MBR methods, this initial rapid screening marks the first step in narrowing down a big design space, that is the identification of potentially interesting genomic modifications that are beneficial for improved cutinase secretion. In the following, some example strains with noticeable phenotypes are discussed in more detail. For one strain (W65-NprE) and its close relatives in terms of genetic characteristics, further in-depth analyses were conducted and will be considered for discussion later.

6.2. Deletion of cg2990-2999 has impact on growth but not on cutinase yield when using NprE signal peptide

The differential analysis of phenotypes for strains W28 and W86 reveals that the deletion of cg2990-2999 impairs growth in combination with cutinase secretion enabled by NprE (W28: μ^{NprE} = 0.16 ± 0.01 h^{-1}, W86: μ^{NprE} = 0.26 ± 0.01 h^{-1}) but not by AmyE (both W28 and W86: μ = 0.30 ± 0.01 h^{-1}). On the other hand, cutinase yields between strains are highly comparable when using either the AmyE (W28: 0.29 ± 0.07 kU g_X^{-1}, W86: 0.29 ± 0.06 kU g_X^{-1}) or the NprE SP (W28: 0.74 ± 0.09 kU g_X^{-1}, W86: 0.71 ± 0.09 kU g_X^{-1}). Strain W28 is characterized by Δ2990-3006 while strain W86 is characterized by Δ3000-3006, meaning that strain W86 shares only a subset of deleted genes compared to strain W28. Both strains were earlier described to show to the same extend of reduced growth rates compared to the wild type strain [55, 222]. This observation is also found when using AmyE SP (but not for the NprE SP) for cutinase secretion, indicating that the deletion of cg2990-2999 in addition to cg3000-3006 does not impact growth and cutinase secretion in that case (i.e., when using AmyE SP). A comparison

of deleted genes in strains W28 and W86 with corresponding annotations is given in Table E.2.

Consequently, the exchange of a rather small genetic feature, that is the switch of SP from AmyE to NprE completely changes the observations made. Thus, conclusions drawn from differential studies on gene characterization subjected to cutinase secretion would classify cluster cg2990-2999 as non-relevant if only the AmyE SP would have been selected to enable secretion. In turn, a completely different gene classification results by using the NprE SP.

6.3. AmyE but not NprE signal peptide impairs growth in strain W134

With respect to the use of either AmyE or NprE Sec SP for cutinase secretion, strain W134 shows a unique behavior. In contrast to the observations described in the previous section, the use of AmyE provokes a heavy growth defect in strain W134, compared to the use of NprE (μ^{AmyE} = 0.19 ± 0.03 h^{-1} vs. μ^{NprE} = 0.26 ± 0.01 h^{-1}). This growth defect is accompanied by an absence of extracellular cutinase activity ($Y_{P/X}^{AmyE}$ = 0.01 ± 0.01 kU g_X^{-1} vs. $Y_{P/X}^{NprE}$ = 0.69 ± 0.17 kU g_X^{-1}). When considering that strain W134-NprE is able to secrete cutinase comparatively efficient as strain WT-NprE, although at a lowered growth rate, loss(es) of cellular functions due to the introduced gene deletions render strain W134 very sensitive to the AmyE SP itself. However, it is not clear whether the negative impact of the AmyE SP occurs at the transcriptional, translational, translocational or any other functional level.

6.4. Deletion of cg0158-0183 results in a lethal growth phenotype for *C. glutamicum* when containing a cutinase secretion plasmid

Another interesting phenotype can be observed for strain W73 when considering plasmid-borne cutinase secretion in addition to growth behavior. After plasmid transformation, strain W73 was not able to grow in defined CgXII medium anymore. This observation was made repeatedly and irrespective of the SP fused to the N-terminus of the cutinase gene. Interestingly, when supplementing the defined CgXII medium with 10% v v^{-1} of complex brain heart infusion (BHI) medium, transformed W73 strains were able to grow again. The addition of IPTG needed for induction of cutinase biosynthesis imposed a strong growth burden on strain W73, irrespectively of the applied Sec SP. More specifically, growth rate for strain W73 (i.e., without plasmid) in supplemented CgXII medium was μ = 0.46 ± 0.02 h^{-1}, while for both W73-AmyE and W73-NprE a growth rate of μ = 0.24 ± 0.03 h^{-1} was determined (cf. Table 6.1), corresponding to a reduction of 48%. When strain W73-NprE was cultivated without IPTG, a growth

rate of μ = 0.37 ± 0.02 h^{-1} was determined, confirming the metabolic burden caused by recombinant gene expression, as the addition of IPTG causes a growth rate reduction of 35% (cf. Table 6.1).

An overview of deleted genes in strain W73 is given in Table E.3. Since these genes were classified as non-essential in previous studies with respect to growth only, the loss of function(s) by deleting cg0158-0183 results in a lethal phenotype in combination with effects provoked by plasmid maintenance and/or recombinant gene expression. However, because cutinase secretion variants of strain W73 could only be cultivated in supplemented CgXII medium, this strain was excluded from further screening since the use of different cultivation media prevents strain evaluation on a fair basis.

Table 6.1.: Comparison of growth rates obtained for strain W73.

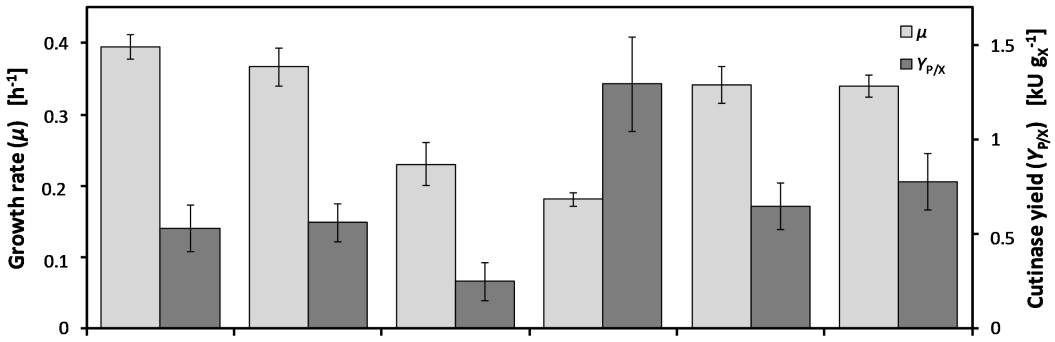
		Growth rate μ [h^{-1}]		
IPTG	**Supplemented medium**	**W73**	**W73-AmyE**	**W73-NprE**
no	no	0.33 ± 0.07 (n = 12) [#]	n.t.	no growth (n = 12)
yes	no	0.38 ± 0.02 (n = 8)	no growth (n = 6)	no growth (n = 18)
no	yes	0.42 ± 0.03 (n = 11)	n.t.	0.37 ± 0.02 (n = 4)
yes	yes	0.46 ± 0.02 (n = 4)	0.24 ± 0.02 (n = 7)	0.24 ± 0.03 (n = 7)

Supplemented medium refers to CgXII mineral medium containing 10% (v v^{-1}) BHI broth. No growth in CgXII medium was also observed for strains W73-Epr, W73-YpjP, and W73-YwmC (all n = 6). *n.t.* not tested. [#] including data from [222].

6.5. Differential phenotyping reveals that the combinatorial deletion of cg3263-3301 and cg3324-3345 is beneficial for improved cutinase yield

In section 3.4.5, an improved method for quantitative phenotyping was introduced and applied on a selection of cutinase secreting GRS employing the NprE SP, namely W31-NprE, W115-NprE, W116-NprE, W127-NprE and W130-NprE (referred to as W116-NprE strain family). In addition, those GRS were independently and repeatedly cultivated in further MBR growth experiments using a standard workflow (cf. section 8.7.2), which confirmed the results obtained in section 3.4.5. Thus, all data concerning these GRS are evaluated collectively in the following paragraph, with focus on the two PIs growth rate μ and cutinase yield $Y_{P/X}$. Corresponding collected data and genotypes of the strains are depicted in Figure 6.3.

Although different induction strategies were used, the values for growth rate and cutinase yield obtained with strain WT-NprE (μ = 0.40 ± 0.02 h^{-1}, $Y_{P/X}$ = 0.53 ± 0.13 kU g_X^{-1}, n = 46) are in accordance with literature (i.e., 500 µM IPTG added at cultivation start [36]) and previous experiments using the



	WT-NprE	W31-NprE	W115-NprE	W116-NprE	W127-NprE	W130-NprE
ΔCGP123		X	X	X	X	X
ΔIS*Cg*12		X	X	X	X	X
Δ2312-2322					X	X
Δ2621-2643					X	X
Δ2663-2686					X	X
Δ2755-2760					X	X
Δ3102-3111			X	X	X	X
Δ3263-3301			X	X		
Δ3324-3345		X		X		X

Figure 6.3.: Differential phenotyping of *C. glutamicum* GRS-NprE for evaluation of impact from deleting cg3325-3345. Top: Growth rates μ and cutinase yields $Y_{P/X}$, as depicted in Table E.1. Bottom: Corresponding combinatorial gene cluster deletions. The combination of deleted clusters responsible for improved cutinase yield are marked by a dotted frame.

MBR method of sequential harvest (100 µM IPTG added at a BS signal of 15 a.u., cf. Table 3.1 as well as results and discussion in section 3.4.5).

The comparison between strains WT-NprE and W31-NprE reveals no prominent differences for growth rate (μ = 0.40 ± 0.02 vs. 0.37 ± 0.03 h^{-1}) and cutinase yield ($Y_{P/X}$ = 0.53 ± 0.13 vs. 0.56 ± 0.10 kU g_X^{-1}), meaning that the deletion of cg3324-3345 in strain W31-NprE has no apparent impact on the PIs determined. This observation agrees with previous results, where the background strain W31 (i.e., without carrying an expression plasmid) was characterized with respect to growth rate and maximum biomass in defined and enriched CgXII mineral medium in comparison to the wild type strain [55]. Furthermore, the deletion of cg3324-3345 in the L-lysine producing *C. glutamicum* strain DM1933 also showed no negative effect on growth rate and only a slightly lower secreted amino acid titer [55, 160]. Conclusively, neither of the gene products encoded in cluster cg3324-3345 seems to be of relevance in *C. glutamicum*, at least under the conditions tested in this and in previous studies.

The comparison between strains W127-NprE and W130-NprE reveals no distinct differences with respect to growth rate (μ = 0.34 ± 0.03 vs. 0.34 ± 0.02 h^{-1}) and cutinase yield ($Y_{P/X}$ = 0.65 ± 0.13 vs.

0.78 ± 0.15 kU g_X^{-1}), also not in comparison to strain WT-NprE, meaning that again neither of the shared genomic deletions in strains W127-NprE and W130-NprE, as well as the additional deletion of cg3324-3345 in strain W130-NprE, have an effect on biological fitness or cutinase secretion performance. Furthermore, it can be concluded that there is no interaction between any gene product encoded by the deleted gene clusters that has impact on biological fitness or cutinase secretion performance, meaning that all of these genes are dispensable in this sense. Also, the genotypes of background strains W127 and W130 were shown not to affect growth rate and biomass yield in comparison to the wild type strain [48, 222].

Strains W115-NprE and W116-NprE share several genomic deletions apart from the other strains, which presumably causes a drastically impaired growth rate in comparison to strain WT-NprE (μ = 0.23 ± 0.03 h^{-1} and 0.18 ± 0.01 h^{-1}, respectively, vs. μ = 0.38 ± 0.01 h^{-1}). Since strain W116-NprE features the genotype of W115-NprE plus Δ3324-3345, the deletion of this additional cluster seems not to be responsible for the impaired growth phenotype of both strains. Indeed, when deleting only cg3324-3345 (i.e., strain W31), no significantly different phenotypes compared to the wild type strain are observed (cf. above). The observation that the combinatorial deletion of gene clusters does impact the biological fitness measured as growth rate while the single deletion of a gene cluster does not, was made for the background strains W115 and W116, as well as for other *C. glutamicum* GRS [48, 55, 222]. In general, strain W115-NprE exhibits substantially lower values for growth rate (μ = 0.23 ± 0.03 h^{-1}) and cutinase yield ($Y_{P/X}$ = 0.25 ± 0.10 kU g_X^{-1}) in comparison to strain WT-NprE, indicating that the combinatorial deletion of clusters cg3072-3091, cg3102-3111 and cg3263-3301 greatly affects the entire metabolism of *C. glutamicum* in a negative way. However, when additionally deleting cg3324-3345 (yielding strain W116-NprE from W115-NprE), growth rate is still significantly lowered (μ = 0.23 ± 0.03 h^{-1}) but cutinase yield is remarkably increased ($Y_{P/X}$ = 1.29 ± 0.25 kU g_X^{-1}) in comparison to strain WT-NprE. Consequently, in the genomic background characterizing strain W115, the additional deletion of cg3324-3345 greatly effects heterologous cutinase secretion in a positive yet unknown way but, at the same time, has no further impact on growth rate.

Comparison of the genotypes of these six *C. glutamicum* strains reveals that the combinatorial deletion of cg3263-3301 and cg3324-3345 seems to be responsible for greatly enhanced cutinase yield, providing new gene targets that may contribute to the regulation or other functions of recombinant protein secretion. Since in total 60 genes were deleted with these two clusters (see Table E.5 for a detailed list), of which many gene products are annotated as putative or hypothetical proteins, effort has to be made in order to identify the function of the gene products encoded by the deleted genes. In the beginning, it will be necessary to evaluate the effect of solely deleting cg3263-3301

and cg3324-3345. Afterwards, depending on these results it is necessary to further narrow down which combination of genes from the two clusters are responsible for the increase cutinase yield, a "divide-and-conquer" strategy of gene deletions seems a promising approach here. However, it may be that also gene deletions leading to the impaired growth phenotype of both W115-NprE and W116-NprE are a prerequisite for the greatly improved cutinase yield due to the combinational deletion of cg3263-3301 and cg3324-3345, so this aspect has also to be carefully evaluated.

Interestingly, the combinatorial deletion of cg3263-3301 and cg3324-3345 in strain W116 is only beneficial in terms of remarkably increase cutinase yield $Y_{P/X}$ when the NprE SP is employed. In case the AmyE SP is used, all the strains discussed in this paragraph show no differences in $Y_{P/X}$ (WT-AmyE: 0.23 ± 0.04 kU g_X^{-1}, W31-AmyE: 0.24 ± 0.07 kU g_X^{-1}, W115-AmyE: 0.28 ± 0.05 kU g_X^{-1}, W116-AmyE: 0.24 ± 0.04 kU g_X^{-1}, W127-AmyE: 0.29 ± 0.06 kU g_X^{-1}, W130-AmyE: 0.23 ± 0.04 kU g_X^{-1}). This confirms that a proper SP enabling the secretion of a target protein of interest has to be carefully chosen and, preferably, this choice has to be evaluated against other SPs [29,36]. In case the phenotyping of *C. glutamicum* GRS with respect to cutinase secretion would have been performed using solely the AmyE SP, the genotype of strain W116 would not have been detected as a potentially interesting one in this sense.

In summary, it becomes clear that high-throughput quantitative phenotyping is indispensable to further improve the understanding when dealing with unexpected and surprising effects on growth or heterologous protein secretion, especially when deleting poorly characterized genes with only predicted or unknown functions. The application of high-throughput MBR methods allows to phenotype such strain mutants quickly and thus, apparent responsible genotypes are also identified quickly, so that targeted analyses can focus on these. In the end, this will contribute to the identification of functional relationships between yet poorly characterized genes and resulting phenotypes.

6.6. Impact of *rrn* operon deletion and neighboring genes on extra- and intracellular phenotypes during heterologous cutinase secretion

As mentioned earlier, strain W65-NprE shows the highest improvement with respect to cutinase yield $Y_{P/X}$ among all GRS tested (cf. Table E.1). Therefore, this and its closely related strains were investigated in more detail.

6.6.1. Extracellular phenotypes: Growth rate and cutinase yield

In Figure 6.4 the construction lineage of the W65 strain families is depicted, with corresponding cutinase yields and growth rates using the AmyE and NprE Sec SP, referred to as W65-AmyE and W65-NprE strain family, respectively. The diagram shows the impact of gene deletions on cutinase yield and growth rate. In general, no prominent improvement in $Y_{P/X}$ phenotype using the AmyE SP is seen compared to strain WT ($Y_{P/X}$ range: 0.23 - 0.45 kU g_X^{-1}) for all the strains, as it is seen for the use of the NprE SP ($Y_{P/X}$ range: 0.49 - 1.62 kU g_X^{-1}). This observation agrees with results for the other strains of the GRS library (cf. section 6.1). Henceforth, the following discussion focuses on phenotyping results obtained with the strains containing the NrpE SP.

The deletion of the prophage regions, yielding strain MB001 from strain WT, does not affect the $Y_{P/X}$ phenotype (both for AmyE and NprE), meaning that strain WT secretes cutinase as effective as MB001 (MB001: $Y_{P/X}^{AmyE}$ = 0.29 ± 0.09 kU g_X^{-1}, $Y_{P/X}^{NprE}$ = 0.52 ± 0.10 kU g_X^{-1}; WT: $Y_{P/X}^{AmyE}$ = 0.23 ± 0.04 kU g_X^{-1}, $Y_{P/X}^{NprE}$ = 0.53 ± 0.13 kU g_X^{-1}). Corresponding growth rates are also very similar, only a slightly lower growth rate is seen for strain MB001 when employing the NprE SP (MB001: μ^{AmyE} = 0.39 ± 0.01 h^{-1}, μ^{NprE} = 0.33 ± 0.02 h^{-1}, WT: μ^{AmyE} = 0.42 ± 0.01 h^{-1}, μ^{NprE} = 0.40 ± 0.02 h^{-1}). Strain MB001 was previously described to be beneficial for heterologous protein production since it exhibits a higher plasmid stability due to the deletion of the restriction system encoded by *cglMRR* (cg1996-1998) located within CGP3 [53]. However, that study reports on the intracellular production of YFP as model protein, while in this study the secretion a heterologous protein (cutinase) is investigated. Strain MB001 is also available employing the very strong T7 expression system (denoted as strain MB001(DE3) [269]), but the expression of NprE cutinase gene fusion using the T7 system in MB001(DE3) did not result in higher cutinase yields compared to the wild type strain (personal communication, Maike Kortmann, IBG-1). This may indicate that plasmid stability and/or expression of the target gene is not the limiting step in secretory cutinase production using *C. glutamicum*.

Starting from MB001 and after deletion of IS*Cg12*, strains W25, W52 and W54 result from deletion of cg2801-2826 [1], *rrnB* and *rrnC*-3298, respectively. For the NprE SP, in comparison to the parental strain WT significantly [2] reduced growth rates are observed, though to different extends, for strains W25-NprE (μ = 0.30 ± 0.01 h^{-1}), W52-NprE (μ = 0.34 ± 0.02 h^{-1}) and W54-NprE (μ = 0.20 ± 0.01 h^{-1}), all $p < 0.001$. In terms of cutinase yield when using the NprE SP and in comparison to strain WT, strains W52-NprE and W25-NprE show no significantly different values ($Y_{P/X}$ = 0.49 ± 0.22 and 0.69 ± 0.14 kU g_X^{-1}, respectively), while the corresponding value for strain W54-NprE is significantly increased to a re-

[1] includes *rrnD* and *rrnE*
[2] according to Welch test

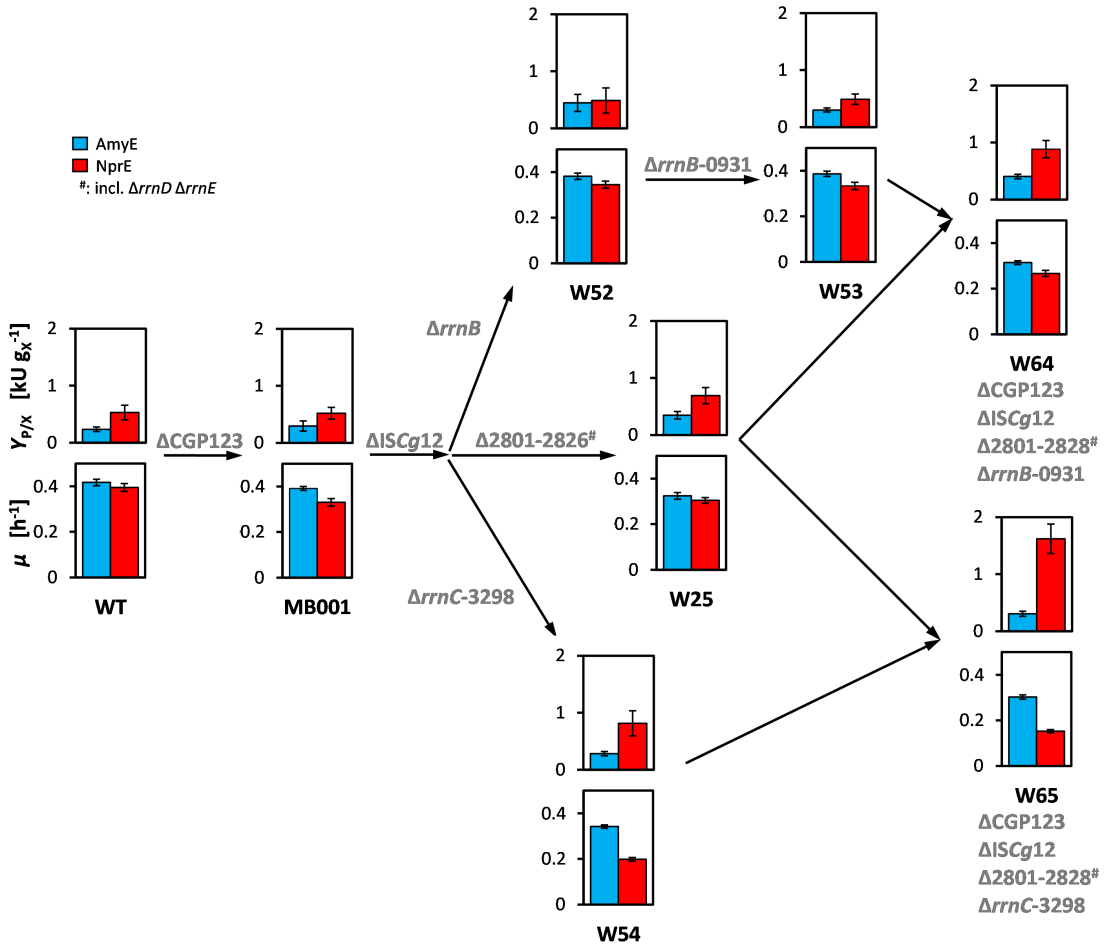


Figure 6.4.: Schematic drawing of construction lineage for the W65 strain family. Arrows indicate relationships between strains with respect to genotype, as indicated in grey font. For each strain, two bar diagrams are shown, of which the upper and lower one depicts the corresponding cutinase yields $Y_{P/X}$ and growth rates μ, respectively. Blue and red bars encode the use of AmyE and NrpE SPs, respectively. Numerical data including errors corresponding to this figure is depicted in Table E.1.

markable value of $Y_{P/X}$ = 0.82 ± 0.22 kU g_X^{-1} ($p < 0.001$). For strain W53-NprE, growth rate μ = 0.33 ± 0.02 h^{-1} and cutinase yield $Y_{P/X}$ = 0.49 ± 0.09 kU g_X^{-1} are highly comparable to the corresponding values obtained with strain W52-NprE.

The combinatorial genotypes of strains W25 and W53 result in strain W64, whose corresponding variant W64-NprE shows a significantly lower growth rate and significantly higher cutinase yield in comparison to strain WT-NprE (all $p < 0.001$). More specifically, for strain W64-NprE a growth rate and cutinase yield of μ = 0.27 ± 0.01 h^{-1} and $Y_{P/X}$ = 0.88 ± 0.15 kU g_X^{-1}, respectively, are observed. Apparently, the combinatorial genomic background of strains W25-NprE and W53-NprE (yielding strain W64-NprE) does not result in a further reduced growth rate μ. In terms of $Y_{P/X}$ the deletion of cg2801-2826 comprising Δ*rrnD* and Δ*rrnE* seems to be dominant since values for $Y_{P/X}$ are more similar between W25-NprE and W64-NprE than between W53-NprE and W64-NprE.

Strain W65-NprE represents the best-performing strain in terms of cutinase yield $Y_{P/X}$ = 1.62 ± 0.26 kU g_X^{-1}) among all strains tested. At the same time, for this strain the lowest growth rate is found (μ = 0.15 ± 0.01 h^{-1}). Its genotype is the combination of genomic deletions characterizing strains W25 and W54. In contrast to strain W64-NprE, the dominating gene deletions in W65-NprE are apparently those stemming from W54-NprE instead of W25-NprE, since for strain W54-NprE both increase in cutinase yield and reduction in growth rate are higher than for strain W25-NprE, in comparison to strain WT-NprE. Strain W65 is characterized by the deletion of three out of six *rrn* operons (Δ*rrnCDE*), whose positions in the chromosome of *C. glutamicum* are depicted in Figure 6.5. Since the corresponding gene products (i.e., rRNA) are indispensable to the cell to build ribosomes, the resulting phenotype of W65-NprE in terms of cutinase yield and growth rate is highly interesting: Why is cell-specific cutinase secretion efficiency remarkably increased when, at the same time, ribosomal capacity is drastically reduced?

Previously, the deletion of up to two *rrn* operons was described as non-critical in terms of biological fitness (i.e., growth rate) [55]. Also, for *E. coli* a reduced growth rate was reported with two or more of its seven *rrn* operons being deleted [270]. The six *rrn* operons in *C. glutamicum* share nearly 100% sequence similarity [271], which suggests that specific effects of the corresponding gene products can be excluded. In addition, *C. glutamicum* was very recently reported to initiate a new round of chromosome replication before the previous replication is finished in a growth rate dependent manner, resulting in a partly polyploid chromosome [272]. Considering the positions of the six *rrn* operons on the chromosome, this would result in a higher number of gene copies for those *rrn* operons which are located closer to the origin of replication (i.e., *rrnC* and *rrnA*, cf. Figure 6.5) for the wild type strain. Furthermore, different expression levels of the six *rrn* operons were reported with

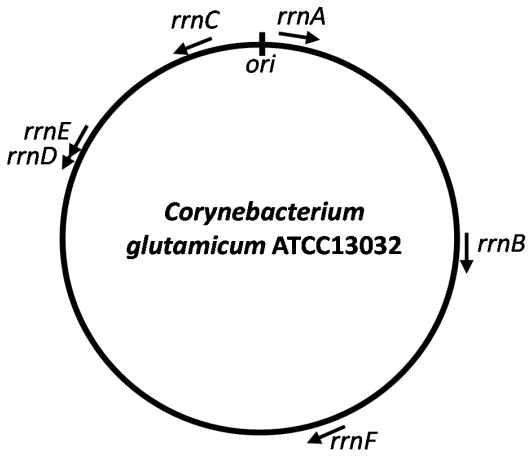


Figure 6.5.: Position of the six operons encoding ribosomal RNA (*rrnA* - *rrnF*) in the circular genome of *C. glutamicum* ATCC13032. Arrows indicate orientation of transcription. *ori*, origin of replication.

rrnA and *rrnF* showing the highest expression, while the remaining four *rrn* operons showed similar expression levels of about 80% the *rrnA* and *rrnF* expression level [273].

Altogether, these findings suggest a general relation between growth rate and copy number of *rrn* operons, as well as a probable effect owed to the proximity of *rrn* operons to the origin of replication. More specifically, the abundance of *rrn* gene products and thus, the importance of the *rrn* operons with respect to ribosomal capacity would be expected as follows: high importance for *rrnA*, *rrnC*, mid-range importance for *rrnB*, *rrnE*, and *rrnF*, as well as low importance for *rrnF*. This ranking corresponds to the distance of the respective *rrn* operons in direction of transcription from the origin of replication, as depicted in Figure 6.5.

Indeed, strains W54-NprE and W65-NprE, both having *rrnC* deleted, exhibit a stronger reduction of growth rates than the other members of the W65-NprE strain family. To investigate this consideration based on the newly acquired data in this study, the observed reduction in growth rates for strains having at least one *rrn* operon deleted are depicted in Figure 6.6 as function of the number of *rrn* operons deleted. More specifically, the correlations are shown for the mentioned strains without a cutinase secretion plasmid (i.e., with biomass formation being the only metabolic burden to the cell, Figure 6.6A), as well as under cutinase secretion conditions enabled by either the AmyE (Figure 6.6B) or NprE SP (Figure 6.6C), so that in these cases any additional "metabolic cost" needed for plasmid maintenance as well as cutinase biosynthesis and secretion is considered.

The assignment of a linear correlation between *rrn* copy number and growth rate reduction, as in-

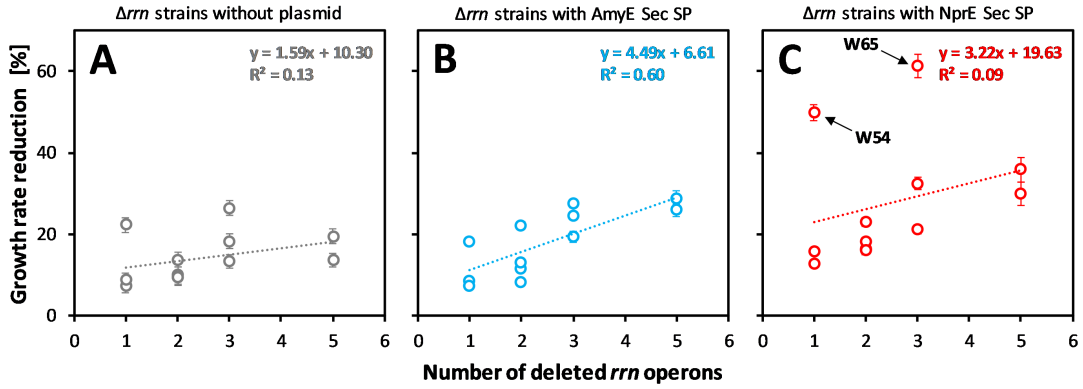


Figure 6.6.: Growth rate reduction of *C. glutamicum* GRS as function of deleted number of *rrn* operons. (A): Strains that were not transformed with an expression plasmid. (B): Cutinase secretion strains with AmyE SP. (C): Cutinase secretion strains with NprE SP. Reduction in growth rate refers to the value for the respective wild type strain. Data covers all tested strains listed in Table E.1 having at least one *rrn* operon deleted (referred to as Δrrn strains). Dotted lines indicate linear ordinary least squares fit with corresponding numerical results depicted in the upper right corner of the diagram. Data on growth rates for strains not containing an expression plasmid are taken from literature [48, 55, 222]. Each data point refers to a single strain. Depicted strains include the W65 strain family, as well as strains rrnBC, rrnBF, rrnCF, rrnBCF, rrnABCEF and rrnBCDEF, cf. Table E.1.

dicated in each panel of Figure 6.6, indicates that such correlation is not significantly different from zero (i.e., no such correlation can be assumed), as corresponding correlation coefficients are very low (R^2 = 0.13, 0.60, and 0.09 for GRS without expression plasmid, AmyE SP enabling cutinase secretion and NprE SP enabling cutinase secretion, respectively, cf. Figure 6.6A, B, and C). In particular, the observed phenotypes of strains W54-NprE and W65-NprE (highlighted in Figure 6.6C), showing a strongly reduced growth rate with one and three, respectively, *rrn* operons deleted, indicate something different to be the reason than solely the number of *rrn* operons deleted. Since both strains share the deletion of *rrnC* as well as 56 further genes, the loss of encoded function(s) by these genes should be considered as primarily causative, suggesting further analyses at the transcriptomic and proteomic level to decipher differently expressed genes for these strains.

6.6.2. Intracellular phenotypes: Transcriptome and proteome

6.6.2.1. Transcriptome- and proteome-wide analysis

From the comparative analyses of the extracellular phenotypes, it can be concluded that a lowered growth rate is necessary to obtain increased cutinase yields using NrpE SP (cf. previous sections), next to other strain properties which need to be unraveled. The reduced growth rate of strain W65 was

speculated to be caused by the loss of three *rrn* operons [55], however, new results from this study indicate that further genomic deletions present in strain W65 need to be taken into account additionally. Moreover, a recent study described specifically an improved co-translational Sec-dependent periplasmatic secretion into the periplasm of *E. coli* due to the deletion of one *rrn* operon [274]. An upregulation of SRP components on the transcriptional level was seen, which was deemed to be the reason for improved secretion, although the nature of the apparent relation between *rrn* deletion and transcriptional regulation of Sec components is not known [274]. Since the Sec pathway is highly conserved between species (cf. section 1.2), maybe such effect occurs in *C. glutamicum* also. Therefore, further investigations were envisaged on the transcriptome of the W65 strain family. Accompanying proteomic analyses were conducted since the proteome defines the extracellular phenotype more directly than the transcriptome.

The differently regulated transcriptomes and proteomes of the members of the W65-NprE strain family are shown in Figure 6.7. Both transcriptomes (Figure 6.7A) and proteomes (Figure 6.7B) are visualized as heatmap, encoding the standardized $\log_2$ value for fold changes of transcript and protein levels, respectively, in comparison to strain WT-NprE. For each strain, $\log_2$ values were standardized, that is centered by the mean of all respective values and scaled by the corresponding standard deviation. The color scales next to the heatmaps indicate down- and upregulation of transcript or protein levels from blue to red.

Furthermore, the figure shows how similar the strains are in terms of their transcriptome and proteome, based on cluster analysis using correlation as distance metric, as suggested for gene expression studies [275]. The similarity of strains is indicated by the dendrograms at the top of the heatmaps, while left to the heatmaps a dendrogram is shown that indicates how similar transcript or protein levels among strains for the specific genes are. The longer the edges of clusters before joining a node in a dendrogram, the higher the dissimilarity between the clusters (i.e., strains). For both heatmaps, four main groups of transcripts and protein levels (i.e., subclusters A to D) are highlighted, which mainly contribute to strain similarity. As additional information, Table E.7 lists the composition of genes assigned to the respective subclusters based on their functional categorization resolved at the transcriptomic and proteomic level (cf. Figure 6.7A and B, respectively). The subclusters A to D indicated at the transcriptome and proteome heatmaps are not composed from the same genes, since the definition of a subcluster was arbitrarily chosen based on the second-level node in the corresponding dendrograms.

Furthermore, it should be noted that the kind of representation provided in Figure 6.7 intends to make big amounts of complex data evaluable by human visual inspection. Visualization of big and

heterogeneous data in order to make them accessible for visual evaluation is a common issue in many areas of research and an active field of research itself [276–279].

Cluster analysis at the transcriptomic level (Figure 6.7A) separates the W65-NprE strain family into two groups, one group comprising strains having one *rrn* operon deleted as well as another group having more than one *rrn* operon deleted. Surprisingly, strain MB001-NprE was also assigned to the latter group. Unexpectedly, strains W52-NprE and W54-NprE are more similar in terms of gene expression than strains W52-NprE and W53-NprE, as determined from cluster analysis and as it is seen by the color-encoded differential transcript levels (Figure 6.7A). When analyzing by visual inspection the color-coding for the W52-NprE/W54-NprE pair of strains in comparison to strain W53-NprE, genes assigned to subcluster A are mostly downregulated to the same extended for these three strains, as indicated by a comparable color scheme. Genes assigned to subcluster B are mostly downregulated for all these three strains, although downregulations are much more pronounced for strain W53-NprE. Furthermore, genes from subcluster C are up- and downregulated at a similar ratio for strain W53-NprE, while those genes are mostly upregulated for strains W52-NprE/W54-NprE. Several strongly up-regulated genes are found in subcluster D for strain W53-NprE (indicated by the dark red bands), which is not the case for strains W52-NprE/W54-NprE where mostly light blue colored bands are seen.

The clear separation between strains W52-NprE and W53-NprE, as well the high similarity between strains W52-NprE and W54-NprE would not have been inferred from the corresponding genotypes. The genomic difference between strains W52-NprE and W53-NprE is, next to shared Δ*rrnB*, the additional deletion of cg0931, which encodes for a Class I aminotransferase, which is only poorly characterized to date. A double deletion mutant Δ(cg0931 *ilvE*) exhibited no altered growth phenotype in defined medium [280]. Strikingly, the absence of one *rrn* operon in strains W52-NprE (Δ*rrnB*) and W54-NprE (Δ*rrnC*) results in highly comparable transcriptomes, meaning that the additional deletion of cg3261-3298 in W54-NprE does not substantially contribute to the transcriptomic phenotype, as inferred from cluster analysis and visual inspection of Figure 6.7A.

The second main group of strains (comprising MB001-NprE, W25-NprE, W64-NprE and W65-NprE), according to the clustering based on transcriptomes, share quite similar moderately downregulated gene transcripts constituting subcluster A, with the exception of strain W25-NprE that exhibits strongly upregulated genes here. Gene expressions found for subclusters B and C are rather similar for the four strains, while the majority of genes assigned to subclusters B and C are identified as up- and downregulated, respectively. Like analyzed by visual inspection for subcluster A, the three strains MB001-NprE, W64-NprE and W65-NprE show comparable regulation patterns at the transcript level for the genes assigned to subcluster D (moderately upregulated), while the majority of those

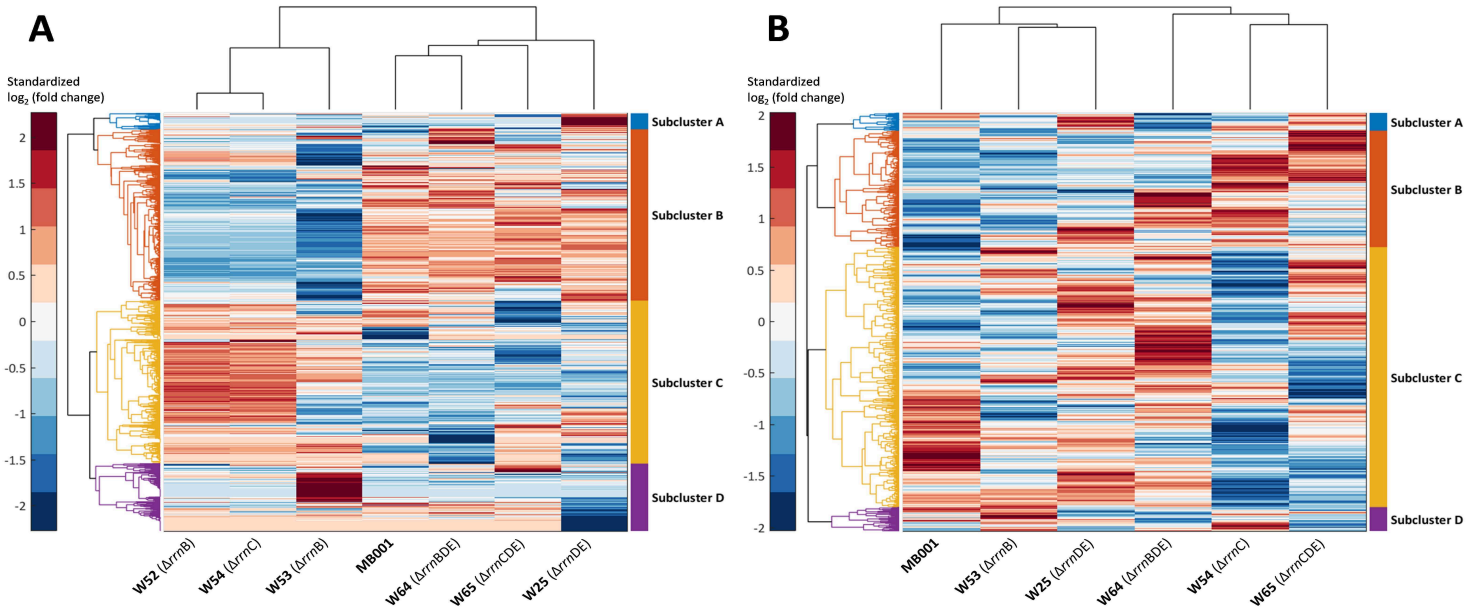


Figure 6.7.: Cluster analysis from transcriptome microarray (A) and proteome (B) data of W65-NprE strain family. Four subclusters (A, B, C and D) of differently regulated mRNA and protein levels are identified from the second-level nodes of dendrograms on the left sides. Subclusters splitting at the second-level nodes are indicated by blue, red, yellow and purple colored parts of the dendrograms. Colorbars on the right sides encode the values for standardized $\log_2$ of mRNA level fold change as ratio of medians or for standardized $\log_2$ ratio of protein levels as ratio of means, in comparison to WT-NprE. Deletions of *rrn* operons in the strains are indicated in parenthesis. Clustering was made using MATLAB's `clustergram` function from the Bioinformatics toolbox, based on correlation as distance metric after standardization, as suggested [275], with further default settings. Strain W52-NprE was not tested for proteome.

genes is found to be strongly downregulated in strain W25-NprE. Conclusively, transcriptomes of the second group of strains (i.e., MB001-NprE, W25-NprE, W64-NprE, W65-NprE) seem to be determined mainly by the deletion of the prophage regions, as their common genetic characteristic is the genotype of MB001 and as they produce similar patterns for the genes assigned to subclusters B and C. The impact of additional deleting cg2801-2828 seems to be restricted to the regulation of genes assigned to subclusters A and D, although the additional deletion of $\triangle$*rrnB*-0931 (i.e., in strain W64-NprE) and $\triangle$*rrnC*-3298 (i.e., in strain W65-NprE) seems to reduce the impact of the W25 genotype on the transcriptome, for unknown reasons.

Overviewing the heatmap for differently regulated proteomes (Figure 6.7B) reveals a more diverse picture than the corresponding transcriptomes (cf. Figure 6.7A) for the members of the W65-NprE strain family. This observation is also reflected in the top dendrogram that separates the strains into clusters. More specifically, the edges of the dendrogram connect to nodes lately (i.e., the edges are rather long, which is, e.g. not the case for the transcriptomes of strains W52-NprE and W54-NprE, cf. Figure 6.7A). Furthermore, clustering of strains is different when looking at the proteomes and the transcriptome. Based on proteome, two main strain groups cluster together, that is MB001-NprE, W53-NprE and W25-NprE, as well as W64-NprE, W54-NprE and W65-NprE.

When inspecting the protein level patterns for subcluster A, a rather balanced up- and downregulation is seen for MB001-NprE and W54-NprE, while mainly down-regulated protein levels are found for strains W53-NprE and W64-NprE, with the latter one showing mostly strongly down-regulated protein levels. Strains W25-NprE and W65-NprE show mostly upregulated and mostly strongly upregulated protein levels in subcluster A, respectively. With respect to the genes assigned to subcluster B, a trend from mainly downregulation to mainly upregulation is observed from strains MB001-NprE to W65-NprE, as the strains are arranged at the bottom of Figure 6.7B. More specifically, the dominating color in subcluster B is dark blue for MB001-NprE. Strain W54-NprE shows mainly light-blue encoded protein levels, and for strain W25-NprE some more protein levels are up-regulated.

In the second group of strains, protein levels in subcluster B are predominantly upregulated, with increasing upregulation from strain W64-NprE over W54-NprE to W65-NprE, as inferred from the increasing number of red-encoded protein levels. The majority of genes are assigned to subcluster C, based on proteome data clustering. Visual pattern analysis for this subcluster shows that almost half of the genes are down- (upper part of subcluster B) or upregulated (lower part of subcluster B) for strain MB001-NprE. For strain W53-NprE, the corresponding color pattern is distributed quite equally, with two focused regions observed of strongly upregulated genes (upper part of subcluster C) and downregulated genes (positioned below the middle part of subcluster C). For strain W25-NprE, genes

are mainly upregulated in terms of proteome, again with a concentrated region in subcluster C is found with strongly upregulated protein levels, indicated by the dark red color codings. The same observation is also found for strain W64-NprE, although the amount of strongly upregulated protein levels is higher and located more in the mid range of subcluster C. In strain W54-NprE most protein levels in subcluster C are strongly downregulated, as indicated by the dominating dark blue color codes, while for strain W65-NprE the distribution of up- and downregulated protein levels in subcluster C is almost the same. Lastly, this applies also for genes assigned to subcluster D in the case of strains MB001-NprE and W25-NprE, while for this subcluster mainly upregulated protein levels are found for strains W53-NprE and W54-NprE. Remaining strains W64-NprE and W65-NprE are characterized by mainly downregulated protein levels in subcluster D.

Clustering of strains, as well as clustering of genes, is different at the transcriptomic and proteomic level. Thus, when reviewing the transcriptomes and proteomes for the W65-NprE strain family (Figure 6.7A and B, respectively), it becomes clear that conclusions drawn from transcriptomic as well as proteomic studies should be compared on an individual basis since the proteome cannot be considered as a proteinaceous mirror of the transcriptome. The same holds true for the relation between transcriptome and genome, for example as seen in Figure 6.7A, where different genotypes of strain W52-NprE and W54-NprE result in highly comparable transcriptomes and vice versa (cf. strains W52-NprE and W53-NprE, which deviate only by $\triangle$0931). Speculatively speaking, different manifestations of transcriptomic phenotypes could result in the same certain proteomic phenotype that is to be maintained, representing a mechanism to compensate for different genetic modifications.

In a more general context of genome reduction, both transcriptomic and proteomic studies can be helpful to identify genome modifications that result in down- or upregulation of gene products that may indicate resource savings or wastings, respectively. For example, genomic modifications, resulting in increased production of mRNAs that encode for cellular functions not needed in the (artificial) bioprocess cultivation environment, can be detected and avoided in order to save the resources needed for elevated mRNA production. The same applies in an analogous manner for proteomic studies, and these considerations apply in the opposite sense for the identification of genomic modifications that save cellular resources due to the lowered production of mRNA and proteins not needed in the bioprocess cultivation environment. Aside these considerations for applied research, such studies generate precious knowledge for basic research since functional relationships between genomes, transcriptome, proteomes and phenotypes can be resolved.

6.6.2.2. Analysis of genes encoding ribosomal proteins at the transcriptomic and proteomic level

Because the transcriptome- and proteome-wide analysis for the members of the W65-NprE strain family did not reveal conclusive results with respect to differential gene expression arising from strain specific genomic deletions, more specific analyses were needed. A common feature of the W65-NprE strain family is the deletion of one to three *rrn* operons, which provide rRNA as essential building block for ribosome formation, next to ribosomal proteins (r-proteins). However, as extensively reviewed for the prokaryotic model organism *E. coli*, ribosome biosynthesis and thus, control of ribosome concentration in the cell is tightly regulated in a complex way. rRNA and r-protein make up more than 50% of the total synthesized RNA [270] and up to one third of the total mass [281] of a cell, respectively. Moreover, modeling approaches of bacterial growth concluded that the growth rate for a given environment (e.g., nutrition medium) reached by a microbial culture is mainly determined by the number of ribosomes [260, 282, 283].

The biosynthesis of r-proteins is regulated on a translational level: Newly synthesized r-proteins are assembled into ribosomes together with rRNA, resulting in low concentrations of free r-proteins. As soon as rRNA production is slowed down due to some reasons, a surplus of r-proteins is produced that is not engaged in ribosome assembly. In this case (i.e., when not assembled in a ribosome), certain r-proteins act as translational repressor by binding to the mRNA of r-protein operons, which, in turn, represses translation [284, 285]. To this end, ribosome biogenesis is controlled by rRNA transcription [286] and, in consequence, the feedback regulation ensures the production of rRNA and r-proteins at stoichiometric amounts needed to assemble ribosomes. The synthesis of stable RNAs (i.e., rRNA and tRNA) remains undisturbed as long as sufficiently high amounts of amino acids are present to be loaded to tRNAs, which is measured by the ratio of aminoacyl-tRNAs and uncharged tRNAs. As soon as amino acid concentrations become limiting, which results in a non-optimal ratio of aminoacyl-tRNAs and uncharged tRNAs, synthesis of stable RNA is rapidly downregulated, which, in turn, causes a downregulation of (r-)protein biosynthesis [283]. Whether in *C. glutamicum* this non-optimal ratio of uncharged tRNAs to aminoacyl-tRNAs is sensed by increased ppGpp levels like in *E. coli* or *B. subtilis* [270, 287] or the uncharged tRNAs themselves are involved in a negative autoregulation of r-protein biosynthesis is not resolved yet [273, 288].

A specific analysis of expression of r-protein genes is given in Figure 6.8 and Figure 6.9, which show differently regulated transcript levels and protein levels, respectively, of r-protein encoding genes for the members of the W65-NprE strain family. The grouping of genes in operons is indicated by the grouping of corresponding genes in panels numbered with small roman numbers. For each gene, strain specific transcript or protein levels (Figure 6.8 or Figure 6.9, respectively) are color-encoded,

as indicated in the figure legends. To facilitate *inter*-operon comparison of r-protein gene expression at the transcript and protein level, axes are scaled to the same range per figure. Although much information at different levels is shown in Figure 6.8 and Figure 6.9, the main purpose of these figures is to facilitate a comparative analysis by visual inspection of r-protein gene expression at the transcriptome and proteome level among the W65-NprE strain family.

Microarray analysis for the transcriptomes of the W65-NprE strain family (Figure 6.8) shows a diverse picture of r-protein gene expressions. For example, the biggest operon of r-protein genes (operon *i* in Figure 6.8) shows nearly no differential regulations for strain MB001-NprE (red bars), while many of the corresponding transcripts for strain W53-NprE (orange bars) are upregulated, sometimes almost by a factor of four (*rpsJ* and *rplP*). Also, an upregulation for several genes from this operon is observed for strain W54-NprE. Interestingly, for the strains with two (W25-NprE) or three (W64-NprE and W65-NprE) *rrn* operon deletions, transcript levels are found to be mainly lowered. At the protein level, shown in Figure 6.9, this operon shows for nearly all genes a downregulation in all strains except MB001-NprE, which is inconspicuous in this case. Other operons are found to be upregulated for all strains (e.g., operon *ix*), or mainly upregulated for all strains but to different extends of the single genes (e.g., operon *iv*).

Some further operons contain other genes than encoding for r-proteins, for example *infA* and *infC*, encoding translation initiation factors IF-1 and IF-3, respectively, are organized with r-protein encoding genes in operons *ii* and *vii* depicted in Figure 6.8 and Figure 6.9. With respect to the transcript levels, *infA* and *infC* are downregulated for most and all strains, respectively, while the remaining r-protein genes of the corresponding operons are mainly upregulated (cf. Figure 6.8). At the protein level, *infA* gene products are found not to be differently regulated with a tendency of downregulation for the remaining genes in this operon. On the other hand, the *infC* gene product is found to be downregulated for all strains, which is also mainly the case for *rpmI* and *rplT* in this operon.

In general, when overviewing the differential gene expressions at transcriptomic and proteomic levels (i.e., when visually comparing Figure 6.8 and Figure 6.9), a higher diversity in up- and downregulation is seen for the transcriptome with several distinct upregulations seen for strains W53-NprE and W54-NprE (e.g., in operons *i*, *v*, *xiv* and *xv* in Figure 6.8). However, for some operons specific patterns are observable, like almost no differential regulation (e.g., operon *x*), a comparable and strong regulation for all strains (operon *vii*) or inconclusive regulation among strains (e.g., operon *vi*).

In contrast, downregulation of r-protein genes at the proteome is dominant in general, with a preference for the strains having at least one *rrn* operon deleted. For several operons, protein levels for strain W65-NprE are more strongly downregulated than for the other strains (e.g., *rplC*, *rplD*, and

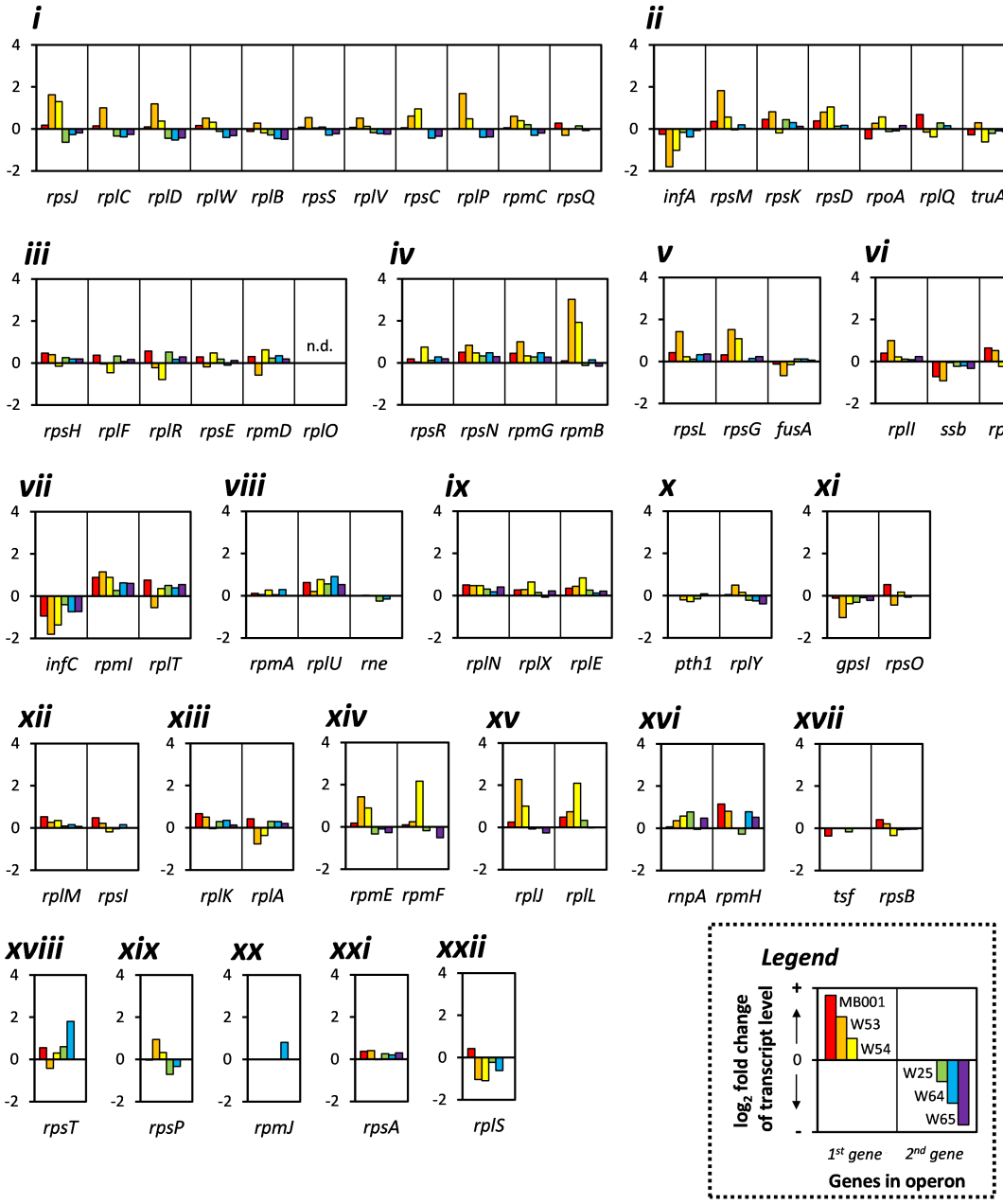


Figure 6.8.: Comparison of transcriptomic phenotypes specifically for ribosomal protein (r-protein) encoding genes for the W65-NprE strain family. Shown are $\log_2$ values of fold changes of transcript levels for the members of the W65-NprE strain family, as indicated in the legend. The organization of r-protein genes into seventeen operons (panels *i* – *xvii*) is represented by grouping of respective genes. Likewise, panels *xviii* – *xxii* show r-protein genes that are not organized with other genes in an operon. Operon organization was retrieved from the online database corynereg.net. *n.d.* not detectable.

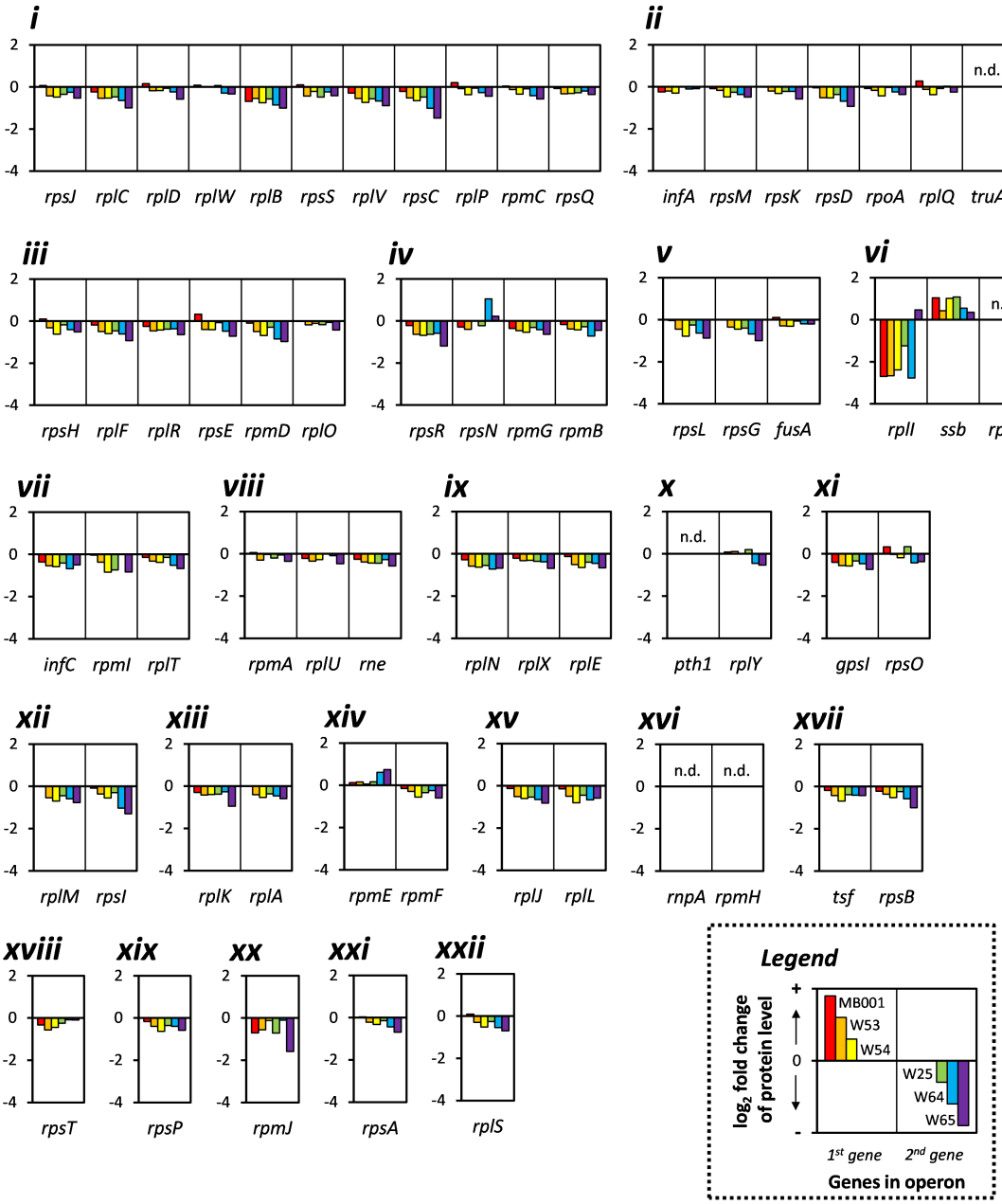


Figure 6.9.: Comparison of proteomic phenotypes specifically for ribosomal protein (r-protein) encoding genes for the W65-NprE strain family. Shown are $\log_2$ values of fold changes of protein levels for the members of the W65-NprE strain family, as indicated in the legend. The organization of r-protein genes into seventeen operons (panels *i* – *xvii*) is represented by grouping of respective genes. Likewise, panels *xviii* – *xxii* show r-protein genes that are not organized with other genes in an operon. Operon organization was retrieved from the online database corynereg.net. *n.d.* not detectable.

rpsC in operon *i* or *eplF*, *rplR*, and *rpmD* in operon *iii*), which could result in the higher growth rate reduction for this strain compared to the other strains. Notably, the strongest downregulation at the protein level is observed for *rplI* in operon *vi* (except for strain W65-NprE), which is a unique pattern among all operons.

As mentioned previously for the global transcriptomes and proteomes, it seems difficult to conclude a proteomic phenotype from transcript microarray analysis in general. However, in case the proteome is deemed more directly causative for a growth phenotype that the transcriptome, the observation of a general lower abundance of r-proteins (and thus, ribosomes) in the W65-NprE strain family compared to WT-NprE agrees well with the observed significantly lowered growth rates compared WT-NprE (all $p < 0.001$ [3], cf. also Table E.1). This is also the case for the corresponding W65-AmyE strain family.

The fact that for strain W65 (without an expression plasmid) an unspecific accumulation of intracellular amino acids was observed [222], may indicate that in strain W65 some regulatory mechanism involving sensing amino acid concentration, ratio of uncharged to aminoacyl-tRNAs, and/or the resulting ratio of actively translating and non-translating ribosomes is disturbed, which finally results in a significantly reduced growth rate. Another speculation would be that the deletion of three out of six *rrn* operons limits the maximal rRNA production rate, which, in turn, limits the production of ribosomes.

Following these considerations, the highly increased biomass specific cutinase yield found in this study may be explained like this: The plasmid-based overproduced NprE-cutinase transcripts may outnumber other mRNAs competing for the (reduced) translational capacity in strain W65-NprE, at the cost of other proteins needed for cell growth. This is supported by the finding that the reduction of growth rate for strain W65-NprE (reduction of 61.3 ± 2.9 %) is much higher than for strain W65 (reduction of 26.3 ± 1.2 %) in comparison to strain WT-NprE than for strain WT, respectively (cf. Figure 6.6, Table E.1). The successful competition of NprE-cutinase transcripts for ribosomal capacity would then be reflected by the remarkably higher cutinase yield of strain W65-NprE (3.1 ± 0.5 fold of WT-NprE level). Strain W65-AmyE reaches a cutinase yield of only 1.3 ± 0.2-fold the level of WT-AmyE, which coincides with a growth rate reduction of 27.4 ± 0.9 %, comparable to strains W65/WT (cf. Figure 6.6, Table E.1).

However, these considerations would also apply for the other GRS-NprE comprising multiple *rrn* operon deletions, which do not exhibit the remarkably increase in cutinase yield as strains W65-NprE does. Therefore, detailed experiments with strains W65-NprE and W65-AmyE are needed, for

[3] according to Welch test

example for the determination of ribosome-binding affinities of AmyE-cutinase and NprE-cutinase mRNAs, copy number estimation of plasmid-borne and chromosomal mRNAs or measurement of intracellular amino acid pools as well as number of ribosomes.

The human visual inspection of figures showing complex data sets with several dimensions (e.g., strains, -omes, phenotypes) is subjected to human bias and thus, conclusions drawn from such analyses may be elusive. Therefore, the focused analysis of gene targets that were subjected to stringent statistical analysis may add another perception on why strain W65-NprE exhibits its remarkably increased biomass-specific cutinase yield.

6.6.2.3. Targeted analysis of the significantly differently regulated proteome of strain W65-NprE

Prior to analyzing results from proteome measurements, corresponding data were subjected to stringent evaluation (cf. section 8.8.9) and several requirements had to be met: From pairwise comparison of detected proteins in members of the W65-NprE strain family and strain WT-NprE, p-values were adjusted to correct for the multiple testing problem (false discovery rate (FDR) controlled at 0.01). Next, protein levels had to be differently regulated by a factor of at least two. Finally, the generation of five to six samples from replicate cultivations in a MBR growth experiment was repeated independently, resulting in eleven to twelve samples per strain. All detected proteins that satisfied these criteria are listed in Table 6.2, together with the corresponding data from microarray analysis.

Table 6.2.: Overview of significantly different protein levels and corresponding transcript levels (given as italics in parentheses) of the W65-NprE strain family in comparison to strain WT-NprE.

			Fold change in GRS-NprE					
Gene	**Name**	**Annotation**	**MB001**	**W25**	**W53**	**W54**	**W64**	**W65**
cg0131		Putative oxidoreductase / aldo-keto reductase						2.2 (*1.8*)
cg0192		Hypothetical protein, conserved				2.7 (*n.d.*)		3.6 (*1.7*)
cg0267	*aroT*	Hypothetical protein, conserved				2.9 (*0.7*)		2.9 (*3.5*)
cg0422	*murA*	UDP-N-acetylglucosamine 1-carboxyvinyl-transferase, horizontally transferred				2.4 (1.1)		2.4 (1.6)
cg0582	*rpsG*	30S ribosomal protein S7						0.5 (*1.2*)

Continued on the next page.

Table 6.2.: Continued.

			Fold change in GRS-NprE					
Gene	**Name**	**Annotation**	**MB001**	**W25**	**W53**	**W54**	**W64**	**W65**
cg0598	*rplB*	50S ribosomal protein L2						0.5 (*0.9*)
cg0601	*rpsC*	30S ribosomal protein S3						0.4 (*0.8*)
cg0674	*rpsI*	30S ribosomal protein S9					0.5 (*0.7*)	0.4 (*0.8*)
cg0693	*groEL*	60 KDa chaperonin (protein CPN60 GroEL), C-terminal fragment, putative pseudogene	0.2 (*n.d.*)					
cg0752		Putative secreted or membrane protein						0.5 (*0.9*)
cg0755	*metY*	O-acetylhomoserine sulfhydrylase, loss causes methionine auxotrophy				2.4 (*2.1*)		2.9 (*3.5*)
cg1049		Putative enoyl-CoA hydratase / isomerase, carnithine racemase				3.2 (*0.4*)		2.4 (*2.0*)
cg1141		Putative protein, homolog of lactam utilization protein B, UPF0271-family, conserved				3.4 (*1.6*)		
cg1216	*nadA*	Quinolinate synthetase						0.5 (*1.3*)
cg1479	*malP*	Maltodextrin phosphorylase				0.4 (*2.2*)		0.4 (*0.3*)
cg1553	*qor2*	Quinone oxidoreductase, involved in disulfide stress response						2.4 (*19.2*)
cg1740		Putative nucleoside-diphosphate-sugar epimerase				3.5 (*0.8*)		2.3 (*3.7*)
cg1811	*cIHF*	Putative integration host factor, conserved					0.5 (*1.2*)	0.5 (*1.0*)
cg1966	*cgpS*	Prophage silencing protein	0.3 (*0.0*)	0.3 (*0.1*)	0.3 (*0.1*)	0.4 (*0.1*)	0.3 (*0.0*)	0.3 (*0.0*)
cg2025		Hypothetical protein CGP3 region		0.2 (*0.0*)	0.2 (*0.1*)	0.2 (*0.1*)	0.2 (*0.0*)	0.3 (*0.0*)
cg2026		Hypothetical protein CGP3 region		0.1 (*n.d.*)	0.2 (*0.1*)	0.2 (*0.1*)		0.2 (*0.0*)

Continued on the next page.

Table 6.2.: Continued.

			Fold change in GRS-NprE				
Gene	**Name**	**Annotation**	**MB001**	**W25**	**W53**	**W54**	**W64**	**W65**
cg2238	*thiS*	Sulfur transfer protein involved in thiamine biosynthesis					0.3 (*0.6*)
cg2444		Hypothetical protein		12.6 (*7.5*)			10.8 (*11.8*)	17.3 (*15.3*)
cg2558		Putative protein, related to aldose 1-epimerase				2.3 (*0.6*)		2.2 (*2.6*)
cg2708	*musK*	ABC-type maltose transport system, ATPase component				0.5 (*1.9*)		
cg2782	*ftn*	Ferritin						2.3 (*1.0*)
cg2847	*mshD*	Acetyltransferease, 1-D-myo-inosityl-2-L-cysteinyl-amido-2-deoxy-α-D-glucopyranoside N-acetyltransferase						2.4 (*3.4*)
cg2958	*butA*	L-2,3-butanediol dehydrogenase / acetoin reductase				2.7 (*2.6*)		3.4 (*3.7*)
cg3079	*clpB*	ATP-dependent protease heat shock protein		4.4 (*7.2*)			3.9 (*6.1*)	4.8 (*15.2*)
cg3186	*cmt2*	Trehalose corynomycolyl transferase				0.5 (*1.0*)		0.5 (*0.5*)
cg3264	*rsmP*	Cytoskeletal protein RsmP, regulates rod-shape morphology, conserved				0.2 (*n.d.*)		0.2 (*n.d.*)
cg3335	*malE*	Malic enzyme ($NADP^+$)				2.2 (*1.3*)		2.1 (*2.8*)
cg3350	*genF*	Fumarylpyruvate hydrolase				4.4 (*1.8*)		4.7 (*5.3*)

Functional annotation of gene products according to [48]. Empty entries indicate no significant differences in protein level. *n.d.* indicates no detectable microarray signal.

One group of differently regulated genes is located within the prophage regions of the wild type strain. These genes are found to be significantly downregulated, as expected since all members of the W65-NprE strain family are based on strain MB001. Thus, the proteome analysis resembles the ΔCGP123 genotype of the strains. From the group of r-proteins, four genes (*rpsG*, *rplB*, *rpsC* and *rpsI*) are significantly downregulated in strain W65-NprE, providing further evidence for the impact on the

observed phenotypes discussed in the previous section.

Another group of genes is regulated the same way for strains sharing specific genomic deletions. For example, only in strains W54-NprE and W65-NprE *aroT* and *metY* are upregulated, while *cmt2* and *rsmP* in both strains are downregulated at the protein (i.e., functional) level. Therefore, genes with the same regulation found for W54-NprE and W65-NprE are deemed to be differently expressed due to the $\triangle$*rrnC* characteristic, which is the only common genomic deletion specifically for these two strains. Likewise, the same regulation at the protein level observed for *clpB* and cg2444 for strains W25-NprE, W64-NprE and W65-NprE is attributed to the deletion of cg2801-2826, which is the exclusively shared genomic deletion in these three strains. Noteworthy, cg2444 is found to be strongly upregulated in these strains (12.6-, 10.8- and 17.3-fold at the protein level, as well as 7.5-, 11.8- and 15.3-fold at the transcript level). Interestingly, the cg2444 gene product (a protein with unknown functional relation to the cytochrome bc_1-aa_3 supercomplex [289]) was also found to be upregulated in an L-lysine producing *C. glutamicum* strain [290]. This might indicate an increased energy demand in strains W25-NprE, W64-NprE and W65-NprE, probably coinciding with the upregulation of ClpB, which creates additional ATP demand.

Some proteins involved in cell envelope formation are differently regulated in strains W54-NprE and W65-NprE: *murA* [291] (upregulated), as well as *cmt2* [292], *rmsP* [293] (both downregulated). Furthermore, some proteins involved in carbohydrate metabolism are also found to be differently regulated (cg2558 and cg1740: both up, *malP* (cg1479) [294]: down). It is not clear how these contribute to the observed phenotypes of the corresponding strains.

The probably most interesting group of proteins comprises those that encode for cellular functions which drain or replenish metabolites from the central carbon metabolism and connected pathways (e.g., amino acid anabolism). The upregulation of malic enzyme specifically in strains W54-NprE and W65-NprE is deemed to $\triangle$*rrnC*, as stated above. Malic enzyme catalyzes the conversion of malate into pyruvate with the generation of an anabolic reducing equivalent NAPDH. Activity of malic enzyme was reported high between growth rates of μ = 0.1 and 0.33 h^{-1} [295], which may indicate a pivotal role of malic enzyme for NADPH replenishment at lowered growth rates.

The activity of fumarylpyruvate hydrolase (*genF*, cg3350) feeds into the pyruvate pool as well as into the malate pool indirectly via fumarate. Its upregulation would therefore result in a higher flux into these metabolite pools, although the origin of the reaction educts for fumarylpyruvate hydrolase remains unclear. The same holds true for the question whether an increased pyruvate demand in strains W54-NprE and W65-NprE exists. Formation of fumarylpyruvate by *genF* is a reaction of the catabolic gentisate and 3-hydroxybenzoate pathway [296]. For strain W65-NprE, also an upregulation

is seen for MshD (cg2847) catalyzing the formation of mycothiol, which is needed as cofactor for the isomerization of maleylpyruvate into fumarylpyruvate [297].

Quinolinate synthetase (*nadA*, cg1216) is also closely connected with the central carbon metabolism. It catalyzes the second reaction in the NAD biosynthesis by converting iminoaspartate (produced from aspartate by *nadB* gene product) and phosphoenolpyruvate (PEP) into quinolinate [298]. Consequently, downregulation of *nadA* results in a lowered drain of PEP from glycolysis, as well as aspartate derived from oxaloactetate.

MetY provides additional methionine synthesizing capacity in *C. glutamicum* via direct sulfhydrylation by omitting the need for cysteine as sulfur-donor [299]. Thus, the observed upregulation of *metY* in strains W54-NprE and W65-NprE may be due to increased methionine demand since *metY* is strongly repressed by low amounts of methionine. Another possible explanation of *metY* upregulation would be a shortage of cysteine, which is needed in the transsulfuration branch of methionine biosynthesis [299]. Interestingly, among the unspecific accumulation of several amino acids in strain W65, no methionine or cysteine was detected [222]. Since MetY utilizes disulfide as sulfur donor [299, 300], this might explain the upregulation of *qor2*, which is involved in disulfide stress response [301]. However, no interrelation on the expression of *metY* and *qor2* has been reported yet.

A class I aminotransferase, *aroT* (cg0267), is also found to be upregulated. A high affinity for L-glutamate as substrate was reported [302] and together with the observation that an Δ(*aroT ilvE*) mutant was auxotrophic for L-leucine, L-isoleucine and L-phenylalanine [280] supports the assumption that *aroT* upregulation maybe due to an increased demand for these amino acids in strains W54-NprE and W65-NprE.

The observed upregulation of *butA* (cg2958) would contribute to an increased production of 2,3-butanediol although its physiological role remains unclear, as well as the reason for the observed upregulation of *butA*. Probably, it is used for recovery of NAD^+ from excess NADH by the reduction of acetoin or diacetyl [303], as excess NADH probably cannot be reduced by oxidative phosphorylation, although upregulation of cg2444 may indicate a correspondingly increased capacity. However, a detoxifying function of *butA* for acetoin can be excluded since *C. glutamicum* was reported to be able to produce acetoin at titers exceeding 90 g L^{-1} [304].

For the highly interesting strain W65-NprE, which shows the highest increase in cutinase yield $Y_{P/X}$ among all GRS tested, the impact of the above described alterations of the proteome is sketched as proposed metabolic scenario in Figure 6.10: Flux into pyruvate pool is increased by the upregulation of *malE* and *genF*. The latter one additionally increases malate availability for *malE* via fumarate. Glutamate is drained from the tricarboxylic acid circle presumably by *aroT*, which is upregulated at

the protein level to compensate for an increased demand for isoleucine, leucine and phenylalanine. A drain of phosphoenolpyruvate is reduced by downregulation of *nadA*, which at the same time lowers the demand for aspartate. Also, a shortage of methionine and/or cysteine is presumed, which leads to upregulation of *metY* for compensation. The physiological purpose of increased production of 2,3-butanediol from acetoin remains unclear in this scenario.

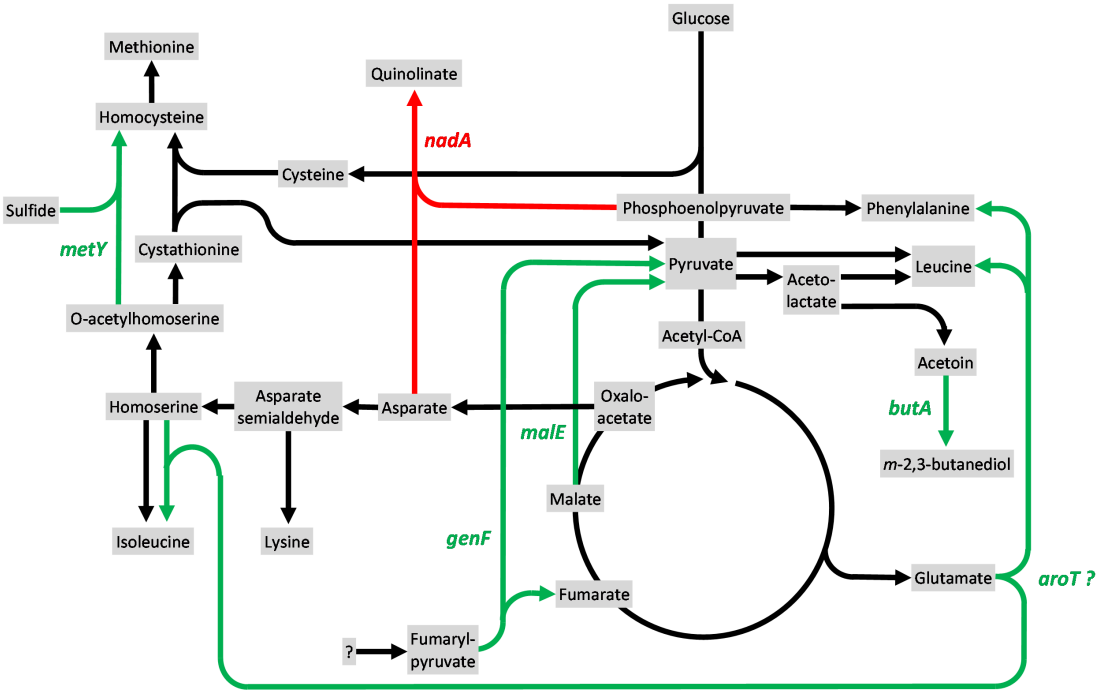


Figure 6.10.: Proposed metabolic scenario resulting from observed differential regulation of protein levels in strain W65-NprE. Arrows indicate known reaction pathways. Green and red arrows indicate a possible increase and decrease, respectively, in flux due to observed protein levels of *aroT*, *butA*, *genF*, *malE*, *metY*, and *nadA* gene products. Presumed in-vivo reactions of AroT are not known, which is indicated by a question mark. Depicted scenario is no complete stoichiometric model and arrows may comprise several reactions.

In summary, the differential regulation of proteins connected with the central carbon metabolism may be a result of the increased energy demand due to heterologous cutinase production in these strains. In a variety of studies, *C. glutamicum* demonstrated a remarkable flexibility by adaption to environmental condition and metabolic perturbations (e.g., [11, 299, 305]). In this sense, the proteomic changes due to deletion of *rrnC*-3298 (yielding strain W54) enable *C. glutamicum* to achieve high cutinase yields, though at the cost of low growth rates. The additional deletion of *rrnD* and *rrnE* covered by Δcg2801-2828 (yielding strain W65) would then be the trigger for the NprE-cutinase transcript to be translated more efficiently than in strain W54-NprE, probably due to more successful

competition for the reduced lowered ribosomal capacity, according to the hypothesis discussed in section 6.6.2.2.

The resulting depicted metabolic scenario in Figure 6.10 does not acknowledge all reactions of current metabolic genome-scale models (GSMs) for *C. glutamicum* (e.g., [306, 307]), but attempts to highlight and visualize possible metabolic perturbations in this strain. To verify or falsify the proposed metabolic scenario, further detailed analyses are needed, including metabolomics and integration of new reactions (e.g., cutinase biosynthesis) into GSMs, which may help to identify metabolic engineering targets affecting heterologous cutinase production.

6.7. Global analysis of cutinase secreting genome reduced *C. glutamicum* strains by a multiple regression approach using dummy variables

Detailed discussions were provided on effects due to the application of different Sec SPs in combination with genomic deletions on extracellular phenotypes (section 6.1), specific strains showing surprising phenotypes (section 6.2 to section 6.4), differential phenotypical analysis of strain families (6.5 and section 6.6), as well as deciphering of transcriptomes and proteomes (section 6.6.2). Finally, a general global analysis in terms of a practical route towards the definition of a streamlined genome with respect to improved cutinase secretion performance (i.e., $Y_{P/X}$) is envisaged. Beside the fundamental knowledge generated in this study, the question for practical application is: Which genotype is needed to obtain the highest cutinase yield by means of genome reduction? To answer this question a "score" needs to be assigned to each genomic deletion that was investigated in this study. The fact that for many strains data on cutinase yield was obtained, for both the AmyE and NprE Sec SP, with accompanying determination of growth rates enables to carry out a multiple regression using dummy variables encoding the genetic characteristics of the GRS tested.

The strain characteristics (i.e., genomic deletions, Sec SPs), which can be described by Boolean variables, were encoded by dummy variables, which take a value of one if the characteristic is present (e.g., gene cluster is deleted) or zero (e.g., gene cluster is not deleted). In total, 33 strain characteristics are taken into account, that is the presence of the AmyE and NprE Sec SP (which cannot take a value of 1 at the same time) and 31 genomic deletions that characterize all strains investigated in this study under the same conditions. The linear model takes the form $y = \beta_0 + \sum_{i=1}^{33} \beta_i x_i + \epsilon$ where each β_i encodes a different genetic characteristic (with $x_i \in \{0, 1\}$) and β_0 is interpreted as response (i.e., growth rate or cutinase yield) without any of the genetic characteristics being present (i.e., $x_i = 0$). Consequently, estimated values for β_1 to β_{33} are interpreted as above mentioned

"scores". Results from multiple regression analysis are shown in Table 6.3.

Table 6.3.: Results from multiple linear regression analysis of growth rate μ and cutinase yield $Y_{P/X}$ as responses in dependence of genetic strain characteristics, encoded by dummy variables.

		Response: μ		**Response: $Y_{P/X}$**	
Genetic characteristic	**Coefficient β_i**	**Estimate (error)**	**p-value**	**Estimate (error)**	**p-value**
	β_0 (Intercept)	0.449 (0.014)	<0.01	-0.093 (0.068)	0.18
AmyE SP	β_1	-0.016 (0.029)	0.59	0.306 (0.145)	0.04
NprE SP	β_2	-0.071 (0.029)	0.02	0.645 (0.145)	<0.01
ΔCGP123 Δ*ISCg12*	β_3	-0.033 (0.028)	0.26	-0.009 (0.141)	0.95
Δ*rrnA*	β_4	0.007 (0.029)	0.80	0.043 (0.145)	0.77
Δ*rrnB*	β_5	-0.008 (0.016)	0.62	0.100 (0.078)	0.21
Δ*rrnC*	β_6	-0.007 (0.017)	0.69	-0.010 (0.086)	0.91
Δ*rrnD*	β_7	0	n/a	0	n/a
Δ*rrnE*	β_8	-0.042 (0.025)	0.10	-0.066 (0.124)	0.59
Δ*rrnF*	β_9	-0.017 (0.017)	0.34	-0.035 (0.086)	0.69
Δ0414-0440	β_{10}	0	n/a	0	n/a
Δ0635-0646	β_{11}	0.091 (0.035)	0.01	-0.044 (0.172)	0.80
Δ0704-0748	β_{12}	0	n/a	0	n/a
Δ0822-0845	β_{13}	0	n/a	0	n/a
Δ0931	β_{14}	-0.012 (0.021)	0.57	-0.060 (0.105)	0.57
Δ1018-1033	β_{15}	0	n/a	0	n/a
Δ1172-1213	β_{16}	0	n/a	0	n/a
Δ1281-1289	β_{17}	-0.095 (0.024)	<0.01	0.076 (0.120)	0.53
Δ1291-1305	β_{18}	0	n/a	0	n/a
Δ1340-1353	β_{19}	-0.127 (0.038)	<0.01	0.034 (0.188)	0.86
Δ1370-1385	β_{20}	-0.153 (0.035)	<0.01	0.026 (0.172)	0.88
Δ2312-2322	β_{21}	0	n/a	0	n/a
Δ2621-2643	β_{22}	0	n/a	0	n/a
Δ2663-2686	β_{23}	0	n/a	0	n/a
Δ2755-2760	β_{24}	0	n/a	0	n/a
Δ(2801-2812 2822-2824 2828)	β_{25}	-0.019 (0.030)	0.53	0.242 (0.151)	0.12
Δ2990-2999	β_{26}	-0.027 (0.029)	0.36	0.009 (0.145)	0.95

Continued on the next page.

Table 6.3.: Continued.

		Response: μ		Response: $Y_{P/X}$	
Genetic characteristic	**Coefficient β_i**	**Estimate (error)**	**p-value**	**Estimate (error)**	**p-value**
Δ3000-3006	β_{27}	-0.115 (0.024)	<0.01	0.118 (0.120)	0.33
Δ3072-3091	β_{28}	0	n/a	0	n/a
Δ3102-3111	β_{29}	-0.002 (0.019)	0.92	0.048 (0.094)	0.61
Δ3261-3262	β_{30}	0	n/a	0	n/a
Δ3263-3298	β_{31}	-0.066 (0.024)	<0.01	0.208 (0.119)	0.09
Δ3299-3301	β_{32}	-0.036 (0.032)	0.26	-0.188 (0.157)	0.24
Δ3324-3345	β_{33}	-0.011 (0.016)	0.48	0.115 (0.079)	0.15
Regression diagnostics					
R^2		0.832		0.769	
adjusted R^2		0.770		0.683	
SSE		0.069		1.695	
RMSE		0.036		0.177	
Standardized residuals [#]		p = 0.015		p < 0.001	

Values for cutinase yields for strains without either AmyE or NprE SP are set to zero. Data on growth rates for strains without cutinase secretion plasmids was published previously [48, 55, 222]. Regressions were calculated using MATLAB's `fitlm` function. *RMSE* root mean squared error. *SSE* sum of squared errors. [#] p-value according to Anderson-Darling test. *n/a* not available.

As stated above, intercepts β_0 represent the growth rate and cutinase yield without further genetic characteristics. For growth rate, a value of 0.45 h^{-1} is determined which is in excellent agreement with literature [48, 55, 218]. For cutinase yield, a value of -0.09 kU g_X^{-1} is determined. Since a negative cutinase yield is not possible and by considering the corresponding p-value of 0.18, this deviation from zero is attributed to the regression method itself.

For both growth rate μ and cutinase yield $Y_{P/X}$, several coefficients in Table 6.3 show negative or positive values. The interpretation would be that when corresponding genetic characteristics are present (i.e., use of a Sec SP or deletion of genes), the resulting response (i.e., μ and $Y_{P/X}$) is decreased or increased, respectively. With respect to calculated error estimates and corresponding p-values for coefficients, for many of them a rather high uncertainty needs to be taken into account. Uncertainties of coefficient values for cutinase yield are often higher than for growth rate, which could be attributed to the generally higher uncertainties of determined cutinase yields in comparison

to growth rate, as seen in Table E.1. However, this regression approach is a purely phenomenological analysis of the collected data without considering any functional effects of genomic deletions and without considering any expert knowledge. For example, only one Sec SP can be used at the same time, as well as the deletion of all *rrn* operons would result in a lethal phenotype. This has to be considered for results interpretation.

By arbitrarily choosing a minimum positive effect on cutinase yield of 0.1 for each genetic characteristic, strain "GRS_Pred1" is proposed with its genetic characteristics (see Table 6.4, compare also Table 6.3). The resulting predictions of growth rate and cutinase yield are μ^{pred} = 0.16 ± 0.06 h^{-1} and $Y_{P/X}^{pred}$ = 1.33 ± 0.30 kU g_X^{-1}, respectively, with bootstrapped standard deviations as metric for prediction uncertainty (cf. Table 6.4). By additionally considering genetic deletions with negligible effects on growth rates (i.e., arbitrarily defined as coefficient values between -0.01 and 0.01 with respect to μ), strain "GRS_Pred2" is proposed in Table 6.4 with μ^{pred} = 0.16 ± 0.07 h^{-1} and $Y_{P/X}^{pred}$ = 1.41 ± 0.35 kU g_X^{-1}. Presumably, such strain exhibits a slightly higher predicted cutinase yield without further compromising biological fitness. Additionally, additional genomic deletions could be concluded that do not affect both growth rate and cutinase yield according to the regression analysis results from Table 6.3: Thereby, proposed strain "GRS_Pred3" would show the same predicted responses as strain GRS_Pred2 (see Table 6.4). Many of these gene clusters were previously classified as non-essential in terms of biological fitness [48, 55]. However, in this study, cutinase yield was an additional evaluation criterion in terms of non-essentiality of gene clusters in *C. glutamicum*.

Table 6.4.: Overview of proposed genotypes of strains with predicted growth rates μ^{pred} and cutinase yields $Y_{P/X}^{pred}$, based on results from multiple regression analysis presented in Table 6.3.

Strain	**Genetic characteristics**	**Predicted response** μ^{pred} **[h^{-1}]**	**Predicted response** $Y_{P/X}^{pred}$ **[kU g_X^{-1}]**
GRS_Pred1	Δ*rrnB* Δ2801-2812 Δ2822-2824 Δ2828 Δ3000-3006 Δ3263-3298 Δ3324-3345 pEKEx2-NprE-cutinase	0.16 ± 0.06	1.33 ± 0.30
GRS_Pred2	GRS_Pred1 Δ*rrnA* Δ*rrnC* Δ3102-3111	0.16 ± 0.07	1.41 ± 0.35
GRS_Pred3	GRS_Pred2 Δ*rrnD* Δ0414-0440 Δ0704-0748 Δ0822-0845 Δ1018-1033 Δ1172-1213 Δ1291-1305 Δ2312-2322 Δ2621-2643 Δ2663-2686 Δ2755-2760 Δ3072-3091 Δ3261-3262	0.16 ± 0.07	1.41 ± 0.35

Predictions represent bootstrapped mean values and standard deviations (n_{BT} = 5000). Parametric bootstraps were calculated using coefficient values for specified genetic characteristics with corresponding errors assuming normality. It should be noted that these errors are not taken as any metric for significance.

Despite its weaknesses, the multiple regression approach using dummy variables represents a strategy that condenses the high amount of data generated in the untargeted GRS screening approach. Thereby, the interpretation of coefficient estimates as scores, assigned to the according genetic characteristics, can be used for decision making in metabolic engineering to generate improved cutinase secretion strains.

6.8. Concluding remarks

In this study, comprehensive data on growth and cutinase secretion at different levels (i.e., extracellular rates, transcriptome and proteome) was collected for a *C. glutamicum* GRS library, which represents a typical task in microbial quantitative phenotyping. For specific strains, detailed insights at the transcriptomic and proteomic level were provided, attempting to explain the resulting observed phenotypes. Also, the collected data were subjected to a regression analysis and, despite the weaknesses of this approach, essential genomic deletions we proposed to improve cutinase yield, together with prediction of resulting phenotypes. Although experimental validation is out of scope in this study, this would be highly interesting and could contribute to further refinement of this approach.

In advance of the present study, the screened GRS library was constructed for a different aim, that is the construction of a genome streamlined *C. glutamicum* chassis that exhibits no impaired biological fitness compared to the wild type strain and thus, gene clusters were classified under this premise. In this study, cutinase secretion efficiency was introduced as another screening objective. In this context, several initial gene classifications could be confirmed, whereas others need to be revised when considering this new screening objective. Also, similar to observations made in the initial genome reduction project of *C. glutamicum*, surprising effects of gene deletions with respect to growth and cutinase secretion phenotypes were seen, which again highlights the need for experimental confirmation of in-silico predictions. Overall, the results obtained in this study may contribute to a more rational design of improved *C. glutamicum* hosts for heterologous protein secretion. Conclusively, the current extensive knowledge about *C. glutamicum* is still rather incomplete, providing option for exciting fundamental and applied research on this microbial, even after more than 60 years of investigations after its discovery.

7. Summary, conclusions and outlook

This chapter provides a summary of this work. Furthermore, a few future aspects outreaching the scope of this work are presented as outlook. These aspects concern the application of *C. glutamicum* as potential alternative hast for heterologous protein production, the demand for further development of microbioreactor systems, as well as the need for smart solutions for warehousing, (re-)processing and interpretation of heterogeneous data to cope with foreseeable increase of information output generated from high-throughput experimentation in combination with powerful analytical methods. Parts of this chapter are based on *Publication V* and *Manuscript II*.

Author contributions:
Johannes Hemmerich wrote the chapter, and prepared the figures. Stephan Noack, Marco Oldiges and Wolfgang Wiechert helped with writing section 7.3.

7.1. Summary

This study concerned the investigation of bioprocesses for secretory production of heterologous protein with *C. glutamicum*, and the outcomes of this work are organized in several chapters in this text.

In Chapter 1, an introduction for the topics covered in this work was given to provide a general background. The MCF *C. glutamicum* and its relevant protein secretion pathways were introduced, as well as the use of SP libraries for optimizing heterologous protein secretion. Furthermore, current progress in the generation of genome-streamline chassis strains was described, followed by a short outline on the development of industrial protein production bioprocesses.

In Chapter 2, to discuss the fundamental need for higher experimental throughput, MBR systems were introduced and the importance of deriving PIs as criteria allowing rapid evaluation of experimental results was highlighted. A method for automated growth rate determination from high-throughput MBR experimentation was presented and discussed in detail, since the growth rate is

probably the most often used metric in biological fitness testing of mutant strain libraries during quantitative microbial phenotyping.

In Chapter 3, the necessity of re-validating established workflows from previous studies for microbial phenotyping as with changing and/or additional optimization objectives was shown. Furthermore, a well-established MBR workflow for determination of time-resolved changes of concentration from MBR cultivation that is accompanied by a loss of throughput was improved. The new workflow combines best of both worlds, that is the whole cultivation throughput of the MBR system is maintained while additionally acquisition of time-resolved substrate consumption and secretory product formation is enabled. As application example, this new workflow was demonstrated on PI determination of several cutinase secreting genome reduced *C. glutamicum* strains.

In Chapter 4, a general iterative workflow for bioprocess optimization was presented where high-throughput MBR cultivation, lab automation and Kriging-based DoE complement each other in a synergistic manner. As illustrating application study, the optimization of a well-known and widely applied cultivation medium used for amino acid production with *C. glutamicum* was conducted to maximize secreted GFP titer as optimization objective. Extracellular GFP titer could be doubled by adjusting other compounds like magnesium or calcium but without changing the available amount of the main carbon source. This highlights again the necessity to re-validate established screening protocols when changing screening objectives, even if these changes are considered as small.

In Chapter 5, the interrelation between Sec SP and bioprocess condition for the optimization of heterologous cutinase secretion with *C. glutamicum* was investigated for the first time. Since the envisaged bioprocess control strategies could not be realized in MBR systems, this was done using laboratory scale bioreactors. Previous knowledge about the fundamental question "Which signal peptide is best suited for a target protein of choice?" was that an optimal specific combination of SP, target protein of choice and expression host cannot be predicted in-silico but needs to be determined each time from scratch. This study adds important aspects about the impact of bioprocess condition to this topic. It was highlighted that careful adjustment of bioprocess conditions is needed to fully maximize the optimization objective (here, cutinase secretion efficiency). It could be concluded that general trends from determined empirical relations can be deduced which point to the direction of fruitful bioprocess optimization. Thus, the big design space comprising bioprocess variables can be narrowed, thereby shortening development timelines.

In Chapter 6, a library of genome reduced *C. glutamicum* strains was screened for heterologous cutinase secretion, addressing the question whether current trends in the construction of "genome streamlined" platform strains are meaningful in terms of biotechnological applications. In general,

this study revealed surprising extracellular phenotypes which would not have been supposed from the genotypes of the different strains. Therefore, the data generated in this work provides new knowledge about the classification of poorly characterized genes in the context of heterologous protein secretion. Also, the wealth of phenotypical data was used in a data-driven approach to determine the essential genomic deletions needed for improved cutinase yield from an application-focused point of view. The fact that two strains with impaired growth phenotypes and completely different genotypes showed remarkably increased cutinase yields when using one specific signal peptide illustrates that the current extensive knowledge on *C. glutamicum* after 60 years of intensive research is still rather incomplete and thus, future studies are expected to still reveal highly interesting results concerning both basic and applied research.

Finally, in this chapter, a few important future aspects arising from the results of this work are discussed in the following, to address questions beyond the scope of this work.

7.2. *C. glutamicum* as alternative protein secretion platform

The great potential of *C. glutamicum* for secretory production of heterologous protein is increasingly recognized since several years now [17, 210]. The fact that it is used as workhorse in industrial production of small molecules like, for example, amino acids makes it attractive due to the accompanied gained industrial experience. This represents a tremendous advantage that is typically lacking when developing new MCFs for industrial scale application. This study adds several new aspects that indicate further potential of *C. glutamicum* as MCF for heterologous protein production beyond current benchmarks.

First, it turned out that nutrition media developed for amino acid production are suboptimal when using *C. glutamicum* for protein secretion. Consequently, further research should focus on the nutritional demand during heterologous protein secretion as well as on how this demand impacts the robustness and flexibility of the metabolism of *C. glutamicum*, also in dependence of different bioprocess control strategies. The expectation is that strains optimized for heterologous protein production can be generated.

Second, genome reduction of *C. glutamicum*, although originally intended to generate chassis strains without comprised biological fitness [48, 55], turned out as powerful tool for assessing poorly characterized genomically encoded cellular functions in terms of heterologous protein secretion. Therefore, this study identified several genetic targets which have been unattended previously for detailed investigation with respect to protein secretion. From both a basic and applied research

point of view, further research would add highly valuable knowledge to advance understanding of *C. glutamicum*.

7.3. Expansion of measurement and control capabilities for microbioreactor systems

In this study, a considerable amount of data was generated during quantitative phenotyping of a *C. glutamicum* GRS library, as well as medium optimization. One key to success was the use of robotics-integrated MBR systems, which enable a higher cultivation throughput and thus, enable the collection of phenotyping data in a high amount. It is foreseeable that MBR systems will become increasingly important in the context of microbial quantitative phenotyping and bioprocess development. This holds true especially with further improvements for such systems concerning measurements that are needed that capture both the microbial phenotype and the bioprocess characteristics.

Next to state-of-the-art monitoring capabilities for biomass and other culture parameters, the incorporation of methods for individual determination of respiration activities from sub-mL cultivations [308, 309] and metabolite quantification by multi-wavelength fluorometric techniques [310], Fourier-transformed infrared spectroscopy (FTIR) [311] or impedance spectroscopy [312] were reported to be valuable extensions for MBR systems. Especially the high information content from 2D fluorescence spectra and FTIR requires chemometric models that rely on training and validation datasets, again highlighting the need for novel data evaluation procedures.

The combination of several measurement techniques to construct soft sensors has been applied several times in bioreactor cultivations [103] and these are also promising tools for MBR cultivations. To make the wide range of soft sensors routinely available for MBR systems, intense research is needed as a driver to the miniaturization of sensing probes providing rich analytical data [313] and increased flexibility to act on the individual cultures by integrated microfluidics [145, 314–316]. In combination with liquid handling systems [112, 122, 158], software modules connecting multiple input data streams with multiple hardware actors will be used to control the individual cultures and, finally, enable the robust interplay of all hardware and software components.

7.4. Library screening and modularized workflows

To facilitate rapid progress in development of biotechnological production processes, the construction of an efficient biocatalyst (i.e., strain) is needed. Preferably, such strain can utilize a wide range of carbon sources, does not waste provided building blocks and energy as by-products, and provides the ability to control growth (i.e., biocatalyst formation) and product synthesis independently [256]. To design such ideal strain on a rational basis, its genetic modifications need to be compiled by elements that haven been characterized earlier. Such approach requires the employment of genetic libraries that are easily configurable as "plug-and-play" modules [7, 31, 35, 72, 73]. By combining several well established methods in metabolic engineering, platform technologies operated at systems levels are aimed for [317].

However, as was seen in this work, the compilation of a genotype from individually characterized modules does not necessarily result in a predictable phenotype of the engineered strain. This fact can be described by the phrase "The whole is greater than the sum of its parts" and makes the characterization of newly designed strain indispensable. This core task of quantitative microbial phenotyping can be accomplished in reasonable timeframes using MBR systems, as seen in this study. However, owing to the combinatorial explosion arising from the use of several genetic libraries and cultivation conditions, a trade-off between experimental throughput and coverage of possible strain variants is needed due to practical reasons. Here, straight-forward and simple modeling approaches [37, 318] can help to develop a feeling on how much experimental effort is needed to catch strain variants with a certain probability out of a library.

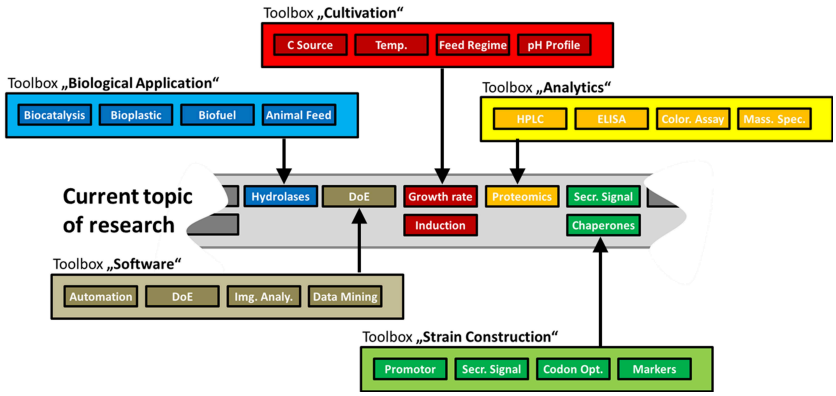


Figure 7.1.: Modularized assembly of methods from several toolboxes. For a current topic of research, previously acquired knowledge on methods is re-used by compilation of individual modules to assemble a new particularly needed method.

Ideally, the newly generated knowledge about the impact of genetic library elements in a new context (e.g., a strain variant) needs to be preserved as additional annotation of the corresponding library items. Importantly, this procedure is not restricted to genetic libraries (e.g., for promotors, ribosome binding sites, strain variants) but should cover also analytical methods, cultivation conditions, or data processing routines. By doing so, the idea is that generated knowledge can accelerate the exploration of future research topics by methods that can be readily compiled from existing and validated component. This is depicted in Figure 7.1, where a current topic of research is elaborated by specific methods that are compiled from modules provided in different toolboxes.

7.5. Continuous knowledgebase improvement

As outlined previously, it can be expected that during each biotechnological development project rich datasets are generated that contribute to a better understanding of genotype-to-phenotype relationships in different platform hosts. Moreover, the integration of further intracellular measurements following proteomics, metabolomics and fluxomics will be the key to faster unravel the complex mechanisms of metabolic regulation. However, to take full advantage of these technologies, it will be important to facilitate the acquirement, (pre-)processing, compressing, filing, managing and evaluation of the resulting complex and heterogeneous datasets. Consequently, further research is needed concerning the establishment of flexible and expandable data warehousing solutions, which should provide convenient software user interfaces, a topic that was put into spotlight again very recently [319, 320].

It will be increasingly important to conserve previously acquired knowledge and, in addition, to enable the re-evaluation of such existing knowledge and raw data in the context of new research. This is summarized by the headline "continuous knowledgebase improvement", which goes beyond the simple filing of data. It is rather about the repeated and continuous analysis of historic raw data in combination with newly acquired raw data and by this, increasing information is transformed into knowledge that is continuously developing. With that, it can be expected that new biotechnology development projects can be conducted within accelerated time lines. This principle of continuous knowledgebase improvement is visualized in Figure 7.2.

To finally conclude, it can be said that research in industrial biotechnology is moving towards an improved knowledge-based strain and bioprocess design. Furthermore, the already existing movement in biotechnology towards small-scale high-throughput experimentation will speed up in the next years [321], making automated and smart routines for data analysis indispensable. This truly

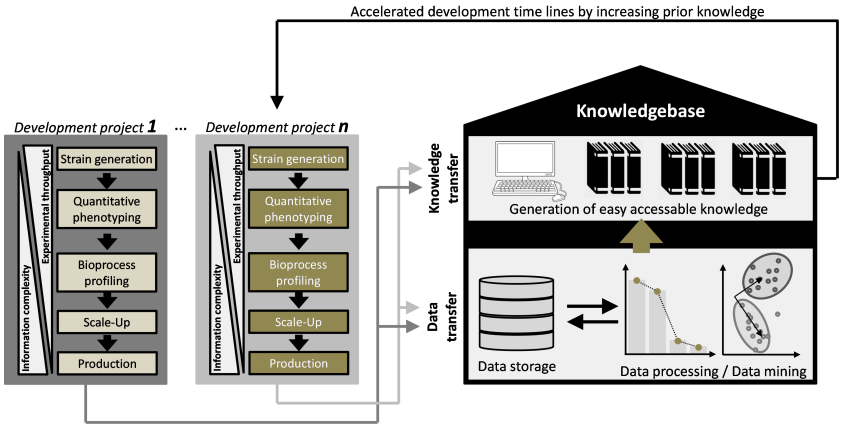


Figure 7.2.: The principle of continuous knowledgebase improvement for accelerated development project time lines. During each development project, substantial amounts of raw data are generated and by analyzing these data, knowledge is produced. The storage of all raw data in combination with continuously developing data processing and data mining routines, the collected raw data can be re-evaluated and re-analyzed. By doing so, previous knowledge can be improved. Finally, the improved knowledge will contribute to accelerated development time lines and thus, accelerates the development of new projects, whose newly generated raw data update the data warehousing.

demands for interdisciplinary work in the fields of biology, biotechnology, information technology, data science, software engineering, hardware and mechanical engineering, providing great opportunities for the upcoming generation of researchers.

8. Material and methods

All chemicals were of analytical grade and purchased from Sigma, Merck or Roth, if not stated otherwise. In case methods were carried out with modifications, this is noted in the text. Data processing was conducted with Microsoft Excel (vers. 2010, 2016) and MATLAB with Statistics Toolbox (vers. 2013a to 2017b). In case other software packages were used for specific tasks, this is noted in the text.

8.1. Strains and plasmids

Table 8.1.: *C. glutamicum* strains and plasmids used in this study.

Name	Relevant characteristics	Reference
Strains		
ATCC 13032	Wild type strain (WT)	[1]
Δ*cglMRR*	WT Δ1996-1998	[53]
MB001	WT ΔCGP123 (Δ1507-1524 Δ1746-1752 Δ1890-2071)	[53]
W25	WT ΔCGP123 Δ*ISCg12* Δ2801-2828	[55]
W28	WT ΔCGP123 Δ*ISCg12* Δ2990-006	[55]
W31	WT ΔCGP123 Δ*ISCg12* Δ3324-3345	[55]
W47	WT ΔCGP123 Δ*ISCg12* Δ1281-1289	[55]
W52	WT ΔCGP123 Δ*ISCg12* Δ*rrnB*	[55]
W53	WT ΔCGP123 Δ*ISCg12* Δ*rrnB*-0931	[55]
W54	WT ΔCGP123 Δ*ISCg12* Δ*rrnC*-3298	[55]
W64	WT ΔCGP123 Δ*ISCg12* Δ2801-2828 Δ*rrnB*-0931	[55]
W65	WT ΔCGP123 Δ*ISCg12* Δ2801-2828 Δ*rrnC*-3298	[55]
W73	WT ΔCGP123 Δ*ISCg12* Δ0158-0183	[222]
W86	WT ΔCGP123 Δ*ISCg12* Δ3000-3006	[222]
W115	WT ΔCGP123 Δ*ISCg12* Δ3102-3111 Δ3263-3301	[222]

Continued on the next page.

Table 8.1.: Continued.

Name	Relevant characteristics	Reference
W116	WT ΔCGP123 ΔIS*Cg*12 Δ3102-3111 Δ3263-3301 Δ3324-3345	[222]
W127	WT ΔCGP123 ΔIS*Cg*12 Δ2312-2322 Δ2621-2643 Δ2663-2686 Δ2755-2760 Δ3102-3111	[48]
W130	WT ΔCGP123 ΔIS*Cg*12 Δ2312-2322 Δ2621-2643 Δ2663-2686 Δ2755-2760 Δ3102-3111 Δ3324-3345	[222]
W134	WT ΔCGP123 ΔIS*Cg*12 Δ0414-0440 Δ0635-0646 Δ0704-0748 Δ0822-0845 Δ1018-1033 Δ1172-1213 Δ1291-1305 Δ1340-1353 Δ1370-1385	[222]
W121.7	WT ΔCGP123 ΔIS*Cg*12 Δ0414-0440 Δ0635-0646 Δ0704-0748 Δ0822-0845 Δ1018-1033 Δ1172-1213 Δ1291-1305 Δ1340-1352 Δ2312-2322 Δ2621-2643 Δ2663-2686 Δ2755-2760 Δ3072-3091 Δ3102-3111	[48]
W127.6	W127 Δ0635-0646 Δ0704-0748 Δ0822-0845 Δ1018-1033 Δ1172-1213 Δ1291-1305	[48]
rrnBC	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnB* Δ*rrnC*	[222]
rrnBF	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnB* Δ*rrnF*	[222]
rrnCF	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnC* Δ*rrnF*	[222]
rrnBCF	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnB* Δ*rrnC* Δ*rrnF*	[222]
rrnABCEF	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnA* Δ*rrnB* Δ*rrnC* Δ*rrnE* Δ*rrnF*	[222]
rrnBCDEF	WT ΔCGP123 ΔIS*Cg*12 Δ*rrnB* Δ*rrnC* Δ*rrnD* Δ*rrnE* Δ*rrnF*	[222]
zwf^{FBR} [#1]	*zwf* A243T point mutation	[214]
Plasmids		
pCGPhoD^{Bs}-GFP	*gfp* gene ligated to *phoD* Tat SP from *B. subtilis* cloned into pEKEx2 plasmid under control of P_{tac}, Km^R	[239]
pEKEx2-SP-cutinase [#2]	Cutinase gene from *F. solani pisi* ligated to Sec SP sequence [*amyE*, *epr*, *nprE*, *ypjP* or *ywmC*] from *B. subtilis*, cloned into pEKEx2 plasmid under control of P_{tac}, Km^R	[30, 249, 322]

[#1] Strain was kindly provided by Jan Marienhagen, IBG-1, Forschungszentrum Jülich. [#2] Plasmids were kindly provided by Sarah Jurischka, IBG-1, Forschungszentrum Jülich

8.2. Cultivation media

BHI medium was prepared from 37 g L^{-1} BHI broth and sterilized by autoclaving. BHIS medium was prepared likewise, but with additionally 91 g L^{-1} sorbitol. CgXII medium [56] for shake flasks and MBR cultivations was prepared using different solutions, listed in Table 8.2. For bioreactor cultivations, CgXII medium was modified as follows: urea, KH_2PO_4, and MOPS were omitted, but 3.027 g L^{-1} $NaH_2PO_4 \cdot 2\,H_2O$ was added, while concentrations of glucose, K_2HPO_4 and $MgSO_4 \cdot 7\,H_2O$ were changed to 10, 5.331 and 0.5 g L^{-1}, respectively. In case strains with plasmids were grown, 25 mg L^{-1} kanamycin was added to the medium to maintain selection pressure.

Table 8.2.: CgXII stock solutions used for shake flask and MBR cultivations.

Name	Addition	Components / preparation	Storage	Sterilization
Salts stock	800 mL L^{-1}	25 g L^{-1} $(NH_4)_2SO_4$, 6.25 g L^{-1} Urea, 1 g L^{-1} KH_2PO_4, 1 g L^{-1} K_2HPO_4, 52.5 g L^{-1} MOPS, to pH 7.0 with 4 M NaOH	RT	Autoclaving
Glucose	40 mL L^{-1}	500 g L^{-1} glucose	RT	Autoclaving
Mg stock	1 mL L^{-1}	250 g L^{-1} $MgSO_4 \cdot 7\,H_2O$	RT	Autoclaving
Ca stock	1 mL L^{-1}	13.25 g L^{-1} $CaCl_2 \cdot 2\,H_2O$	RT	Autoclaving
Trace elements	1 mL L^{-1}	10 g L^{-1} $FeSO_4 \cdot 7\,H_2O$, 10 g L^{-1} $MnSO_4 \cdot H_2O$, 1 g L^{-1} $ZnSO_4 \cdot 7\,H_2O$, 0.31 g L^{-1} $CuSO_4 \cdot 5\,H_2O$, 0.02 g L^{-1} $NiCl_2 \cdot 6\,H_2O$, a few drops of 1 M HCl	RT	Filtration
PCA	1 mL L^{-1}	30 g L^{-1}protocatechuic acid in 1 M NaOH	Aliquots at -20 °C	Filtration
Biotin	1 mL L^{-1}	0.2 g L^{-1} in 1 M NaOH	Aliquots at -20 °C	Filtration
Kanamycin	1 mL L^{-1}	30 mg L^{-1} kanamycin $\cdot H_2SO_4$	Aliquots at -20 °C	Filtration

8.3. Agar plates

Agar plates were prepared as the corresponding liquid media, but with additional 20 g L^{-1} agar agar (Difco). Additional compounds were added aseptically from concentrated stock solutions after sterilization by autoclaving and the media have been cooled to being lukewarm, prior to pouring the plates. Additional compounds used were kanamycin (25 mg L^{-1}) to establish selection pressure, IPTG (100 µM) and Tween 20 (1% v v^{-1}).

8.4. Further solutions

Stock solutions of isopropyl-β-D-1-thiogalactopyranoside (IPTG, 500 mM) and kanamycin (30 g L^{-1} kanamycin sulphate) were sterilized by filtration and stored in aliquots at 20 °C. Saline contained 9 g L^{-1} NaCl and was sterilized by autoclaving. phosphate buffered saline (PBS) with a pH of 7.4 ± 0.1 contained 8 g L^{-1} NaCl, 0.2 g L^{-1} KCl, 1.78 g L^{-1} $NaH_2PO_4 \cdot 2\ H_2O$, and 0.27 g L^{-1} K_2HPO_4 and was sterilized by autoclaving. Freezing solution was prepared by combining 1 volume PBS and 1 volume of glycerol solution (500 g L^{-1}). TG buffer was prepared from 1 mL L^{-1} of a 1 M tris-(2-hydroxyethyl)-ammonium chloride solution and 101.26 g L^{-1} glycerol and was adjusted to pH 7.5 using 1 M NaOH prior to sterilization by autoclaving.

8.5. Construction of cutinase secretion strains

8.5.1. Competent cells

For the generation of competent *C. glutamicum* cells, the respective strain was streaked out on BHIS agar plates and incubated for 24 - 48 h at 30 °C. A single colony was used to inoculate a 15 or 50 mL BHIS overnight shake flask culture. The next morning, 1.5 mL of this culture served as inoculum for a 100 mL BHIS shake flask culture. As soon as an optical density (OD_{600}) of 1.5 was reached, the cells were separated by centrifugation (4500 g, 20 min, 4 °C). Resulting supernatant was discarded and the pellet was resuspended in 6 mL ice cold TG buffer and again centrifuged (4500 g, 10 min, 4 °C). This step was repeated once. Afterwards, the cells were washed twice using 6 mL ice cold 100 g L^{-1} glycerol solution (4500 g, 10 min, 4 °C). Finally, the cells are resuspended in 1 mL ice cold 100 g L^{-1} glycerol solution and stored in 150 µL aliquots at -80 °C. All steps from culture harvest on were carried out on ice.

8.5.2. Transformation

Transformation of competent cells was done by electroporation [323]. Competent cells to be transformed were thawed on ice. In a 1.5 mL microtube, 50 µL of competent cells were mixed with 1 µL plasmid solution. As control, one transformation mix without plasmid was prepared and treated equally throughout the whole procedure. The transformation mixes were transferred into ice cold electroporation cuvettes and 800 µL ice cold 100 g L_{-1} glycerol solution was overlaid carefully. Immediately after electroporation in a Gene Pulser Xcell device (Biorad, Hercules/USA, 25 µF, 2500 V, 12.5 kV cm^{-1}, 200 Ω), the transformation mix is poured into 15 mL tubes with 4 mL BHIS medium, pre-

warmed to 46 °C, and incubated at 46 °C for 6 min with subsequent cooling on ice for a few minutes. After further 1.5 h of incubation at 30 °C with shaking (170 rpm), cells were centrifuged (3000 g, 10 min, 20 °C) and the resulting pellet was plated on selective BHIS-Km agar plates. For the transformation mixes with plasmid, single colonies needed to appear after two days of incubation at 30 °C, while no colonies should appear for the control transformations. To confirm cutinase secretion, transformed cell material was plated on indicator agar plates containing IPTG to induce cutinase secretion and Tween 20 which is hydrolyzed by cutinase, visible by halo formation.

8.6. Strain maintenance

8.6.1. Master cell bank

From each transformation, a single colony was spread on a new agar plate for generation of a sufficient amount of isogenic cell mass. After incubation for two days at 30 °C, the cell material was removed using sterile inoculation loops, resuspended into freezing solution and stored in 2-mL aliquots at -80 °C.

8.6.2. Working cell bank

A few hundred µL from a master cell bank (MCB) aliquot was used to inoculate an overnight shake flask culture with 50 mL CgXII medium containing 10% (v v^{-1}) BHI medium and, if appropriate, 25 mg L^{-1} kanamycin. After grown to saturation, one volume of the cell suspension was combined with 1 volume of sterile glycerol solution (500 g L^{-1}) and aliquots were stored at -80 °C as working cell bank. Strain *C. glutamicum* pCGPhoD^{Bs}-GFP was grown to saturation in 50 mL BHI medium with 25 mg L^{-1} kanamycin using a MCB aliquot [1] and 1 volume of resulting cell suspension was combined with 1 volume of glycerol solution (500 g L^{-1}) and stored in aliquots at -80 °C.

8.7. Cultivations

8.7.1. Shake flask cultivations

Unless otherwise specified, *C. glutamicum* shake flask cultivations were carried out at an orbital shaking frequency of 250 rpm in a shaking incubator ("Multitron", Infors HT, Einsbach/DE) with 25 mm shaking diameter. For a cultivation volume of 15, 50 or 100 mL, baffled shake flasks with a nominal

[1] kindly provided by Katja Schöler, IBG-1, Forschungszentrum Jülich

volume of 100, 500 or 1000 mL, respectively, were used. Flasks were equipped with a sampling port (except 100 mL flask) and covered with an aluminum cap. Cultivation temperature was 30 °C.

8.7.2. Microbioreactor cultivations

8.7.2.1. Pre-culturing

For MBR cultivation experiments with cutinase secretion *C. glutamicum* strains, two sequential pre-cultures in shake flasks were conducted. The first pre-culture was inoculated from a frozen working cell bank (WCB) aliquot and cultivated in 15 mL BHI medium. After approx. 6 h, 300 µL from the first pre-culture served as inoculum for the second pre-culture (50 mL CgXII medium, approx. 16 h incubation). Then, the pre-culture was centrifuged (4000 g, 5 min, 4 °C) and resuspended in sterile PBS and was used to inoculate the MBR culture to an initial OD_{600} of approx. 0.2. For MBR cultivation experiments with *C. glutamicum* pCGPhoD^{Bs}-GFP, a pre-culture of 50 mL BHI medium, inoculated with one WCB aliquot, was grown for approx. 5 h until an OD_{600} of 3 - 4 was reached. Then, the pre-culture was used at 5% (v v^{-1}) of the final MBR cultivation for inoculation.

8.7.2.2. Automated media preparation

Automated media preparation was conducted for media optimization studies. For the preparation of different medium variants according to the respective experimental plan, a LHS, ("Janus", Whatman/USA [36]) was used. Depending on the target concentrations of selected medium components, different volumes of stock solutions were pipetted into the wells of a cultivation MTP. Concentrations of stock solution were adjusted to realize pipetting volumes between 10 and 1000 µL. The cumulative volume of all solutions in each MTP well was 95% of the final volume, leaving 5% of inoculum volume to be added by the LHS. For each MBR media optimization experiment, at least 5 cultures with the reference medium composition were grown, serving for normalization of results for each experiment.

8.7.2.3. Main cultivations

Main cultivations were carried out in MBR devices with 48 well flower shaped microplates ("BioLector" and "Flowerplate", m2p-labs, Baesweiler/DE), integrated in different LHS ("Freedom Evo", Tecan, Männedorf/CH [206] or "Janus"). Cultivation conditions were as follows: CgXII medium with 20 g L^{-1} glucose, 800 µL per well, 1300 rpm at a shaking diameter of 3 mm, 30 °C. For cultivations with *C. glutamicum* pCGPhoD^{Bs}-GFP 10 g L^{-1} glucose, 1000 µL per well and 1200 rpm was applied. The integrated

BioLector devices are capable of quasi-continuous monitoring of GFP fluorescence (488/520 nm), biomass concentration via BS measurement [199], as well as pH and DO via fluorescence sensor spots (optodes) integrated at the bottom of each cultivation well. Flowerplates were covered with as sterile barrier acting self-adhesive sealing films that reduce evaporation while ensuring sufficient gas exchange and robotic access to the culture (F-GP-10 and F-GPRS48-10, m2p-labs, Baesweiler/DE).

8.7.2.4. Automated harvesting and sampling procedures

Automated harvest and sampling procedures were based on pre-defined triggers, which utilize the online monitored MBR process data [158]. Triggers are defined using a supervising software ("RoboLector Agent", m2p-labs, Baesweiler/DE), which activates the integrated LHS to act on the individual MBR cultures by writing pipetting lists ("handshake files"). The LHS devices run a program containing all necessary steps (pipetting, washing, plate movement, etc.) in a loop. A new iteration of this loop is started as soon as a new handshake file is written by the supervising software. After completion of all LHS steps, the handshake file is deleted as last step in the loop, which is the signal of the supervising software to continue with the MBR cultivation.

Automated harvest of MBR cultures was based on the dynamics of the online monitored DO signal to detect the end of the exponential growth phase by a sharp increase in the DO signal. This detection was implemented by two sequential conditions that had to be fulfilled. First, the DO signal had to fall below 50% air saturation (a.s.). After that, the DO had to rise above 80% a.s., which is the trigger condition to cause the pipetting of cell suspension out of the BioLector device. With that, the supervising software ordered the opening of the cover of the BioLector's incubation chamber and the LHS to aspirate the cell suspension 1 mm above the bottom of the cultivation MTP placed in the BioLector device. Per culture, 700 µL cell suspension was removed and pipetted into a 96 deep well plate ("Riplate", Ritter, Schwabmünchen/DE). Afterwards, the cover of the BioLector was closed and sent to the next measurement cycle. In parallel, the 96 deep well plates containing the cell suspensions was placed into a LHS accessible centrifuge ("Ixion", SIAS, Hombrechtikon/CH) to pellet the cells for 10 min at 2000 g. Resulting supernatant was afterwards transferred to another 96 deep well plate, covered with a self-adhesive aluminum foil ("SILVERseal", Sigma Aldrich, Steinheim/DE) and cooled to 4 °C. After the completion of the MBR cultivation experiment, supernatants were stored at -20 °C until analysis.

Automated repeated low-volume sampling from MBR cultures was based on the online monitored BS signal, which served as trigger signal to sample repeatedly (six times) 20 µL of cell suspension from each well. More specifically, for each of the six 20 µL samples to be removed a specific backscatter

signal was predefined. Expected amplitude of BS signals during growth was determined in preliminary reference cultivations for the specific BioLector devices using *C. glutamicum* WT in CgXII medium containing 20 g L^{-1} glucose. The aspiration of cell suspension was carried out as described above and was pipetted into standard 96-well MTPs (Greiner, Frickenhausen/DE), which were covered with a self-adhesive aluminium foil ("SILVERseal", Greiner) and an additional plastic foil ("X Pierce", Excel scientific, Victorville, CA/USA). The sample receiving plates were placed in cooling carriers controlled at 4 °C on the worktable of the LHS. Directly after that, 230 µL PBS was added to the taken samples, resulting in a 1:12.5 dilution of the samples. After the cultivation run, the pre-diluted samples were centrifuged as described above. Afterwards, 200 µL of the resulting supernatants were transferred to new 96-well MTP, which were stored at -20 °C until analysis.

8.7.3. Bioreactor cultivations

Cultivations were carried out in 2 L lab-scale stirred tank reactors (DASGIP, Jülich/DE), equipped with two six-bladed Rushton turbines, pH-electrodes (Mettler-Toldeo, Urdorf/CH) and DO electrodes (Hamilton, Bonaduz/CH). DO was maintained above 30% a.s. by adjustment of stirrer speed between 400 and 1200 rpm with constant gassing of air at 60 sL h^{-1}. The pH value was maintained at 7.0 using 18% w w-1 NH_4OH. Exponential feeding profiles were calculated to control the culture at a desired specific growth rate (μ_{Set}), according to (8.1). Feed medium was 400 g L^{-1} glucose and initial reaction volume ($V_{R,0}$) was 0.8 L. Cultivation temperature was set to 30 °C. Process control was conducted by DASGIP Control 4.0 software (DASGIP, Jülich/DE). Bioreactors were inoculated with 5 mL WCB aliquots to an OD_{600} of approx. 0.05, resulting in a batch growth for approx. 18 h on 10 g L^{-1} glucose. After complete consumption of the batch glucose, indicated by a sharp increase in DO signal, fed-batch phase was initiated by starting the exponential feeding profile and IPTG was added to induce recombinant cutinase secretion. IPTG concentration was varied between 0, 50, 250 and 1000 µM as indicated in results section. In the fed-batch phase, samples were drawn every hour to obtain an expression profile resulting in eight sampling points

Reference batch cultivations have been carried out likewise, with the exception that after the depletion of the initial batch glucose indicated by DO increase, a second batch phase was initiated by adding glucose and IPTG to a final concentration of 40 g L^{-1} and 1000 µM, respectively.

$$F(t) = \left(\frac{\mu_{Set}}{Y_{X/S}} + m_S\right) \cdot \frac{c_{S,0} V_{R,0} Y_{X/S}}{c_{SF}} \cdot e^{\mu_{Set}(t - t_{Feed})} \tag{8.1}$$

8.7.3.1. Determination of culture performance indicators

From the induced fed-batch phases with controlled growth rate, three PIs are derived from experimentally data obtained from regular sampling: (1) Observed specific growth rate μ_{Exp}, since the adjusted growth rate μ_{Set}, cf. (8.1), is a parameter defining the exponentially increasing feed rate and does not necessarily match the observed growth rate μ_{Exp}. (2) Biomass specific cutinase yield $Y_{P/X}$ and (3) biomass specific cutinase productivity q_P. The first two PIs are calculated using regression methods [100], covering the whole time course of the induced cutinase secretion. In the case of μ_{Exp}, values of m_X were ln-transformed for ordinary least squares (OLS) regression, cf. (1.12). Values of m_X are obtained from cell dry weight measurements, according to (8.2). For estimation of $Y_{P/X}$, a linear orthogonal total least squares (TLS) regression was applied [324], cf. (1.13), and values for m_P are calculated from taken samples according to (8.3). Since cutinase activity as product concentration $c_{P,Sup}$ is determined from the supernatant of the cultivation volume, it is necessary to correct the measurement value for the volume fraction that is occupied by the cell mass, which is done by the incorporation of the correlation factor $k_{c_X \rightarrow c_{BioVol}}$, see section 8.7.3.2. Values for q_P are calculated according to (1.14). All cultivations were carried out in biological duplicates and μ_{Exp}, $Y_{P/X}$ as well as q_P are reported as mean with min/max values as errors.

$$m_X = c_X \cdot (V_{R,0} + V_{Feed} + V_{Base} - V_{Sample}) \tag{8.2}$$

$$m_P = c_P \cdot (V_{R,0} + V_{Feed} + V_{Base} - V_{Sample}) \cdot (1 - c_X \cdot k_{c_X \rightarrow c_{BioVol}}) \tag{8.3}$$

8.7.3.2. Biomass / biovolume conversion factor

To account for the volume fraction of the cultivation suspension which is occupied by the cells, an empirical conversion factor $k_{c_X \rightarrow c_{BioVol}}$ was determined to deduce the biovolume from cell dry weight. Biovolume was measured in technical triplicates for each sample using a MultiSizer 3 (Beckman Coulter, Krefeld/DE) particle counter [325]. Samples were adjusted to an optical density (OD_{600}) of approx. 0.1, and 500 µL were diluted in 9.5 mL CASYton buffer (Schärfe Systems, Reutlingen/DE). All particle counts between 0.9 and 4 µm were considered as cells. The conversion factor $k_{c_X \rightarrow c_{BioVol}}$ was calculated as mean ratio of biovolume per cell dry weight from the samples.

8.8. Analytics

8.8.1. Optical density

Optical density was measured at 600 nm (OD_{600}) against PBS. Samples were diluted with PBS to a range of approximately 0.5 - 0.05 OD_{600}.

8.8.2. Cell dry weight

Cell dry weight was determined gravimetrically. Samples of 1 mL were pipetted into pre-dried (80 °C, min. 48 h) and pre-weighted 2 mL tubes and centrifuged at max. speed for 10 min at using benchtop centrifuges ("BIOFUGE pico", Heraeus, Hanau/DE). Resulting supernatants were collected and stored at -20 °C. Cell pellets were washed once with PBS, centrifuged again and the supernatant was discarded. The pellets were dried as before with subsequent weighting to calculate the cell dry weight c_X [g_X L^{-1}].

8.8.3. Backscatter calibration

To convert backscatter reading from the BioLector MBR cultivations into cell dry weight, calibration series were performed for different *C. glutamicum* strains. A sufficient amount of cell suspension was produced from 100 mL shake flask cultivations in CgXII grown to saturation and resulting biomass was washed and resuspended to different cell concentrations in CgXII medium without glucose. For each dilution step, cell dry weight from 2 or 10 mL of cell suspension was determined in three analytical replicates, and backscatter was determined in three technical replicates by filling three times 800 µL of each dilution step into a Flowerplate, which was measured in the BioLector devices using the cultivation protocol. Afterwards, values for backscatter and cell dry weight were correlated by linear regression.

8.8.4. Glucose assay

Glucose from cultivation supernatants were quantified by enzymatic assay in 96 well MTPs, which relies on the formation of glucose-6-phosphate from glucose by hexokinase (HK) with subsequent formation of NADH by glucose-6-phosphate dehydrogenase (G6P-DH). Two dilutions from each sample were prepared using 50 mM phosphate buffer, pH 8, and 20 µL of each dilution (1:12.5 and 1:50) was transferred to 96 well MTPs. Next, three identical glucose dilution series for calibration were prepared from a glucose aliquot (20 g L^{-1}, stored at -20 °C) using 50 mM phosphate buffer, pH 8, as

diluent. HK activity was verified to be not inhibited by the applied phosphate content in the dilution buffer. Into each measuring plate, 20 µL of each dilution series was copied to determine a glucose calibration for each measuring plate in technical triplicates. Afterwards, 180 µL of reaction mix was added and after incubation of 45 min at RT, the absorption at 340 nm was recorded. Reaction mix contained 4.4 mL L^{-1} enzyme mix (containing 340 U mL^{-1} HK and 170 U mL^{-1} G6P-DH, Roche Diagnostics, Mannheim/DE), 4.4 mM $MgSO_4$, 1 g L^{-1} NAD · 2 H_2O and 0.68 g L^{-1} ATP · 2 Na in 100 mM Tris maleate buffer, pH 6.8. Finally, the glucose concentration in the sample was calculated, according to equation (8.4). Only measurements in the linear range of the calibration series were considered.

$$c_{Glc} = \frac{A_{340} - b_{Glc}}{a_{Glc}} \cdot f_D \tag{8.4}$$

8.8.5. Cutinase activity assay

Cutinase activity from cultivation supernatants was determined using *p*-nitrophenylpalmitate (pNPP) as substrate analogue with the anion of pNP as spectrophotometrically detectable reaction product [326]. Samples were diluted appropriately using PBS and 20 µL of the diluted samples were copied into three 96 well MTPs, serving as analytical triplicates for cutinase activity measurements. In each MTP, three replicates of a pNP dilution series were pipetted (20 µL each dilution step) to convert absorption readings a 410 nm into micromoles of pNP formed during pNPP hydrolysis by cutinase. Enzymatic reaction was started by rapid addition of 200 µL reaction mix with subsequent transfer into a MTP reader pre-warmed to 37 °C, and absorption at 410 nm was recorded in 25 sec intervals. The resulting linear increase over time ($\Delta A_{410\ nm}$, R^2 > 0.99) was used for cutinase activity determination, according to equation (8.5). Reaction mix consisted of 1 volume substrate solution (30 mg pNPP in 10 mL isopropanol) and 10 volumes reaction buffer (2.3 g L^{-1} Na-desoxycholate, 1.1 g L^{-1} gum arabic in 55 mM K-P_i buffer, pH 8).

$$EA = \Delta A_{410} \cdot a_{pNP} \cdot f_D \tag{8.5}$$

8.8.6. Protein quantification by Bradford assay

From cultivation supernatants, protein content was determined by colorimetric Bradford assay. From each sample, 30 µL were pipetted into 96 well MTPs, as well as 30 µL from a dilution series of bovine serum albumin (BSA) as calibration standard (0 – 200 mg L^{-1}) in triplicates. After addition of 200 µL ready-to-use Bradford reagent (Sigma, Taufkirchen/DE) and an incubation time of 5 min, absorption at 595 nm was measured and protein concentration was calculated according to equation (8.6).

$$c_{Protein} = A_{595} \cdot a_{Bradford} + b_{Bradford} \tag{8.6}$$

8.8.7. GFP specific fluorescence of cultivation supernatants

To distinguish between GFP specific fluorescence of GFP secreting *C. glutamicum* cell suspensions and cell free supernatants, 800 µL of each cultivation suspension were transferred to 1.5 mL reaction tubes to obtain supernatants by centrifugation, as described in section 8.8.2. Then, 200 µL of resulting supernatants were transferred to black MTPs with transparent bottom ("µclear", Greiner, Frickenhausen/DE) for GFP specific fluorescence measurement.

8.8.8. SDS-Page

Visualization of the amount of secreted GFP was performed by denaturing SDS-Page using 12% Bis-Tris gels according to the supplier's instructions ("TruPage", Life Technologies, Darmstadt/DE). For staining of protein bands in the gel, a microwave-based protocol was used [327].

8.8.9. Proteome analysis

Cell pellets for untargeted proteome analyses were generated from at replicate cultivations as indicated. Cells from 1 mL samples were pelleted using a bench top centrifuge ("himac", Hitachi Koki, Tokio/JP, 1 min, 13000 rpm, 4 °C), washed once with 1 mL ice cold PBS and stored at -20 °C until analysis. Sample preparation and data processing was essentially carried out as described elsewhere [11]. Briefly, cell pellets were disrupted mechanically and tryptically digested after addition of ^{15}N-labeled cell extract as internal standard. Quantification was performed on a LC-MS/MS system (Agilent Infinity 1260 HPLC with Q-ToF 6600 mass spectrometer, AB Sciex, Darmstadt/DE). Peak integration, and protein identification and quantification were done using PeakView and ProteinPilot software packages, respectively (both AB Sciex). Differences in protein levels (ration of means) were calculated using MarkerView (AB Sciex) and p-values were calculated using a Welch test. Resulting p-values were adjusted for multiple testing to control the FDR [328, 329]. Briefly, fold-changes from all m protein levels were listed by their corresponding p-values in ascending order and the resulting ranked-ordered p-values were denoted as p_0 to p_m. Then, the p-value of highest rank, p_k, was determined for which the condition $p_K \leq \frac{k}{m} \cdot \alpha, 1 \leq k \leq m$ holds, with α denoting the desired FDR to be controlled. Finally, for all fold-changes up to rank k, H_0 is rejected. Additionally, protein levels need to show a at least two-fold regulation to be considered as significantly different.

8.8.10. Transcriptome analysis

Samples of 1 mL were pelleted using a bench top centrifuge (30 sec, 13000 rpm, 4 °C) and, after rapidly discarding the supernatant, snap-frozen in liquid nitrogen. Cell pellets were stored at -80 °C until preparation of RNA using a kit system according to the manufacturer's instructions ("RNeasy", Qiagen, Hilden/DE). Fluorescently labeled cDNA was produced from isolated RNA using reverse transcriptase ("Superscript III", Life Technologies, Darmstadt/DE). Purified cDNA samples were pooled together in the form of reference against sample. The prepared two-color samples were hybridized at 65 °C while rotating for 17 h using Agilent's (Santa Clara, CA/USA) gene expression Hybridization Kit, hybridization oven and hybridization chamber with subsequent washing using Agilent's Wash Buffer Kit according to the manufacturer's instructions. Hybridized microarrays were read for fluorescence emission at 532 and 635 nm (cy3 and cy5 dye, respectively) using GenePix hard- and software ("4000B laser" and "Pro 7.0", Molecular Devices, Sunnyvale, CA/USA). Obtained raw images were saved as TIFF-files for quantitative analyses using GenePix image analysis software with GPR-files as output. These were further processed for background correction, ratio calculation and normalization with the help of Bioconductor R-packages `limma` and `marray` [330]. Fold-changes of mRNA levels are reported as ratio of median values.

References

[1] Kinoshita S, Udaka S, Shimono M: **Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms**. *Journal of General and Applied Microbiology* 1957, 3:193–205.

[2] Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L, Goesmann A, Hartmann M, Huthmacher K, Krämer R, Linke B, McHardy AC, Meyer F, Möckel B, Pfefferle W, Pühler A, Rey DA, Rückert C, Rupp O, Sahm H, Wendisch VF, Wiegräbe I, Tauch A: **The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of l-aspartate-derived amino acids and vitamins**. *Journal of Biotechnology* 2003, 104:5–25.

[3] Eggeling L, Bott M, editors: *Handbook of Corynebacterium glutamicum*. Boca Raton: Taylor & Francis, 2005.

[4] Wendisch VF: **Microbial production of amino acids and derived chemicals: Synthetic biology approaches to strain development**. *Current Opinion in Biotechnology* 2014, 30:51–58.

[5] Wendisch VF, Jorge JMP, Pérez-García F, Sgobba E: **Updates on industrial production of amino acids using *Corynebacterium glutamicum***. *World Journal of Microbiology & Biotechnology* 2016, 32:105.

[6] Knoll A, Bartsch S, Husemann B, Engel P, Schroer K, Ribeiro B, Stöckmann C, Seletzky J, Büchs J: **High cell density cultivation of recombinant yeasts and bacteria under non-pressurized and pressurized conditions in stirred tank bioreactors**. *Journal of Biotechnology* 2007, 132:167–179.

[7] Yim SS, An SJ, Kang M, Lee J, Jeong KJ: **Isolation of fully synthetic promoters for high-level gene expression in *Corynebacterium glutamicum***. *Biotechnology and Bioengineering* 2013, 110:2959–2969.

[8] Yim SS, Choi JW, Lee RJ, Lee YJ, Lee SH, Kim SY, Jeong KJ: **Development of a new platform for secretory production of recombinant proteins in *Corynebacterium glutamicum***. *Biotechnology and Bioengineering* 2015, .

[9] Käß F, Hariskos I, Michel A, Brandt HJ, Spann R, Junne S, Wiechert W, Neubauer P, Oldiges M: **Assessment of robustness against dissolved oxygen/substrate oscillations for *C. glutamicum* DM1933 in two-compartment bioreactor**. *Bioprocess and Biosystems Engineering* 2014, 37:1151–1162.

[10] Limberg MH, Pooth V, Wiechert W, Oldiges M: **Plug flow versus stirred tank reactor flow characteristics in two-compartment scale-down bioreactor: Setup-specific influence on the metabolic phenotype and bioprocess performance of *Corynebacterium glutamicum***. *Engineering in Life Sciences* 2016, 16:610–619.

[11] Limberg MH, Schulte J, Aryani T, Mahr R, Baumgart M, Bott M, Wiechert W, Oldiges M: **Metabolic profile of 1,5-diaminopentane producing *Corynebacterium glutamicum* under scale-down conditions: Blueprint for robustness to bioreactor inhomogeneities**. *Biotechnology and Bioengineering* 2016, 114:560–575.

[12] Liu L, Yang H, Shin Hd, Chen RR, Li J, Du G, Chen J: **How to achieve high-level expression of microbial enzymes: Strategies and perspectives**. *Bioengineered* 2013, 4:212–223.

[13] Liu L, Yang H, Shin Hd, Li J, Du G, Chen J: **Recent advances in recombinant protein expression by *Corynebacterium*, *Brevibacterium*, and *Streptomyces*: From transcription and translation regulation to secretion pathway selection**. *Applied Microbiology and Biotechnology* 2013, 97:9597–9608.

[14] Ravasi P, Braia M, Eberhardt F, Elena C, Cerminati S, Peirú S, Castelli ME, Menzella HG: **High-level production of *Bacillus cereus* phospholipase C in *Corynebacterium glutamicum***. *Journal of Biotechnology* 2015, 216:142–148.

[15] Watanabe K, Teramoto H, Suzuki N, Inui M, Yukawa H: **Influence of SigB inactivation on *Corynebacterium glutamicum* protein secretion**. *Applied Microbiology and Biotechnology* 2013, 97:4917–4926.

[16] Freudl R: **Leaving home ain't easy: Protein export systems in Gram-positive bacteria**. *Research in Microbiology* 2013, 164:664–674.

[17] Freudl R: **Beyond amino acids: Use of the *Corynebacterium glutamicum* cell factory for the secretion of heterologous proteins**. *Journal of Biotechnology* 2017, 258:101–109.

[18] Lee HC, Bernstein HD: **The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal**. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98:3471–3476.

[19] Pugsley AP: **The complete general secretory pathway in Gram-negative bacteria.** *Microbiology Reviews* 1993, 57:50–108.

[20] Tsirigotaki A, de Geyter J, Šoštaric N, Economou A, Karamanou S: **Protein export through the bacterial Sec pathway**. *Nature Reviews: Microbiology* 2017, 15:21–36.

[21] Tjalsma H, Bolhuis A, Jongbloed JDH, Bron S, van Dijl JM: **Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome**. *Microbiology and Molecular Biology Reviews* 2000, 64:515–547.

[22] Tjalsma H, Antelmann H, Jongbloed, Jan D H, Braun PG, Darmon E, Dorenbos R, Dubois JYF, Westers H, Zanen G, Quax WJ, Kuipers OP, Bron S, Hecker M, van Dijl JM: **Proteomics of protein secretion by *Bacillus subtilis*: Separating the secrets of the secretome**. *Microbiology and Molecular Biology Reviews* 2004, 68:207–233.

[23] Beckwith J: **The Sec-dependent pathway**. *Research in Microbiology* 2013, 164:497–504.

[24] Huber D, Rajagopalan N, Preissler S, Rocco MA, Merz F, Kramer G, Bukau B: **SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria**. *Molecular Cell* 2011, 41:343–353.

[25] Huber D, Jamshad M, Hanmer R, Schibich D, Döring K, Marcomini I, Kramer G, Bukau B: **SecA cotranslationally interacts with nascent substrate proteins in vivo**. *Journal of Bacteriology* 2017, 199.

[26] Pool MR: **Signal recognition particles in chloroplasts, bacteria, yeast and mammals (Review)**. *Molecular Membrane Biology* 2005, 22:3–15.

[27] Halic M, Becker T, Pool MR, Spahn CM, Grassucci RA, Frank J, Beckmann R: **Structure of the signal recognition particle interacting with the elongation-arrested ribosome**. *Nature* 2004, 808–814.

[28] Kempf G, Wild K, Sinning I: **Structure of the complete bacterial SRP Alu domain**. *Nucleic Acids Research* 2014, 42:12284–12294.

[29] Zalucki YM, Power PM, Jennings MP: **Selection for efficient translation initiation biases codon usage at second amino acid position in secretory proteins**. *Nucleic Acids Research* 2007, 35:5748–5754.

[30] Brockmeier U, Caspers M, Freudl R, Jockwer A, Noll T, Eggert T: **Systematic screening of all signal peptides from *Bacillus subtilis*: A powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria**. *Journal of Molecular Biology* 2006, 362:393–402.

[31] Caspers M, Brockmeier U, Degering C, Eggert T, Freudl R: **Improvement of Sec-dependent secretion of a heterologous model protein in *Bacillus subtilis* by saturation mutagenesis of the N-domain of the AmyE signal peptide**. *Applied Microbiology and Biotechnology* 2010, 86:1877–1885.

[32] Zhang W, Yang M, Yang Y, Zhan J, Zhou Y, Zhao X: **Optimal secretion of alkali-tolerant xylanase in *Bacillus subtilis* by signal peptide screening**. *Applied Microbiology and Biotechnology* 2016, 100:8745–8756.

[33] Mathiesen G, Sveen A, Brurberg MB, Fredriksen L, Axelsson L, Eijsink VG: **Genome-wide analysis of signal peptide functionality in *Lactobacillus plantarum* WCFS1**. *BMC Genomics* 2009, 10:425.

[34] Watanabe K, Tsuchida Y, Okibe N, Teramoto H, Suzuki N, Inui M, Yukawa H: **Scanning the *Corynebacterium glutamicum* R genome for high-efficiency secretion signal sequences**. *Microbiology (Reading, England)* 2009, 155:741–750.

[35] Degering C, Eggert T, Puls M, Bongaerts J, Evers S, Maurer KH, Jäger KE: **Optimization of protease secretion in *Bacillus subtilis* and *Bacillus licheniformis* by screening of homologous and heterologous signal peptides**. *Applied and Environmental Microbiology* 2010, 76:6370–6376.

[36] Rohe P, Venkanna D, Kleine B, Freudl R, Oldiges M: **An automated workflow for enhancing microbial bioprocess optimization on a novel microbioreactor platform**. *Microbial Cell Factories* 2012, 11:144.

[37] Hemmerich J, Rohe P, Kleine B, Jurischka S, Wiechert W, Freudl R, Oldiges M: **Use of a Sec signal peptide library from *Bacillus subtilis* for the optimization of cutinase secretion in *Corynebacterium glutamicum***. *Microbial Cell Factories* 2016, 15:208.

[38] Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: Discriminating signal peptides from transmembrane regions**. *Nature Methods* 2011, 8:785–786.

[39] Xavier JC, Patil KR, Rocha I: **Systems biology perspectives on minimal and simpler cells**. *Microbiology and Molecular Biology Reviews* 2014, 78:487–509.

[40] Martinez-Garcia E, de Lorenzo V: **The quest for the minimal bacterial genome**. *Current Opinion in Biotechnology* 2016, 42:216–224.

[41] Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, Smith HO, Venter JC: **Essential genes of a minimal bacterium**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103:425–430.

[42] Forster AC, Church GM: **Towards synthesis of a minimal cell**. *Molecular Systems Biology* 2006, 2:45.

[43] Hohmann HP, van Dijl JM, Krishnappa L, Pragai Z: **Host Organisms: *Bacillus subtilis***. In *Industrial biotechnology: Microorganisms*, edited by Wittmann C, Liao JC, Advanced biotechnology. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2017, 221–298.

[44] Juhas M, Reuß DR, Zhu B, Commichau FM: ***Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering**. *Microbiology* 2014, 160:2341–2351.

[45] Choe D, Cho S, Kim SC, Cho BK: **Minimal genome: Worthwhile or worthless efforts toward being smaller?** *Biotechnology Journal* 2016, 11:199–211.

[46] Reuß DR, Altenbuchner J, Mäder U, Rath H, Ischebeck T, Sappa PK, Thürmer A, Guérin C, Nicolas P, Steil L, Zhu B, Feussner I, Klumpp S, Daniel R, Commichau FM, Völker U, Stülke J: **Large-scale reduction of the *Bacillus subtilis* genome: Consequences for the transcriptional network, resource allocation, and metabolism**. *Genome Research* 2017, 27:289–299.

[47] Tanaka K, Henry CS, Zinner JF, Jolivet E, Cohoon MP, Xia F, Bidnenko V, Ehrlich SD, Stevens RL, Noirot P: **Building the repertoire of dispensable chromosome regions in *Bacillus subtilis* entails major refinement of cognate large-scale metabolic model**. *Nucleic Acids Research* 2013, 41:687–699.

[48] Baumgart M, Unthan S, Kloß R, Radek A, Polen T, Tenhaef N, Müller MF, Küberl A, Siebert D, Brühl N, Marin K, Hans S, Krämer R, Bott M, Kalinowski J, Wiechert W, Seibold G, Frunzke J, Rückert C, Wendisch VF, Noack S: ***Corynebacterium glutamicum* chassis C1*: Building and testing a novel platform host for synthetic biology and industrial biotechnology**. *ACS Synthetic Biology* 2017, 7:132–144.

[49] Pósfai G, Plunkett G, Fehér T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M, Burland V, Harcum SW, Blattner FR: **Emergent properties of reduced-genome *Escherichia coli***. *Science* 2006, 312:1044–1046.

[50] Martinez-Garcia E, Jatsenko T, Kivisaar M, de Lorenzo V: **Freeing *Pseudomonas putida* KT2440 of its proviral load strengthens endurance to environmental stresses**. *Environmental Microbiology* 2015, 17:76–90.

[51] Lieder S, Nikel PI, de Lorenzo V, Takors R: **Genome reduction boosts heterologous gene expression in *Pseudomonas putida***. *Microbial Cell Factories* 2015, 14:23.

[52] Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H: **Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107:2646–2651.

[53] Baumgart M, Unthan S, Rückert C, Sivalingam J, Grünberger A, Kalinowski J, Bott M, Noack S, Frunzke J: **Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology**. *Applied and Environmental Microbiology* 2013, 79:6006–6015.

[54] Choi JW, Yim SS, Kim MJ, Jeong KJ: **Enhanced production of recombinant proteins with *Corynebacterium glutamicum* by deletion of insertion sequences (IS elements)**. *Microbial Cell Factories* 2015, 14:207.

[55] Unthan S, Baumgart M, Radek A, Herbst M, Siebert D, Brühl N, Bartsch A, Bott M, Wiechert W, Marin K, Hans S, Krämer R, Seibold G, Frunzke J, Kalinowski J, Rückert C, Wendisch VF, Noack S: **Chassis organism from *Corynebacterium glutamicum* - a top-down approach to identify and delete irrelevant gene clusters**. *Biotechnology Journal* 2015, 10:290–301.

[56] Keilhauer C, Eggeling L, Sahm H: **Isoleucine synthesis in *Corynebacterium glutamicum*: Molecular analysis of the *ilvB-ilvN-ilvC* operon**. *Journal of Bacteriology* 1993, 175:5595–5603.

[57] Adrio JL, Demain AL: **Microbial enzymes: Tools for biotechnological processes**. *Biomolecules* 2014, 4:117–139.

[58] Li S, Yang X, Yang S, Zhu M, Wang X: **Technology prospecting on enzymes: Application, marketing and engineering**. *Computational and Structural Biotechnology Journal* 2012, 2:e201209017.

[59] Choi S, Song CW, Shin JH, Lee SY: **Biorefineries for the production of top building block chemicals and their derivatives**. *Metabolic Engineering* 2015, 28:223–239.

[60] Choi JM, Han SS, Kim HS: **Industrial applications of enzyme biocatalysis: Current status and future aspects**. *Biotechnology Advances* 2015, 33:1443–1454.

[61] Erickson B, Nelson, Winters P: **Perspective on opportunities in industrial biotechnology in renewable chemicals**. *Biotechnology Journal* 2012, 7:176–185.

[62] Lee SY, Mattanovich D, Villaverde A: **Systems metabolic engineering, industrial biotechnology and microbial cell factories**. *Microbial Cell Factories* 2012, 11:156.

[63] Zheng X, Xing XH, Zhang C: **Targeted mutagenesis: A sniper-like diversity generator in microbial engineering**. *Synthetic and Systems Biotechnology* 2017, 2:75–86.

[64] de Lorenzo V, Schmidt M: **Biological standards for the knowledge-based BioEconomy: What is at stake**. *New Biotechnology* 2017, 40:170–180.

[65] Schallmey M, Frunzke J, Eggeling L, Marienhagen J: **Looking for the pick of the bunch: High-throughput screening of producing microorganisms with biosensors**. *Current Opinion in Biotechnology* 2014, 26:148–154.

[66] Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH: **A CRISPR-Cas9 system for genetic engineering of filamentous fungi**. *PloS one* 2015, 10:e0133085.

[67] Cleto S, Jensen JV, Wendisch VF, Lu TK: ***Corynebacterium glutamicum* metabolic engineering with CRISPR interference (CRISPRi)**. *ACS Synthetic Biology* 2016, 5:375–385.

[68] Stovicek V, Borodina I, Forster J: **CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains**. *Metabolic Engineering Communications* 2015, 2:13–22.

[69] AMFEP: **List of commercial enzymes (Update April 2014)**, 2014. URL `www.amfep.org`.

[70] Carneiro S, Ferreira EC, Rocha I: **Metabolic responses to recombinant bioprocesses in *Escherichia coli***. *Journal of Biotechnology* 2013, 164:396–408.

[71] Spadiut O, Capone S, Krainer F, Glieder A, Herwig C: **Microbials for the production of monoclonal antibodies and antibody fragments**. *Trends in Biotechnology* 2014, 32:54–60.

[72] Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, Lin-Cereghino J, Kovar K, Cregg JM, Glieder A: **Promoter library designed for fine-tuned gene expression in *Pichia pastoris***. *Nucleic Acids Research* 2008, 36:e76.

[73] Zhang B, Zhou N, Liu YM, Liu C, Lou CB, Jiang CY, Liu SJ: **Ribosome binding site libraries and pathway modules for shikimic acid synthesis with *Corynebacterium glutamicum***. *Microbial Cell Factories* 2015, 14:71.

[74] Gräslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, dhe Paganon S, Park Hw, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, Prilusky J, Sussman JL, Stevens RC, Lesley SA, Wilson IA, Joachimiak A, Collart F, Dementieva I, Donnelly MI, Eschenfeldt WH, Kim Y, Stols L, Wu R, Zhou M, Burley SK, Emtage JS, Sauder JM, Thompson D, Bain K, Luz J, Gheyi T, Zhang F, Atwell S, Almo SC, Bonanno JB, Fiser A, Swaminathan S, Studier FW, Chance MR, Sali A, Acton TB, Xiao R, Zhao L, Ma LC, Hunt JF, Tong L, Cunningham K, Inouye M, Anderson S, Janjua H, Shastry R, Ho CK, Wang D, Wang H, Jiang M, Montelione GT, Stuart DI, Owens RJ, Daenke S, Schütz A, Heinemann U, Yokoyama S, Büssow K, Gunsalus KC: **Protein production and purification**. *Nature Methods* 2008, 5:135–146.

[75] Braun P, LaBaer J: **High throughput protein production for functional proteomics**. *Trends in Biotechnology* 2003, 21:383–388.

[76] Kleman GL, Strohl WR: **Developments in high cell density and high productivity microbial fermentation**. *Current Opinion in Biotechnology* 1994, 5:180–186.

[77] Öztürk S, Çalık P, Özdamar TH: **Fed-batch biomolecule production by *Bacillus subtilis*: A state of the art review**. *Trends in Biotechnology* 2016, 34:329–345.

[78] Neubauer P, Junne S: **Scale-down simulators for metabolic analysis of large-scale bioprocesses**. *Current Opinion in Biotechnology* 2010, 21:114–121.

[79] Takors R: **Scale-up of microbial processes: Impacts, tools and open questions**. *Journal of Biotechnology* 2012, 160:3–9.

[80] Jain J, Sapna, Singh B: **Characteristics and biotechnological applications of bacterial phytases**. *Process Biochemistry* 2016, 51:159–169.

[81] Juturu V, Wu JC: **Microbial cellulases: Engineering, production and applications**. *Renewable and Sustainable Energy Reviews* 2014, 33:188–203.

[82] Yamada H, Kobayashi M: **Nitrile hydratase and its application to industrial production of acrylamide**. *Bioscience, Biotechnology, and Biochemistry* 1996, 60:1391–1400.

[83] Smidt H, Fischer A, Fischer P, Schmid R: **Preparation of optically pure chiral amines by lipase-catalyzed enantioselective hydrolysis of N-acyl-amines**. *Biotechnology Techniques* 1996, 10.

[84] Colomera A, Kuilderd H: **Biotechnological washing of denim jeans**. In *Denim: Manufacture, finishing and applications*, edited by Paul R, Woodhead Publishing series in textiles. Amsterdam and Cambridge: Elsevier and Woodhead Publ, 2015, 357–403.

[85] Ferralli P, Egan JD, Erickson FL: **Making Taq DNA polymerase in the undergraduate biology laboratory**. *BIOS* 2007, 78:69–74.

[86] White JS: **Sucrose, HFCS, and fructose: History, manufacture, composition, applications, and production**. In *Fructose, high fructose corn syrup, sucrose and health*, edited by Rippe JM, Nutrition and Health. New York: Humana Press, 2014, 13–33.

[87] Mahmoodi M, Najafpour GD, Mohammadi M: **Production of pectinases for quality apple juice through fermentation of orange pomace**. *Journal of Food Science and Technology* 2017, 54:4123–4128.

[88] Javed S, Azeem F, Hussain S, Rasul I, Siddique MH, Riaz M, Afzal M, Kouser A, Nadeem H: **Bacterial lipases: A review on purification and characterization**. *Progress in Biophysics and Molecular Biology* 2017, .

[89] Vojcic L, Pitzler C, Körfer G, Jakob F, Ronny M, Maurer KH, Schwaneberg U: **Advances in protease engineering for laundry detergents**. *New Biotechnology* 2015, 32:629–634.

[90] Black TD, Briggs BS, Evans R, Muth WL, Vangala S, Zmijewski MJ: ***o*-Phthalyl amidase in the synthesis of loracarbef: Process development using this novel biocatalyst**. *Biotechnology Letters* 1996, 18:875–880.

[91] Ogawa J, Shimizu S: **Industrial microbial enzymes: Their discovery by screening and use in large-scale production of useful chemicals in Japan**. *Current Opinion in Biotechnology* 2002, 13:367–375.

[92] Liang J, Lalonde J, Borup B, Mitchell V, Mundorff E, Trinh N, Kochrekar DA, Nair Cherat R, Pai GG: **Development of a biocatalytic process as an alternative to the (—)-DIP-Cl-mediated asymmetric reduction of a key intermediate of Montelukast**. *Organic Process Research & Development* 2010, 14:193–198.

[93] Suurnäkki A, Tenkanen M, Buchert J, Viikari L: **Hemicellulases in the bleaching of chemical pulps**. In *Biotechnology in the Pulp and Paper Industry*, edited by Scheper T, Eriksson KEL, Babel W, Blanch HW, Cooney CL, Enfors SO, Fiechter A, Klibanov AM, Mattiasson B, Primrose SB, Rehm HJ, Rogers PL, Sahm H, Schügerl K, Tsao GT, Venkat K, Villadsen J, von Stockar U, Wandrey C, volume 57. Berlin, Heidelberg: Springer, 1997, 261–287.

[94] Monod J: **The growth of bacterial cultures**. *Annual Review of Microbiology* 1949, 3:371–394.

[95] Luedeking R, Piret EL: **A kinetic study of the lactic acid fermentation. Batch process at controlled pH**. *Biotechnology and Bioengineering* 1959, 1:393–412.

[96] Glick BR: **Metabolic load and heterologous gene expression**. *Biotechnology Advances* 1995, 13:247–261.

[97] Schaepe S, Kuprijanov A, Simutis R, Lübbert A: **Avoiding overfeeding in high cell density fed-batch cultures of *E. coli* during the production of heterologous proteins**. *Journal of Biotechnology* 2014, 192:146–153.

[98] Yamane T, Shimizu S: **Fed-batch techniques in microbial processes**. *Advances in Biochemical Engineering/Biotechnology* 1984, 30:147–194.

[99] Gnoth S, Jenzsch M, Simutis R, Lübbert A: **Control of cultivation processes for recombinant protein production: A review**. *Bioprocess and Biosystems Engineering* 2008, 31:21–39.

[100] Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar K: **Cultivation strategies to enhance productivity of *Pichia pastoris*: A review**. *Biotechnology Advances* 2015, 33:1177–1193.

[101] Jenzsch M, Simutis R, Lübbert A: **Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations**. *Journal of Biotechnology* 2006, 122:483–493.

[102] Jenzsch M, Simutis R, Lübbert A: **Optimization and control of industrial microbial cultivation processes**. *Engineering in Life Sciences* 2006, 6:117–124.

[103] Randek J, Mandenius CF: **On-line soft sensing in upstream bioprocessing**. *Critical Reviews in Biotechnology* 2018, 38:106–121.

[104] Hemmerich J, Noack S, Wiechert W, Oldiges M: **Microbioreactor systems for accelerated bioprocess development**. *Biotechnology Journal* 2018, 13:1700141.

[105] Doig SD, Baganz F, Lye GJ: **High-throughput screening and process optimisation**. In *Basic biotechnology*, edited by Ratledge C, Kristiansen B. Cambridge University Press, 2006, 289–306.

[106] Hsu WT, Aulakh RPS, Traul DL, Yuk IH: **Advanced microscale bioreactor system: A representative scale-down model for bench-top bioreactors**. *Cytotechnology* 2012, 64:667–678.

[107] Moses S, Manahan M, Ambrogelly A, Ling WLW: **Assessment of AMBR™ as a model for high-throughput cell culture process development strategy**. *Advances in Bioscience and Biotechnology* 2012, 3:918–927.

[108] Rameez S, Mostafa SS, Miller C, Shukla AA: **High-throughput miniaturized bioreactors for cell culture process development: Reproducibility, scalability, and control**. *Biotechnology Progress* 2014, 30:718–727.

[109] Ratcliffe E, Glen KE, Workman VL, Stacey AJ, Thomas RJ: **A novel automated bioreactor for scalable process optimisation of haematopoietic stem cell culture**. *Journal of Biotechnology* 2012, 161:387–390.

[110] Velez-Suberbie ML, Betts JP, Walker KL, Robinson C, Zoro B, Keshavarz-Moore E: **High throughput automated microbial bioreactor system used for clone selection and rapid scale-down process optimisation**. *Biotechnology Progress* 2018, 34:58–56.

[111] Funke M, Diederichs S, Kensy F, Müller C, Büchs J: **The baffled microtiter plate: Increased oxygen transfer and improved online monitoring in small scale fermentations**. *Biotechnology and Bioengineering* 2009, 103:1118–1128.

[112] Hemmerich J, Adelantado N, Barrigon JM, Ponte X, Hormann A, Ferrer P, Kensy F, Valero F: **Comprehensive clone screening and evaluation of fed-batch strategies in a microbioreactor and lab scale stirred tank bioreactor system: Application on *Pichia pastoris* producing *Rhizopus oryzae* lipase**. *Microbial Cell Factories* 2014, 13:36.

[113] Jensen SI, Lennen RM, Herrgard MJ, Nielsen AT: **Seven gene deletions in seven days: Fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress**. *Scientific Reports* 2015, 5:17874.

[114] Motta Dos Santos LF, Coutte F, Ravallec R, Dhulster P, Tournier-Couturier L, Jacques P: **An improvement of surfactin production by *B. subtilis* BBG131 using design of experiments in microbioreactors and continuous process in bubbleless membrane bioreactor**. *Bioresource Technology* 2016, 218:944–952.

[115] Mühlmann M, Kunze M, Ribeiro J, Geinitz B, Lehmann C, Schwaneberg U, Commandeur U, Büchs J: **Cellulolytic RoboLector - towards an automated high-throughput screening platform for recombinant cellulase expression**. *Journal of Biological Engineering* 2017, 11:1.

[116] Wewetzer SJ, Kunze M, Ladner T, Luchterhand B, Roth S, Rahmen N, Kloss R, Costa E Silva A, Regestein L, Büchs J: **Parallel use of shake flask and microtiter plate online measuring devices (RAMOS and BioLector) reduces the number of experiments in laboratory-scale stirred tank bioreactors**. *Journal of Biological Engineering* 2015, 9:9.

[117] Wittgens A, Tiso T, Arndt TT, Wenk P, Hemmerich J, Müller C, Wichmann R, Kupper B, Zwick M, Wilhelm S, Hausmann R, Syldatk C, Rosenau F, Blank LM: **Growth independent rhamnolipid production from glucose using the non-pathogenic *Pseudomonas putida* KT2440**. *Microbial Cell Factories* 2011, 10:80.

[118] Käß F, Prasad A, Tillack J, Moch M, Giese H, Büchs J, Wiechert W, Oldiges M: **Rapid assessment of oxygen transfer impact for *Corynebacterium glutamicum***. *Bioprocess and Biosystems Engineering* 2014, 37:2567–2577.

[119] Lennen RM, Nilsson Wallin AI, Pedersen M, Bonde M, Luo H, Herrgard MJ, Sommer MOA: **Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects**. *Nucleic Acids Research* 2016, 44:e36.

[120] Hortsch R, Stratmann A, Weuster-Botz D: **New milliliter-scale stirred tank bioreactors for the cultivation of mycelium forming microorganisms**. *Biotechnology and Bioengineering* 2010, 106:443–451.

[121] Knepper A, Heiser M, Glauche F, Neubauer P: **Robotic platform for parallelized cultivation and monitoring of microbial growth parameters in microwell plates**. *Journal of Laboratory Automation* 2014, 19:593–601.

[122] Puskeiler R, Kaufmann K, Weuster-Botz D: **Development, parallelization, and automation of a gas-inducing milliliter-scale bioreactor for high-throughput bioprocess design (HTBD)**. *Biotechnology and Bioengineering* 2005, 89:512–523.

[123] Riedlberger P, Brüning S, Weuster-Botz D: **Characterization of stirrers for screening studies of enzymatic biomass hydrolyses on a milliliter scale**. *Bioprocess and Biosystems Engineering* 2013, 36:927–935.

[124] Schmideder A, Weuster-Botz D: **High-performance recombinant protein production with *Escherichia coli* in continuously operated cascades of stirred-tank reactors**. *Journal of Industrial Microbiology & Biotechnology* 2017, 44:1021–1029.

[125] Wellenbeck W, Mampel J, Naumer C, Knepper A, Neubauer P: **Fast-track development of a lactase production process with *Kluyveromyces lactis* by a progressive parameter-control workflow**. *Engineering in Life Sciences* 2016, 17:1185–1194.

[126] Schmideder A, Severin TS, Cremer JH, Weuster-Botz D: **A novel milliliter-scale chemostat system for parallel cultivation of microorganisms in stirred-tank bioreactors**. *Journal of Biotechnology* 2015, 210:19–24.

[127] Begot C, Desnier I, Daudin JD, Labadie JC, Lebert A: **Recommendations for calculating growth parameters by optical density measurements**. *Journal of Microbiological Methods* 1996, 25:225–232.

[128] Friedman AJ, Blecher K, Schairer D, Tuckman-Vernon C, Nacharaju P, Sanchez D, Gialanella P, Martinez LR, Friedman JM, Nosanchuk JD: **Improved antimicrobial efficacy with nitric oxide releasing nanoparticle generated S-nitrosoglutathione**. *Nitric Oxide: Biology and Chemistry* 2011, 25:381–386.

[129] Medina A, Lambert RJW, Magan N: **Rapid throughput analysis of filamentous fungal growth using turbidimetric measurements with the Bioscreen C: A tool for screening antifungal compounds**. *Fungal Biology* 2012, 116:161–169.

[130] Reed JL, Patel TR, Chen KH, Joyce AR, Applebee MK, Herring CD, Bui OT, Knight EM, Fong SS, Palsson BO: **Systems approach to refining genome annotation**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103:17480–17484.

[131] Tang W, Xing Z, Li C, Wang J, Wang Y: **Molecular mechanisms and in vitro antioxidant effects of *Lactobacillus plantarum* MA2**. *Food Chemistry* 2017, 221:1642–1649.

[132] Linger JG, Hobdey SE, Franden MA, Fulk EM, Beckham GT: **Conversion of levoglucosan and cellobiosan by *Pseudomonas putida* KT2440**. *Metabolic Engineering Communications* 2016, 3:24–29.

[133] Klein M, Islam ZU, Knudsen PB, Carrillo M, Swinnen S, Workman M, Nevoigt E: **The expression of glycerol facilitators from various yeast species improves growth on glycerol of *Saccharomyces cerevisiae***. *Metabolic Engineering Communications* 2016, 3:252–257.

[134] Strucko T, Buron LD, Jarczynska ZD, Nodvig CS, Molgaard L, Halkier BA, Mortensen UH: **CASCADE, a platform for controlled gene amplification for high, tunable and selection-free gene expression in yeast**. *Scientific Reports* 2017, 7:41431.

[135] Swinnen S, Klein M, Carrillo M, McInnes J, Nguyen HTT, Nevoigt E: **Re-evaluation of glycerol utilization in *Saccharomyces cerevisiae*: Characterization of an isolate that grows on glycerol without supporting supplements**. *Biotechnology for Biofuels* 2013, 6:157.

[136] Swinnen S, Fernandez-Nino M, Gonzalez-Ramos D, van Maris AJA, Nevoigt E: **The fraction of cells that resume growth after acetic acid addition is a strain-dependent parameter of acetic acid tolerance in *Saccharomyces cerevisiae***. *FEMS Yeast Research* 2014, 14:642–653.

[137] Betts J, Warr S, Finka GB, Uden M, Town M, Janda JM, Baganz F, Lye GJ: **Impact of aeration strategies on fed-batch cell culture kinetics in a single-use 24-well miniature bioreactor**. *Biochemical Engineering Journal* 2014, 82:105–116.

[138] Chen A, Chitta R, Chang D, Amanullah A: **Twenty-four well plate miniature bioreactor system as a scale-down model for cell culture process development**. *Biotechnology and Bioengineering* 2009, 102:148–160.

[139] Isett K, George H, Herber W, Amanullah A: **Twenty-four-well plate miniature bioreactor high-throughput system: Assessment for microbial cultivations**. *Biotechnology and Bioengineering* 2007, 98:1017–1028.

[140] Kim DY, Chaudhry MA, Kennard ML, Jardon MA, Braasch K, Dionne B, Butler M, Piret JM: **Fed-batch CHO cell t-PA production and feed glutamine replacement to reduce ammonia production**. *Biotechnology Progress* 2013, 29:165–175.

[141] Beckers S, Noor F, Müller-Vieira U, Mayer M, Strigun A, Heinzle E: **High throughput, non-invasive and dynamic toxicity screening on adherent cells using respiratory measurements**. *Toxicology in Vitro* 2010, 24:686–694.

[142] Glauche F, John GT, Arain S, Knepper A, Neubauer A, Goelling D, Lang C, Violet N, King R, Neubauer P: **Toward microbioreactor arrays: A slow-responding oxygen sensor for monitoring of microbial cultures in standard 96-well plates**. *Journal of Laboratory Automation* 2015, 20:438–446.

[143] Kensy F, John GT, Hofmann B, Büchs J: **Characterisation of operation conditions and online monitoring of physiological culture parameters in shaken 24-well microtiter plates**. *Bioprocess and Biosystems Engineering* 2005, 28:75–81.

[144] Köster M, Krause C, Paffenhöfer G: **Time-series measurements of oxygen consumption of copepod nauplii**. *Marine Ecology Progress Series* 2008, 353:157–164.

[145] Gebhardt G, Hortsch R, Kaufmann K, Arnold M, Weuster-Botz D: **A new microfluidic concept for parallel operated milliliter-scale stirred tank bioreactors**. *Biotechnology Progress* 2011, 27:684–690.

[146] Büchs J: **Out-of-phase operating conditions, a hitherto unknown phenomenon in shaking bioreactors**. *Biochemical Engineering Journal* 2001, 7:135–141.

[147] Kensy F, Zimmermann HF, Knabben I, Anderlei T, Trauthwein H, Dingerdissen U, Büchs J: **Oxygen transfer phenomena in 48-well microtiter plates: Determination by optical monitoring of sulfite oxidation and verification by real-time measurement during microbial growth**. *Biotechnology and Bioengineering* 2005, 89:698–708.

[148] Hermann R, Lehmann M, Büchs J: **Characterization of gas-liquid mass transfer phenomena in microtiter plates**. *Biotechnology and Bioengineering* 2002, 81:178–186.

[149] Hermann R, Walther N, Maier U, Büchs J: **Optical method for the determination of the oxygen-transfer capacity of small bioreactors based on sulfite oxidation**. *Biotechnology and Bioengineering* 2001, 74:355–363.

[150] Duetz WA, Witholt B: **Oxygen transfer by orbital shaking of square vessels and deepwell microtiter plates of various dimensions**. *Biochemical Engineering Journal* 2004, 17:181–185.

[151] Duetz WA, Witholt B: **Effectiveness of orbital shaking for the aeration of suspended bacterial cultures in square-deepwell microtiter plates**. *Biochemical Engineering Journal* 2001, 7:113–115.

[152] Dürauer A, Hobiger S, Walther C, Jungbauer A: **Mixing at the microscale: Power input in shaken microtiter plates**. *Biotechnology Journal* 2016, 11:1–11.

[153] Lattermann C, Funke M, Hansen S, Diederichs S, Büchs J: **Cross-section perimeter is a suitable parameter to describe the effects of different baffle geometries in shaken microtiter plates**. *Journal of Biological Engineering* 2014, 8:18.

[154] Doig SD, Pickering SCR, Lye GJ, Baganz F: **Modelling surface aeration rates in shaken microtitre plates using dimensionless groups**. *Chemical Engineering Science* 2005, 60:2741–2750.

[155] Morschett H, Schiprowski D, Rohde J, Wiechert W, Oldiges M: **Comparative evaluation of phototrophic microtiter plate cultivation against laboratory-scale photobioreactors**. *Bioprocess and Biosystems Engineering* 2017, 40:663–673.

[156] Duetz WA: **Microtiter plates as mini-bioreactors: Miniaturization of fermentation methods**. *Trends in Microbiology* 2007, 15:469–475.

[157] Kirk TV, Szita N: **Oxygen transfer characteristics of miniaturized bioreactor systems**. *Biotechnology and Bioengineering* 2013, 110:1005–1019.

[158] Hemmerich J, Kensy F: **Automation of microbioreactors: Operating 48 parallel fed-batch fermentations at microscale.** *BioProcess International* 2013, 11:68–76.

[159] Mühlmann M, Forsten E, Noack S, Büchs J: **Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures**. *Microbial Cell Factories* 2017, 16:220.

[160] Unthan S, Radek A, Wiechert W, Oldiges M, Noack S: **Bioprocess automation on a Mini Pilot Plant enables fast quantitative microbial phenotyping**. *Microbial Cell Factories* 2015, 14:32.

[161] Heux S, Poinot J, Massou S, Sokol S, Portais JC: **A novel platform for automated high-throughput fluxome profiling of metabolic variants**. *Metabolic Engineering* 2014, 25:8–19.

[162] Huber R, Ritter D, Hering T, Hillmer AK, Kensy F, Müller C, Le Wang, Büchs J: **Robo-Lector – a novel platform for automated high-throughput cultivations in microtiter plates with high information content**. *Microbial Cell Factories* 2009, 8:42.

[163] Cruz Bournazou MN, Barz T, Nickel DB, Lopez Cardenas DC, Glauche F, Knepper A, Neubauer P: **Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities**. *Biotechnology and Bioengineering* 2017, 114:610–619.

[164] Gorochowski TE, van den Berg E, Kerkman R, Roubos JA, Bovenberg RAL: **Using synthetic biological parts and microbioreactors to explore the protein expression characteristics of *Escherichia coli***. *ACS Synthetic Biology* 2014, 3:129–139.

[165] Baumann P, Bluthardt N, Renner S, Burghardt H, Osberghaus A, Hubbuch J: **Integrated development of up- and downstream processes supported by the Cherry-Tag™ for real-time tracking of stability and solubility of proteins**. *Journal of Biotechnology* 2015, 200:27–37.

[166] Dörr M, Fibinger MPC, Last D, Schmidt S, Santos-Aberturas J, Böttcher D, Hummel A, Vickers C, Voss M, Bornscheuer UT: **Fully automatized high-throughput enzyme library screening using a robotic platform**. *Biotechnology and Bioengineering* 2016, 113:1421–1432.

[167] Freier L, Hemmerich J, Schöler K, Wiechert W, Oldiges M, von Lieres E: **Framework for Kriging-based iterative experimental analysis and design: Optimization of secretory protein production in *Corynebacterium glutamicum***. *Engineering in Life Sciences* 2016, 16:538–549.

[168] Kwon YC, Jewett MC: **High-throughput preparation methods of crude extract for robust cell-free protein synthesis**. *Scientific Reports* 2015, 5:8663.

[169] Glauche F, Pilarek M, Bournazou MNC, Grunzel P, Neubauer P: **Design of experiments-based high-throughput strategy for development and optimization of efficient cell disruption protocols**. *Engineering in Life Sciences* 2017, 17:1166–1172.

[170] Konstantinidis S, Goh HY, Bufájer JMM, de Galbert P, Parau M, Velayudhan A: **Flexible and accessible automated operation of miniature chromatography columns on a liquid handling station**. *Biotechnology Journal* 2018, 13:1700390.

[171] Zimmermann S, Gretzinger S, Zimmermann PK, Bogsnes A, Hansson M, Hubbuch J: **Cell separation in aqueous two-phase systems - influence of polymer molecular weight and tie-line length on the resolution of five model cell lines**. *Biotechnology Journal* 2018, 13:1700250.

[172] Baumgartner K, Galm L, Nötzold J, Sigloch H, Morgenstern J, Schleining K, Suhm S, Oelmeier SA, Hubbuch J: **Determination of protein phase diagrams by microbatch experiments: Exploring the influence of precipitants and pH**. *International Journal of Pharmaceutics* 2015, 479:28–40.

[173] Baumann P, Hahn T, Hubbuch J: **High-throughput micro-scale cultivations and chromatography modeling: Powerful tools for integrated process development**. *Biotechnology and Bioengineering* 2015, 112:2123–2133.

[174] Baumann P, Hubbuch J: **Downstream process development strategies for effective bioprocesses: Trends, progress, and combinatorial approaches**. *Engineering in Life Sciences* 2017, 17:1142–1158.

[175] Chhatre S, Titchener-Hooker NJ: **Review: Microscale methods for high-throughput chromatography development in the pharmaceutical industry**. *Journal of Chemical Technology & Biotechnology* 2009, 84:927–940.

[176] Schmideder A, Hensler S, Lang M, Stratmann A, Giesecke U, Weuster-Botz D: **High-cell-density cultivation and recombinant protein production with *Komagataella pastoris* in stirred-tank bioreactors from milliliter to cubic meter scale**. *Process Biochemistry* 2016, 51:177–184.

[177] Islam RS, Tisi D, Levy MS, Lye GJ: **Scale-up of *Escherichia coli* growth and recombinant protein expression conditions from microwell to laboratory and pilot scale based on matched k_La**. *Biotechnology and Bioengineering* 2008, 99:1128–1139.

[178] Kensy F, Engelbrecht C, Büchs J: **Scale-up from microtiter plate to laboratory fermenter: Evaluation by online monitoring techniques of growth and protein expression in *Escherichia coli* and *Hansenula polymorpha* fermentations**. *Microbial Cell Factories* 2009, 8:68.

[179] Koepff J, Keller M, Tsolis KC, Busche T, Rückert C, Hamed MB, Anné J, Kalinowski J, Wiechert W, Economou A, Oldiges M: **Fast and reliable strain characterization of *Streptomyces lividans* through micro-scale cultivation**. *Biotechnology and Bioengineering* 2017, 114:2011–2022.

[180] Šiurkus J, Panula-Perälä J, Horn U, Kraft M, Rimšeliene R, Neubauer P: **Novel approach of high cell density recombinant bioprocess development: Optimisation and scale-up from microlitre to pilot scales while maintaining the fed-batch cultivation mode of *E. coli* cultures**. *Microbial Cell Factories* 2010, 9:35.

[181] Marques MPC, Cabral JMS, Fernandes P: **Bioprocess scale-up: Quest for the parameters to be used as criterion to move from microreactors to lab-scale**. *Journal of Chemical Technology & Biotechnology* 2010, 85:1184–1198.

[182] Islam RS, Tisi D, Levy MS, Lye GJ: **Framework for the rapid optimization of soluble protein expression in *Escherichia coli* combining microscale experiments and statistical experimental design**. *Biotechnology Progress* 2007, 23:785–793.

[183] Olsson IM, Johansson E, Berntsson M, Eriksson L, Gottfries J, Wold S: **Rational DoE protocols for 96-well plates**. *Chemometrics and Intelligent Laboratory Systems* 2006, 83:66–74.

[184] Mandenius CF, Brundin A: **Bioprocess optimization using design-of-experiments methodology**. *Biotechnology Progress* 2008, 24:1191–1203.

[185] Solle D, Hitzmann B, Herwig C, Pereira Remelhe M, Ulonska S, Wuerth L, Prata A, Steckenreiter T: **Between the poles of data-driven and mechanistic modeling for process operation**. *Chemie Ingenieur Technik* 2017, 89:542–561.

[186] Kunze M, Roth S, Gartz E, Büchs J: **Pitfalls in optical on-line monitoring for high-throughput screening of microbial systems**. *Microbial Cell Factories* 2014, 13:53.

[187] Hecht A, Endy D, Salit M, Munson MS: **When wavelengths collide: Bias in cell abundance measurements due to expressed fluorescent proteins**. *ACS Synthetic Biology* 2016, 5:1024–1027.

[188] Surribas A, Resina D, Ferrer P, Valero F: **Rivoflavin may interfere with on-line monitoring of secreted green fluorescence protein fusion proteins in *Pichia pastoris***. *Microbial Cell Factories* 2007, 6:15.

[189] Gernaey KV, Baganz F, Franco-Lara E, Kensy F, Krühne U, Luebberstedt M, Marx U, Palmqvist E, Schmid A, Schubert F, Mandenius CF: **Monitoring and control of microbioreactors: An expert opinion on development needs**. *Biotechnology Journal* 2012, 7:1308–1314.

[190] Neubauer P, Cruz N, Glauche F, Junne S, Knepper A, Raven M: **Consistent development of bioprocesses from microliter cultures to the industrial scale**. *Engineering in Life Sciences* 2013, 13:224–238.

[191] Jacques P, Bechet M, Bigan M, Caly D, Chataigne G, Coutte F, Flahaut C, Heuson E, Leclere V, Lecouturier D, Phalip V, Ravallec R, Dhulster P, Froidevaux R: **High-throughput strategies for the discovery and engineering of enzymes for biocatalysis**. *Bioprocess and Biosystems Engineering* 2016, 40:161–180.

[192] Noack S, Voges R, Gätgens J, Wiechert W: **The linkage between nutrient supply, intracellular enzyme abundances and bacterial growth: New evidences from the central carbon metabolism of *Corynebacterium glutamicum***. *Journal of Biotechnology* 2017, 258:13–24.

[193] Yordanov B, Dalchau N, Grant PK, Pedersen M, Emmott S, Haseloff J, Phillips A: **A computational method for automated characterization of genetic components**. *ACS Synthetic Biology* 2014, 3:578–588.

[194] Ladner T, Mühlmann M, Schulte A, Wandrey G, Büchs J: **Prediction of *Escherichia coli* expression performance in microtiter plates by analyzing only the temporal development of scattered light during culture**. *Journal of Biological Engineering* 2017, 11:469.

[195] Hughes D, Andersson DI: **Evolutionary consequences of drug resistance: Shared principles across diverse targets and organisms**. *Nature Reviews: Genetics* 2015, 16:459–471.

[196] Hall BG, Acar H, Nandipati A, Barlow M: **Growth rates made easy**. *Molecular Biology and Evolution* 2014, 31:232–238.

[197] Sprouffske K, Wagner A: **Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves**. *BMC Bioinformatics* 2016, 17:172.

[198] Schöning L: **Croissance, a tool for estimating growth rates in growth curves**, 2017. URL `github.com/biosustain/croissance`.

[199] Kensy F, Zang E, Faulhammer C, Tan RK, Büchs J: **Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates**. *Microbial Cell Factories* 2009, 8:31.

[200] Hemmerich J, Wiechert W, Oldiges M: **Automated growth rate determination in high-throughput microbioreactor systems**. *BMC Research Notes* 2017, 10:617.

[201] Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R, Estadilla J, Teisan S, Schreyer HB, Andrae S, Yang TH, Lee SY, Burk MJ, van Dien S: **Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol**. *Nature Chemical Biology* 2011, 7:445–452.

[202] Eiteman MA, Altman E: **Overcoming acetate in *Escherichia coli* recombinant protein fermentations**. *Trends in Biotechnology* 2006, 24:530–536.

[203] Samorski M, Müller-Newen G, Büchs J: **Quasi-continuous combined scattered light and fluorescence measurements: A novel measurement technique for shaken microtiter plates**. *Biotechnology and Bioengineering* 2005, 92:61–68.

[204] Bruder S, Reifenrath M, Thomik T, Boles E, Herzog K: **Parallelised online biomass monitoring in shake flasks enables efficient strain and carbon source dependent growth characterisation of *Saccharomyces cerevisiae***. *Microbial Cell Factories* 2016, 15:127.

[205] Ude C, Schmidt-Hager J, Findeis M, John GT, Scheper T, Beutel S: **Application of an online-biomass sensor in an optical multisensory platform prototype for growth monitoring of biotechnical relevant microorganism and cell lines in single-use shake flasks**. *Sensors* 2014, 14:17390–17405.

[206] Radek A, Tenhaef N, Müller MF, Brüsseler C, Wiechert W, Marienhagen J, Polen T, Noack S: **Miniaturized and automated adaptive laboratory evolution: Evolving *Corynebacterium glutamicum* towards an improved d-xylose utilization**. *Bioresource Technology* 2017, 245:1377–1385.

[207] Huth I, Schrader J, Holtmann D: **Microtiter plate-based cultivation to investigate the growth of filamentous fungi**. *Engineering in Life Sciences* 2017, 17:1064–1070.

[208] Radek A, Müller MF, Gätgens J, Eggeling L, Krumbach K, Marienhagen J, Noack S: **Formation of xylitol and xylitol-5-phosphate and its impact on growth of d-xylose-utilizing *Corynebacterium glutamicum* strains**. *Journal of Biotechnology* 2016, 231:160–166.

[209] Rohe P: *Prozessnahe Hochdurchsatzoptimierung der heterologen Proteinproduktion in alternativen Wirtsorganismen*. Ph.D. thesis, RWTH Aachen, 2012.

[210] Vertès AA: **Protein secretion systems of *Corynebacterium glutamicum***. In *Corynebacterium glutamicum: Biology and biotechnology*, edited by Yukawa H, Inui M, volume 23. Berlin, Heidelberg: Springer, 2013, 351–389.

[211] Li W, Zhou X, Lu P: **Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis***. *Research in Microbiology* 2004, 155:605–610.

[212] Hemmerich J, Kensy F: **Automated microbioreactor systems for pharmaceutical bioprocessing: Profiling of seeding and induction conditions in high-throughput fermentations**. *Pharmaceutical Bioprocessing* 2014, 2:227–235.

[213] Hedderich J, Sachs L: *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.

[214] Becker J, Klopprogge C, Herold A, Zelder O, Bolten CJ, Wittmann C: **Metabolic flux engineering of L-lysine production in *Corynebacterium glutamicum* - overexpression and modification of G6P dehydrogenase**. *Journal of Biotechnology* 2007, 132:99–109.

[215] Büchs J: **Introduction to advantages and problems of shaken cultures**. *Biochemical Engineering Journal* 2001, 7:91–98.

[216] Tillack J, Paczia N, Nöh K, Wiechert W, Noack S: **Error propagation analysis for quantitative intracellular metabolomics**. *Metabolites* 2012, 2:1012–1030.

[217] Henderson AR: **The bootstrap: A technique for data-driven statistics. Using computer-intensive analyses to explore experimental data**. *Clinica Chimica Acta* 2005, 359:1–26.

[218] Grünberger A, van Ooyen J, Paczia N, Rohe P, Schiendzielorz G, Eggeling L, Wiechert W, Kohlheyer D, Noack S: **Beyond growth rate 0.6: *Corynebacterium glutamicum* cultivated in highly diluted environments**. *Biotechnology and Bioengineering* 2013, 110:220–228.

[219] Krömer JO, Bolten CJ, Heinzle E, Schröder H, Wittmann C: **Physiological response of *Corynebacterium glutamicum* to oxidative stress induced by deletion of the transcriptional repressor McbR**. *Microbiology* 2008, 154:3917–3930.

[220] Frunzke J, Engels V, Hasenbein S, Gätgens C, Bott M: **Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2**. *Molecular Microbiology* 2008, 67:305–322.

[221] Voges R, Noack S: **Quantification of proteome dynamics in *Corynebacterium glutamicum* by ^{15}N-labeling and selected reaction monitoring**. *Journal of Proteomics* 2012, 75:2660–2669.

[222] Unthan S: *Robot-assisted phenotyping of genome-reduced Corynebacterium glutamicum strain libraries to draft a chassis organism*. Ph.D. thesis, RWTH Aachen University, 2015.

[223] Fisher RA: *The design of experiments*. Edinburgh: Oliver & Boyd, 1935.

[224] Freier L, Wiechert W, von Lieres E: **Kriging with trend functions nonlinear in their parameters: Theory and application in enzyme kinetics**. *Engineering in Life Sciences* 2017, 17:916–922.

[225] Morschett H, Freier L, Rohde J, Wiechert W, von Lieres E, Oldiges M: **A framework for accelerated phototrophic bioprocess development: Integration of parallelized microscale cultivation, laboratory automation and Kriging-assisted experimental design**. *Biotechnology for Biofuels* 2017, 10:26.

[226] Freier L, von Lieres E: **Kriging based iterative parameter estimation procedure for biotechnology applications with nonlinear trend functions**. *IFAC-PapersOnLine* 2015, 48:574–579.

[227] Teramoto H, Watanabe K, Suzuki N, Inui M, Yukawa H: **High yield secretion of heterologous proteins in *Corynebacterium glutamicum* using its own Tat-type signal sequence**. *Applied Microbiology and Biotechnology* 2011, 91:677–687.

[228] Huber R, Roth S, Rahmen N, Büchs J: **Utilizing high-throughput experimentation to enhance specific productivity of an *E. coli* T7 expression system by phosphate limitation**. *BMC Biotechnology* 2011, 11:22.

[229] Kottmeier K, Müller C, Huber R, Büchs J: **Increased product formation induced by a directed secondary substrate limitation in a batch *Hansenula polymorpha* culture**. *Applied Microbiology and Biotechnology* 2010, 86:93–101.

[230] Kennedy MJ, Krouse D: **Strategies for improving fermentation medium performance: A review**. *Journal of Industrial Microbiology & Biotechnology* 1999, 23:456–475.

[231] Jones R, Gadd G: **Ionic nutrition of yeast - physiological mechanisms involved and implications for biotechnology**. *Enzyme and Microbial Technology* 1990, 12:402–418.

[232] Zhang J, Greasham R: **Chemically defined media for commercial fermentations**. *Applied Microbiology and Biotechnology* 1999, 51:407–421.

[233] Freier L, von Lieres E: **Kriging toolKit (KriKit)**, January 2017. URL `github.com/modsim/KriKit`.

[234] Hemmerich J, Freier L, Wiechert W, von Lieres E, Oldiges M: **Generic protocol for optimization of heterologous protein production using automated microbioreactor technology**. *Journal of Visualized Experiments* 2017, 130:e56234.

[235] Lu C, Bentley WE, Rao G: **A high-throughput approach to promoter study using green fluorescent protein**. *Biotechnology Progress* 2004, 20:1634–1640.

[236] Zanzotto A, Boccazzi P, Gorret N, van Dyk TK, Sinskey AJ, Jensen KF: **In situ measurement of bioluminescence and fluorescence in an integrated microbioreactor**. *Biotechnology and Bioengineering* 2006, 93:40–47.

[237] Weuster-Botz D, Kelle R, Frantzen M, Wandrey C: **Substrate controlled fed-batch production of L-lysine with *Corynebacterium glutamicum***. *Biotechnology Progress* 1997, 13:387–393.

[238] Zimmerman DL: **Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction**. *Environmetrics* 2006, 17:635–652.

[239] Meissner D, Vollstedt A, van Dijl JM, Freudl R: **Comparative analysis of twin-arginine (Tat)-dependent protein secretion of a heterologous model protein (GFP) in three different Gram-positive bacteria**. *Applied Microbiology and Biotechnology* 2007, 76:633–642.

[240] Zimmer M: **Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior**. *Chemical Reviews* 2002, 102:759–782.

[241] Silhavy TJ, Kahne D, Walker S: **The bacterial cell envelope**. *Cold Spring Harbor Perspectives in Biology* 2010, 2:a000414.

[242] Umakoshi H, Nishida M, Suga K, Huong T, Bui T, Shimanouchi R, Kuboi R: **Characterization of green fluorescent protein using aqueous two-phase systems**. *Solvent Extraction Research and Development, Japan* 2009, 16:145–150.

[243] Adachi T, Yamagata H, Tsukagoshi N, Udaka S: **Repression of the cell wall protein gene operon in *Bacillus brevis* 47 by magnesium and calcium ions**. *Journal of Bacteriology* 1991, 173:4243–4245.

[244] Rahmen N, Fulton A, Ihling N, Magni M, Jaeger KE, Büchs J: **Exchange of single amino acids at different positions of a recombinant protein affects metabolic burden in *Escherichia coli***. *Microbial Cell Factories* 2015, 14:10.

[245] Rahmen N, Schlupp CD, Mitsunaga H, Fulton A, Aryani T, Esch L, Schaffrath U, Fukuzaki E, Jaeger KE, Büchs J: **A particular silent codon exchange in a recombinant gene greatly influences host cell metabolic activity**. *Microbial Cell Factories* 2015, 14:156.

[246] Saez NJ, Nozach H, Blemont M, Vincentelli R: **High throughput quantitative expression screening and purification applied to recombinant disulfide-rich venom proteins produced in *E. coli***. *Journal of Visualized Experiments* 2014, 89:e51464.

[247] Scheidle M, Jeude M, Dittrich B, Denter S, Kensy F, Suckow M, Klee D, Büchs J: **High-throughput screening of *Hansenula polymorpha* clones in the batch compared with the controlled-release fed-batch mode on a small scale**. *FEMS Yeast Research* 2010, 10:83–92.

[248] Freier L, von Lieres E: **Multi-objective global optimization (MOGO): Algorithm and case study in gradient elution chromatography**. *Biotechnology Journal* 2017, 12:1600613.

[249] Egmond M: ***Fusarium solani pisi* cutinase**. *Biochimie* 2000, 82:1015–1021.

[250] Chen S, Su L, Chen J, Wu J: **Cutinase: Characteristics, preparation, and application**. *Biotechnology Advances* 2013, 31:1754–1767.

[251] Koch-Koerfges A, Pfelzer N, Platzen L, Oldiges M, Bott M: **Conversion of *Corynebacterium glutamicum* from an aerobic respiring to an aerobic fermenting bacterium by inactivation of the respiratory chain**. *Biochimica et Biophysica Acta* 2013, 1827:699–708.

[252] Baneyx F: **Recombinant protein expression in *Escherichia coli***. *Current Opinion in Biotechnology* 1999, 10:411–421.

[253] Hou J, Tyo KEJ, Liu Z, Petranovic D, Nielsen J: **Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae***. *FEMS Yeast Research* 2012, 12:491–510.

[254] Sun Y, Guo W, Wang F, Peng F, Yang Y, Dai X, Liu X, Bai Z: **Transcriptome and multivariable data analysis of *Corynebacterium glutamicum* under different dissolved oxygen conditions in bioreactors**. *PloS one* 2016, 11:e0167156.

[255] Heyland J, Fu J, Blank LM, Schmid A: **Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production**. *Biotechnology and Bioengineering* 2010, 107:357–368.

[256] Klamt S, Mahadevan R, Hädicke O: **When do two-stage processes outperform one-stage processes?** *Biotechnology Journal* 2018, 13:1700539.

[257] Cheng LC, Hor LI, Wu JY, Chen TL: **Effect of specific growth rate on the production of a recombinant nuclease by *Escherichia coli***. *Biochemical Engineering Journal* 2003, 14:101–107.

[258] Heyland J, Blank LM, Schmid A: **Quantification of metabolic limitations during recombinant protein production in *Escherichia coli***. *Journal of Biotechnology* 2011, 155:178–184.

[259] Borkowski O, Goelzer A, Schaffer M, Calabre M, Mader U, Aymerich S, Jules M, Fromion V: **Translation elicits a growth rate-dependent, genome-wide, differential protein production in *Bacillus subtilis***. *Molecular Systems Biology* 2016, 12:870.

[260] Klumpp S, Zhang Z, Hwa T: **Growth rate-dependent global effects on gene expression in bacteria**. *Cell* 2009, 139:1366–1375.

[261] Bienick MS, Young KW, Klesmith JR, Detwiler EE, Tomek KJ, Whitehead TA: **The interrelationship between promoter strength, gene expression, and growth rate**. *PloS one* 2014, 9:e109105.

[262] Collins T, Azevedo-Silva J, da Costa A, Branca F, Machado R, Casal M: **Batch production of a silk-elastin-like protein in *E. coli* BL21(DE3): Key parameters for optimisation**. *Microbial Cell Factories* 2013, 12:21.

[263] Claes JE, Geeraerd AH, van Impe JF: **Heuristic feed rate profiles for optimal yield and productivity of fed-batch bioprocesses**. *Chemical Engineering Communications* 1999, 172:189–216.

[264] Jadot F, Bastin G, van Impe JF: **Optimal adaptive control of a bioprocess with yield–productivity conflict**. *Journal of Biotechnology* 1998, 65:61–68.

[265] Freier L, von Lieres E: **Robust multi-objective global optimization of stochastic processes with a case study in gradient elution chromatography**. *Biotechnology Journal* 2018, 13:1700268.

[266] Logist F, Houska B, Diehl M, van Impe JF: **A toolkit for multi-objective optimal control in bioprocess engineering**. *IFAC Proceedings Volumes* 2010, 43:281–286.

[267] Logist F, Telen D, Houska B, Diehl M, van Impe J: **Multi-objective optimal control of dynamic bioprocesses using ACADO Toolkit**. *Bioprocess and Biosystems Engineering* 2013, 36:151–164.

[268] Zhou YH, Titchener-Hooker NJ: **The application of a Pareto optimisation method in the design of an integrated bioprocess**. *Bioprocess and Biosystems Engineering* 2003, 25:349–355.

[269] Kortmann M, Kuhl V, Klaffl S, Bott M: **A chromosomally encoded T7 RNA polymerase-dependent gene expression system for *Corynebacterium glutamicum*: Construction and comparative evaluation at the single-cell level**. *Microbial Biotechnology* 2015, 8:253–265.

[270] Condon C, Liveris D, Squires C, Schwartz I, Squires CL: **rRNA operon multiplicity in *Escherichia coli* and the physiological implications of *rrn* inactivation**. *Journal of Bacteriology* 1995, 177:4152–4156.

[271] Martin JF, Barreiro C, González-Lavado E, Barriuso M: **Ribosomal RNA and ribosomal proteins in corynebacteria**. *Journal of Biotechnology* 2003, 104:41–53.

[272] Böhm K, Meyer F, Rhomberg A, Kalinowski J, Donovan C, Bramkamp M: **Novel chromosome organization pattern in Actinomycetales - overlapping replication cycles combined with diploidy**. *mBio* 2017, 8:e00511–17.

[273] Tauch A, Wehmeier L, Götker S, Pühler A, Kalinowski J: **Relaxed *rrn* expression and amino acid requirement of a *Corynebacterium glutamicum rel* mutant defective in (p)ppGpp metabolism**. *FEMS Microbiology Letters* 2001, 201:53–58.

[274] Lee YJ, Lee R, Lee SH, Yim SS, Jeong KJ: **Enhanced secretion of recombinant proteins via signal recognition particle (SRP)-dependent secretion pathway by deletion of *rrsE* in *Escherichia coli***. *Biotechnology and Bioengineering* 2016, 113:2453–2461.

[275] Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns**. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95:14863–14868.

[276] Chen Ch, Härdle WK, Unwin A, editors: *Handbook of data visualization*. Springer Handbooks of Computational Statistics. Berlin, Heidelberg: Springer, 2008.

[277] Noack S, Wahl A, Qeli E, Wiechert W: **Visualizing regulatory interactions in metabolic networks**. *BMC Biology* 2007, 5:46.

[278] Thorvaldsdóttir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration**. *Briefings in Bioinformatics* 2013, 14:178–192.

[279] Keim DA: **Information visualization and visual data mining**. *IEEE Transactions on Visualization and Computer Graphics* 2002, 8:1–8.

[280] McHardy AC, Tauch A, Rückert C, Pühler A, Kalinowski J: **Genome-based analysis of biosynthetic aminotransferase genes of *Corynebacterium glutamicum***. *Journal of Biotechnology* 2003, 104:229–240.

[281] Milo R, Jorgensen P, Moran U, Weber G, Springer M: **BioNumbers - the database of key numbers in molecular and cell biology**. *Nucleic Acids Research* 2010, 38:D750–3.

[282] Bosdriesz E, Molenaar D, Teusink B, Bruggeman FJ: **How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization**. *The FEBS journal* 2015, 282:2029–2044.

[283] Scott M, Klumpp S, Mateescu EM, Hwa T: **Emergence of robust growth laws from optimal regulation of ribosome synthesis**. *Molecular Systems Biology* 2014, 10:747.

[284] Nomura M: **Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: Diversity and common principles**. *Journal of Bacteriology* 1999, 181:6857–6864.

[285] Kaczanowska M, Rydén-Aulin M: **Ribosome biogenesis and the translation process in *Escherichia coli***. *Microbiology and Molecular Biology Reviews* 2007, 71:477–494.

[286] Paul BJ, Ross W, Gaal T, Gourse RL: **rRNA transcription in *Escherichia coli***. *Annual Review of Genetics* 2004, 38:749–770.

[287] Potrykus K, Murphy H, Philippe N, Cashel M: **ppGpp is the major source of growth rate control in *E. coli***. *Environmental Microbiology* 2011, 13:563–575.

[288] Brockmann-Gretza O, Kalinowski J: **Global gene expression during stringent response in *Corynebacterium glutamicum* in presence and absence of the *rel* gene encoding (p)ppGpp synthase**. *BMC Genomics* 2006, 7:230.

[289] Niebisch A, Bott M: **Purification of a cytochrome bc_1-aa_3 supercomplex with quinol oxidase activity from *Corynebacterium glutamicum*: Identification of a fourth subunit of cytochrome aa_3 oxidase and mutational analysis of diheme cytochrome c_1**. *Journal of Biological Chemistry* 2003, 278:4339–4346.

[290] Schluesener D, Rögner M, Poetsch A: **Evaluation of two proteomics technologies used to screen the membrane proteomes of wild-type *Corynebacterium glutamicum* and an L-lysine-producing strain**. *Analytical and Bioanalytical Chemistry* 2007, 389:1055–1064.

[291] Burkovski A: **Cell envelope of corynebacteria: Structure and influence on pathogenicity**. *ISRN Microbiology* 2013, 2013:935736.

[292] Brand S, Niehaus K, Pühler A, Kalinowski J: **Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: The genes *cop1*, *cmt1*, and *cmt2* can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope**. *Archives of Microbiology* 2003, 180:33–44.

[293] Fiuza M, Letek M, Leiba J, Villadangos AF, Vaquera J, Zanella-Cléon I, Mateos LM, Molle V, Gil JA: **Phosphorylation of a novel cytoskeletal protein (RsmP) regulates rod-shaped morphology in *Corynebacterium glutamicum***. *Journal of Biological Chemistry* 2010, 285:29387–29397.

[294] Seibold GM, Wurst M, Eikmanns BJ: **Roles of maltodextrin and glycogen phosphorylases in maltose utilization and glycogen metabolism in *Corynebacterium glutamicum***. *Microbiology* 2009, 155:347–358.

[295] Cocaign-Bousquet M, Guyonvarch A, Lindley ND: **Growth rate-dependent modulation of carbon flux through central metabolism and the kinetic consequences for glucose-limited chemostat cultures of *Corynebacterium glutamicum***. *Applied and Environmental Microbiology* 1996, 62:429–436.

[296] Chao H, Zhou NY: **GenR, an IclR-type regulator, activates and represses the transcription of gen genes involved in 3-hydroxybenzoate and gentisate catabolism in *Corynebacterium glutamicum***. *Journal of Bacteriology* 2013, 195:1598–1609.

[297] Feng J, Che Y, Milse J, Yin YJ, Liu L, Rückert C, Shen XH, Qi SW, Kalinowski J, Liu SJ: **The gene *ncgl2918* encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in *Corynebacterium glutamicum***. *Journal of Biological Chemistry* 2006, 281:10778–10785.

[298] Teramoto H, Suda M, Inui M, Yukawa H: **Regulation of the expression of genes involved in NAD *de novo* biosynthesis in *Corynebacterium glutamicum***. *Applied and Environmental Microbiology* 2010, 76:5488–5495.

[299] Hwang BJ, Yeom HJ, Kim Y, Lee HS: ***Corynebacterium glutamicum* utilizes both transsulfuration and direct sulfhydrylation pathways for methionine biosynthesis**. *Journal of Bacteriology* 2002, 184:1277–1286.

[300] Hwang BJ, Park SD, Kim Y, Kim P, Lee HS: **Biochemical analysis on the parallel pathways of methionine biosynthesis in *Corynebacterium glutamicum***. *Journal of Microbiology and Biotechnology* 2007, 17:1010–1017.

[301] Ehira S, Ogino H, Teramoto H, Inui M, Yukawa H: **Regulation of quinone oxidoreductase by the redox-sensing transcriptional regulator QorR in *Corynebacterium glutamicum***. *Journal of Biological Chemistry* 2009, 284:16736–16742.

[302] Marienhagen J, Kennerknecht N, Sahm H, Eggeling L: **Functional analysis of all aminotransferase proteins inferred from the genome sequence of *Corynebacterium glutamicum***. *Journal of Bacteriology* 2005, 187:7639–7646.

[303] Radoš D, Carvalho AL, Wieschalka S, Neves AR, Blombach B, Eikmanns BJ, Santos H: **Engineering *Corynebacterium glutamicum* for the production of 2,3-butanediol**. *Microbial Cell Factories* 2015, 14:171.

[304] Mao Y, Fu J, Tao R, Huang C, Wang Z, Tang YJ, Chen T, Zhao X: **Systematic metabolic engineering of *Corynebacterium glutamicum* for the industrial-level production of optically pure d -(−)-acetoin**. *Green Chemistry* 2017, 19:5691–5702.

[305] Sahm H, Eggeling L, de Graaf AA: **Pathway analysis and metabolic engineering in *Corynebacterium glutamicum***. *Biological Chemistry* 2000, 381:899–910.

[306] Zelle E, Nöh K, Wiechert W: **Growth and production capabilities of *Corynebacterium glutamicum*: Interrogating a genome-scale metabolic network model**. In *Corynebacterium glutamicum*, edited by Burkovski A. Norfolk, UK: Caister Academic Press, 2015, 39–55.

[307] Zhang Y, Cai J, Shang X, Wang B, Liu S, Chai X, Tan T, Zhang Y, Wen T: **A new genome-scale metabolic model of *Corynebacterium glutamicum* and its application**. *Biotechnology for Biofuels* 2017, 10:169.

[308] Ladner T, Held M, Flitsch D, Beckers M, Büchs J: **Quasi-continuous parallel online scattered light, fluorescence and dissolved oxygen tension measurement combined with monitoring of the oxygen transfer rate in each well of a shaken microtiter plate**. *Microbial Cell Factories* 2016, 15:206.

[309] Flitsch D, Krabbe S, Ladner T, Beckers M, Schilling J, Mahr S, Conrath U, Schomburg WK, Büchs J: **Respiration activity monitoring system for any individual well of a 48-well microtiter plate**. *Journal of Biological Engineering* 2016, 10:14.

[310] Ladner T, Beckers M, Hitzmann B, Büchs J: **Parallel online multi-wavelength (2D) fluorescence spectroscopy in each well of a continuously shaken microtiter plate**. *Biotechnology Journal* 2016, 11:1605–1616.

[311] Kosa G, Shapaval V, Kohler A, Zimmermann B: **FTIR spectroscopy as a unified method for simultaneous analysis of intra- and extracellular metabolites in high-throughput screening of microbial bioprocesses**. *Microbial Cell Factories* 2017, 16:147.

[312] Luchterhand B, Nolten J, Hafizovic S, Schlepütz T, Wewetzer SJ, Pach E, Meier K, Wandrey G, Büchs J: **Newly designed and validated impedance spectroscopy setup in microtiter plates successfully monitors viable biomass online**. *Biotechnology Journal* 2015, 10:1259–1268.

[313] Gruber P, Marques MPC, Szita N, Mayr T: **Integration and application of optical chemical sensors in microbioreactors**. *Lab on a Chip* 2017, 17:2693–2712.

[314] Funke M, Buchenauer A, Mokwa W, Kluge S, Hein L, Müller C, Kensy F, Büchs J: **Bioprocess control in microscale: Scalable fermentations in disposable and user-friendly microfluidic systems**. *Microbial Cell Factories* 2010, 9:86.

[315] Funke M, Buchenauer A, Schnakenberg U, Mokwa W, Diederichs S, Mertens A, Müller C, Kensy F, Büchs J: **Microfluidic BioLector - microfluidic bioprocess control in microtiter plates**. *Biotechnology and Bioengineering* 2010, 107:497–505.

[316] Faust G, Janzen NH, Bendig C, Römer L, Kaufmann K, Weuster-Botz D: **Feeding strategies enhance high cell density cultivation and protein expression in milliliter scale bioreactors**. *Biotechnology Journal* 2014, 9:1293–1303.

[317] Barton NR, Burgard AP, Burk MJ, Crater JS, Osterhout RE, Pharkya P, Steer BA, Sun J, Trawick JD, van Dien SJ, Yang TH, Yim H: **An integrated biotechnology platform for developing sustainable chemical processes**. *Journal of Industrial Microbiology & Biotechnology* 2015, 42:349–360.

[318] Reetz MT, Kahakeaw D, Lohmer R: **Addressing the numbers problem in directed evolution**. *Chembiochem: A European Journal of Chemical Biology* 2008, 9:1797–1804.

[319] Craig T, Holland R, D'Amore R, Johnson JR, McCue HV, West A, Zulkower V, Tekotte H, Cai Y, Swan D, Davey RP, Hertz-Fowler C, Hall A, Caddick M: **Leaf LIMS: A flexible laboratory information management system with a synthetic biology focus**. *ACS Synthetic Biology* 2017, 6:2273–2280.

[320] Morrell WC, Birkel GW, Forrer M, Lopez T, Backman TWH, Dussault M, Petzold CJ, Baidoo EEK, Costello Z, Ando D, Alonso-Gutierrez J, George KW, Mukhopadhyay A, Vaino I, Keasling JD, Adams PD, Hillson NJ, Garcia Martin H: **The Experiment Data Depot: A web-based software tool for biological experimental data storage, sharing, and visualization**. *ACS Synthetic Biology* 2017, 12:2248–2259.

[321] Pollard J, McDonald P, Hesslein A: **Lessons learned in building high-throughput process development capabilities**. *Engineering in Life Sciences* 2016, 16:93–98.

[322] Eikmanns BJ, Thum-Schmitz N, Eggeling L, Lüdtke KU, Sahm H: **Nucleotide sequence, expression and transcriptional analysis of the *Corynebacterium glutamicum gltA* gene encoding citrate synthase**. *Microbiology* 1994, 140:1817–1828.

[323] van der Rest ME, Lange C, Molenaar D: **A heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogeneic plasmid DNA**. *Applied Microbiology and Biotechnology* 1999, 52:541–545.

[324] Petras I, Bednarova D: **Total least squares approach to modeling: A Matlab toolbox**. *Acta Montanistica Slovaca* 2010, 15:158–170.

[325] Morschett H, Wiechert W, Oldiges M: **Automation of a Nile red staining assay enables high throughput quantification of microalgal lipid production**. *Microbial Cell Factories* 2016, 15:34.

[326] Winkler UK, Stuckmann M: **Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*.** *Journal of Bacteriology* 1979, 138:663–670.

[327] Wong C, Sridhara S, Bardwell JCA, Jakob U: **Heating greatly speeds Coomassie blue staining and destaining**. *BioTechniques* 2000, 28:426–8, 430, 432.

[328] Benjamini Y, Hochberg Y: **Controlling the false discovery rate: A practical and powerful approach to multiple testing**. *Journal of the Royal Statistical Society, Series B* 1995, 57:289–300.

[329] Limberg MH, Joachim M, Klein B, Wiechert W, Oldiges M: **pH fluctuations imperil the robustness of *C. glutamicum* to short term oxygen limitation**. *Journal of Biotechnology* 2017, 259:248–260.

[330] Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M: **Orchestrating high-throughput genomic analysis with Bioconductor**. *Nature Methods* 2015, 12:115–121.

List of Figures

List of Tables

Appendices

A. Supporting material for Chapter 2

The presented MATLAB code for automated determination of growth rates from high-throughput MBR cultivations (cf. Figure A.1) was used to calculate growth rates from a BioLector raw data set of a previously reported characterization of a library of genome reduced *C. glutamicum* strains derived from wildtype ATCC1303 [55] or lysine producer DM1933 [160], detailed results are listed in Table A.1. When comparing the deviation of the automatically calculated growth rates to the data from literature, $Dev.\,[\%] = \left|1 - \frac{\mu_{autom.}}{\mu_{Lit.}}\right|$, the following is found: In most cases (13 out of 23) a deviation of 10% or less is seen, few values (8 out of 23) deviated by 15% or less, and only two growth rates show a higher deviation of 16% and 21%. Consequently, growth rates calculated from automated detection of the corresponding growth phase agree very well with literature, where manual effort was required, for example, to remove backscatter readings form the stationary phase before data fitting.

Table A.1.: Automated calculation of growth rates from different genome reduced *C. glutamicum* strains, with comparison to literature.

	Growth rate μ [h^{-1}]	
Strain	**Literature**	**Re-calculated from automated procedure (95% CI)**
DM1933	0.32 ± 0.01	0.30 (0.29 – 0.30)
DM1933 $\triangle$CGP1	0.30 ± 0.03	0.30 (0.29 – 0.31)
DM1933 $\triangle$CGP2	0.31 ± 0.03	0.32 (0.31 – 0.32)
DM1933 $\triangle$CGP3	0.31 ± 0.02	0.32 (0.31 – 0.32)
DM1933 $\triangle$CGP123	0.33 ± 0.03	0.33 (0.32 – 0.34)
GRLP16	0.31 ± 0.02	0.35 (0.34 – 0.35)
GRLP23	0.33 ± 0.02	0.31 (0.30 – 0.32)
GRLP41	0.31 ± 0.01	0.29 (0.28 – 0.30)
GRLP42	0.36 ± 0.06	0.30 (0.29 – 0.31)
GRLP46	0.30 ± 0.04	0.32 (0.32 – 0.33)
WT[#]	0.43 ± 0.04	0.46 (0.45 – 0.46)
		0.46 (0.45 – 0.46)
MB001	0.43 ± 0.04	0.48 (0.47 – 0.48)
GRS16	0.44 ± 0.03	0.49 (0.48 – 0.50)
GRS23	0.46 ± 0.02	0.47 (0.47 – 0.48)
GRS37	0.45 ± 0.03	0.51 (0.51 – 0.52)
GRS40	0.44 ± 0.03	0.50 (0.49 – 0.51)
GRS41	0.41 ± 0.04	0.47 (0.46 – 0.47)
GRS42	0.39 ± 0.01	0.47 (0.46 – 0.48)
GRS46	0.43 ± 0.03	0.50 (0.49 – 0.50)
GRS45	0.31 ± 0.02	0.34 (0.34 – 0.34)
GRS47	0.41 ± 0.03	0.48 (0.47 – 0.49)
GRS48	0.45 ± 0.01	0.49 (0.49 – 0.50)
GRS53	0.44 ± 0.03	0.47 (0.47 – 0.48)

Data originate from a BioLector raw data set and results are compared to previously reported data collections from literature [55, 160]. [#] For strain WT, two data sets from growth duplicates were processed. Data from literature is given as mean ± standard deviation. *CI* confidence interval.

```
1   ** Calculate BS blank from first BS signal with cell concentration below
2   limit of detection (LOD)
3   Blank := mean(BS[1:3])

4   ** Estimate measurement error for BS signal from first BS readings.
5   BS_Std := std(BS[1:3])

6   ** Find first measurement cycle where BS signal exceeds user defined
7   threshold for limit of quantification (LOQ)
8   for (i := 1 to n_Last) {
9     if (BS[i]) >= BSThres) {
10      n_Min := i
11      break
12    }
13  }

14  ** Set n_End := n_Last. Calculate growth rate from BS signals between
15  measurement cycle n_Min and n_End. If StopCriteria are not fulfilled, set
16  n_End := n_End-1 and repeat procedure.
17  n_End := n_Last
18  while (n_End > n_Min) {
19    BSSubset := BS[n_Min] to BS[n_End]
20    Blank BSSubset and make log-transformation
21    BSBlanked_Std := sqrt(BS_Std^2 + BS_Std^2)  ** Error propagation
22    Calculate growth rate by weighted linear regression (WLR) using
23      corresponding inverse squared BSBlanked_Stds
24    if StopCriteria
25      break
26    else
27      n_End := n_End-1
28  }

29  StopCriteria {
30    R2adj > 0.99
31    (BS[n_End] - BS[n_End-1]) > (BS[n_End-1] - BS[n_End-2])
32    (BS[n_End] - BS[n_End-1]) > 0
33    (BS[n_End-1] - BS[n_End-2]) > 0
34  }
```

Figure A.1.: Pseudocode for automated calculation of growth rates, using backscatter signal as example biomass signal. The procedure starts with the calculation of the blank (zero) value from first BS signals, where the cell concentration is below the limit of detection (LOD) (lines 1 to 3). These first BS signals below the LOD are also used to calculate the BS measurement error which is considered to be additive for all BS signals (lines 4 to 5). Next, the measurement cycle is identified where the BS signal exceeds the user defined limit of quantification limitation of quantification (LOQ) (lines 6 to 13). To detect the exponential growth phase from the complete BS data set, a BS subset is extracted ranging from measurement cycles where the BS signal reaches the LOQ to the last one. For this BS subset, the growth rate is calculated by WLR after ln-transformation and several stopping criteria are evaluated. If these criteria are not fulfilled, a new BS subset is evaluated from which the last measurement is removed. This procedure is repeated until the stopping criteria are fulfilled. Please note: this is implemented as for-loop in the MATLAB function (lines 14 to 28). Three stopping criteria are defined: A certain adjusted R^2 from the regression has to be reached. The last biomass increase has to be higher than the previous one, and these two increases need to have a positive sign (lines 29 to 34). Figure reprinted from [200].

B. Supporting material for Chapter 3

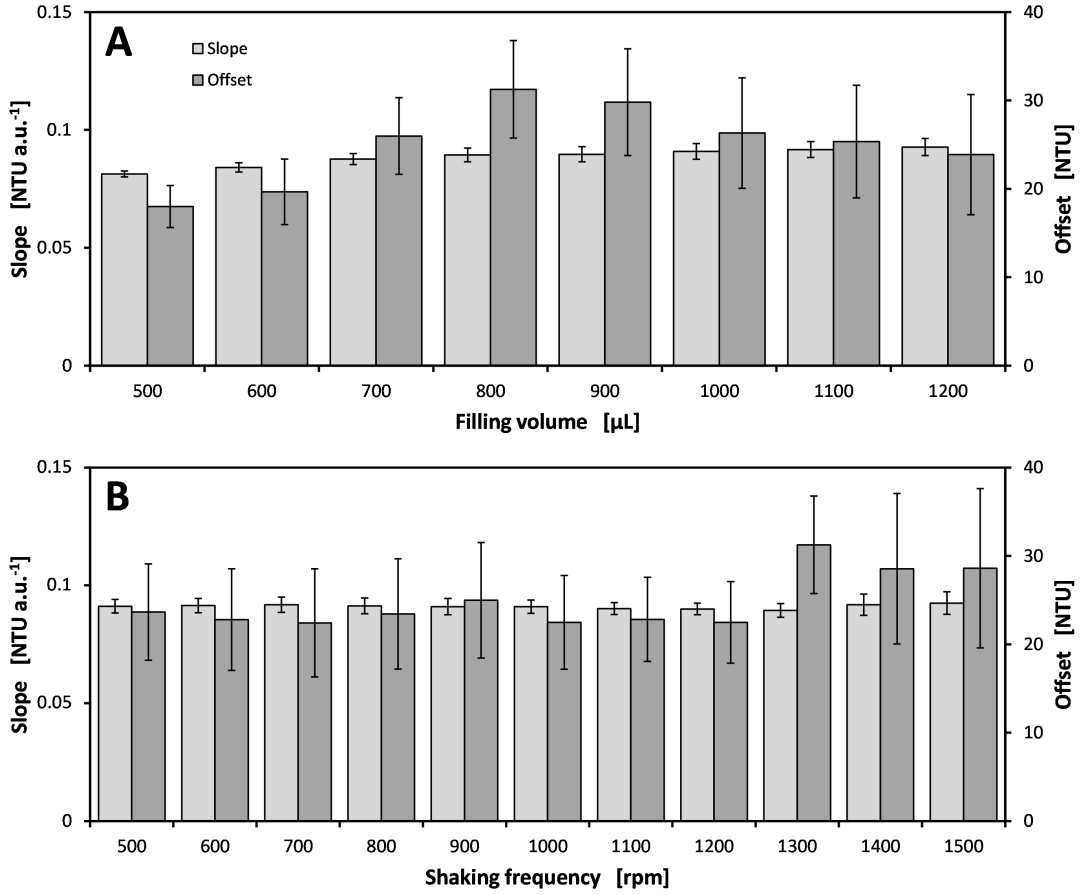


Figure B.1.: Check for linearity of BS signal for different filling volumes and shaking frequencies. (A): Variation of filling volume at a shaking frequency of 1300 rpm. (B): Variation of shaking frequencies at a volume of 800 µL. For each condition, a turbidity standard dilution series (100, 200, 500, 1000, 2000, and 4000 NTU) was applied and a OLS linear regression was calculated, resulting R^2 was always > 0.99.

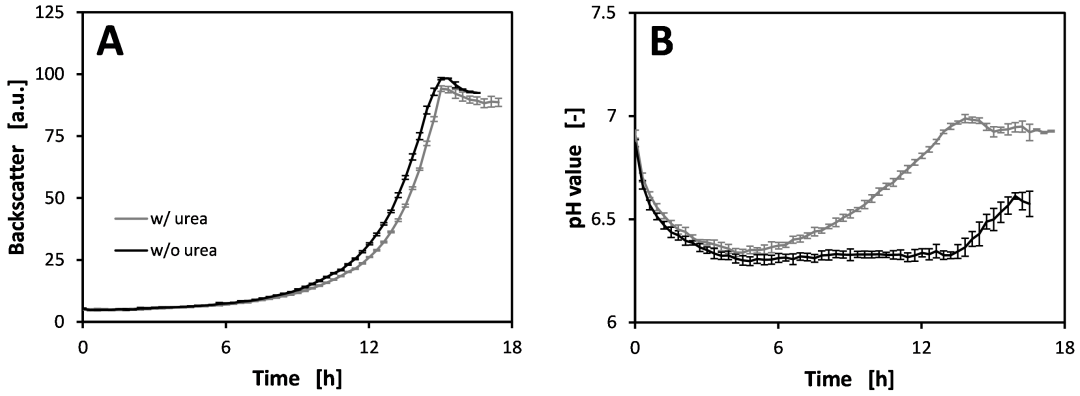


Figure B.2.: Impact of urea on biomass formation and pH development during growth of strain WT-NprE in CgXII medium. (A): Time courses of biomass growth. (B): Corresponding pH time courses. Cultures were induced with 100 µM IPTG upon reaching a BS value of 15 a.u., using an automated MBR protocol. Data represent mean and standard deviations from eight replicate cultivations, of which two were harvested 1, 2, 4 and 6 h after induction.

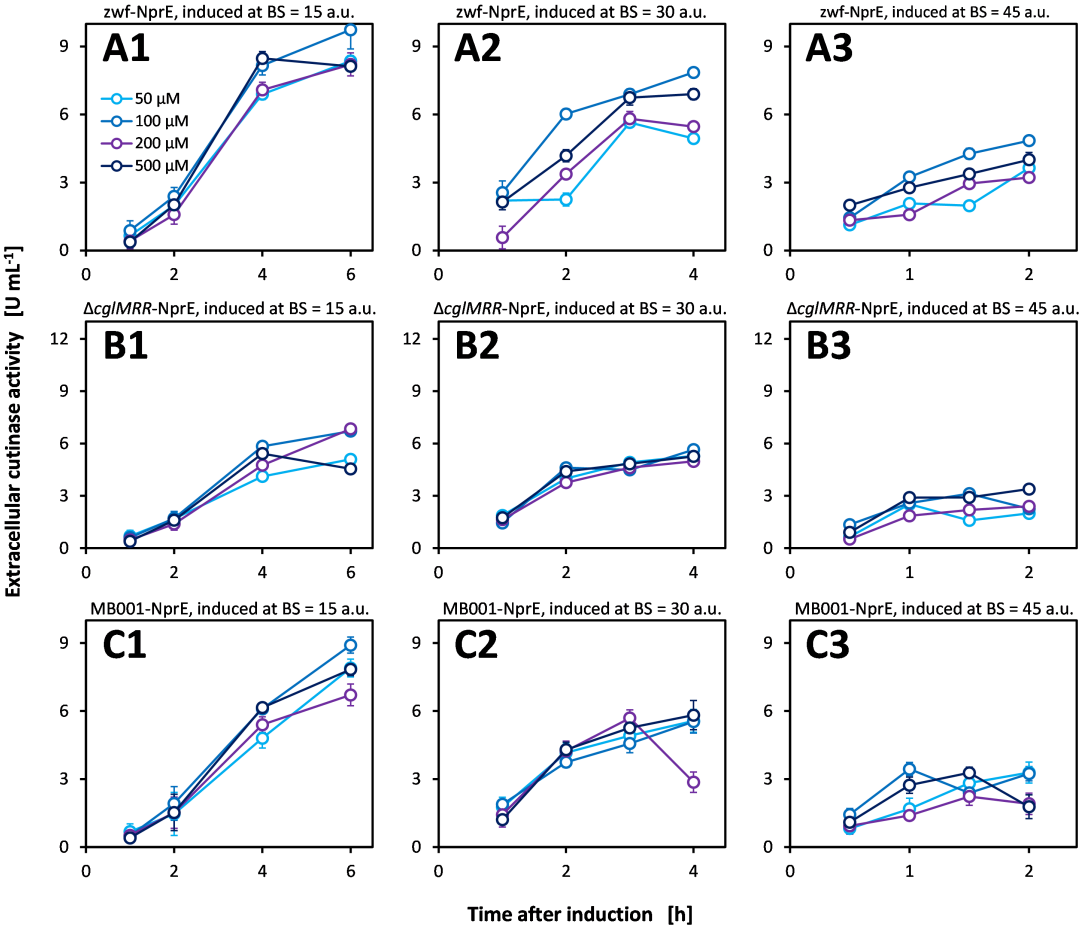


Figure B.3.: Comparative induction profiling for *C. glutamicum* cutinase secretion strains zwf^{FBR}-NprE, Δ*cglMRR*-NprE and MB001-NprE. Each panel shows the extracellular cutinase activity over time after induction, determined from BioLector MBR cultivations. Induction profiling conditions copmrise three strains, four IPTG concentrations and three different growth stages at induction. The latter parameter is represented by a BS signal of 15 a.u. (early growth phase), 30 a.u. (mid-time growth phase), and 45 a.u. (late growth phase), which serves as trigger signal for IPTG addition. Tested IPTG concentrations are 50, 100, 200 and 500 µM. Panels A, B, and C show time profiles of cutinase activities obtained with strains zwf^{FBR}-NprE, Δ*cglMRR*-NprE and MB001-NprE, respectively, while panels 1, 2, and, 3 show results for these strains induced at early, mid-time, and late growth phase, respectively. In each panel (i.e., for a specific combination of strain and growth stage where IPTG was added, as indicated at the top of each panel), results for the four tested IPTG concentrations are shown. The corresponding color coding is given in panel A1, which applies to all panels. For clarity reasons, biomass signals are not shown. Error bars indicate standard deviations from replicates (n = 3) of cutinase activity determination from each sample.

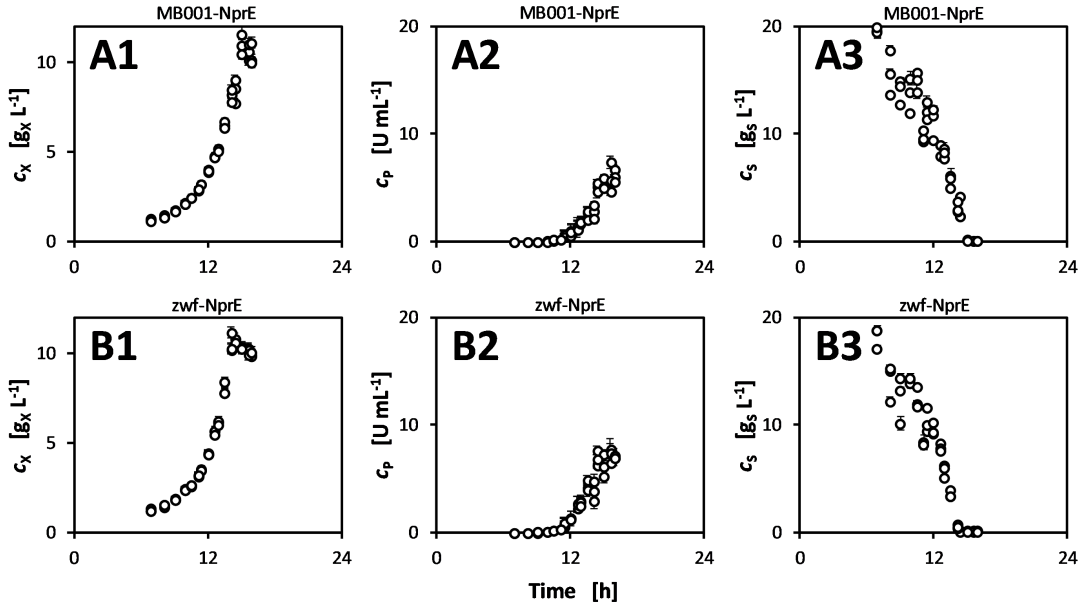


Figure B.4.: High-resolution time courses for cell dry weight, glucose and extracelluar cutinase activity (panels 1, 2 and 3, respectively) for strains MB001-NprE and zwf^{FBR}-NprE (panels A and B, respectively). For both strains, the NprE SP is used. For induction of each culture replicate, 100 µM IPTG was added upon reaching a BS signal of 15 a.u., and each strain was cultivated in 48 replicates. Cell dry weight is calculated from BS signal at time point of harvest via linear calibration. c_X cell dry weight. c_P cutinase activity. c_S glucose concentration.

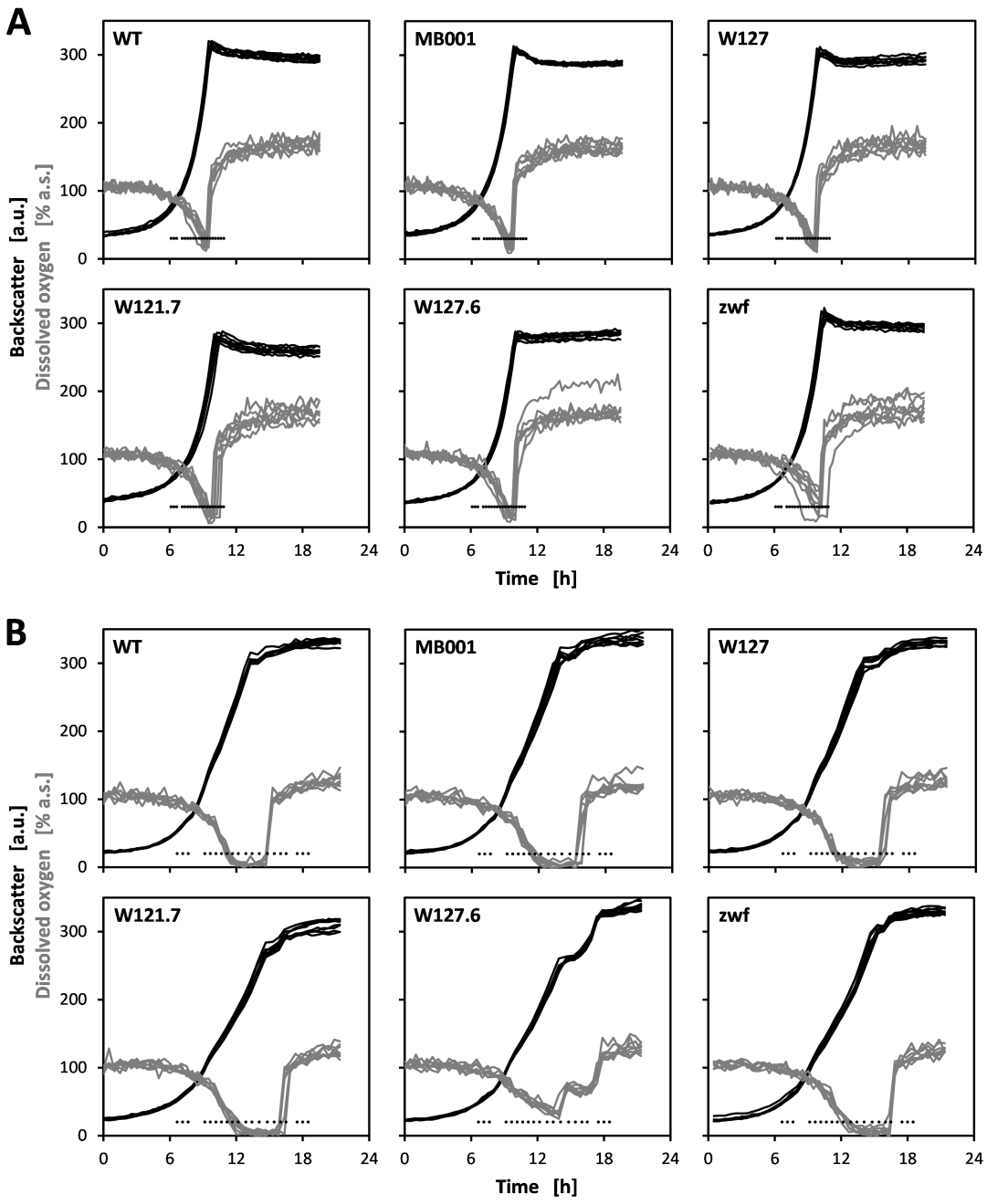


Figure B.5.: Impact of interrupted shaking on growth of *C. glutamicum* WT, MB001, W127, W121.7, W127.6, and zwf^{FBR} during repeated low-volume sampling. (A): Cultivations with stopped shaking for sample taking. (B): Cultivations where sampling was conducted without interruption of shaking. Black dots indicate time points of sampling.

C. Supporting material for Chapter 4

C.1. Step-by-step protocol

Part A: Conceiving the study and definition of methods (Figure 4.1, first box)

Note: *Definition of the optimization objective: Is a time course of product formation needed or is only a limited time interval or even a fixed time point relevant? Also, consider potential issues such as stability, effort of analytical quantification, or cultivation time. As an alternative to final protein titer, other objectives could be considered such as biomass or cultivation time. Biomass is considered by biomass specific product yield, while cultivation time is considered by space-time-yield. (Minimum) product quality could also be an aim. Multi-objective optimization may be required in certain situations, as discussed elsewhere [248]. In this study, GFP titer after 17 h of cultivation was chosen as the optimization objective. GFP fluorescence can be followed online using the equipment available in this study, which greatly simplifies determining the concentration of the model protein.*

Note: *Definition of the parameters to be optimized: CgXII medium consists of 16 individual components [56] and investigating all of these in a full factorial design would result in 2^{16} (i.e., > 65000) experiments. Consequently, the search space needs to be reduced on a rational and experience-driven basis. The selection of media components considered for optimization can be supported by available expert knowledge or literature data.*

1. **Decide which medium component concentrations should not be varied.** For the optimization of CgXII medium, the following components were chosen to be fixed:
    1.1. Glucose is fixed at 10 g L^{-1}. This optimization is trivial because more glucose yields more GFP secreting biomass. The aim was to reveal non-intuitive medium effects.
    1.2. 3-(N-morpholino)propanesulfonic acid (MOPS) is fixed at 42 g L^{-1}. This provides sufficient buffering capacity during batch cultivation and is not metabolized.
    Note: *The addition of MOPS at this final concentration ensures a starting value of pH 7, even with the different volumes of the other stock solutions added, which are not at pH 7. However, to exclude any deviation in starting pH values, the pH should be checked for medium compositions where all stock solutions are added at their maximum and minimum volumes.*
    1.3. KH_2PO_4 and K_2HPO_4 are fixed at 1 g L^{-1} each. These provide also buffering capacity, and serve as a phosphate source.

1.4. Urea is fixed at 5 g L^{-1}. This serves as a basal nitrogen source and pH-stabilizing agent, sufficient to prevent N-limitation.

1.5. Biotin is fixed at 0.2 mg L^{-1}. *C. glutamicum* ATCC13032 is auxotrophic for biotin.

1.6. Protocatechuic acid (PCA) is fixed at 30 mg L^{-1}. It serves as an iron chelating agent.

1.7. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) is fixed at 100 µM. It induces expression of the *gfp* gene.

2. **Select high and low concentrations of media components for initial sensitivity screenings.** Here, components from three typical groups of media components were chosen. The investigated low and high concentrations for each component are:

2.1. Standard nitrogen source: $(NH_4)_2SO_4$ (8 - 32 g L^{-1}); the variation of nitrogen source was reported to be promising for optimization of biomass-specific GFP signal [36].

2.2. Trace elements: $FeSO_4 \cdot 7\,H_2O$ (4 - 16 mg L^{-1}), $MnSO_4 \cdot H_2O$ (4 - 16 mg L^{-1}), $ZnSO_4 \cdot 7\,H_2O$ (0.4 - 1.6 mg L^{-1}), $CuSO_4 \cdot 5\,H_2O$ (125 - 501 µg L^{-1}), $NiCl_2 \cdot 6\,H_2O$ (8 - 32 µg L^{-1}) and $CoCl_2 \cdot 6\,H_2O$ (52 - 208 µg L^{-1}). The composition of trace elements is inherited from the first publication of CgXII medium [56]. In addition, $Na_2MoO_4 \cdot H_2O$ (26 - 104 µg L^{-1}) and H_3BO_3 (20-80 µg L^{-1}) were included as trace elements, as they are used in a published variant of CgXII medium [237].

2.3. Macro elements: $MgSO_4 \cdot 7\,H_2O$ (0.2 - 0.4 g L^{-1}), $CaCl_2 \cdot 2\,H_2O$ (5.3 - 21.2 mg L^{-1}). These were found among other divalent cations to enhance secretory GFP production in another study [227], where *C. glutamicum* strain R was used with a different medium background and the CgR0949 Tat signal peptide.

3. **Preparation of materials and definition procedures**

Note: *It is advised to define and fix the experimental methods at a detailed level. This standardization ensures that different results can be traced back to intrinsic biological variability or different medium compositions.*

Note: *Media stock solutions: Prepare individual stock solutions for all investigated media components. Prepare highly concentrated stocks to allow for the dilution of different volumes for all components. This means that after combining all stock solutions at their highest volume according to the design plan, this cannot exceed the final cultivation volume, cf. paragraph "Generation of medium stock solution pipetting list". If the addition of a highly concentrated solution results in a very low addition volume, e.g. < 1/100th of the final cultivation volume, dilute the stock solution accordingly. For instance, in this study a pipetting volume of less than 10 µL was defined to be infeasible. Typically, trace element solutions are concentrated as high as 1000-fold with respect to final standard concentration in the medium. It is advantageous to concentrate media components as multiples (x-fold) with respect to concentrations in the reference recipe. By doing so, infeasible pipetting volumes of fractional decimals are avoided. Detailed recipes for all stock solutions of CgXII medium are given in section C.3*

3.1. Working cell bank (WCB)

Note: *For each growth experiment, one WCB aliquot is used. If more aliquots are needed, use*

100 mL BHI medium in 1000 mL baffled shake flasks or inoculate multiple shake flasks, which are pooled before the addition of glycerol solution.

3.1.1. Prepare single colonies of the recombinant expression strain *C. glutamicum* pCGPhoD^{Bs}-GFP [239] on agar plates (37 g L^{-1} BHI powder, 20 g L^{-1} agar, 25 mg L^{-1} kanamycin). Plating material can either come from fresh transformation or from a cryopreserved aliquot. Incubate at 30 °C until the appearance of single colonies; this takes usually one to two days.

3.1.2. Inoculate a shaker flask culture (50 mL BHI medium with 25 mg L^{-1} kanamycin, 500 mL baffled flask, 250 rpm, 25 mm shaking diameter, 30 °C) with colony material and incubate overnight (approximately 16 h).

3.1.3. Combine one volume of resulting cell suspension with one volume of 500 g L^{-1} sterile glycerol solution and distribute in 2 mL aliquots in sterile cryopreservation vials. Store at -80 °C.

3.2. MBR cultivation protocol

Note: *For the employed BioLector MBR system in this study, scattered light (biomass) and GFP fluorescence are intensity measurements, which need a certain gain value assigned. The higher the gain, the higher the optical signal is amplified; this is also why scattered light and fluorescence are measured in arbitrary units (a.u.). Determine suitable gain values for biomass and GFP detection in preliminary experiments, along with suitable shaking frequency and filling volume to avoid oxygen limitation at higher biomass concentrations during later process stages. The employed cultivation conditions ("Flowerplates", i.e., flower-shaped 48-well MTPs, shaking frequency of 1200 rpm, filling volume of 1000 µL, 10 g L^{-1} glucose as main carbon source) ensured oxygen unlimited cultivations. For maximum oxygen transfer rates resulting from other combinations of filling volumes and shaking frequencies in flower-shaped 48-well MTPs, datasheets from the supplier are available. Also, growth defects of individual cultures may occur due to low amounts of secondary substrates (e.g., nitrogen, trace elements). Therefore, check the biomass concentrations at the end of cultivations. In this study, no such growth effects were observed.*

3.2.1. Define the cultivation protocol for the MBR system ("BioLector") as follows:

3.2.1.1. Filterset 1: Biomass, gain 14. Filterset 2:pH, gain is preset. Filterset 3: pO2, gain is preset. Filterset 4: GFP, gain 80.

3.2.1.2. Shaking frequency: 1200 rpm.

3.2.1.3. Temperature: 30 °C.

3.2.1.4. Cycle time: 15 min.

3.2.1.5. Experiment time: manual (Cultivation will not be stopped automatically).

3.3. Generation of medium stock solution pipetting list

Note: *Almost any liquid handling robot system is capable to read pipetting actions from external files. Basically, the minimal information needed is the transfer volume and position of reagent source, as well as the destination of each represented by a labware position on the robotic deck and a specific cavity within the labware. However, different syntaxes for different liquid handling systems must be considered. Figure C.1 shows an example file structure for the employed pipetting*

system in this study.

3.3.1. Four types of stock solutions are pipetted into each cultivation well:

3.3.1.1. Stock solutions of media components that are varied ("*Variation Stocks*").

3.3.1.2. Water to compensate for different cumulative volumes of above stocks.

3.3.1.3. A stock solution ("*Rest Stock*") containing all components that are fixed. This stock solution can be composed from stocks containing the different components that are, e.g., stored at different temperatures or sterilized by different methods.

3.3.1.4. Inoculum, which should be added as the last component and put onto the worktable just before addition to avoid settling of the cells.

3.3.2. Per cultivation well, calculate the volumes to be transferred for all media components i according to: $V_i = V_{tot} \cdot \frac{c_{i,Stock}}{c_{i,Target}}^{-1}$. The volume to be added for certain *Variation Stocks* maybe zero, i.e., when this specific component is omitted.

3.3.3. Revise all calculated volumes V_i for suitable numbers. Pipetting increments in volume steps should not be too small. If necessary, adjust the concentrations of *Variation Stocks*. For instance, the minimal pipetting volume here was defined as 10 µL, and the minimal increment was defined as 5 µL. In general, these volumes should be defined based on experimentally determined precision and accuracy for the used liquid handling station.

3.3.4. Calculate the volume of *Rest Stock* containing the fixed components j for all wells as follows: $V_i = V_{tot} \cdot \frac{c_{i,Stock}}{c_{i,Target}}^{-1}$, where the maximum cumulative volume of *Variation Stock* $\left(\sum_i V_i\right)_{max}$ is used. Consequently, calculate the required concentrations of the fixed components in *Rest Stock* as follows: $c_{j,Rest} = c_j \cdot \frac{V_{tot}}{V_{RestStock}}$. Prepare *Rest Stock* accordingly.

3.3.5. Per cultivation well, calculate the volume of water to be added as follows: $V_{H_2O} = V_{tot} - V_{RestStock} - V_{Inok} - \sum_i V_i$. Per cultivation well, list all volumes to be added in the following order of addition: V_{H_2O}, $V_{RestStock}$, V_i, V_{Inok}. Format the pipetting list according to the specifications of the liquid handling station, as shown for an example in Figure C.1.

4. Seed culture, automated media preparation, and start of main culture

4.1. Create a protocol for the liquid handling robot. See Figure C.2 and Figure **??** for an example protocol for a "Janus" system, implemented in corresponding "WinPREP" software. The protocol should consider the following aspects:

4.1.1. Include a sufficient runtime of sterile housing prior to media preparation.

4.1.2. Include an initial excessive cleaning and washing of all pipetting tips and tubing.

4.1.3. Choose appropriate labware containers for stock solutions. Here, deep well plates with 12 columns are suitable for storage of Variation Stocks, as all eight pipetting tips can dip in a parallel arrangement into the well columns. This greatly speeds up media preparation. If 15 mL or 50 mL reagent tubes are used as reservoirs, only one pipetting tip can dip into it at once. For other stocks like water and *Rest Stock*, 100 mL troughs are used, as these stock solutions require higher volumes in total.

4.1.4. Provide a sufficient total volume for each stock solution to compensate for waste volumes, height pipetting offsets, etc.

4.1.5. Insert a user prompt before the inoculation step, ensuring that this step is carried out immediately prior to the seed culture procedure.

4.2. Prepare all stock solutions in a sterile manner and store until use, in appropriate containers, e.g., sterile 15 mL and 50 mL test tubes.

4.3. Sterilize the deep well plates for stock solution storage on a worktable, e.g., by wiping with 70% ethanol and subsequent drying in a laminar flow hood.

4.4. Start the seed culture by inoculating 50 mL BHI medium containing 25 mg L^{-1} kanamycin with one aliquot from the WCB. Before the MBR cultivation, prepare fresh, vital inoculum from exponentially growing seed cultures in a sufficient amount.

4.5. Place all necessary labware on the robotic worktable and pour stock solutions in the corresponding labware.

4.6. Start the robotic workflow for media preparation, so that the last step (inoculation), is reached in time with the start of seed culture. The total runtime of the robotic workflow needs to be evaluated previously.

4.7. Sample the seed culture after approximately 2 h, then each hour, to monitor the growth by optical density (OD_{600}). After approximately 5 h, the culture reaches 3-4 OD_{600} and is used to inoculate the main cultures.

4.8. Place the seed culture on the liquid handler worktable and continue the media preparation protocol. Seal cultivation MTP after inoculation.

4.9. Place the sealed cultivation MTP in the BioLector device and start the pre-defined cultivation protocol.

4.10. Dispose the remaining stock solutions, according to biosafety regulations if necessary, and seed the culture from the robotic worktable. Clean the re-usable labware and start the liquid handler decontamination protocol.

5. Product quantification and preprocessing of raw data for analysis

5.1. Stop the main culture after a 17 h runtime.

5.2. Transfer the measurement data from the MBR device to the connected computer using the associated software package according to the manufacturer's user manual.

5.3. Use the "Data Management -> Transform Data" function of the software package that accompanies the MBR system to convert the raw data file into an easy access spreadsheet format. Copy the GFP signal data for all cultivation wells from the timestamp column closest to 17 h. Identify the GFP signals from the reference cultivation wells and the average. Normalize all remaining GFP signal by the average reference signal.

5.4. Quantify the GFP titer using additional methods (if required)

5.4.1. Transfer the cell suspensions from all cultivation wells into prelabeled reaction tubes and

obtain the cell-free supernatant after centrifugation for 10 min at maximum speed using a benchtop centrifuge.

5.4.2. Transfer 200 µL of each cell-free supernatant into a black 96-well MTP with transparent bottom, and read the GFP specific fluorescence at 488/520 nm using an appropriate microplate reader. Normalize the GFP signal from all wells with average signal from reference cultivations.

Note: *The absolute GFP fluorescence signals cannot be compared for different measurement devices. Therefore, the 5 reference strains from all main cultivations serve as an internal standard. The improvement of GFP titer can be expressed in relation to the internal standard. This allows for comparison of results from different experiments and different GFP quantification methods (see next two steps).*

5.4.3. Determine the protein content of the cell-free supernatants with standard protocols, e.g., Bradford or BCA assay. In combination with the results on GFP signal, the determination of a protein specific GFP titer is possible.

Note: *After fluorescence measurement, use the sample directly from this microplate. Use of multi-channel pipettes greatly facilitates the necessary liquid handling steps.*

5.4.4. Perform an SDS-Page visualization of cell-free supernatants following standard protocols to verify that most of the increase protein content is due to increased GFP secretion.

Note: *SDS-Page is more laborious than the previously mentioned methods, especially with a high sample load. For verification purposes, it is often sufficient to run SDS-Pages of cell-free supernatants from reference medium and final optimized medium cultivations [167].*

Part B: Sensitivity analysis (Figure 4.1, second box)

Note: *The goal of this part is to identify the important factors that have a significant effect on the objective.*

1. Choose an initial concentration range of media components. The reference medium composition should lie inside the chosen concentration ranges.
2. Choose an appropriate DoE. Such designs can be found in literature, e.g., from the NIST/SEMATECH e-Handbook of Statistical Methods [1]. The chosen design depends on number of media components of interest (here, 11) and the number of performed experiments (here, 32). In the given example, design "2_{IV}^{11-6}" was chosen that results in 32 experiments and allows estimation of the effect of increasing the concentration of one of the eleven components on the objective of interest. To be more specific, the main effects are not confounded with pair-wise factor interactions. They can be confounded with higher order interactions that are, however, most likely not significant.
3. Use the remaining wells (here, 16) for performing multiple replicates with the reference medium to assess the reproducibility of the process. Replicates should be equally spread over the plate to discover

[1] available at www.itl.nist.gov/div898/handbook/pri/section3/pri3347.htm

positional effects. The mean value of the measured output from the reference experiments is used for normalization. That is, each measured output of the sensitivity analysis is divided by the mean value of the reference output.

4. Calculate the main and, if possible, the combinatorial effects using appropriate statistics software. The classic design of experiment is based on a polynomial approximation [223]. For instance, the MATLAB function `mvregress` can be used for the estimation of the polynomial coefficients. The `mvregress` function is part of the statistics and machine learning toolbox.
5. Identify the relevant effects of various media components on the objective using a *t*-test. In the example, NH_4^+ has a significant negative effect, and Ca^{2+} and Mg^{2+} show the strongest positive effects (Figure C.4). Because GFP signal was normalized, the coefficients represent the average relative change in the GFP signal when increasing concentration of the particular component from its center value, $\frac{Max-Min}{2} + Min$, to its maximum value. Fixed components without a relevant effect are at their reference value during the optimization.

Part C: Iterative optimization (Figure 4.1, third box)

Note: *The MATLAB tool KriKit was used for the interpretation and statistical data analysis [233]. KriKit allows the user to construct a data-driven Kriging model. This Kriging model predicts the functional relationship between the media components and the objective. It also provides information on the prediction uncertainty. High uncertainty indicates noisy data and/or insufficient data density.*

1. Design of Experiments

   Note: *New experiments are iteratively designed, based on the results of the previous run.*

   1.1. In the first iteration, design new experiments for detailed investigation of identified media components of interest. Experimental results from Part B cannot be transferred to the iterative optimization (Part C), as the concentrations of components without relevant effect are now fixed to the respective reference value. Consequently, experiments from the sensitivity analysis and the iterative optimization are not comparable.

   1.2. Otherwise, follow the scheme illustrated in Figure 4.1, frame "Iterative optimization". If the potential optimum lies inside the defined concentration range, new experiments are designed using the expected improvement [225, 248]. The experimental design based on the expected improvement is integrated in the KriKit toolbox. If the optimum lies on the boundary, expand the concentration range.

2. Perform experiments at the designed sample points according to experimental methods defined in Part A.
3. Statistical analysis

   3.1. Construct a Kriging model using the combined data from all iterations, including the current.

   3.2. Investigate the model output by visualization using the comprehensive tools of KriKit (2/3D interpolation, movie analysis, screening plot, etc.).

4. If an optimal area was estimated with sufficient accuracy, stop optimization. If not, continue.
   Note: *Figure C.5 illustrates the iterative optimization for the test study. The concentration range was successively expanded until a plateau was found (iterations 1 to 6). The seventh iteration was used for exploring the boundaries in more detail.*

Part D: Verification of results (Figure 4.1, fourth box)

Note: *After finishing the iterative optimization, the initial assumptions need to be checked for validity.*

1. Redo sensitivity analysis for optimal medium composition. That is, the concentrations of components that are investigated in Figure 4.1, third box, are fixed to their optimal values. Concentrations of other components are varied according to an appropriate DoE.
   Note: *Similar results in both screenings indicate that the concentration levels of the investigated components do not alter the effect of the other components. If the screening shows significant differences to the results from part B, components with a changed effect should be added to the pool of investigated components and part C should be repeated.*
2. In the case of indirect measurement (e.g., fluorescence as indicator for product concentration), apply orthogonal measurement approaches (e.g., activity assay, Bradford or BCA protein quantification, SDS page) to confirm the change in the objective of interest by comparing results from the reference medium and the optimized medium [167].

C.2. Additional figures

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		Sln-01	Sln-02	Sln-03	Sln-04	Sln-05	Sln-06	Sln-07	Sln-08	Sln-09	Sln-10	Sln-11	Sln-12	Sln-13	Sln-14	Sln-15
2	ID	Stock	H2O	Fe-40X	Mn-40X	Zn-40X	Cu-40X	NH4-20X	betain-40X	Ni-40X	Co-40X	Mo-40X	BO-40X	Ca-40X	Mg-40X	Inoculum
3	#1	300	425	25	25	25	25	50	0	25	0	0	0	25	25	50
4	#2	300	310	10	40	40	40	80	0	40	40	0	0	40	10	50
5	#3	300	390	40	10	40	10	20	0	40	40	0	40	10	10	50
6	#4	300	425	25	25	25	25	50	0	25	0	0	0	25	25	50
7	#5	300	370	10	40	40	40	20	0	40	0	0	40	10	40	50
8	#6	300	425	25	25	25	25	50	0	25	0	0	0	25	25	50
9	#7	300	340	10	40	10	40	20	40	10	40	40	40	10	10	50
10	#8	300	425	25	25	25	25	50	0	25	0	0	0	25	25	50
11	#9	300	290	10	10	40	10	80	40	40	0	40	40	40	10	50
12	#10	300	425	25	25	25	25	50	0	25	0	0	0	25	25	50

Figure C.1.: Screenshot from the volume pipetting list for sensitivity analysis. Entries in the first column assign a unique identifier to all volumes of a row; this identifier is the MTP well number of the target cultivation MTP on the liquid handler worktable, cf. Figure C.3C. Remaining columns encode volumes for different solutions ("Sln-01" to "Sln-15") to be pipetted. The cumulative volume of one row corresponds to the final cultivation volume of the corresponding well.

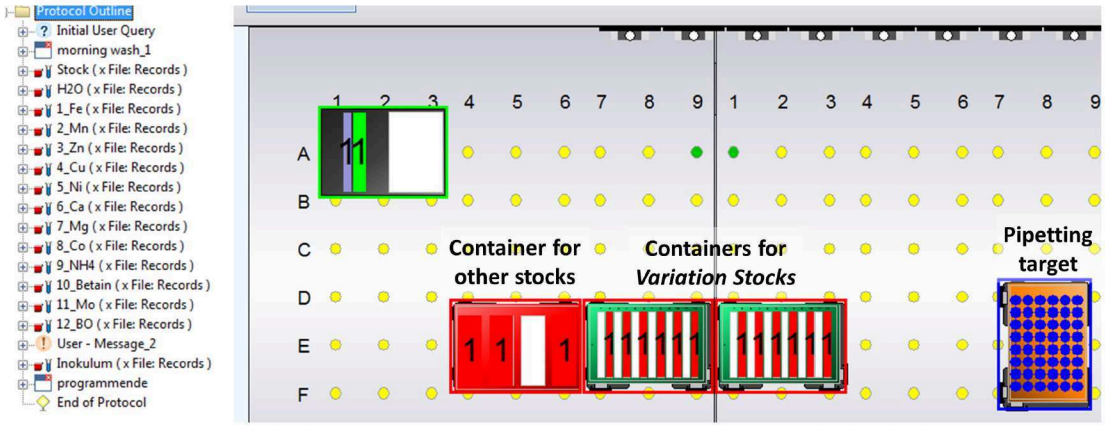


Figure C.2.: Screenshot from the liquid handling control software "WinPREP". Left: Row-ordered commands, including a transfer command for each stock solution to be pipetted. Before the final command for inoculum addition, a user prompt is inserted to ensure the seed culture is placed at the table just in time. Right: Schematic of the worktable, including the source labware for *Variation Stocks* (two deep well plates with 12 column-like wells), the reagent trough for *Rest Stock*, water and inoculum, and the media preparation target cultivation MTP.

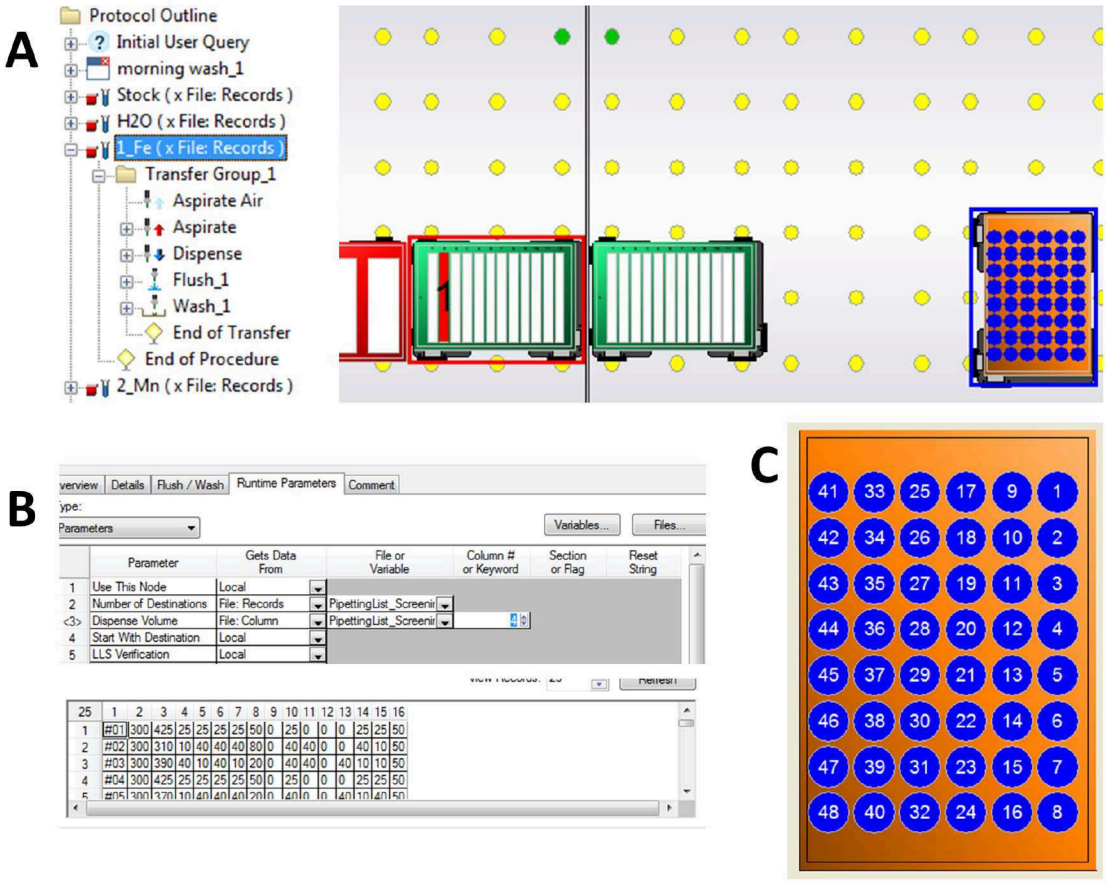


Figure C.3.: Compilation of detailed screenshots for setup of pipetting of a stock solution. (A): Unwrapped command for pipetting of Fe stock. Source labware and source well within are marked on the worktable by the read frame and red column of the corresponding deep well plate. Destination labware and destination wells within are marked by the blue frame around and blue wells of the target cultivation MTP. (B): Detailed example view on assignment of pipetting volumes for this step (Fe stock solution). Number of destinations is read from the pipetting list, which has 48 rows. The dispense volumes for all destination wells for Fe stock solution is found in column 4 in the pipetting list. Note that the first column in the pipetting list contains identifiers and not volumes to be transferred, see Figure C.1. (C): Details on destination well numbering. Volumes written in the row #1 of the corresponding pipetting list will be pipetted into well marked as #1, and so on. Wells #01, #08, #41 and #48 correspond to wells A01, A08, F01 and F08 for the alpha-numeric coding, which is also printed into the cultivation MTP itself.

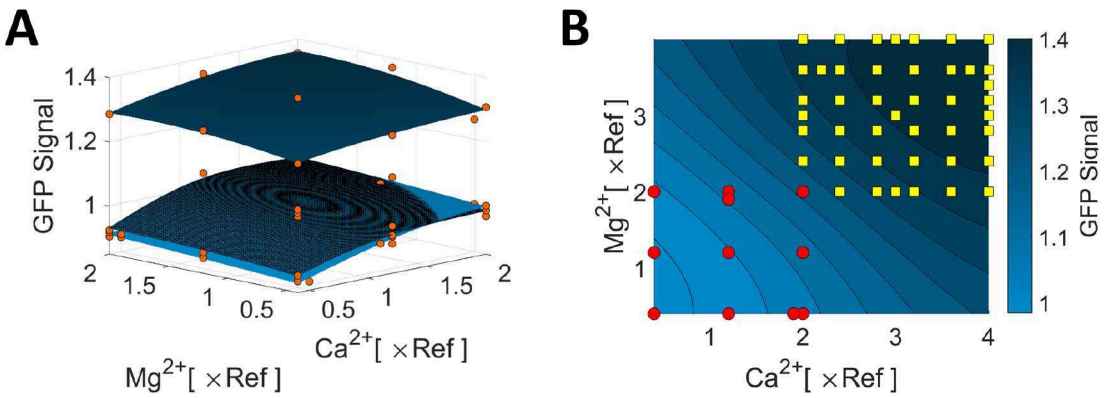


Figure C.4.: Detailed results from the first iteration. (A): Kriging interpolation based on experimental data of iteration 1. Red dots indicate the data set. For comparison, all three interpolation surfaces are overlayed in one plot (dark blue: 0x Ref, checkered: 1x Ref, light blue: 2x Ref). An alternative representation of the results can be found elsewhere [167]. (B): Kriging interpolation based on the experiments performed in iteration 1 (red dots) and iteration 2 (yellow squares). Parts of the data presented in this figure have been previously published [167].

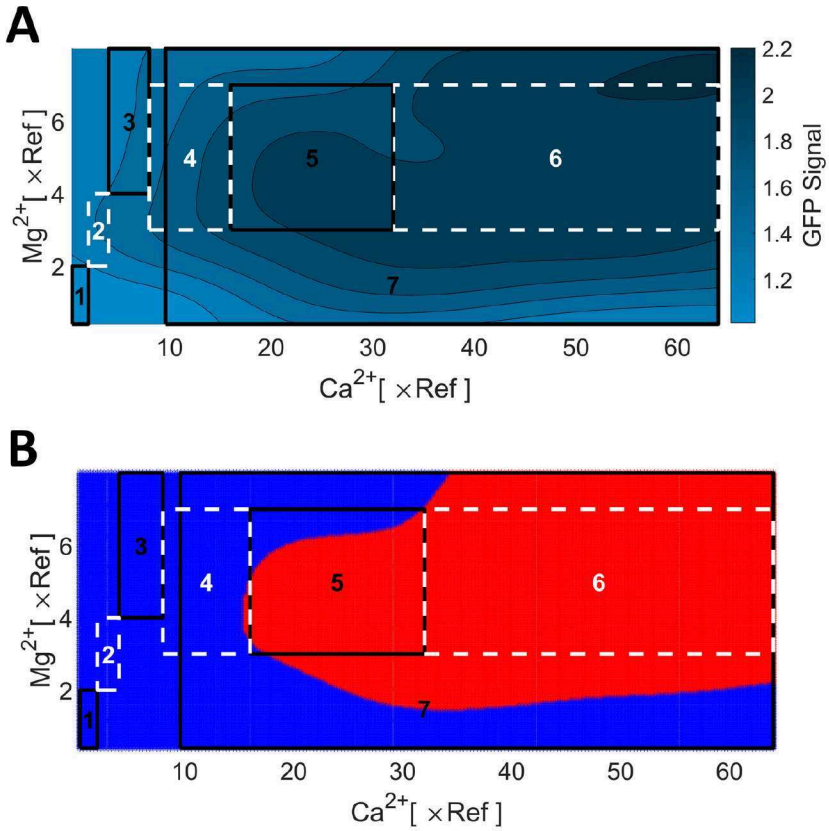


Figure C.5.: Summary of results from iterative workflow. (A): Final Kriging model prediction. (B): Statistical identification of optimal area (red) based on the statistical *z*-test, which is provided by KriKit. Boxes indicate successive steps of iterative design and execution of experiments. Parts of the data presented in this figure have been previously published [167].

C.3. Component stock solutions for automated media preparation

C.3.1. Stock solutions for preparation of "*Rest Stock*"

Biotin stock	**1000X**
Biotin	0.02 g
dH_2O	ad 100 mL
Sterilize by filtration, store in aliquots at -20 °C	

Glucose stock	**50X**
Glucose	250 g
dH_2O	ad 500 mL
Sterilize by autoclaving, store at RT	

IPTG stock	**200X**
IPTG	1.19 g
dH_2O	ad 50 mL
Sterilize by filtration, store in aliquots at -20 °C	

Kanamycin stock	**1000X**
Kanamycin	1.25 g
dH_2O	ad 50 mL
Sterilize by filtration, store in aliquots at -20 °C	

PCA stock	**500X**
PCA	0.75 g
dH_2O	ad 50 mL
Sterilize by filtration, store in aliquots at -20 °C	

Salts stock	**4X**
Urea	10 g
KH_2PO_4	2 g
K_2HPO_4	2 g
MOPS	84 g
4 M NaOK	adjust pH to 7
dH_2O	ad 500 mL
Sterilize by autoclaving, store at RT	

Rest stock	**3.33X**
Salts stock 4X	75 mL
PCA stock 500X	0.6 mL
Kanamycin stock 1000X	0.3 mL
IPTG stock 200X	1.5 mL
Biotin stock 1000X	0.3 mL
Glucose stock 50X	6 mL
dH_2O, sterile	ad 90 mL

C.3.2. Stock solutions for varied components ("*Variation Stocks*")

Betaine stock	**40X**
Betaine	1 g
dH_2O	ad 50 mL
Sterilize by filtration, store at RT	

BO stock	**40X**
H_3BO_3	1 g
dH_2O	ad 1000 mL, yields 1000X BO solution
1000X BO solution	16 mL
dH_2O	84 mL, yields 160X BO solution
160X BO solution	25 mL
dH_2O	75 mL, yields 40X BO solution
Sterilize by filtration, store at RT	

Ca stock	**40X**
$CaCl_2 \cdot 2\,H_2O$	0.994 g
dH_2O	ad 75 mL, yields 1000X Ca solution
1000X Ca solution	16 mL
dH_2O	84 mL, yields 160X Ca solution
160X Ca solution	25 mL
dH_2O	75 mL, yields 40X Ca solution
Sterilize by filtration, store at RT	

Co stock	**40X**
$CoCl_2 \cdot 2\,H_2O$	0.013 g
dH_2O	ad 100 mL, yields 1000X Co solution
1000X Co solution	16 mL
dH_2O	84 mL, yields 160X Co solution
160X Co solution	25 mL
dH_2O	75 mL, yields 40X Co solution
Sterilize by filtration, store at RT	

Cu stock	**40X**
$CuSO_4 \cdot 5\,H_2O$	0.0313 g
dH_2O	ad 100 mL, yields 1000X Cu solution
1000X Cu solution	16 mL
dH_2O	84 mL, yields 160X Cu solution
160X Cu solution	25 mL
dH_2O	75 mL, yields 40X Cu solution
Sterilize by filtration, store at RT	

Fe stock	**40X**
$FeSO_4 \cdot 7\ H_2O$	0.04 g
dH_2O	ad 100 mL
Sterilize by filtration, store at RT	

Mg stock	**40X**
$MgSO_4 \cdot 7\ H_2O$	18.74 g
dH_2O	ad 75 mL, yields 1000X Mg solution
1000X Mg solution	16 mL
dH_2O	84 mL, yields 160X Mg solution
160X Mg solution	25 mL
dH_2O	75 mL, yields 40X Mg solution
Sterilize by filtration, store at RT	

Mn stock	**40X**
$MnSO_4 \cdot H_2O$	0.04 g
dH_2O	ad 100 mL
Sterilize by filtration, store at RT	

Mo stock	**40X**
$Na_2MoO_4 \cdot 2\ H_2O$	0.065 g
dH_2O	ad 1000 mL, yields 1000X Mo solution
1000X Mo solution	16 mL
dH_2O	84 mL, yields 160X Mo solution
160X Mo solution	25 mL
dH_2O	75 mL, yields 40X Mo solution
Sterilize by filtration, store at RT	

NH_4^+ stock	**40X**
$(NH_4)_2SO_4$	40 g
dH_2O	ad 100 mL
Sterilize by filtration, store at RT	

Ni stock	**40X**
$NiCl_2 \cdot 6\ H_2O$	0.02 g
dH_2O	ad 1000 mL, yields 1000X Ni solution
1000X Ni solution	16 mL
dH_2O	84 mL, yields 160X Ni solution
160X Ni solution	25 mL
dH_2O	75 mL, yields 40X Ni solution
Sterilize by filtration, store at RT	

Zn stock	**40X**
$ZnSO_4 \cdot 7\ H_2O$	0.04 g
dH_2O	ad 100 mL, yields 400X Zn solution
400X Zn solution	10 mL
dH_2O	90 mL, yields 40X Zn solution
Sterilize by filtration, store at RT	

D. Supporting material for Chapter 5

Table D.1.: Comparison of observed growth rates μ_{Exp} for five cutinase secretion strains denoted by the applied Sec SP (AmyE, Epr, NprE, YpjP, YwmC) for different bioprocess conditions.

	Observed specific growth rate μ_{Exp} [h^{-1}] for *C. glutamicum* strains				
Bioprocess condition	**AmyE**	**Epr**	**NprE**	**YpjP**	**YwmC**
Batch, 1000 µM IPTG	0.34 ± 0.00	0.33 ± 0.01	0.32 ± 0.01	0.32 ± 0.01	0.34 ± 0.02
μ_{Set} = 0.10, 1000 µM IPTG	0.09 ± 0.00	0.09 ± 0.00	0.08 ± 0.01	0.10 ± 0.00	0.09 ± 0.00
μ_{Set} = 0.15, 1000 µM IPTG	0.14 ± 0.00	0.14 ± 0.00	0.13 ± 0.01	0.14 ± 0.01	0.14 ± 0.00
μ_{Set} = 0.20, 1000 µM IPTG	0.20 ± 0.01	0.21 ± 0.00	0.20 ± 0.00	0.18 ± 0.00	0.19 ± 0.00
μ_{Set} = 0.25, 1000 µM IPTG	0.25 ± 0.01	0.24 ± 0.00	0.24 ± 0.00	0.24 ± 0.00	0.24 ± 0.00
μ_{Set} = 0.10, 250 µM IPTG	0.08 ± 0.00	0.08 ± 0.00	0.08 ± 0.00	0.09 ± 0.00	0.08 ± 0.00
μ_{Set} = 0.15, 250 µM IPTG	0.14 ± 0.00	0.13 ± 0.00	0.13 ± 0.02	0.14 ± 0.00	0.13 ± 0.00
μ_{Set} = 0.20, 250 µM IPTG	0.20 ± 0.00	0.19 ± 0.00	0.20 ± 0.00	0.18 ± 0.01	0.19 ± 0.00
μ_{Set} = 0.25, 250 µM IPTG	0.24 ± 0.00	0.24 ± 0.00	0.24 ± 0.00	0.22 ± 0.00	0.24 ± 0.01
μ_{Set} = 0.10, 50 µM IPTG	0.10 ± 0.00	0.07 ± 0.00	0.07 ± 0.00	0.09 ± 0.01	0.10 ± 0.00
μ_{Set} = 0.15, 50 µM IPTG	0.14 ± 0.00	0.13 ± 0.00	0.13 ± 0.00	0.14 ± 0.01	0.14 ± 0.01
μ_{Set} = 0.20, 50 µM IPTG	0.21 ± 0.01	0.18 ± 0.00	0.19 ± 0.00	0.18 ± 0.00	0.19 ± 0.00
μ_{Set} = 0.25, 50 µM IPTG	0.25 ± 0.00	0.23 ± 0.00	0.23 ± 0.00	0.24 ± 0.00	0.23 ± 0.00

Tested bioprocess conditions comprise batch cultivations as well as fed-batch cultivations with glucose added in a growth-limiting manner at different exponentially increasing feeding rates for growth rate adjustment (μ_{Set} = 0.10, 0.15, 0.20 and 0.25 h^{-1}). Fed-batch bioprocesses were conducted without IPTG added (strains WT, NprE and YwmC) and IPTG concentrations of 50, 250 and 1000 µM (all strains), as indicated. In batch cultivations, 1000 µM IPTG was added. Data represent mean with deviations from min/max values as errors from two biological replicates. *n.t.* not tested.

Table D.2.: Comparison of observed cutinase yields $Y_{P/X}$ for five cutinase secretion strains, denoted by the applied Sec SP (AmyE, Epr, NprE, YpjP, YwmC), for different bioprocess conditions.

	Biomass specific cutinase yield $Y_{P/X}$ [kU g_X^{-1}] for *C. glutamicum* strains				
Bioprocess condition	**AmyE**	**Epr**	**NprE**	**YpjP**	**YwmC**
Batch, 1000 µM IPTG	0.33 ± 0.01	0.92 ± 0.09	0.86 ± 0.15	0.56 ± 0.03	0.37 ± 0.02
μ_{Set} = 0.10, 1000 µM IPTG	2.01 ± 0.07	3.06 ± 0.06	3.42 ± 0.26	2.68 ± 0.06	1.65 ± 0.01
μ_{Set} = 0.15, 1000 µM IPTG	1.58 ± 0.02	2.12 ± 0.15	2.39 ± 0.01	2.23 ± 0.18	1.29 ± 0.08
μ_{Set} = 0.20, 1000 µM IPTG	1.17 ± 0.05	1.61 ± 0.02	1.66 ± 0.05	1.48 ± 0.02	1.14 ± 0.00
μ_{Set} = 0.25, 1000 µM IPTG	0.89 ± 0.09	1.40 ± 0.08	1.50 ± 0.11	1.18 ± 0.08	1.06 ± 0.07
μ_{Set} = 0.10, 250 µM IPTG	2.46 ± 0.04	2.87 ± 0.04	2.96 ± 0.00	2.72 ± 0.06	1.42 ± 0.02
μ_{Set} = 0.15, 250 µM IPTG	1.72 ± 0.07	2.23 ± 0.07	2.02 ± 0.14	2.05 ± 0.05	1.19 ± 0.03
μ_{Set} = 0.20, 250 µM IPTG	1.17 ± 0.02	1.63 ± 0.01	1.67 ± 0.02	1.50 ± 0.03	1.13 ± 0.01
μ_{Set} = 0.25, 250 µM IPTG	1.03 ± 0.05	1.29 ± 0.12	1.51 ± 0.12	1.22 ± 0.03	1.00 ± 0.08
μ_{Set} = 0.10, 50 µM IPTG	2.13 ± 0.12	3.39 ± 0.12	4.16 ± 0.07	2.56 ± 0.34	1.91 ± 0.04
μ_{Set} = 0.15, 50 µM IPTG	1.67 ± 0.03	2.23 ± 0.06	3.00 ± 0.11	2.09 ± 0.18	1.72 ± 0.21
μ_{Set} = 0.20, 50 µM IPTG	1.06 ± 0.03	1.54 ± 0.05	1.94 ± 0.27	1.26 ± 0.03	1.22 ± 0.02
μ_{Set} = 0.25, 50 µM IPTG	0.90 ± 0.06	1.33 ± 0.07	2.00 ± 0.35	1.04 ± 0.10	0.98 ± 0.03

Tested bioprocess conditions comprise batch cultivations as well as fed-batch cultivations with glucose added in a growth-limiting manner at different exponentially increasing feeding rates for growth rate adjustment (μ_{Set} = 0.10, 0.15, 0.20 and 0.25 h^{-1}). Fed-batch bioprocesses were conducted without IPTG added (strains NprE and YwmC) and IPTG concentrations of 50, 250 and 1000 µM (all strains), as indicated. In batch cultivations, 1000 µM IPTG was added. Data represent mean with deviations from min/max values as errors from two biological replicates. *n.t.* not tested.

Table D.3.: Comparison of cutinase productivities q_P for five cutinase secretion strains, denoted by the applied Sec SP (AmyE, Epr, NprE, YpjP, YwmC), for different bioprocess conditions.

	Biomass specific cutinase productivity q_P [U g_X^{-1} h^{-1}] for *C. glutamicum* strains				
Bioprocess condition	**AmyE**	**Epr**	**NprE**	**YpjP**	**YwmC**
Batch, 1000 µM IPTG	112 ± 1	303 ± 38	274 ± 43	180 ± 7	126 ± 11
μ_{Set} = 0.10, 1000 µM IPTG	182 ± 10	260 ± 2	280 ± 3	264 ± 12	142 ± 6
μ_{Set} = 0.15, 1000 µM IPTG	214 ± 2	296 ± 11	312 ± 13	309 ± 13	179 ± 16
μ_{Set} = 0.20, 1000 µM IPTG	230 ± 1	333 ± 10	329 ± 3	272 ± 5	218 ± 1
μ_{Set} = 0.25, 1000 µM IPTG	219 ± 17	341 ± 14	363 ± 27	283 ± 20	250 ± 18
μ_{Set} = 0.10, 250 µM IPTG	195 ± 0	240 ± 8	236 ± 6	251 ± 15	109 ± 1
μ_{Set} = 0.15, 250 µM IPTG	235 ± 8	299 ± 4	257 ± 17	293 ± 1	153 ± 5
μ_{Set} = 0.20, 250 µM IPTG	229 ± 4	312 ± 5	327 ± 4	269 ± 7	213 ± 1
μ_{Set} = 0.25, 250 µM IPTG	245 ± 13	304 ± 26	363 ± 28	273 ± 8	242 ± 15
μ_{Set} = 0.10, 50 µM IPTG	205 ± 4	252 ± 8	308 ± 7	235 ± 10	183 ± 7
μ_{Set} = 0.15, 50 µM IPTG	240 ± 8	287 ± 12	375 ± 15	289 ± 9	242 ± 18
μ_{Set} = 0.20, 50 µM IPTG	219 ± 4	280 ± 7	362 ± 50	233 ± 3	233 ± 3
μ_{Set} = 0.25, 50 µM IPTG	226 ± 16	303 ± 14	470 ± 83	248 ± 23	227 ± 6
μ_{Set} = 0.25, no IPTG	n.t.	n.t.	22 ± 3	n.t.	11 ± 1

Tested bioprocess conditions comprise batch cultivations as well as fed-batch cultivations with glucose added in a growth-limiting manner at different exponentially increasing feeding rates for growth rate adjustment (μ_{Set} = 0.10, 0.15, 0.20 and 0.25 h^{-1}). Fed-batch bioprocesses were conducted without IPTG added (strains NprE and YwmC) and IPTG concentrations of 50, 250 and 1000 µM (all strains), as indicated. In batch cultivations, 1000 µM IPTG was added. Data represent mean with deviations from min/max values as errors from two biological replicates. *n.t.* not tested.

Table D.4.: Results from multiple regression analysis using the full model comprising an intercept term, linear terms, and interaction terms between the different predictor variables.

Description	Model parameter	Estimate	SE	p-value
Intercept	β_0	2.10e-15	3.04e-01	1.00e+00
Linear terms				
AmyE	β_1	3.11e+00	3.64e-01	1.74e-13
Epr	β_2	4.22e+00	3.64e-01	3.86e-20
NprE	β_3	4.77e+00	3.64e-01	2.53e-23
YpjP	β_4	3.70e+00	3.64e-01	5.24e-17
YwmC	β_5	2.20e+00	3.64e-01	2.78e-08
μ_{Set}	β_6	-1.80e-14	1.66e+00	1.00e+00
IPTG	β_7	-2.89e-04	2.25e-04	2.01e-01
Interaction terms				
AmyE * Epr	$\beta_{1,2}$	0	0	n/a
AmyE * NprE	$\beta_{1,3}$	0	0	n/a
AmyE * YpjP	$\beta_{1,4}$	0	0	n/a
AmyE * YwmC	$\beta_{1,5}$	0	0	n/a
AmyE * μ_{Set}	$\beta_{1,6}$	-9.08e+00	1.96e+00	1.14e-05
AmyE * IPTG	$\beta_{1,7}$	6.15e-06	1.85e-04	9.74e-01
Epr * NprE	$\beta_{2,3}$	0	0	n/a
Epr * YpjP	$\beta_{2,4}$	0	0	n/a
Epr * YwmC	$\beta_{2,5}$	0	0	n/a
Epr * μ_{Set}	$\beta_{2,6}$	-1.23e+01	1.96e+00	1.05e-08
Epr * IPTG	$\beta_{2,7}$	5.02e-05	1.85e-04	7.86e-01
NprE * YpjP	$\beta_{3,4}$	0	0	n/a
NprE * YwmC	$\beta_{3,5}$	0	0	n/a
NprE * μ_{Set}	$\beta_{3,6}$	-1.30e+01	1.96e+00	1.94e-09
NprE * IPTG	$\beta_{3,7}$	-2.39e-04	1.85e-04	1.98e-01
YpjP * YwmC	$\beta_{4,5}$	0	0	n/a
YpjP * μ_{Set}	$\beta_{4,6}$	-1.10e+01	1.96e+00	2.13e-07
YpjP * IPTG	$\beta_{4,7}$	2.16e-04	1.85e-04	2.44e-01
YwmC * μ_{Set}	$\beta_{5,6}$	-4.86e+00	1.96e+00	1.51e-02
YwmC * IPTG	$\beta_{5,7}$	0	0	n/a

Continued on the next page.

Table D.4.: Continued.

μ_{Set} * IPTG	$\beta_{6,7}$	1.13e-03	1.05e-03	2.81e-01
Regression diagnostics				
R^2 (adjusted R^2)	0.914 (0.900)			
RMSE	0.262			

Data were fitted against the model $Y_{P/X} = \beta_0 + \sum_{i=1}^{7} \beta_i x_i + \sum_{i=1}^{6} \sum_{j=2}^{7} \beta_{i,j} x_i x_j, i \neq j$ with the categorical predictor variables $x_1 ... x_5$ encoding the used SP (i.e., AmyE, Epr, NprE, YpjP, and YwmC) and the predictor variables x_6 and x_7 encoding the adjusted growth rate and applied IPTG concentration, respectively. Data from batch cultivations were not incorporated since for this kind of bioprocess no growth rate can be adjusted. Calculations were made using MATLAB's `fitlm` function. *SE* standard error. *RMSE* root mean squared error. *n/a* not available.

E. Supporting material for Chapter 6

Table E.1.: Summary of extracellular phenotypes (μ: specific growth rate, $Y_{P/X}$: biomass specific cutinase yield) from microbioreactor screenings of *C. glutamicum* strains with either AmyE SP or NprE SP for cutinase secretion.

	AmyE SP			NprE SP		
Background strain	μ [h^{-1}]	$Y_{P/X}$ [kU g_X^{-1}]	**n**	μ [h^{-1}]	$Y_{P/X}$ [kU g_X^{-1}]	**n**
WT	0.42 ± 0.01	0.23 ± 0.04	20	0.40 ± 0.02	0.53 ± 0.13	46
MB001	0.39 ± 0.01	0.29 ± 0.09	12	0.33 ± 0.02	0.52 ± 0.10	26
W25	0.32 ± 0.01	0.35 ± 0.07	12	0.30 ± 0.01	0.69 ± 0.14	27
W28	0.30 ± 0.01	0.29 ± 0.07	12	0.16 ± 0.01	0.74 ± 0.09	9
W31	0.40 ± 0.01	0.24 ± 0.07	14	0.37 ± 0.03	0.56 ± 0.10	22
W47	0.29 ± 0.01	0.37 ± 0.12	12	0.28 ± 0.01	0.51 ± 0.12	12
W52	0.38 ± 0.01	0.45 ± 0.15	12	0.34 ± 0.02	0.49 ± 0.22	26
W53	0.39 ± 0.01	0.30 ± 0.04	12	0.33 ± 0.02	0.49 ± 0.09	33
W54	0.34 ± 0.01	0.28 ± 0.04	12	0.20 ± 0.01	0.82 ± 0.22	26
W64	0.31 ± 0.01	0.40 ± 0.04	12	0.27 ± 0.01	0.88 ± 0.15	28
W65	0.30 ± 0.01	0.31 ± 0.05	12	0.15 ± 0.01	1.62 ± 0.26	23
W73 [#]	0.24 ± 0.02	n.t.	7	0.24 ± 0.03	n.t.	7
W86	0.30 ± 0.01	0.29 ± 0.06	12	0.26 ± 0.01	0.71 ± 0.09	11
W115	0.32 ± 0.01	0.28 ± 0.05	14	0.23 ± 0.03	0.25 ± 0.10	25
W116	0.33 ± 0.01	0.24 ± 0.04	14	0.18 ± 0.01	1.29 ± 0.25	36
W127	0.42 ± 0.01	0.29 ± 0.06	14	0.34 ± 0.03	0.65 ± 0.13	22
W130	0.41 ± 0.01	0.23 ± 0.04	12	0.34 ± 0.02	0.78 ± 0.15	14
W134	0.19 ± 0.03	0.01 ± 0.01	10	0.26 ± 0.01	0.69 ± 0.17	18
W121.7	0.34 ± 0.01	0.30 ± 0.08	11	0.33 ± 0.01	0.47 ± 0.06	12
W127.6	0.39 ± 0.01	0.34 ± 0.09	12	0.33 ± 0.01	0.55 ± 0.08	12
rrnBC	0.38 ± 0.01	0.39 ± 0.09	12	0.33 ± 0.01	0.63 ± 0.27	15
rrnBF	0.37 ± 0.01	0.44 ± 0.09	12	0.32 ± 0.02	0.51 ± 0.22	16
rrnCF	0.36 ± 0.02	0.24 ± 0.03	11	0.33 ± 0.03	0.36 ± 0.14	17
rrnBCF	0.34 ± 0.02	0.24 ± 0.04	11	0.31 ± 0.02	0.40 ± 0.12	26
rrnABCEF	0.31 ± 0.02	0.35 ± 0.11	10	0.28 ± 0.03	0.40 ± 0.13	19
rrnBCDEF	0.30 ± 0.02	0.25 ± 0.04	10	0.25 ± 0.02	0.37 ± 0.08	31

Results are given as mean with standard deviation from n replicate cultivations, as indicated, and originate from two to five independently carried out BioLector growth experiments. [#] Strain W73 was cultivated in CgXII defined medium containing 10% (v v^{-1}) BHI medium. *n.t.* not tested.

Table E.2.: Overview of deleted genes with corresponding annotations in strains W28 and W86, according to [48, 55].

			Deleted in stain	
CDS	**Name**	**Annotation**	**W28**	**W86**
cg2991		Putative membrane protein	x	
cg2992		Putative secreted protein	x	
cg2993		Conserved hypothetical protein	x	
cg2994		Putative secreted or membrane protein	x	
cg2995	*ptsX*	Phosphotransferase system (PTS), disfunctional enzyme IIC component, putative pseudogene	x	
cg2996	*ptsX*	Phosphotransferase system (PTS), disfunctional enzyme IIAB component, putative pseudogene	x	
cg2997		Hypothetical protein	x	
cg2998		Putative ferredoxin reductase	x	
cg2999		Putative thiosulfate sulfurtransferase	x	
cg3000		Putative transcriptional regulator, MarR-family	x	x
cg3001	*cps*	Putative L-aminoadipate-semialdehyde dehydrogenase large subunit	x	x
cg3002	*gabD1*	Succinate-semialdehyde dehydrogenase ($NAD(P)^+$)	x	x
cg3003		Hypothetical protein	x	x
cg3004		Hypothetical protein	x	x
cg3005	*speE*	Putative spermidine synthase with a transmembrane domain	x	x
cg3006		Putative membrane protein	x	x

Table E.3.: Overview of deleted genes with corresponding annotations in strain W73, according to [48, 55].

CDS	Name	Annotation
cg0158		Putative membrane transport protein, MFS-family
cg0159		Hypothetical protein
cg0160		Hypothetical protein
cg0161		Putative membrane protein
cg0162		Putative membrane spanning protein
cg0163		Putative N-acetylglucosaminyltransferase
cg0165		Putative ABC-2-type transporter
cg0166		Putative Ankyrin repeat containing protein, conserved
cg0167		Putative membrane protein, DUF81-family
cg0168		Putative secondary chloramphenicol transporter, drug/metabolite transporter DMT superfamily
cg0170		Putative transmembrane protein
cg0171		Putative secreted protein
cg0172	*PanD*	Aspartate 1-decarboxylase precursor
cg0173		Hypothetical protein, conserved
cg0174		Putative transport protein, conserved
cg0175		Putative secreted protein, signal peptide
cg0176		Putative membrane protein, DUF81-family
cg0177		Hypothetical protein
cg0178	*HrpB*	Probable ATP-dependent RNA helicase protein
cg0179		Putative membrane protein
cg0180	*Maa*	Maltose O-acetyltransferase
cg0181	*AlkB*	Alkylated DNA repair protein
cg0182	*TagA2*	DNA-3-methyladenine glycosylase I protein
cg0183		Putative LysE-type translocator, threonine efflux transporter, resistance to homoserine/threonine RhtB-family

Table E.4.: Overview of deleted genes with corresponding annotations in strain W134, according to [48, 55].

CDS	Name	Annotation
cg0414	*wzz*	Cell surface polysaccharide biosynthesis/chain length determinant
cg0415	*ptpA2*	Putative protein-tyrosine-phosphatase
cg0416		Putative secreted protein, carrying a eukaryotic domain
cg0417	*capD*	Putative dTDP-glucose-4,6-dehydratase, transmembrane protein
cg0418		Putative aminotransferase, involved in cell wall biosynthesis
cg0419		Putative glycosyltransferase
cg0420		Putative glycosyltransferase
cg0421	*wzx*	Putative PST O-antigen protein, multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) translocase
cg0422	*murA*	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
cg0423	*murB*	UDP-N-acetylenolpyruvoylglucosamine reductase
cg0424		Putative glycosyltransferase
cg0426	*tnp17a* (IS*Cg*17a)	Transposase, putative pseudogene
cg0427	*tnp17b* (IS*Cg*17a)	Transposase, putative pseudogene
cg0428	*tnp17c* (IS*Cg*17a)	Transposase, putative pseudogene
cg0431		Putative membrane protein, involved in polysaccharide acetylation
cg0432		Putative lipopolysaccharide modification acyltransferase
cg0433		Hypothetical protein
cg0434		Hypothetical protein
cg0435	*udgA1*	UDP-glucose-6-dehydrogenase
cg0436		Hypothetical protein
cg0437	*wzy*	Putative membrane protein, involved in polysaccharide polymerization
cg0438		Putative glycosyltransferase
cg0439		Putative acetyl transferase
cg0440		Hypothetical protein
cg0635		Putative NAD-dependent aldehyde dehydrogenase
cg0636		Putative membrane protein
cg0637	*betB*	Putative betaine aldehyde dehydrogenase (BADH)
cg0638		Hypothetical protein
cg0639		Putative ferredoxin reductase
cg0640	*fdxB*	2Fe-2S ferredoxin
cg0641	*fabG2*	Putative secreted short-chain dehydrogenase
cg0642		Conserved hypothetical protein, probably DNA-binding

Continued on the next page.

Table E.4.: Continued.

CDS	Name	Annotation
cg0644		Pyruvate phosphate dikinase, PEP/pyruvate-binding
cg0645	*cytP*	Putative cytochrome P450
cg0646		Putative transcriptional regulator, IclR-family
cg0704		Conserved hypothetical protein
cg0705		Hypothetical protein
cg0706		Conserved putative membrane protein
cg0707	*cgtS7*	Two-component system, sensory histidine kinase
cg0709	*cgtR7*	Two-component system, transcriptional response regulator
cg0710		Putative membrane protein
cg0711		Putative membrane protein
cg0712		Putative secreted protein
cg0713		Conserved hypothetical protein
cg0714		Hypothetical protein
cg0715		Putative secreted protein
cg0716		Conserved hypothetical protein
cg0717	*crtEb*	Lycopene elongase
cg0718	*crtYf*	C50 carotenoid ε-cyclase
cg0719	*crtYe*	C50 carotenoid ε-cyclase
cg0720	*crtI2*	Phytoene dehydrogenase (desaturase)
cg0721	*crtB2*	Phytoene synthetase
cg0722		Putative multidrug efflux protein, resistance-nodulation-cell division (RND) superfamily
cg0723	*crtE*	Geranylgeranyl pyrophosphate synthase
cg0725		Putative transcriptional regulator, MarR-family
cg0726		Putative secreted lipoprotein
cg0727		Putative nucleoside-diphosphate-sugar epimerase
cg0728	*phr*	Deoxyribodipyrimidine photo-lyase
cg0730		Putative glycosyl transferase
cg0731		Hypothetical protein
cg0732		ABC-type transporter, permease subunit
cg0733		ABC-type transporter, ATPase subunit
cg0735	*metI*	ABC-type methionine transporter, permease subunit
cg0736	*metN*	ABC-type methionine transporter, ATPase subunit
cg0737	*metQ*	ABC-type methionine transporter, substrate-binding lipoprotein
cg0738	*dnaE2*	Putative DNA polymerase III, α-chain
cg0739		Putative integral membrane protein
cg0740		Putative membrane protein

Continued on the next page.

Table E.4.: Continued.

CDS	Name	Annotation
cg0741	*sirR*	Transcriptional repressor, DtxR-family
cg0742		Putative integral membrane protein
cg0745		Putative NAD-dependent protein deacetylase, SIR2-family
cg0747		Conserved hypothetical protein
cg0748		ABC-type putative iron-siderophore transporter, substrate-binding lipoprotein
cg0822		Conserved hypothetical protein
cg0823	*ntaA*	Putative nitrilotriacetate monooxygenase, subunit A
cg0824	*tnp5a* (IS*Cg*5a)	Transposase
cg0825		Putative short chain dehydrogenase related to 3-oxoacyl-(acyl-carrier protein) reductase
cg0826		Putative membrane protein
cg0828		Putative dihydrofolate reductase
cg0829		Conserved hypothetical protein, glyoxylase-family
cg0830		Putative membrane protein
cg0831	*tusG*	ABC-type trehalose transporter, permease subunit
cg0832	*tusF*	ABC-type trehalose transporter, permease subunit
cg0833		Conserved hypothetical protein
cg0834	*tusE*	ABC-type trehalose transporter, substrate-binding lipoprotein
cg0835	*tusK*	ABC-type trehalose transporter, substrate-binding lipoprotein
cg0836		Hypothetical protein
cg0837		Hypothetical protein
cg0838		Putative helicase
cg0839		Hypothetical protein
cg0840		Conserved hypothetical protein
cg0841		Conserved hypothetical protein
cg0842		Putative DNA helicase
cg0843		Putative helicase
cg0844		Type II restriction enzyme, methylase subunit
cg0845		DNA/RNA helicase, superfamily II
cg1018		Putative ATP-dependent DNA helicase
cg1019		Conserved hypothetical protein, probably metal-dependent hydrolase
cg1020		Hypothetical protein
cg1021		Hypothetical protein, probable esterase/lipase
cg1022	*tnp6a* (IS*Cg*6a)	Transposase

Continued on the next page.

Table E.4.: Continued.

CDS	Name	Annotation
cg1023	*tnp6b* (IS*Cg*6a)	Transposase
cg1024	*tnp7a* (IS*Cg*7a)	Transposase
cg1025		Hypothetical protein
cg1027	*dld*	D-lactate dehydrogenase
cg1028		Putative restriction-modification system methylase
cg1030	*tnp6c* (IS*Cg*6c)	Transposase
cg1031	*tnp6d* (IS*Cg*6c)	Transposase
cg1032		Putative transcriptional regulator, ArsR-family
cg1033		Putative secondary Cd^{2+} transporter, cadmium resistance (CadD) family
cg1172		Conserved hypothetical protein, putative plasmid maintenance system killer protein
cg1173		Conserved hypothetical protein, putative plasmid maintenance system antidote protein
cg1174	*arcB*	Putative ornithine carbamoyltransferase
cg1175		Putative acetyltransferase, GNAT-family
cg1176		Putative short chain dehydrogenase
cg1178	*tnp9a* (IS*Cg*9a)	Transposase
cg1179		Putative membrane protein
cg1180		Glycosyltransferase, probably involved in cell wall biogenesis
cg1181		Glycosyltransferase, probably involved in cell wall biogenesis
cg1182		Putative membrane protein
cg1183		Hypothetical protein
cg1184	*tnp10c* (IS*Cg*10a)	Transposase, putative pseudogene
cg1185	*tnp10b* (IS*Cg*10a)	Transposase, putative pseudogene
cg1187	*tnp10a* (IS*Cg*10a)	Transposase, putative pseudogene
cg1189		Hypothetical protein
cg1190		Hypothetical protein
cg1191		Hypothetical protein
cg1192		Putative oxidoreductase, aldo/keto reductase family
cg1193		Conserved hypothetical protein, putative carboxymuconolactone decarboxylase

Continued on the next page.

Table E.4.: Continued.

CDS	Name	Annotation
cg1194		Predicted nucleoside-diphosphate-sugar epimerase
cg1195		Putative permease, sulfate permease (SulP) family
cg1197		ABC-type putative lipoprotein release transporter, ATPase subunit
cg1198		Hypothetical protein
cg1199		Hypothetical protein
cg1200		Hypothetical protein
cg1201		Hypothetical protein
cg1202		Conserved hypothetical protein
cg1203		Putative magnesium chelatase, ChlI subunit
cg1204		Conserved hypothetical protein, similar to 2,3-PDG dependent phosphoglycerate mutase
cg1205		Conserved hypothetical protein
cg1206		Conserved hypothetical protein
cg1207		ABC-type transporter, ATPase subunit
cg1208		Conserved hypothetical protein
cg1209		Putative Zn-ribbon-containing protein, involved in phosphonate metabolism
cg1210		Putative membrane protein
cg1211		Putative transcriptional regulator, MarR-family
cg1212		Putative antibiotic efflux permease, MFS-type
cg1213	*tnp1a* (IS*Cg*1a)	Transposase
cg1291		Putative membrane protein
cg1292		Flavin-containing monooxygenase 3
cg1293		Putative secreted protein
cg1294		Putative esterase, α-β-hydrolase superfamily
cg1295		Putative hydrolase/acyltransferase, α-β-hydrolase superfamily
cg1296		Conserved hypothetical protein, putative non-ribosomal peptide synthetase module
cg1297		Hypothetical protein
cg1298	*cydC*	ABC-type putative multidrug/protein/lipid transporter, ATPase and permease subunit
cg1299	*cydD*	ABC-type putative multidrug/protein/lipid transporter, ATPase and permease subunit
cg1300	*cydB*	Cytochrome d ubiquinol oxidase subunit II
cg1301	*cydA*	Cytochrome d ubiquinol oxidase subunit I
cg1302		Putative nuclease, HKD-family
cg1303		NTP pyrophosphohydrolase
cg1304		Putative secreted protein
cg1305		Amino acid permease
cg1340	*arnR*	Transcriptional regulator ArsR-family

Continued on the next page.

Table E.4.: Continued.

CDS	Name	Annotation
cg1341	*narI*	Respiratory nitrate reductase 2, γ-chain
cg1342	*narJ*	Respiratory nitrate reductase 2, δ-chain
cg1343	*narH*	Respiratory nitrate reductase 2, β-chain
cg1344	*narG*	Respiratory nitrate reductase 2, α-chain
cg1345	*narK*	Putative nitrate/nitrite permease, MFS-type
cg1346	*mog*	Putative molybdopterin biosynthesis protein Mog
cg1347		Putative secreted or membrane protein
cg1348		Putative membrane protein, containing a CBS domain
cg1349		Putative membrane protein, containing a CBS domain
cg1350	*mob*	Putative molybdopterin-guanine dinucleotide biosynthesis protein
cg1351	*moeA3*	Molybdopterin biosynthesis protein MoeA
cg1352	*moaA*	Molybdopterin biosynthesis protein A
cg1353	*fadD4*	Putative acyl-CoA synthetase
cg1370		Conserved hypothetical protein
cg1371		Putative nuclease, RecB-family
cg1372		Conserved hypothetical protein
cg1373		Putative glyoxalase
cg1374		Conserved hypothetical protein
cg1375		Putative thioredoxin
cg1376	*ssuD1*	$FMNH_2$-dependent aliphatic sulfonate monooxygenase
cg1377	*ssuC*	ABC-type aliphatic sulfonate transporter, permease subunit
cg1379	*ssuB*	ABC-type aliphatic sulfonate transporter, ATPase subunit
cg1380	*ssuA*	ABC-type aliphatic sulfonate transporter, substrate-binding lipoprotein
cg1381	*glgB*	1,4-α-glucan branching enzyme
cg1382	*glgE*	Putative α-amylase
cg1383		ABC-type putative molybdenum transporter, ATPase subunit
cg1384		Putative NUDIX hydrolase
cg1385		Putative SAM-dependent methyltransferase

Table E.5.: Overview of deleted genes with corresponding annotations in strains W31, W115, W116, W127 and W130, according to [48, 55].

			Deleted in strain				
CDS	**Name**	**Annotation**	**W31**	**W115**	**W116**	**W127**	**W130**
cg2312	*gip*	Hydroxypyruvate isomerase				x	x
cg2313	*idhA3*	Myo-inositol 2-dehydrogenase				x	x
cg2314		Putative transcriptional regulator, LacI-family				x	x
cg2315		ABC-type putative iron-III dicitrate transporter, ATPase subunit				x	x
cg2317		ABC-type putative iron-III dicitrate transporter, permease subunit				x	x
cg2318		ABC-type putative iron-III dicitrate transporter, substrate-binding lipoprotein				x	x
cg2320		Putative transcriptional regulator, ArsR-family				x	x
cg2321		DNA-directed DNA polymerase, DNA polymerase III εsubunit				x	x
cg2322		Conserved hypothetical protein				x	x
cg2621		Conserved hypothetical protein				x	x
cg2622	*pcaJ*	Putative 3-oxoadipate CoA-transferase				x	x
cg2623	*pcaI*	Putative 3-oxoadipate CoA-transferase				x	x
cg2624	*pcaR*	Transcriptional activator of the protocatechuate branch of the β-ketoadipate pathway, IclR-family				x	x
cg2625	*pcaF*	Putative β-ketoadipyl CoA thiolase				x	x
cg2626	*pcaD*	Putative 3-oxoadipate enol-lactonase				x	x
cg2627	*pcaO*	Putative transcriptional regulator, LuxR-family				x	x
cg2628	*pcaC*	Putative 4-carboxymuconolactone decarboxylase				x	x
cg2629	*pcaB*	Putative 3-carboxy-cis,cis-muconate cycloisomerase				x	x
cg2630	*pcaG*	Protocatechuate 3,4-dioxygenase, α-subunit				x	x
cg2631	*pcaH*	Protocatechuate 3,4-dioxygenase, β-subunit				x	x
cg2633		Putative restriction endonuclease				x	x
cg2634	*catC*	Putative muconolactone delta-isomerase				x	x
cg2635	*catB*	Putative muconate cycloisomerase				x	x
cg2636	*catA*	Catechol 1,2-dioxygenase				x	x
cg2637	*benA*	Putative benzoate 1,2-dioxygenase, α-subunit				x	x
cg2638	*benB*	Putative benzoate 1,2-dioxygenase, β-subunit				x	x
cg2639	*benC*	Putative benzoate 1,2-dioxygenase, electron transfer subunit				x	x
cg2640	*benD*	Putative 1,6-dihydroxycyclohexa-2,4-diene-1-carboxylate dehydrogenase				x	x
cg2641	*benR*	Putative transcriptional regulator, LuxR-family				x	x
cg2642	*benK*	Putative MFS-type benzoate permease				x	x

Continued on the next page.

Table E.5.: Continued.

CDS	Name	Annotation	Deleted in strain				
			W31	W115	W116	W127	W130
cg2643	*benE*	Putative secondary benzoate symporter, benzoate:H^+ symporter (BenE) family				x	x
cg2663		Hypothetical protein				x	x
cg2664		Putative type IV restriction endonuclease				x	x
cg2665		Hypothetical protein				x	x
cg2666		Hypothetical protein				x	x
cg2667		Hypothetical protein				x	x
cg2668	*'crtI*	Putative phytoene dehydrogenase, putative pseudogen (C-terminal fragment)				x	x
cg2670	*crtI'*	Putative phytoene dehydrogenase, putative pseudogen (N-terminal fragment)				x	x
cg2672	*crtB*	Geranylgeranyl-diphosphate geranylgeranyltransferase				x	x
cg2673		Putative multidrug efflux permease, MFS-type				x	x
cg2674		Alkylhydroperoxidase, AhpD-family core domain				x	x
cg2675		ATPase component of ABC-type transport system, contains duplicate				x	x
cg2676		ABC-type dipeptide/oligipeptide/nickel transport system, permease component				x	x
cg2677		ABC-type dipeptide/oligipeptide/nickel transport system, permease component				x	x
cg2678		ABC-type dipeptide/oligipeptide/nickel transport system, secreted component				x	x
cg2679		Hypothetical protein				x	x
cg2680		Aminotransferase, AT class II				x	x
cg2683		Conserved hypothetical protein				x	x
cg2684		Conserved hypothetical membrane protein, DedA-family				x	x
cg2685		Putative short-chain dehydrogenase/reductase				x	x
cg2686		Putative transcriptional regulator, TetR-family				x	x
cg2755		Conserved hypothetical protein				x	x
cg2756		Conserved hypothetical protein				x	x
cg2757	*tnp15a* (IS*Cg*15a)	Transposase				x	x
cg2758		Conserved putative membrane protein				x	x
cg2759	*tnp15b* (IS*Cg*15b)	Transposase				x	x
cg2760		Conserved hypothetical protein				x	x

Continued on the next page.

Table E.5.: Continued.

			Deleted in strain				
CDS	**Name**	**Annotation**	**W31**	**W115**	**W116**	**W127**	**W130**
cg3102		Secreted nucleoside phosphorylase		x	x	x	x
cg3103		Conserved hypothetical protein		x	x	x	x
cg3104		Conserved hypothetical protein		x	x	x	x
cg3105		Hypothetical protein		x	x	x	x
cg3106		Conserved hypothetical protein		x	x	x	x
cg3107	*adhA*	Alcohol dehydrogenase		x	x	x	x
cg4002		Hypothetical protein		x	x	x	x
cg3108		Putative secreted protein		x	x	x	x
cg3109		Putative membrane protein		x	x	x	x
cg3110		Hypothetical protein, slightly similar to aquaporin		x	x	x	x
cg3111		Putative secreted protein		x	x	x	x
cg3263		Conserved hypothetical protein		x	x		
cg3264		Conserved hypothetical protein		x	x		
cg3266	*tnp5c* (IS*Cg5c*)	Transposase		x	x		
cg3267		Putative membrane protein, putative pseudogene (C-terminal fragment)		x	x		
cg3268		Putative membrane protein, putative pseudogene		x	x		
cg3269		Putative membrane protein, putative pseudogene		x	x		
cg3270		Putative membrane protein, putative pseudogene (N-terminal fragment)		x	x		
cg3271		SAM-dependent methyltransferase		x	x		
cg3272		Putative membrane protein		x	x		
cg3273		Hypothetical protein		x	x		
cg4011		Transposon Tn501 resolvase		x	x		
cg3274		Putative DNA invertase, putative pseudogene		x	x		
cg3275	*fdxA*	Putative ferredoxin		x	x		
cg3277		Hypothetical protein, containing double-stranded beta-helix domain		x	x		
cg3278	*tnp20a* (IS*Cg20a*)	Transposase, putative pseudogene		x	x		
cg3279		Putative dehydrogenase, putative pseudogene		x	x		
cg3280		Putative secreted protein		x	x		
cg3281		Putative Cu^{2+} transporting P-type ATPase		x	x		
cg3282		Putative Cu^{2+} transporting P-type ATPase		x	x		
cg3283		Hypothetical protein		x	x		
cg3284	*cgtS9*	Two-component system, sensory histidine kinase		x	x		

Continued on the next page.

Table E.5.: Continued.

CDS	Name	Annotation	Deleted in strain				
			W31	W115	W116	W127	W130
cg3285	*cgtR9*	Two-component system, transcriptional response regulator		x	x		
cg3286		Putative secreted protein		x	x		
cg3287		Putative secreted multicopper oxidase		x	x		
cg3288		Hypothetical protein		x	x		
cg3289		Putative Thiol-disulfide isomerase or thioredoxin		x	x		
cg3290		Putative oxidoreductase		x	x		
cg3291		Putative transcriptional regulator, Crp-family		x	x		
cg3292		Putative heavy-metal ion transporting P-type ATPase		x	x		
cg3293		Hypothetical protein		x	x		
cg3294		Hypothetical protein		x	x		
cg3295		Putative Cd^{2+} transporting P-type ATPase		x	x		
cg3296	*tnp19c* (IS*Cg*19a)	Transposase, putative pseudogene		x	x		
cg3297	*tnp19b* (IS*Cg*19a)	Transposase, putative pseudogene		x	x		
cg3298	*tnp19a* (IS*Cg*19a)	Transposase, putative pseudogene		x	x		
cg3299	*trxB1*	Thioredoxin (TRX)		x	x		
cg3300		Putative Cu^{2+} transporting P-type ATPase		x	x		
cg3301		Putative sugar/metabolite permease, MFS-type		x	x		
cg3324		Putative secreted protein	x		x		x
cg3325		Conserved hypothetical protein	x		x		x
cg3326		Hypothetical protein	x		x		x
cg3327	*dps*	Putative starvation-induced DNA protecting protein	x		x		x
cg3328	*mutM2*	DNA-formamidopyrimidine glycosylase	x		x		x
cg3329		Conserved hypothetical protein	x		x		x
cg3330		Putative secreted protein	x		x		x
cg3331	*ogt*	Methylated-DNA–[protein]-cysteine S-methyltransferase	x		x		x
cg3332	*qor*	NADPH:quinone reductase	x		x		x
cg3333		Hypothetical protein	x		x		x
cg3334		Putative sugar permease, MFS-type	x		x		x
cg3335	*malE*	Malic enzyme ($NADP^+$)	x		x		x
cg3336	*gntK*	Gluconokinase	x		x		x
cg3337		Putative membrane protein	x		x		x

Continued on the next page.

Table E.5.: Continued.

			Deleted in strain				
CDS	**Name**	**Annotation**	**W31**	**W115**	**W116**	**W127**	**W130**
cg3338		Putative membrane protein	x		x		x
cg3339	*merA*	Putative Hg^{2+} reductase	x		x		x
cg3340	*dadA*	D-Amino-acid dehydrogenase	x		x		x
cg3341		Putative membrane protein	x		x		x
cg3342		Putative secreted protein	x		x		x
cg3343		Putative secreted membrane protein	x		x		x
cg3344		Putative nitroreductase	x		x		x
cg3345		Hypothetical protein	x		x		x

Table E.6.: Overview of deleted genes with corresponding annotations in strains W25, W52, W53, W54, W64 and W65, according to [48, 55].

			Deleted in strain					
CDS	**Name**	**Annotation / function**	**W25**	**W52**	**W53**	**W54**	**W64**	**W65**
	rrnB	rRNA operon		x	x		x	
cg0931		Putative aminotransferase, AT class I			x		x	
cg2801	*crcB'*	Putative membrane protein, CrcB-like protein, putative pseudogene (N-terminal fragment)	x				x	x
cg2802	*'crcB*	Putative membrane protein, CrcB-like protein, putative pseudogene (C-terminal fragment)	x				x	x
cg2803		Conserved hypothetical protein	x				x	x
cg2804	*tnp21a* (ISC*g*21a)	Transposase	x				x	x
cg2805	*psp4*	Putative secreted protein	x				x	x
cg2806		Putative membrane protein	x				x	x
cg2807	*tnp11a* (ISC*g*11a)	Transposase, putative pseudogene	x				x	x
cg2808	*tnp13a* (ISC*g*13a)	Transposase	x				x	x
cg2809		Putative membrane protein	x				x	x
cg2810	*cynT*	Cyst(e)ine importer	x				x	x
cg2811		ABC-type lipoprotein release transporter, permease subunit	x				x	x
cg2812		ABC-type lipoprotein release transporter, ATPase subunit	x				x	x
	rrnD	rRNA operon	x				x	x
cg2819		Hypothetical membrane protein	x				x	x
cg2822		Sugar phosphate isomerase/epimerase	x				x	x
cg2823		Hypothetical protein, dehydrogenase or related protein	x				x	x
cg2824		SAM-dependent methyltransferase	x				x	x
	rrnE	rRNA operon	x				x	x
cg2828		Putative membrane protein	x				x	x
	rrnC	rRNA operon				x		x
cg3261		Putative transcriptional regulator, GntR-family				x		x
cg3263		Conserved hypothetical protein				x		x
cg3264		Conserved hypothetical protein				x		x
cg3266	*tnp5c* (ISC*g*5c)	Transposase				x		x
cg3267		Putative membrane protein, putative pseudogene (C-terminal fragment)				x		x

Continued on the next page.

Table E.6.: Continued.

CDS	Name	Annotation / function	W25	W52	W53	W54	W64	W65
cg3268		Putative membrane protein, putative pseudogene				x		x
cg3269		Putative membrane protein, putative pseudogene				x		x
cg3270		Putative membrane protein, putative pseudogene (N-terminal fragment)				x		x
cg3271		SAM-dependent methyltransferase				x		x
cg3272		Putative membrane protein				x		x
cg3273		Hypothetical protein				x		x
cg4011		Transposon Tn501 resolvase				x		x
cg3274		Putative DNA invertase, putative pseudogene				x		x
cg3275	*fdxA*	Putative ferredoxin				x		x
cg3277		Hypothetical protein, containing double-stranded β-helix domain				x		x
cg3278	*tnp20a* (IS*Cg*20a)	Transposase, putative pseudogene				x		x
cg3279		Putative dehydrogenase, putative pseudogene				x		x
cg3280		Putative secreted protein				x		x
cg3281		Putative Cu^{2+} transporting P-type ATPase				x		x
cg3282		Putative Cu^{2+} transporting P-type ATPase				x		x
cg3283		Hypothetical protein				x		x
cg3284	*cgtS9*	Two-component system, sensory histidine kinase				x		x
cg3285	*cgtR9*	Two-component system, transcriptional response regulator				x		x
cg3286		Putative secreted protein				x		x
cg3287		Putative secreted multicopper oxidase				x		x
cg3288		Hypothetical protein				x		x
cg3289		Putative Thiol-disulfide isomerase or thioredoxin				x		x
cg3290		Putative oxidoreductase				x		x
cg3291		Putative transcriptional regulator, Crp-family				x		x
cg3292		Putative heavy-metal ion transporting P-type ATPase				x		x
cg3293		Hypothetical protein				x		x
cg3294		Hypothetical protein				x		x
cg3295		Putative Cd^{2+} transporting P-type ATPase				x		x
cg3296	*tnp19c* (IS*Cg*19a)	Transposase, putative pseudogene				x		x
cg3297	*tnp19b* (IS*Cg*19a)	Transposase, putative pseudogene				x		x

Continued on the next page.

Table E.6.: Continued.

CDS	Name	Annotation / function	W25	W52	W53	W54	W64	W65
cg3298	*tnp19a* (IS*Cg*19a)	Transposase, putative pseudogene				x		x

Table E.7.: Composition of subclusters shown in Figure 6.7 based on functional categories for transcriptome data and proteome data, as assigned by cluster analyses.

	Distribution of genes assigned to subclusters							
	Transcriptome				**Proteome**			
Functional categorization	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Metabolism								
Central carbon metabolism	2.1%	2.6%	1.5%	1.8%	4.0%	5.7%	3.3%	1.5%
Carbon source transport and metabolism	10.6%	8.3%	6.3%	5.8%	18.0%	6.6%	8.3%	2.9%
Anaerobic metabolism	0%	0.5%	0.1%	0.5%	0%	0.6%	0.5%	0%
Respiration and oxidative phosphorylation	0%	1.9%	1.3%	0.5%	0%	2.4%	2.1%	1.5%
Amino acid transport and metabolism	6.4%	8.2%	7.2%	3.3%	6.0%	11.0%	10.1%	7.4%
Nucleotide transport and metabolism	2.1%	3.1%	2.8%	2.3%	6.0%	4.8%	4.9%	2.9%
Lipid transport and metabolism	0%	0.6%	1.0%	1.0%	0	1.2%	1.2%	0%
Coenzyme transport and metabolism	3.2%	4.8%	4.6%	2.8%	6.0%	6.0%	5.0%	10.3%
Inorganic ion transport, metabolism, and storage	0%	6.6%	4.0%	6.5%	2.0%	4.2%	3.7%	2.9%
Transport and metabolism of further metabolites	5.3%	5.7%	4.4%	4.3%	2.0%	5.7%	3.7%	5.9%
Cell wall/membrane/envelope biogenesis	5.3%	4.0%	3.4%	5.5%	4.0%	4.2%	4.7%	5.9%
Information storage and processing								
DNA replication, recombination, repair, and degradation	3.2%	3.1%	4.5%	5.5%	2.0%	3.0%	4.5%	4.4%
Cell division, chromosome partitioning	0%	0.7%	1.3%	1.3%	2.0%	0.6%	1.3%	0%
Transcription including sigma factors, RNA processing and modification	1.1%	1.5%	2.5%	1.8%	2.0%	1.5%	3.2%	0%
Translation, ribosomal structure and biogenesis	4.3%	5.0%	5.3%	3.3%	6.0%	2.7%	12.5%	5.9%
Prophage genes	0%	0%	0%	0%	0%	1.2%	1.8%	1.5%
Cellular processes and signaling								
Protein secretion	3.2%	0.6%	0.8%	0.5%	0%	0	0.4%	0%
Protein turnover and chaperones	1.1%	1.9%	2.8%	1.3%	0	5.1%	2.1%	2.9%
Posttranslational modification	2.1%	1.3%	1.4%	1.3%	4.0%	1.5%	0.7%	0%
Signal transduction mechanisms	5.3%	6.2%	6.5%	6.0%	6.0%	5.1%	4.7%	4.4%
Poorly characterized								
General function prediction only	16.0%	9.6%	13.5%	15.5%	8.0%	11.6%	8.4%	14.7%
Unknown function	28.7%	23.8%	25.0%	29.8%	22.0%	15.5%	12.8%	25.0%

Functional categorization according to [48]. Definition of subclusters corresponds to Figure 6.7. Transcriptome subclusters A, B, C and D comprise 91, 976, 928 and 384 genes, respectively. Proteome subclusters A, B, C and D comprise 49, 322, 719 and 66 genes, respectively.