

DEVELOPMENT OF AN INTEGRATED PLATFORM COMBINING UP- AND DOWNSTREAM PROCESSING FEATURES

zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

der Fakultät für Chemieingenieurwesen und Verfahrenstechnik des
Karlsruher Instituts für Technologie (KIT)
vorgelegte

genehmigte
DISSERTATION

von
Dipl.-Ing. Pascal Baumann
aus Rheinhausen

Referent: Prof. Dr. Jürgen Hubbuch
Korreferent: Prof. Dr. rer. nat. Christoph Syldatk
Tag der mündlichen Prüfung: 08.09.2015

Acknowledgments

Over the past years I have received support from a large number of people. I want to thank with all my heart:

My doctoral advisor Prof. Dr. Jürgen Hubbuch for giving me the opportunity to work on this doctoral project in the Biomolecular Separation Engineering Group. His professional expertise and high standards were always inspiring and helped for unimpeded development for this thesis.

Prof. Dr. rer. nat. Christoph Syldatk for being the second referee of my thesis and for reviving my interest in industrial biotechnology and microbiology during my bioengineering studies at the Karlsruhe Institute of Technology.

All project members who supported my work not only in material and financial respect but in discussions with all their knowledge and expertise. Special thanks I would like to express for Dr. Cedric Szpirer from Delphi Genetics for giving me the opportunity to work with the pStaby and pSCherry systems as well as for Carsten Müller and Pamela Felden from m2p-Labs for the great support using the BioLector[®] technology.

Tobias Hahn and Thiemo Huuk for the great collaboration projects including the modeling software ChromX, as well as for the spectacular sailing trip. Thank you Marc Hoffmann for the great time in the lab and the kind introduction into cell culture technologies and Dr.-Ing. Anna Osberghaus for the guidance during my doctoral thesis.

All members of the Biomolecular Separation Engineering Group not only for being great colleagues but also for having become good friends. I would like to express special thanks to Kai Baumgartner and Frank Haemmerling for having become best friends and for openly accepting my former hair style without prejudices right from the beginning. Also I would like to express my gratitude to Margret Meixner and Susanna Suhm for always helping me out and being a great support. Finally, I would like to acknowledge my office colleagues Frieder Kröner, Sarah Zimmermann, and Christopher Ladd Effio for making my day during some stressful hours in the office. In this context I would like to especially acknowledge my great friend and bureau colleague Josefine Morgenstern who always knew how to cheer me up at stressful times.

Lisa Eiden, Sarah Renner, Ursula Simon, Nicolai Bluthardt, Hannah Burghardt and Claudia Weigel for helping me out with the experiments as Bachelor, Master and Diploma students.

The power circle at the Venice Beach fitness club and my marathon running group in Karlsruhe for being a perfect balance after stressful days and weeks although some colleagues at work were worried I would take it too far some time (I guess they were right at the end). Thank you Carsten Radtke for finally letting me win one Karlsruhe Schlossparklauf in 2014. I also thank all friends I did not mention by name here but who went through all the time with me together.

I would like to give my loving thanks to the most important people in my life. Thanks to my family, always standing by my side and giving me loving support. Without you I would not be where I am today. Finally, I would like to thank my girlfriend Lea. Thanks for being just as perfect as you are and for the wonderful time we are having. Thanks for your loving support and for tolerating my endless stories from work. I love you!

Finally, I would like to finish with the words of Christoph Daum:
'Wenn der Kopf richtig funktioniert, dann ist er das dritte Bein.'

Abstract

Biopharmaceutical process development is complex and based on a vast number of sub-processes to be optimized. Nowadays, many approaches for process design are still expert- and experience-oriented and follow a certain hierarchy. Upstream and downstream processes are still optimized separately, creating figuratively spoken an 'Over the Wall' workflow. Consequently, the upstream conditions of highest product titers are selected for subsequent purification strategies even though further parameters like the concentration levels of critical impurities may pose a problem in downstream processes. Choosing an integrated process development approach instead of a strict separation in the workflow may increase the overall process performance. Another challenge during industrial process development is caused by the 'Time to Market' demand, resulting in insufficient screening and optimization.

This thesis deals with strategies to tackle both industrial issues by the development of an integrated platform combining upstream and downstream processing features coupled with the latest technologies in high-throughput screening (HTS) systems and mechanistic modeling. To prevent an analytical bottleneck created by high-throughput (HT) technologies another main focus is the development of HTS compatible analytical tools, meeting the quality attributes for creating mechanistic models.

The first section of this thesis is concerned with the development and evaluation of sophisticated analytical technologies and their use in HTS and modeling applications. The first approach makes use of a protein fusion tag with special optical properties for quantitative real-time and in-line tracking of proteins in up- and downstream applications. The Cherry-TagTM resembles the heme group of cytochrome c and can be detected in the spectral VIS region. It was proven that destabilizing conditions result in a loss of the tag's chromophore function, making the Cherry-TagTM an ideal tool in screening for protein solubility and stability. In HT upstream process development in 48-well format using the BioLector[®] (m2p-Labs, Germany) the tag helped identifying conditions of non-native protein formation as inclusion bodies in *Escherichia coli* SE1, producing glutathione-S-transferase. During downstream processing the red-colored enzyme fusion was identified to be the active species as it was effectively bound to an affinity column, based on a binding mechanism which relies on the catalytic activity of the enzyme. The Cherry-TagTM proved to be a reliable and exact analytical tool for real-time tracking of protein stability and solubility in diverse fields of biopharmaceutical process development. Additionally, the study revealed that the optimized conditions for soluble product formation in the upstream apply for the product without the additional tag. Alternatively, an effective enzymatic tag cleavage reaction was presented for final removal of the colored fusion protein.

As a second approach for HTS compatible analytical tools, multi-variate data analysis of protein spectra was employed. The method exploits differences in protein mid-UV (200 - 300 nm) spectra, derived by properties of peptide bonds and amino acid residues. These absorption differences in spectra of distinct protein species enable generating empirical models (partial least squares - projection to latent structures (PLS) regression) based on calibration samples of known concentration and composition which can be used for unknown samples in the calibrated region. The methodology was successfully used for deconvolution of high-throughput multi-component isotherm data of model proteins at

different pH levels and salt concentrations. The generated data were of modeling quality, resulting in predictive semi-mechanistic models. Besides the accuracy of the analytical method, its fast and non-invasive nature makes it an ideal tool for implementation into the HTS workflow of robotic workstations without the need of sampling.

The third proposed approach for HTS analytics deals with high-throughput capillary gel electrophoresis (HT-CGE) which was used as the central analytics for establishing a high-throughput platform for the production of virus-like particles (VLPs) in *E. coli* Rosetta. The presented work on VLPs can be seen as the interface between the analytical process development discussed above and the up- and downstream process development strategies proposed below. As mentioned above, the HT upstream optimization was performed in 48-well format in the BioLector[®] system (m2p-Labs, Germany). An *Escherichia coli* strain, regulated by a tac-promoter, was investigated for producing the murine polyomavirus capsid protein (VP1). As a large number of cultivation conditions for the intracellular product were screened including different cultivation temperatures, shaking speeds, induction conditions and cultivation media, a fast and effective cell disruption process needed to be developed. The released product was analyzed by HT-CGE which was shown to be a highly HTS compatible analytical tool for characterizing the soluble and insoluble protein fraction of all tested setups. Based on the results of the screening a well performing system was scaled up and the VP1 product was purified using an anion-exchange chromatography membrane adsorbent, followed by a size exclusion chromatography polishing procedure. Using HT-CGE also the downstream process was well characterized, enabling even high-resolution fractionation procedures. Finally, the purified VP1 proteins were successfully assembled into empty capsids and characterized using transmission electron microscopy (TEM).

The topic of the second section of the thesis follows up downstream process intensification strategies by developing systematic methods for different chromatographic unit-operations. A first strategy is proposed for purifying salt-intolerant proteins using ion-exchange (IEX) chromatography which might be a challenging task due to increased salt levels during product elution. Product losses are not only based on insufficient optimization of purification setups but also product sensitivity to process conditions. Thus, the methodology includes both stability investigations in HT, as well as a global optimization approach for optimal operational conditions for product capturing under stabilized conditions. The stability experiments cover short- and long-term pH stability, salt tolerance tests and additive screenings. With the boundary conditions for product stability being set the crude feedstock is then analyzed by pH gradient chromatography for identifying the operational pH window for IEX chromatography (deterministic experimental process development). Finally, the product's salt stability determines the mode of action for product capturing, being salt elution mode for salt-tolerant protein species and pH elution mode for salt-intolerant species respectively. The proposed systematic strategy was successfully applied in a case study for capturing the salt-intolerant enzyme human α -Galactosidase A from *Pichia pastoris*. The study enabled an effective low salt capture step under stable conditions by well-chosen prior screening experiments.

A second approach for downstream process intensification dealt with finding a similar deterministic experimental approach for the setup of hydrophobic interaction chromatography (HIC). The main focus was to correlate the protein binding behavior to the protein solubility and the binding pH. Solubility screening and chromatographic column

experiments were performed on a fully automated robotic workstation for proteins of acidic, neutral, and alkaline isoelectric points to cover a broad spectrum of biomolecules. In analogy to the strategy for ion-exchange chromatography discussed above, conditions of instant protein precipitation during the solubility screening were identified and eliminated from the miniaturized robotic chromatography column experiments. The screening results showed a direct correlation between the dynamic binding capacity and the operational pH as well as the overall solubility. Interestingly, process conditions of decreased product solubility were found to be beneficial for binding to the HIC column. As an inverse trend of protein binding and the binding kinetics was observed, a re-orientation process was assumed to be the determining factor for increasing the dynamic binding capacity in HIC. Overall, the dynamic binding capacity could be increased up to 30% compared to the standard purification procedure performed at neutral pH. Structural changes during the binding process were finally excluded by performing a principal component analysis (PCA) of protein spectra of feed and elution samples as another elegant analytical tool based on multi-variate data analysis, as discussed above.

The third section of this work describes a straightforward methodology for mechanistic chromatography modeling based on process UV data without any knowledge about the feed composition or concentrations. Industrial chromatography steps are mainly controlled based on process UV data, making this methodology highly relevant for diverse applications in terms of feed variability and pooling decisions. With the 'Quality by Design' guideline being close to be extended to *in silico* methods there is even more need for such straight-forward modeling tools. For UV-based mechanistic modeling of a complex *E. coli* feed stream presented here, the model equations including the transport dispersive model as well as the binding terms were modified to allow modeling based on boundary conditions in UV absorbance units. The molar concentration based terms were thus remodeled by including absorption coefficients so standard parameter estimation procedures based on inverse modeling (fitting the chromatogram to the measured data by altering parameters) were applicable. An anion-exchange chromatography case study of a target protein expressed in *E. coli* with diverse lumped impurity peaks demonstrated the practical applicability of the method. The product concentration from the feed material was estimated from solely UV-based chromatogram data and verified by HT-CGE.

In a final setup the above-mentioned tools for high-throughput process development and mechanistic modeling were merged into a project for integrated upstream and downstream process development. As upstream processes are highly complex and global models for the large variety of process conditions are hardly applicable, the upstream optimization was based on HT cultivation screenings in micro-scale format. The cell lysate characterization of the different cell lysates was performed *in silico* using the above-mentioned procedure of mechanistic UV-based inverse modeling of chromatography data. The created models were highly predictive and were used for optimizing IEX capture steps *in silico*. The validation experiments of the proposed optimal setups showed that shallow and steep gradients from the calibration runs could correctly predict setups which were unknown to the model. The overall integrated up- and downstream performance was determined by an *in silico* sampling of all possible elution pooling setups and subsequent Pareto optimization (multi-objective optimization), based on product recovery (initial titer \times purification yield) and purity. The so determined Pareto fronts of the different lysates resulted in a global set of optimal system points which need to be chosen according to the

value of the product. In an *E. coli* case study this proposed integrated approach showed large potential for concerted up- and downstream optimization, as was the main focus of the presented thesis. As upstream conditions were found to have a strong impact on the lysates' impurity profiles the 'Over the Wall' workflow of a strict separation of up- and downstream process design should be questioned.

Zusammenfassung

Die Prozessentwicklung in der biopharmazeutischen Industrie ist komplex und basiert auf der Optimierung einer Vielzahl an Teilprozessschritten. Heutzutage beruht ein Großteil der Ansätze zur Prozessentwicklung auf Meinungen von Experten oder Erfahrungsberichten und folgt einer strengen Hierarchie. Prozessstrategien zur Fermentation und Produktaufreinigung werden meist getrennt voneinander optimiert. Kultivierungsprozesse werden dabei hauptsächlich auf das Erzielen hoher Produkttiter ausgelegt. Weitere wichtige Parameter, wie z.B. die Reduktion von fr die Aufreinigung kritischen Verunreinigungen, werden meist vernachlässigt. Eine integrierte Betrachtung beider Arbeitsgebiete birgt das Potential die Gesamtprozessleistung zu erhöhen. Eine weitere Herausforderung bei der Entwicklung industrieller Prozesse ist der 'Time to Market' Anspruch (ein Prozess muss aus finanzieller Sicht schnellstmöglich marktfähig werden), sodass ausreichende Prozessoptimierung und -charakterisierung stark eingeschränkt ist.

Diese Dissertation befasst sich mit Strategien, um die genannten Herausforderungen der Industrie zu erfüllen. Dies wird durch die Entwicklung einer integrierten Plattform für kombinierte Prozessentwicklung im Kultivierungs- und Aufarbeitungsbereich (Up- und Downstream Processing) unter Verwendung moderner Technologie im Bereich der Hochdurchsatz-Verfahren sowie mechanistischer Modelle realisiert. Um einem durch Hochdurchsatz-Experimente entstehenden analytischen Engpass entgegenzuwirken, liegt ein weiterer Fokus der Arbeit in der Entwicklung von Hochdurchsatz kompatiblen Analytikmethoden, welche gleichzeitig eine ausreichend hohe Datenqualität für den Einsatz fr mechanistische Modelle bieten.

Der erste Abschnitt der Dissertation beschäftigt sich mit der Entwicklung und Untersuchung solcher hoch entwickelter analytischer Verfahren und deren Anwendung für Hochdurchsatz-Experimente und Modellierung. Das erste Konzept beschäftigt sich mit dem Einsatz eines Protein-Fusionstags mit besonderen optischen Eigenschaften für den quantitativen in-line Nachweis von Proteinen in Kultivierungs- und Aufarbeitungsanwendungen in Echtzeit. Der Cherry-TagTM ähnelt der Hämgruppe von Cytochrom C und kann im VIS-Bereich detektiert werden. Es wurde gezeigt, dass Protein destabilisierende Bedingungen zu einem Verlust der chromophoren Eigenschaften des Tags führen, sodass der Cherry-TagTM ein geeignetes Werkzeug zur Untersuchung von Proteinlöslichkeit und -stabilität darstellt. Unter Verwendung des Tags in Hochdurchsatz-Kultivierungsexperimenten zur Prozessentwicklung im 48-Well Format im BioLector[®] System (m2p-Labs, Deutschland) konnten Bedingungen für *Escherichia coli* SE1 identifiziert werden, bei der die Bildung des Produktes Glutathione-S-Transferase (GST) in nicht nativer Form von Einschlusskörperchen (Inclusion Bodies) erfolgte. In Aufreinigungsexperimenten konnte das rot gefärbte Enzym als aktive Proteinspezies identifiziert werden, da GST effektiv an eine Affinitätssäule gebunden werden konnte, deren Bindemechanismus auf der katalytischen Aktivität des Enzyms beruht. Der Cherry-TagTM erwies sich als verlässliches und genaues Hilfsmittel für analytische Anwendungen zur Echtzeit-Detektion von Proteininstabilität und -löslichkeit in verschiedenen Teilgebieten der biopharmazeutischen Prozessentwicklung. Zusätzlich konnte gezeigt werden, dass die optimalen Prozessbedingungen der Kultivierung zur Erzeugung großer Mengen an löslichem Produkt direkt übertragbar sind für das Produkt ohne Tag. Zusätzlich wurde eine effektive enzymatische Tag-Spaltreaktion vorgestellt, um den Fusionstag final zu entfernen.

Als zweites Konzept Hochdurchsatz kompatibler analytischer Methoden wurde multivariate Datenanalyse (MVDA) von Proteinspektren angewandt. Diese Methode nutzt Unterschiede im Spektrum des mittleren UV-Bereichs verschiedener Proteine, hervorgerufen durch Eigenschaften der Peptidbindungen und Aminosäurereste. Diese Unterschiede in den Absorptionsspektren verschiedener Proteine ermöglichen die Erstellung empirischer Modelle (Projection to Latent Structures (PLS) Regression) basierend auf einer Kalibrierung mit Proben bekannter Konzentration und Zusammensetzung, die für unbekannte Proben im kalibrierten Bereich genutzt werden können. Diese Methodik wurde erfolgreich für die Spektrenentfaltung von Hochdurchsatz-Multikomponenten-Isothermen Daten von Modellproteinen bei verschiedenen pH-Werten und Salzkonzentrationen verwendet. Die erzeugten Daten waren von ausreichender Qualität für Modellierungsanwendungen und ermöglichten die Erstellung von prädiktiven semi-mechanistischen Chromatographie Modellen. Neben der Genauigkeit der analytischen Methode kann sie insbesondere durch ihre schnelle und nicht invasive Durchführbarkeit überzeugen und eignet sich somit zum Einsatz in Hochdurchsatz-Prozessen auf roboterbasierten Plattformen ohne zusätzliche Probenentnahme.

Der dritte vorgeschlagene Ansatz einer Hochdurchsatz-Analytik beschäftigt sich mit Hochdurchsatz-Kapillargelelektrophorese (HT-CGE), die als zentrale Analytik für die Entwicklung einer Hochdurchsatz basierten Plattform für die Produktion virusähnlicher Partikel (VLPs) in *Escherichia coli* Rosetta verwendet wurde. Die Arbeit über VLPs kann als Schnittstelle zwischen der bisher behandelten analytischen Methodenentwicklung und den nachfolgenden Arbeiten über Prozessentwicklung im Kultivierungs- und Aufarbeitungsbe- reich betrachtet werden. Die Kultivierungsoptimierung im Hochdurchsatz erfolgte im 48-Well Format im BioLector[®] Mikrokultivierungssystem (m2p-Labs, Germany). Ein *Esche- richia coli* Rosetta Stamm, reguliert durch einen Tac-Promotor, wurde hinsichtlich hoher Ausbeuten des murinen Polyomavirus Kapsid Proteins (VP1) untersucht. Aufgrund einer hohen Zahl an untersuchten Kultivierungsbedingungen (Temperatur, Schüttlerdrehzahl, Induktionsbedingungen und Kulturmedien) musste eine schnelle und effektive Zellauf- schlussmethode entwickelt werden. HT-CGE erwies sich als kompatible Analytik zur Charakterisierung der löslichen und nicht löslichen Proteinfraktion aller untersuchter Kultivierungsbedingungen. Anhand der Ergebnisse der Kultivierungsexperimente wurde das optimale System auf einen großen Maßstab skaliert und durch Anionenaustausch- Chromatographie mit einem Membranadsorber und anschließender Größenausschluss- Chromatographie aufgereinigt. Auch der Aufreinigungsprozess konnte durch HT-CGE genau charakterisiert werden und ermöglichte eine hochauflösende Fraktionsanalytik. Als letzter Prozessschritt wurden die VP1-Proteine erfolgreich zu leeren Kapsidhüllen assem- bliert und über Transmissionselektronen-Mikroskopie (TEM) nachgewiesen.

Der zweite Abschnitt der Dissertation beschreibt Strategien zur Intensivierung von Auf- arbeitsprozessen durch Entwicklung systematischer Strategien für verschiedene Chro- matographiearten. Ein erster Ansatz behandelt die Aufreinigung salztoleranter Prote- ine durch Ionenaustauschchromatographie was aufgrund erhöhter Salzkonzentration bei der Produktelution eine Herausforderung sein kann. Da Produktverluste nicht nur auf suboptimale Prozessführung sondern ebenfalls auf Anfälligkeit des Proteins auf Prozess- bedingungen zurückzuführen sind, beinhaltet die Methode Stabilitätsuntersuchungen im Hochdurchsatz sowie eine globale Strategie zur Bestimmung optimaler Prozessbedingun- gen zur Proteinaufreinigung unter stabilen Bedingungen. Die Stabilitätsuntersuchungen

beinhalten Kurz- und Langzeitstudien bezüglich pH- und Salzeinflüssen, sowie Untersuchungen zum Einsatz stabilisierender Additive. Nach Festlegung der prozessrelevanten Grenzen und Sicherstellung der Proteinstabilität im untersuchten Bereich kann das entsprechende Zellysate mit pH-Gradienten-Chromatographie zur Feststellung des optimalen pH Fensters zur Aufreinigung analysiert werden (deterministischer experimenteller Ansatz). Letztendlich entscheidet das Salzstabilitätsverhalten des Proteins, ob eine Aufreinigung durch traditionelle Elution im Salzgradienten oder durch pH Shift für salzintolerante Spezies erfolgen muss. Die vorgeschlagene Methodik wurde erfolgreich in einer Fallstudie für das salzintolerante Enzym α -Galactosidase A von *Pichia pastoris* angewandt. So konnte durch sinnvoll gewählte Vorexperimente ein effektiver Aufreinigungsschritt unter produktschonenden Niedersalz-Bedingungen entwickelt werden.

Ein zweiter Ansatz zur Intensivierung von Aufarbeitungsprozessen beschäftigte sich mit der Entwicklung eines ähnlichen deterministisch experimentellen Ansatzes für hydrophobe Interaktionschromatographie (HIC). Das Hauptziel war die Korrelation des Proteinbindeverhaltens mit der Proteinlöslichkeit und dem pH Wert der Umgebung. Sowohl für die Löslichkeitsstudien als auch für die Säulenchromatographie-Experimente wurde eine voll automatisierte Roboterplattform verwendet. Es wurden Proteine mit saurem, neutralem und basischem isoelektrischen Punkt verwendet um ein breites Spektrum an Biomolekülspezies abzudecken. In Anlehnung an die zuvor erläuterte Strategie für Ionenaustauschchromatographie wurden produktschädigende Bedingungen vorab in Löslichkeitsstudien von den Säulenexperimenten ausgeschlossen. Die Studie zeigte eine direkte Korrelation zwischen dynamischer Bindekapazität und pH Wert der Umgebung sowie der Proteinlöslichkeit. Insbesondere Bedingungen verminderter Proteinlöslichkeit erwiesen sich als geeignet für erhöhtes Bindeverhalten bei HIC. Der festgestellte gegenläufige Trend von Bindeverhalten und der Bindekinetik ließ eine Umorientierung an der Adsorberoberfläche während der Beladung vermuten. Die maximale Bindekapazität konnte im Vergleich zu den Standardbedingungen bei neutralem pH um bis zu 30% gesteigert werden. Strukturveränderungen durch den Bindeprozess wurden durch eine Hauptkomponentenanalyse (PCA) an Proteinspektren des Ausgangsmaterials und der Elutionsproben ausgeschlossen. Hier erwies sich die multi-variate Datenanalyse (MVDA), wie zuvor im analytischen Teil erwähnt, als geeignete analytische Methode.

Der dritte Abschnitt dieser Arbeit beschreibt eine elegante Methode für mechanistische Modellierung von Chromatographie-Prozessen unter direkter Verwendung von UV-Prozessdaten ohne genaue Kenntnis der Zusammensetzung und Konzentration des Ausgangsmaterials. Industrielle Chromatographieprozesse werden hauptsächlich über UV-Prozessdaten gesteuert, sodass die entwickelte Methode hohe Relevanz für verschiedene Anwendungen betreffend Variabilität des Ausgangsmaterials und Festlegung der Elutionsgrenzen hat. Durch die möglicherweise zeitnahe Erweiterung der 'Quality by Design (QbD)' Richtlinie um computergestützten Methoden wird der Bedarf an unkomplizierten Modellierungswerkzeugen weiter steigen. Zur UV-basierten mechanistischen Modellierung des hier verwendeten komplexen *E. coli* Lysats wurden die Modellgleichungen des 'Transport Dispersive Models' sowie der Adsorptionsgleichungen (Isothermen) modifiziert, sodass die Randbedingungen ebenfalls in UV-Absorptionseinheiten verwendet werden konnten. Die Terme, basierend auf molaren Konzentrationen, wurden durch eingeführte Absorptionskoeffizienten abgewandelt, sodass eine Standard-Parameterbestimmung durch die inverse Modellierungsmethode (Fitten des Chromatogramms an die Messdaten durch

Veränderung der Parameter) erfolgen konnte. Eine Anionenaustausch-Fallstudie mit einem in *E. coli* exprimierten Zielprotein und verschiedenen überlagerten Kontaminantenpeaks bestätigte die praktische Anwendbarkeit der Methodik. Die Produktkonzentration des Ausgangsmaterials konnte allein anhand UV-basierten Daten des Chromatogramms geschätzt und durch Hochdurchsatz-Kapillargelelektrophorese bestätigt werden.

Der letzte Abschnitt der vorgelegten Dissertation beschreibt die Verknüpfung der zuvor erläuterten Werkzeuge der Hochdurchsatz-Prozessentwicklung und der mechanistischen Modellierung zur integrierten Optimierung von Kultivierungs- und Aufarbeitungsprozessen. Aufgrund der hohen Komplexität und Vielfalt von Kultivierungsprozessen sind globale Modelle für die Vielfalt an Prozessparametern kaum anwendbar, sodass die Kultivierungsoptimierung anhand von Hochdurchsatz-Experimenten durchgeführt wurde. Die Charakterisierung der großen Vielfalt an Zelllysaten erfolgte *in silico* unter Anwendung der zuvor erläuterten UV-basierten inversen mechanistischen Modellierungsmethode für Prozess-Chromatographiedaten. Die generierten Modelle erwiesen sich als höchst prädiktiv und wurden zur *in silico* Optimierung des Ionenaustausch-Aufreinigungsschritts verwendet. Die Validierungsexperimente der vorgeschlagenen Prozessabläufe zeigten, dass sowohl flache als auch steile Elutionsgradienten anhand der Kalibrierungsexperimente korrekt vorhergesagt werden konnten, obwohl die Aufreinigungsschritte dem System unbekannt waren. Die globale Leistungsfähigkeit des integrierten Ansatzes für Kultivierung und Aufreinigung wurde durch *in silico* Fraktionierung unter Betrachtung aller möglicher Elutionspools und anschließender Pareto-Optimierung (mehrkriterielle Optimierung) durchgeführt. Die betrachteten Parameter waren hierbei die Gesamtproduktausbeute (Anfangskonzentration \times Ausbeute in der Aufreinigung) und die Produktreinheit. Die so bestimmten Pareto-Fronten der verschiedenen Zelllysate ergaben eine Auswahl an optimalen Systempunkten die anschließend je nach Wert des Produktes gewählt werden müssen. In einer *E. coli* Fallstudie zeigte dieser integrierte Ansatz ein großes Potential für gemeinsame Optimierungsprozesse unter Berücksichtigung von Kultivierung und Aufarbeitung, was das Hauptziel der hier vorgestellten Dissertation war. Da gezeigt wurde, dass die Kultivierungsbedingungen einen großen Einfluss auf die Kontaminantenprofile der Zelllysate haben, sollte die strikte Trennung von Kultivierungs- und Aufarbeitungsprozessen zukünftig in Frage gestellt werden.

Contents

1	Introduction	15
1.1	Upstream Process (USP) Development	16
1.1.1	Expression Hosts	16
1.1.2	<i>E. coli</i> Expression System	17
1.1.3	Codon Usage	20
1.1.4	Process Management	20
1.2	Miniaturized USP Systems	22
1.3	Downstream Process (DSP) Development	24
1.4	Protein Purification Process	26
1.4.1	Heuristic Approaches & Platform Processes	27
1.4.2	Experimental Approaches	27
1.4.3	Model-based Approaches	30
1.5	Integrated USP - DSP Approach	33
2	Research Proposal	35
3	Publications & Manuscripts	37
	Integrated Development of Up- and Downstream Processes Supported by the Cherry-Tag TM for Real-time Tracking of Stability and Solubility of Proteins	41
	Deconvolution of High-throughput Multi-component Isotherms Using Multi- variate Data Analysis of Protein Spectra	65
	Establishment of a High-throughput Platform for the Production of Virus-like Particles in <i>Escherichia coli</i>	79
	Systematic Purification of Salt-intolerant Proteins in Ion-exchange Chromatog- raphy: The Example of Human α -Galactosidase A	107
	Influence of Binding pH and Protein Solubility on the Dynamic Binding Capac- ity in Hydrophobic Interaction Chromatography	131
	UV Absorption-based Inverse Modeling of Protein Chromatography	151
	High-throughput Micro-scale Cultivations and Chromatography Modeling: Pow- erful Tools for Integrated Process Development	167
4	Conclusion & Outlook	189
5	Abbreviations	193
6	References	197
7	Akademischer Lebenslauf/Curriculum Vitae	209

CONTENTS

1 Introduction

The biopharmaceutical industry has changed rapidly in the last couple of decades. The variety of products was expanded from the production of small molecules to more complex pharmaceuticals such as enzymes, antibodies, protein vaccines (e.g. virus-like particles), as well as hormones and growth factors (Grgacic and Anderson [2006], Walsh [2005]). Whereas larger molecules were initially extracted from natural sources like the human blood stream or specific cell tissues, genetic engineering enabled the directed production of complex pharmaceuticals in large quantities starting from the 1970s (Walsh and Murphy [1999]). Nowadays, the biopharmaceutical industry is at another turning point going from large quantity blockbuster production of therapeutic agents ('one drug fits all approach') to a more personalized medicine (Jain [2009]).

All stages of process development for biopharmaceutics are shifted from straightforward routines like antibody platform processes (Shukla *et al.* [2007]) to individual and tailored productions. As a consequence of this new trend and the 'Time to Market' demands, fast, reliable and predictive screening tools need to be developed in all stages of biopharmaceutical production. In that context the 'Quality by Design (QbD)' approach (Chhatre *et al.* [2011], ICH [2011]) demands a deeper process understanding not only for existing but also for new processes for gaining a more profound insight of the parameters' impact on the overall process. Possible approaches exploit systematic exploration of the design space based on Design of Experiments (DoE) and high-throughput process development (HTPD) (Kelley *et al.* [2008]). This leads to a miniaturization of experimental setups, a reduction in time and materials and allows for parallelization and automation of screening experiments e.g. on robotic workstations (Oelmeier *et al.* [2011], Wiendahl *et al.* [2008]). However, the generation of large data sets poses major challenges on subsequent analytical technologies creating a bottleneck in the process workflow. New analytical technologies still need to be of highest quality for producing accurate experimental results that can at best be used for predictive *in silico* models.

Such new technologies and tools can be implemented in all stages of biopharmaceutical process development as in upstream processing - USP (strain construction/ screening and cultivations), in downstream processing - DSP (cell disruption, clarification and purification), as well as in stability studies (crystallization, solubility screenings). The high information content might even enable predictions of influences of one certain unit operation on the following process step. Today, e.g. USP and DSP are still optimized separately, meaning that the cultivations of highest product titers are chosen for subsequent downstream process development. Concentration levels of critical impurities for DSP (e.g. product variants), however, can be regulated up or down based on different cultivation conditions (Gawlitzeck *et al.* [1995]). Thus, a compromise between product titer and depletion or reduction of certain impurities might lead to an overall superior process performance. High-throughput screening (HTS) tools and sophisticated models open up new opportunities for process design and even concerted process optimization strategies.

1.1 Upstream Process (USP) Development

The development of a successful biopharmaceutical upstream process is dependent on a large number of variables. The expression host, the expression vector, the inserted gene of interest as well as the cultivation conditions need to be scored and selected for maximizing the overall product concentration.

1.1.1 Expression Hosts

The first step in designing a process for new or existing biopharmaceuticals is selecting a suited expression system. Not only the choice of host cells (e.g. bacteria, yeasts, insect cells, mammalian cells, etc.) but also the expression vectors need to be screened and evaluated. Currently, recombinant proteins are predominantly produced in *Escherichia coli* - 40% of the recombinant protein market - and chinese hamster ovary cells (CHOs) - 50% of the recombinant protein market (Demain and Vaishnav [2009]).

Bacteria (e.g. *Escherichia coli* from the family of Enterobacteriaceae) as hosts have the advantage of short cultivation times and cost-effective media compared to e.g. mammalian cells like CHOs which are often used for antibody production. Protein expression in bacterial cells can be easily controlled as there exists a large variety of commercially available cloning vectors. The protein of interest may make up to 50% of the total soluble protein (TSP) in *E. coli* (de Marco *et al.* [2005]). However, this overproduction often leads to a bottleneck in protein biosynthesis leading to misfolded or incomplete product species without the desired biological activity. Additionally, bacteria are not capable of many post-translational modifications, e.g. glycosylations, disulfide bond formations, amidations and others. These incapacibilities introduce another source of non-active product species and in many cases formation of insoluble product agglomerates called inclusion bodies (IBs) (Demain and Vaishnav [2009]). For small protein species IBs can be transformed into the active form of the product, however, this process is rather complex starting with unfolding and different sequential refolding procedures (Berg *et al.* [2012], Vallejo and Rinas [2004]). For larger proteins such a process becomes more and more complex and expensive additives like folding enhancers (e.g. the major *E. coli* chaperonin GroEL - Vallejo and Rinas [2004]) need to be included into the reaction mixtures. Another drawback (especially in *E. coli*) is the high content of endotoxins which can trigger strong immune responses in humans and mammals in general (Voedisch *et al.* [2005]).

Yeast cells, being eucaryotic microorganisms, present a number of advantages. As for bacterial cells, cultivations of yeasts are easy to handle due to their low shear sensitivity and cultivation media are inexpensive compared to those for mammalian cells. In addition, yeast cell lines enable product secretion and some posttranslational modifications such as disulfide bond formations and glycosylation patterns (Wink [2011]). Those glycosylation patterns can, however, vary strongly from those produced by mammalian cells and hyperglycosylation was reported (Demain and Vaishnav [2009]). Cultivation durations, being in a range of several days, are much longer than those for bacteria, being not longer than 24 h in most setups.

In contrast to bacteria and yeasts, mammalian cell cultures as CHOs enable the production of recombinant proteins with identical activity and structure as proteins produced in the human body with almost all human translational modifications (Voedisch *et al.* [2005]). However, mammalian cell cultivations are very complex and cost intensive and

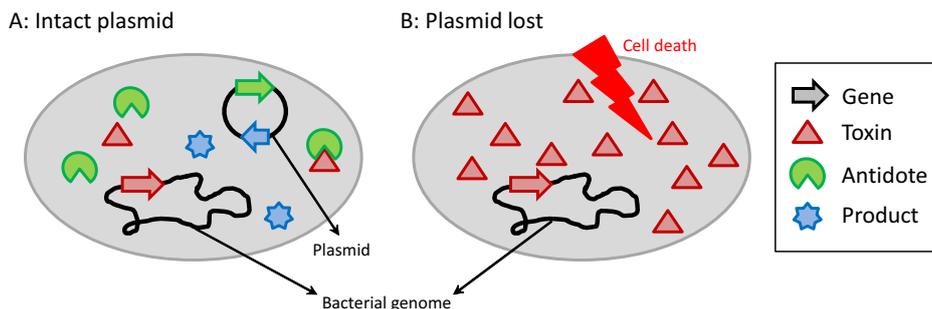


Figure 1: Schematic illustration of the StabyExpress™ dot - antidote technology, introduced by Delphi Genetics, Belgium (Stieber *et al.* [2008]). The toxin (red triangle) is constitutively produced as an element in the bacterial genome. A: Cell including an intact plasmid producing the antidote and the product. B: Cell without or lost plasmid, unable to produce the antidote leading to cell death.

suffer from poor recombinant protein secretion. Also, processes using mammalian cells as hosts carry the risk of contamination by viruses (Demain and Vaishnav [2009]). The lack of a cell wall makes mammalian cells very shear-sensitive leading to challenges in process control (Chmiel [2011]).

Besides bacteria, yeasts and mammalian cells, many other expression systems, i.a. insect cell lines, filamentous fungi, plant cell lines etc. exist (Demain and Vaishnav [2009]) but will not be discussed here.

1.1.2 *E. coli* Expression System

The following chapters will focus on *E. coli* as an expression host, being the central cell system of this PhD thesis. When a suited cell system is selected being capable of all needed requirements for producing the protein of interest, there are different possibilities of protein expression and process optimization in *E. coli*. Nowadays, a large variety of restriction enzymes, vectors, bacterial host systems and even straight-to-use cloning kits exist. Besides phages, cosmids and artificial chromosomes the most frequently used vectors are bacterial plasmids allowing insertion of up to 10 kb cDNA being sufficient for most recombinantly expressed proteins in the biopharmaceutical sector (Wink [2011]). Such vectors can then be transformed into the cells (e.g. by electroporation and chemical treatment - Shrivastava [2013]) as additional genetic elements being replicated independently from the host genome. Some plasmids can also be integrated into the bacterial genome. In general, plasmids are equipped with several basic elements (Wink [2011]):

- Origin of replication (ori) for increasing the copy numbers inside the transformed cells (e.g. derived from different plasmids like pBR322-ori or pUC-ori)
- Marker sequence as a selection marker (e.g. antibiotics resistance like ampicillin Amp^R / kanamycin Kan^R or dot - antidote systems like *ccdA*/*ccdB* system, etc.)
- Regulated promoter (RNA polymerase binding site including operator response elements for regulation) for transcription (e.g. *lac*, *tac*, *trp* etc. - de Boer *et al.* [1983])
- Multiple cloning site (MCS)/ Polylinker with different recognition sites for restriction enzymes to include the gene sequence of the product

Promoter	Origin & References	Expression Characteristics
araBAD	arabinose metabolic operon (Guzman <i>et al.</i> [1995])	Inducible by arabinose Glucose/ Fucose repressed
lac	lac-operon (de Boer <i>et al.</i> [1983])	Inducible by IPTG Basal expression
pL/pR	Bacteriophage λ (Elvin <i>et al.</i> [1990])	Temperature regulated
SP6	SP6 bacteriophage (Kang and Wu [1987])	Constitutive with SP6-RNA-Polymerase
T7	Bacteriophage T7 (Kang and Wu [1987]) (Tabor and Richardson [1984])	Constitutive with T7-RNA-Polymerase
T7-lac	Bacteriophage T7 and lac-operon (Tabor and Richardson [1984]) (de Boer <i>et al.</i> [1983])	Inducible by IPTG T7-RNA-Polymerase controlled by lac Almost no basal expression
tac	Hybrid of lac and trp (de Boer <i>et al.</i> [1983])	Inducible by IPTG
trp	tryptophan-operon (de Boer <i>et al.</i> [1983])	Repressible by tryptophan

Table 1: Overview of frequently used bacterial promoters including origin, references and expression characteristics

The marker sequence helps selecting successfully transformed cells from the wild type species. Orthogonal antibiotic strategies, meaning that two different antibiotic resistance genes are included into the host genome and into the plasmid, help overcoming these shortcomings. Employed cell types can be selected from contaminating species by an antibiotic resistance integrated into the bacterial genome. The second antibiotic resistance is included in the plasmid so that transformed cells can be distinguished from the wild type species (Wink [2011]). A non antibiotic based alternative was introduced by e.g. Delphi Genetics (Belgium), being based on a dot - anti-dote (*ccdA/ ccdB*) system illustrated in Fig. 1 (Stieber *et al.* [2008]). The *ccdB*-gen is encoding for a bacterial toxin and is integrated into the bacterial genome. The genetic code of the antidote (*ccdA*) is included into the plasmid sequence leading to cell death when the plasmid is lost. Consequently, the system cannot only be used as a selection marker but also for enhancing plasmid stability inside the transformed cells.

The choice of the promoter system is one of the most central aspects when searching for a suited vector system. The employed systems should be strictly controlled without any undesired basal activity and expression rates should be high. A selection of frequently used bacterial promoters including origin and expression characteristics are given in Table 2 showing different possibilities for induction, repression and overall regulation. For example the T7 promoter derived from the bacteriophage T7 is transcribed very strongly by a special T7 RNA Polymerase but not by the host RNA Polymerase (Kang and Wu [1987], Tabor and Richardson [1984]) and is a very commonly used system for overproduction of recombinant proteins. Cell growth and the production phase can thus be decoupled

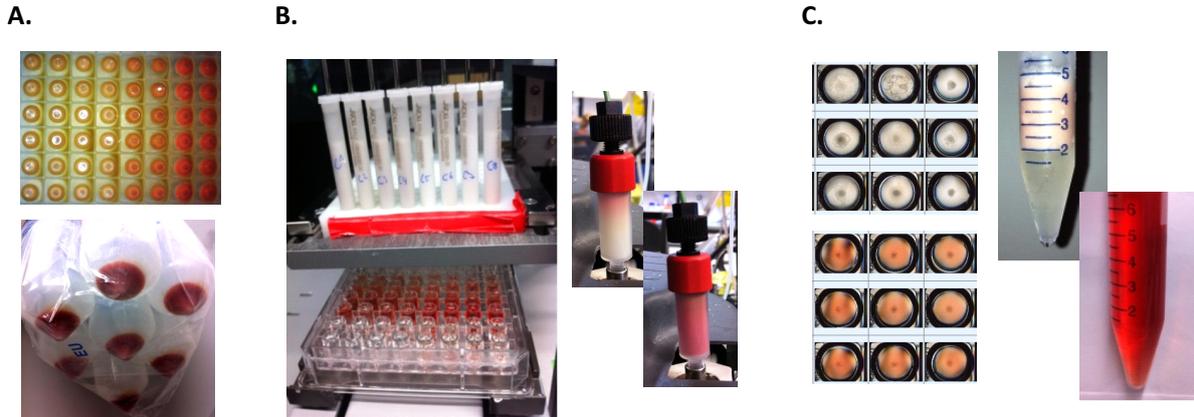


Figure 2: Overview over different applications of the Cherry-TagTM technology. A: Upstream applications for finding conditions of soluble protein expression and scale-up validation B: Downstream applications for product tracking C: Stability studies for increasing product solubility and for identifying conditions of instant product denaturation

by putting the gene encoding for the T7 RNA polymerase under the control of a second promoter e.g. lac (de Boer *et al.* [1983], Tabor and Richardson [1984]). The first growth phase can then be used for biomass accumulation as the gene for the T7 RNA polymerase is repressed by the lac repressor followed by the recombinant protein production phase after addition of IPTG which inactivates the lac repressor (Fig. 3). In straight-to-use cloning kits such sophisticated promoter fusions are realized by integrating the T7 RNA polymerase gene into the host genome under control of the lac promoter whereas the protein of interest is cloned into a bacterial plasmid and controlled by the T7 promoter as shown in Fig. 3. As a result the host strain can be used for all combinations of products and only the plasmid needs to be changed.

Very strong promoters like T7 can, however, result in a metabolic burden which might decrease cell growth and might yield in inclusion body formation (Huber *et al.* [2009]). A reasonable induction strategy is therefore one of the central tasks to tackle (section 1.1.4). For some products choosing a weaker promoter like tac is another option.

Conveniently, straight-to-use cloning kits are available including suitable cell lines for transformation and bacterial vectors as described above. Many of these employed plasmids include additional genetic elements encoding for fusion proteins which are appended to the product sequence. Such tags are mostly solubility/ stability enhancers (e.g. Cherry-TagTM - Azhar and Somashekhar [2014], Padmanabha Das *et al.* [2009], Halo-Tag[®]- Ohana *et al.* [2009]), visibility markers (e.g. Cherry-TagTM - Padmanabha Das *et al.* [2009], GFP-Tag - Cabantous *et al.* [2005]) and affinity sequences for simplifying the purification process (e.g. His-Tag - Ramos *et al.* [2004], GST-Tag - Lipin *et al.* [2008], Halo-Tag[®]- Ohana *et al.* [2009]). An overview over different applications of the Cherry-TagTM technology is shown in Fig. 2. In upstream applications it helps finding suitable conditions for soluble protein expression and scale-up validation (Fig. 2A), in downstream applications the product can be tracked throughout the purification process (Fig. 2B). For stability studies (Fig. 2C), the Cherry-TagTM helps increasing the product solubility and can be used for identifying conditions of instant product denaturation.

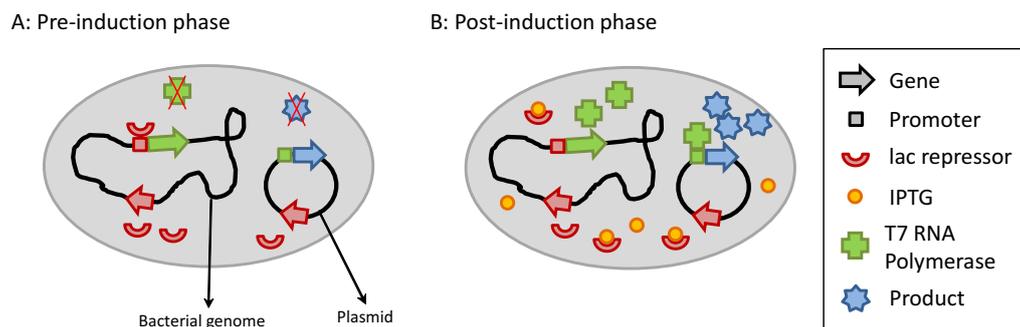


Figure 3: Schematic illustration of the T7-lac system of the pStaby plasmid by Delphi Genetics, Belgium. The lac repressor (red) is constitutively produced as an element in the bacterial genome and additionally on the plasmid. A: The lac repressor hinders the production of the T7 RNA polymerase until induction with IPTG (orange). As a result the T7 promoter (green) is not transcribed and the recombinant protein (blue) is not expressed. B: Addition of the inducer (IPTG) results in a removal of the lac repressor and the T7 RNA polymerase (green) is expressed. As a consequence the T7 promoter is transcribed and the protein of interest (blue) is produced.

1.1.3 Codon Usage

However, even with a suited *E. coli* strain and a stable plasmid of high copy number the product expression level might be low. One of the main reasons for such behavior is the codon (base triplet) usage of bacterial cells. Different codons encode for the same amino acid and expression hosts have distinct codon usage (Baneyx [2004]). Some genetic elements from e.g. human-derived genes include base triplets that are translated by a low copy number of transfer RNAs (tRNAs) during protein biosynthesis (rare codons) as shown in Table 2 (Makrides [1996]). This results in a biochemical bottleneck which may induce low expression, premature translational termination and changes in the amino acid sequence (Baneyx [2004]). A possible solution is to replace such rare elements by codons commonly used in *E. coli* in a codon optimization. This results in a modified messenger RNA (mRNA) which is not effecting the amino acid sequence and can improve protein expression in *E. coli* as was shown for short-chain dehydrogenase/reductase genes (SDRs) (Burgess-Brown *et al.* [2008]) and a recombinant malaria candidate vaccine (Zhou *et al.* [2004]). A coexpression of rare tRNAs is another alternative for eliminating these shortcomings in protein expression (Makrides [1996]).

1.1.4 Process Management

High yields of expressed protein in *E. coli* are not only influenced by the expression system itself but by the cultivation strategy used. The following section will focus on process management in batch cultivation mode. Other modes of operation (e.g. fed-batch and continuous cultivations) will be excluded.

1.1.4.1 Media & Additives

Choosing suitable media combinations and additives can strongly enhance protein solubility, stability and activity. Not only a medium type (e.g. minimal, complex, defined media) needs to be chosen but those media types also need to be evaluated among each other.

Amino acids	Rare Codons
Arginine	AGA, AGG, CGA, CGG
Cysteine	UGU, UGC
Glycine	GGA, GGG
Isoleucine	AUA
Leucine	CUA, CUC
Proline	CCC, CCU, CCA
Serine	UCA, AGU, UCG, UCC
Threonine	ACA

Table 2: Rare codons in *E. coli* (Makrides [1996])

Losen *et al.* [2004] used two different rich complex media and enhanced the benzoylformate decarboxylase productivity in *E. coli* about factor 10 when comparing LB and TB media. Changing the pH level of the employed media can also enhance protein production as was shown by Kopetzki *et al.* [1989] for α -Glucosidase under lowered initial cultivation pH. Also, different levels of trace elements and additives affected productivity. Huber *et al.* [2011] showed that the production of a fluorescent protein was enhanced by introducing a phosphate limitation. In the study of Gill *et al.* [2000] addition of DTT before induction yielded in enhanced productivity of GFP-chloramphenicol-acetyltransferase. Each cell type in combination with different expression vectors and products have different requirements concerning media compositions and can differ majorly in product titers as well as impurity profiles. Consequently, screening for such compositions is a crucial factor for resulting in a successful upstream process.

1.1.4.2 Process Conditions

Besides the media composition, the process conditions need to be set carefully for enhancing the productivity of cells. As an example, the cultivation temperature influences cell growth, metabolic pathways and protein folding kinetics. Different studies indicated that a reduction in temperature is beneficial for some proteins for soluble and correctly folded product species. These findings were shown by e.g. Schein and Noteborn [1988] for human interferon- α 2, human interferon- γ , and the interferon-induced murine protein Mx and Shirano and Shibata [1990] for rice lipoxygenase L-2 in *E. coli*.

As a second factor the oxygen transfer rate was identified as a central factor for soluble protein formation. Dikshit *et al.* [1990] found an increase in *Vitreoscilla* globin production in *E. coli* under oxygen limitations. Similar results were obtained by Losen *et al.* [2004] when expressing benzoylformate decarboxylase from *Pseudomonas putida* in *E. coli*. Increased product concentrations were investigated for oxygen limited cultures, whereas the overall cell density followed an inverse trend. Overall, introducing stress factors can help for enhancing product formation. Nevertheless, a minimum oxygen supply must be guaranteed to prevent cell death and product degradation.

1.1.4.3 Growth State

The growth state of the cell often determines whether or in which form a protein is produced. The typical bacterial growth curve in batch cultivation mode can be divided into four different phases (Barton [2005]):

- Lag-phase in which bacteria adapt to a new environment and the growth rate is very slow
- Exponential or log-phase in which bacteria have the highest reproduction rate
- Stationary phase in which the number of new cells produced equals the number of dying cells due to limitations
- Death phase in which a limitation or toxic environment leads to cell death

When the sequence of the protein of interest is coupled with an inducible promoter the bacterial growth and production phase can be decoupled. It is important to screen for induction conditions e.g. the induction time (dependent on cell density) and the inducer concentration for enhancing the productivity of cells. Whereas in traditional setups the induction time is set to $OD_{600\ nm} = 0.7 - 0.8$ AU being the early exponential phase (Wink [2011]), newer studies indicate that an induction at higher cell densities can be highly beneficial. Induction in the early growth phase can result in a metabolic burden which hinders cell growth and might lead to bottlenecks during protein biosynthesis (Huber *et al.* [2009]). Wrongly folded proteins as inclusion bodies might be the consequence as discussed earlier. Ou *et al.* [2003] reported the enhancement of soluble expression of RNA binding protein in *E. coli* when performing the induction in the early stationary phase. Similar results were obtained by Galloway *et al.* [2003] for human proteins.

In a study by Huber *et al.* [2009] an extended DoE approach of different induction times and inducer concentrations was performed using a robotic platform for producing high concentrations of the fluorescent protein EcFbFP. An induction during the early log-phase was found beneficial for productivity and the inducer concentration showed different trends depending on the induction time. Low inducer concentrations were beneficial for early induced cells, whereas the inducer concentration was negligible for late induced cells.

1.2 Miniaturized USP Systems

The previously presented literature review of different possibilities in process management implies the need for screening tools allowing for investigating large numbers of process parameters. Industrial cultivations in fully instrumented bioreactors have the highest process information available but cannot be used for high-throughput experimentation. A large variety of miniaturized high-throughput cultivation systems were described in literature, mostly lacking of accessible process information (Betts and Baganz [2006]). A compromise between basic process control and HTS capabilities offer small-scale stirred vessels in milliliter-scale run in parallel, imitating industrial scale fermenters (Akkoc *et al.* [2011], Lamping *et al.* [2003], Puskeiler *et al.* [2005], Wollerton *et al.* [2006]). Altenbach-Rehm *et al.* [1999] introduced a HTS compatible system of small-scale bubble column reactors. Improvements have also been made for standard shake flask cultures where

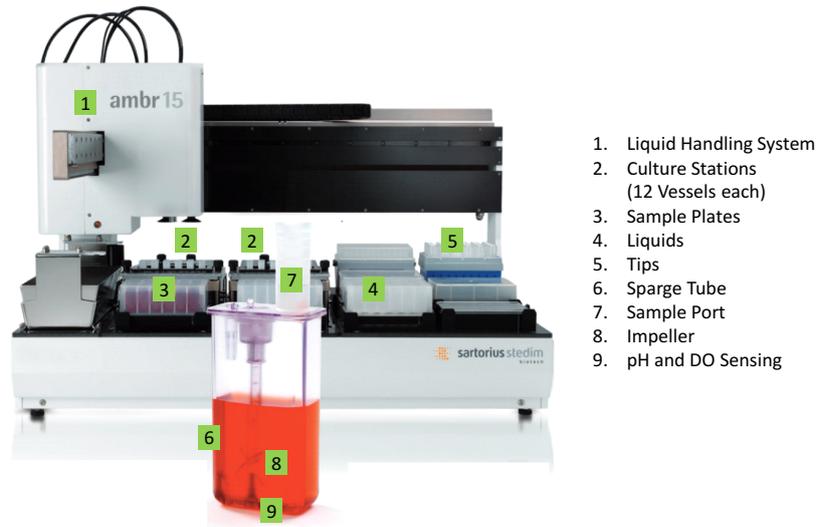


Figure 4: Illustration of the ambr[®]15 cell culture system by Sartorius Stedim Biotech, Germany integrated into a robotic workstation (illustration taken from: www.sartorius.us - Poster Publication No: SBI1536-e150301 by Tait et al.).

oxygen and carbodioxide rates can e.g. be monitored using the Respiration Activity Monitoring System (RAMOS[®]-Technology) and pH values can be determined by optical measurements using fluorescent dyes (Anderlei *et al.* [2004], Scheidle *et al.* [2007]). In terms of high-throughput, micro-scale cultivations in microtiter plates are optimal for exploring the experimental design space as shown by Duetz [2007]. However such a simplification often results in a pronounced reduction in accessible process information and control and can produce results that are not transferable for scale-up.

An ideal HTS tool for upstream process development must combine both, a high amount of process information and control, as well as low material consumption and low culture volume. An example for such a sophisticated technology is the Advanced Microscale Bioreactor (ambr[®]) by Sartorius Stedim Biotech, Germany (Fig. 4). It mimics lab-scale fermenters using 24 or 48 parallel bioreactors of 10 - 15 mL, operated on fully automated robotic workstations (www.sartorius.us - Poster Publication No: SBI1536-e150301 by Tait et. al). Each block of 12 bioreactors can be set to a separate temperature and shaking speed and for each reactor the oxygen uptake as well as the pH can be monitored continuously using optical sensors (Ratcliffe *et al.* [2012]). An additional impeller at the reactor bottom enables good oxygen uptake and mixing. The combination with a pipetting robot enables fed-batch cultures and sampling for determination of cell densities and product levels. This new technology was successfully applied for optimizing monoclonal antibody processes in CHO cells (Hsu and Aulakh [2012], Rameez *et al.* [2011]) and for scale-down experiments of haematopoietic stem cell suspension cultures (Ratcliffe *et al.* [2012]).

The BioLector[®] system by m2p-Labs (Germany), as was used in this PhD thesis, is another elegant alternative for HT upstream screenings. The cultivation experiments are carried out in sub-milliliter scale in continuously shaken 48-well plates equipped with an optical bottom. Cell growth can be detected quasi-continuously via scattered light signals without the need of sampling as necessary in shake flask experiments and the

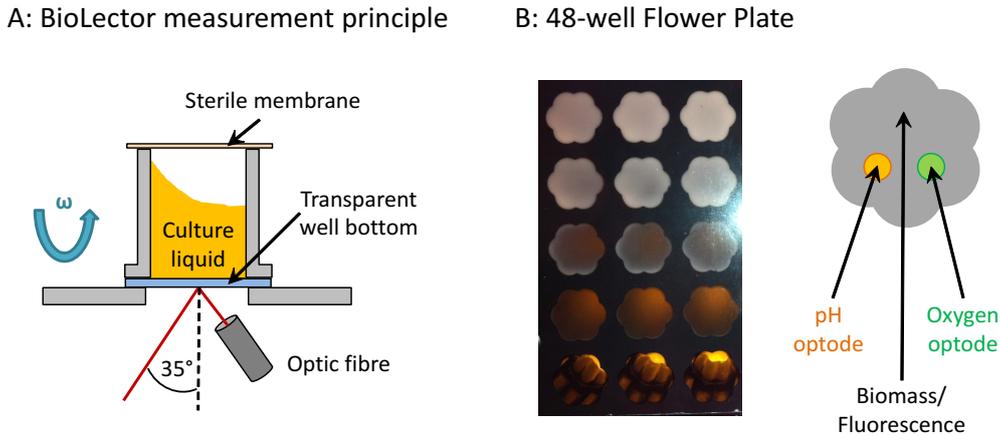


Figure 5: Schematic overview over the BioLector[®] system by m2p-Labs, Germany. A: BioLector[®] measurement principle based on scattered light determinations with a defined angle of the excitation light beam of the optical fiber according to Kensy *et al.* [2009a]. The optical fiber is fixed to the shaken parts making measurements possible throughout the process without pausing the system. B: Photograph of an inoculated 48-well FlowerPlate[®] (left) and a schematic drawing of the optical bottom with additional optodes for pH and oxygen level measurements (right).

ambr[®] technology (Kensy *et al.* [2009b], Kottmeier *et al.* [2009]) as illustrated in Fig. 5A. Different additional filters enable the determination of different colorimetric products such as GFP or labeled proteins. So called Flower Plates were developed, designed with an optimal geometry to counteract inhomogeneous mixing and low oxygen uptake in microtiter plates (Funke *et al.* [2009]). Optionally, FlowerPlates[®] with additional optodes fixed to the well bottom of the plate are accessible for pH and oxygen uptake measurements (Wenk *et al.* [2012]) as shown in Fig. 5B. Changes in the chemical environment in the cultivation well result in a change in fluorescence of the dyes, as was discussed earlier for shake flask cultures.

Further improvements of the BioLector[®] system was realized by fusion with a fully automated robotic station called RoboLector[®] (Kensy *et al.* [2012]). By using this technology, fully automated screenings for media, additives and induction profiles can be realized (Huber *et al.* [2009]). Nowadays, even a microfluidic version of the BioLector[®] exists allowing for feeding strategies and pH adjustment (Funke *et al.* [2010]). The BioLector[®] system has been employed for different studies, e.g. for identifying *C. glutamicum* succinate production strains (Kinast *et al.* [2012]), optimizing the secretory production of a cutinase in *C. glutamicum* (Rohe *et al.* [2012]) and optimizing the production of the recombinant parathyroid hormone (rPTH) fragment 1-34 in *Hansenula polymorpha* (Mueller *et al.* [2013]). As a straightforward scale-up strategy based on constant oxygen uptake rate (the sulfite oxidation method) exists (Hermann *et al.* [2003], Kensy *et al.* [2009a]), optimization results are directly transferable for production scale making these HTS technologies compatible with the bioprocess development workflow.

1.3 Downstream Process (DSP) Development

In *E. coli* recombinant proteins are mostly expressed intracellularly demanding for a cell disruption strategy after cell harvest. In contrast to extracellular proteins as produced in

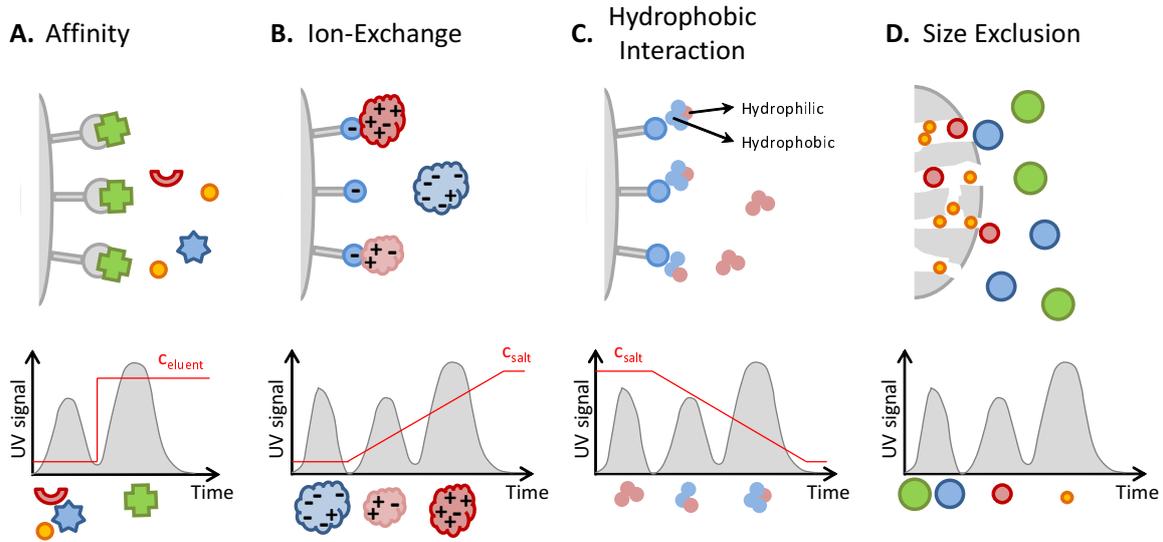


Figure 6: Schematic overview of different chromatographic purification modes: A: Affinity chromatography with a receptor binding the product selectively. Elution is performed by a shift in chemical environment or competitive ligand. B: Ion-exchange chromatography with a ligand binding oppositely charged proteins. Proteins are eluted by increasing the salt concentration or changing the buffer pH. C: Hydrophobic interaction chromatography with a hydrophobic ligand binding hydrophobic surface patches of proteins. Proteins are bound at high salt conditions and are eluted by decreasing the salt concentration. D: Size exclusion chromatography with chromatography media of a certain pore size distribution. Proteins of different size have different access to these pores and thus have differing residence times inside the chromatography column.

many cell systems like the methylotrophic yeast *P. pastoris* (Wink [2011]) the crude feed stock from disrupted cells is much more complex. It contains high amounts of impurities like cytosolic proteins, nucleic acids, substances from the cultivation process, as well as many other macromolecules like sugars and fatty acids. Such a composition represents a challenge for the purification process in order to meet the requirements stated by the authorities like the Food and Drug Administration (FDA - <http://www.fda.gov/>) in the USA or the European Medicines Agency (EMA - <http://www.ema.europa.eu>). New technologies in protein purification in terms of strategies, materials and additional statistical and *in silico* methods resulted in major advancements in the biopharmaceutical industry for designing new or improving existing drug substances and meeting the given regulatory standards. Typical protein purification procedures include selective precipitation/ crystallization of the product or contaminants, extractions using aqueous two-phase systems (ATPS), filtration technologies (e.g. ultra-/ diafiltration), and predominantly chromatographic procedures (Ahuja [2000]).

Chromatography exploits different interactions of molecule species (proteins, nucleic acids etc.) in the mobile phase with a certain type of stationary phase resulting in distinct residence times or binding behavior (Janson [2011]). A schematic overview of different chromatographic purification modes used in this thesis is shown in Fig. 6. Affinity chromatography (Fig. 6A) is based on a specific interaction between the protein of interest and a ligand (e.g. Protein A resins for monoclonal antibodies, glutathione resins for proteins with GST-tag etc.) (Carta and Jungbauer [2010]). In theory, only the desired protein species is bound to the ligand whereas all impurities can be found in the flow-through

fraction. Nevertheless, especially for low column loading some product-resemblant contaminants and product binding species (backpack proteins) can result in certain levels of impurities in the elution fraction. The elution can be performed by weakening the binding affinity after changing the chemical environment (e.g. pH shift) or by introducing a species that has a higher affinity than the protein of interest. Ion-exchange chromatography (Fig. 6B) exploits the electrostatic binding between charged ligands (positive for anion-exchange/ negative for cation-exchange) and charged protein patches (Roos [1999], Walsh [2002]). Under low salt conditions proteins of identical charge as the ligand are washed through, whereas oppositely charged molecules are bound. The elution is then performed by increasing the salt concentration or changing the buffer pH and proteins elute according to their binding strength. Hydrophobic interaction chromatography (Fig. 6C) employs hydrophobic ligands that bind hydrophobic surface patches of proteins (Walsh [2002]). For the binding procedure a high salt condition is used which enhances hydrophobic interactions and highly hydrophilic proteins are washed through the column. The elution procedure follows a decreasing salt concentration whereas weakly hydrophobic protein species elute first. For size exclusion chromatography (Fig. 6D) no special ligands are used but porous media of a certain distribution of pores (Carta and Jungbauer [2010]). Dependent on the protein size these pores are mostly accessible or inaccessible yielding in different residence times within the chromatography column and thus a different elution time based on size. Small molecules like salts have access to all pores and elute as the final species. Additionally, mixed mode interaction chromatography resins are available e.g. combining both hydrophobic and electrostatic interactions (Zhao *et al.* [2009]) which were successfully employed for e.g. an alternative antibody capturing process (Pezzini *et al.* [2011]).

1.4 Protein Purification Process

The protein purification process after cell disruption of different biopharmaceuticals can be divided into four process steps (Ahuja [2000]):

- Solid-liquid separation/ Clarification
- Product isolation/ Capturing
- Product purification/ Intermediate purification
- Product polishing

Whereas solid-liquid separation using precipitation, centrifugation and filtration steps are rather straightforward for most cell extracts, it will not be discussed here. However, subsequent unit operations can be challenging. During product isolation (Capturing) the main focus is to reduce the sample volume by removing large quantities of water. Also critical impurities such as proteases are removed for product stabilization and isolation (Wen *et al.* [2015]). Predominantly used unit operations include affinity or ion-exchange chromatography, precipitation steps and also extractions using ATPS.

The main objective of the subsequent product purification (intermediate purification) step is the removal of impurities with similar physico-chemical properties as the product. At this stage orthogonal downstream strategies (unit operations exploiting different protein specific properties than the capturing step) must be considered (Wen *et al.* [2015]).

Typical unit operations include ion-exchange chromatography (IEX), hydrophobic interaction chromatography (HIC) or selective precipitation.

In a final step (Product polishing) product-related contaminants such as charge variants or aggregates and additional trace contaminants are removed (Janson [2011]). Additionally, the product is transferred into a stabilizing environment using unit operations for buffer exchange. Polishing procedures include ultra- and diafiltration procedures, as well as further chromatographic procedures such as size exclusion chromatography (Wen *et al.* [2015]).

The selection and combination of the above-mentioned separation technologies in combination with different modes of operation demands tools for process understanding, optimization and exploring the experimental design space. Especially with the 'Time to Market' demands the need for a basic understanding of the parameters' influence on the process performance is of major importance. Three fundamentally distinct approaches can be used for designing purification processes for proteins (Guiochon and Beaver [2011], Nfor *et al.* [2009]): expert-based approaches (heuristics and platform processes), experimental approaches (high-throughput process development and statistical design of experiments) and model-based approaches.

1.4.1 Heuristic Approaches & Platform Processes

Heuristic or knowledge-based approaches make use of expert knowledge and conservative rules of thumb. Collections of such rational design strategies were captured in databases in so called expert systems including facts about main impurities encountered (Leser and Asenjo [1992], Lienqueo and Asenjo [2000]). Whereas such rules may help for finding suitable starting points still the design space for finding an overall optimum is rather limited resulting in suboptimal process conditions (Nfor *et al.* [2009]). Additionally, some ideas suffer from being based on outdated laboratory equipment and questionable scientific theories (Guiochon and Beaver [2011]).

Platform processes, in contrast, are sophisticated downstream strategies which are known as highly effective processes for a certain type of molecule and might be used as a template for additional resemblant products (Guiochon and Beaver [2011]). Such platforms are known to result in more effective starting points than employing heuristics for new products of a similar type. A very well-known example is the monoclonal antibody (mAB) platform process being based on a Protein A affinity chromatography capture step (selective binding of the mAB F_C region) and two additional purification steps for removal of residual host cell proteins (HCPs), DNA and product-related impurities such as aggregates (Shukla and Thoemmes [2010]).

1.4.2 Experimental Approaches

Designing processes for new products and expression paths is rather difficult and heuristic approaches and platform processes are based on knowledge derived from past studies and similar existing processes. In cases of no founded knowledge about a certain bioprocess or protein a straightforward solution can be experimental approaches, based on a large number of screening experiments for exploring the design space. High-throughput process development becomes the method of choice for investigating large sets of parameters in downstream processing (Kelley *et al.* [2008]). Often such miniaturized technologies are

coupled with statistical design of experiments (DoE), optimal experimental design (OED) or genetic algorithms (GA) for gaining a more profound process understanding and for reducing the number of relevant experiments (Nfor *et al.* [2009]).

High-throughput systems for downstream processing unit operations are widely employed in biotechnological research, as well as in biopharmaceutical companies. A large variety of HTS strategies exist for almost all unit operations employed during protein release and purification which are reviewed in the following chapter.

1.4.2.1 HT Cell Lysis

Cell disruption methods in terms of chemical treatment are straightforward for HTS applications. Cell lysis can be achieved by alkaline pH, chaotropes like guanidinium hydrochloride or urea, detergents, solvents etc. (Harrison [1991]) in standard 96-well plates. Analogously, temperature treatment, osmotic shocks, freeze drying and enzymatic strategies are easy to use methodologies. Cell disruption by mechanical treatment can be realized using beating beads in combination with agitation, as was used for the analysis of bacteria from digesta and fecal samples (Yu and Morrison [2004]) and salivara bacteria (Lazarevic *et al.* [2013]) using the NucleoSpin[®] bead system (Machery-Nagel, Germany). Nowadays, also more sophisticated sonication devices were adapted for HTS applications, e.g. a 24-well HTS sonication device (Hohnadel *et al.* [2014]) which was successfully employed for 15 different microbial cell types (*Staphylococcus*, *Pseudomonas*, *Salmonella*, *Aspergillus*, *Candida* etc.) including gram negative and gram positive bacteria. Poudineh *et al.* [2014] introduced three-dimensional, sharp-tipped electrodes for cell disruption by electric fields in parallel microchannels shown for the extraction of biomarkers from *E. coli* cells.

1.4.2.2 HT Solubility Screenings

HTS technologies are also widely applied for screenings on colloidal stability of proteins and investigating protein phase behavior. Wiendahl *et al.* [2009] developed a fast 96-well method for the determination of protein solubility lines on a robotic liquid handling station for lysozyme and different insulin variants. The experimental setup includes system preparation (pipetting steps), equilibration and several consecutive steps of optical analysis and liquid evaporation until crossing the solubility line. The Rock Imager by Formulatrix (USA) is an automated imaging system which cannot only distinguish between the soluble and insoluble state but can also identify e.g. aggregates, crystals, gelation states etc. This technology was proven to be highly useful for investigating protein phase phenomena as was shown for an antibody (Rakel *et al.* [2014]) and lysozyme, glucose oxidase as well as glucose isomerase (Baumgartner *et al.* [2015]).

1.4.2.3 HT Inclusion Body Refolding

As discussed earlier the formation of inclusion bodies is a major problem during upstream processing. Nevertheless, for some proteins effective unfolding and refolding procedures exist making such processes highly feasible. High-throughput technologies can be used in screening for the large number of variables for inclusion body refolding processes. Berg *et al.* [2012] optimized an HTS refolding process for lysozyme on a fully automated robotic

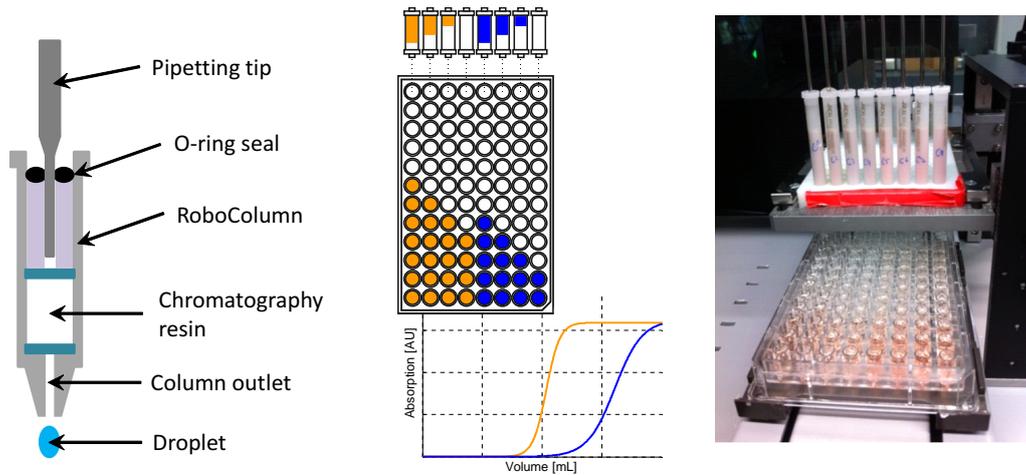


Figure 7: Schematic illustration of the RoboColumn[®] system by Atoll (Germany) according to Wiendahl *et al.* [2008]. A cross-section of the RoboColumn[®] is shown on the left, droplet collection in 96-well UV plates and spectral evaluation in the middle and a laboratory photograph during breakthrough experiments on the right.

workstation by varying solute pH and applying different additives like sodium chloride, ammonium sulfate, dithiothreitol and glutathione. Vincentelle *et al.* [2004] developed a similar process for different proteins expressed in *E. coli* considering additional additives like folding enhancers (chaperones), glycerol, polyethylene glycol, co-factors and redox components. Commercially available HTPD refolding kits include the Pro-Matrix[™] system (Thermo Scientific, USA) and the QuickFold[™] system (Athena Enzyme Systems, USA) (Bhambure *et al.* [2011]).

1.4.2.4 HT Aqueous Two-phase Systems

HTPD in screening for aqueous two-phase systems (ATPS) was introduced by Bensch *et al.* [2007] with a fast methodology for spectrophotometric determination of binodal curve and conodes using dyes. The method was successfully applied on bovine serum albumine (BSA) as a model protein as well as on a complex mixture of a monoclonal antibody and host cell proteins using a PEG4000 - phosphate system (Oelmeier *et al.* [2011]).

1.4.2.5 HT Chromatography

For liquid chromatography a large variety of HTS equipment exists allowing for static as well as dynamic binding experiments. Batch chromatography in filter plates can be used for optimizing binding and elution conditions, determination of isotherm data as well as kinetic studies. Such systems have been applied for the angiotensin-II-generating enzyme from a raw extract from renal tissue using strong cation- and anion-exchange resins (Thiemann *et al.* [2004]), α -amylase from *B. subtilis* and scFv- β -Lactamase fusion protein from *E. coli* (Rege *et al.* [2005]), as well as for mABs and F_c fusion proteins testing different IEX and HIC resins (Coffman *et al.* [2008]). The MediaScout[®] ResiQuot system (Atoll, Germany) was introduced by Herrmann *et al.* [2006] for generating equal amounts of adsorber volumes (adsorber plaques) using a special vacuum device. An al-

ternative for batch chromatography to all before mentioned strategies offer PreDicator™ filter plates prepacked with a defined adsorber type and volume by GE Healthcare Life Sciences (USA). PhyNexus Inc. (USA) introduced a system of pipetting tips, prepacked with a certain type and volume of adsorber. Binding, washing and elution studies can than be easily performed by aspirating and dispensing the respective liquids. Such adsorber tips were used for fast affinity isolation of Fab-fragments (Wassaf *et al.* [2006]) and as a pre-processing step for oligosaccharide mapping of glycoproteins (Prater *et al.* [2007]). All those described strategies lack in mimicking real dynamic binding processes in chromatography columns. To eliminate these shortcomings special small-scale columns for robotic workstations (RoboColumn® by Atoll, Germany) were developed (Wiendahl *et al.* [2008]). The column inlet of RoboColumns® is of conical shape for the direction of the pipetting needle of the liquid handling system (LHS) and an O-ring assures tight sealing as shown in Fig. 7 left. The flow of liquids is solely based on pumped liquids by the LHS and is stopped as soon as the pipetting needle leaves the column. The outlet of the column is not connected to a tubing and the leaving droplets are collected in 96-well UV plates for analysis in a spectrophotometer (Fig. 7 middle/right). The RoboColumn® system thus describes a quasi-continuous setup which was proven to be highly effective for scale-down models for multi-stage monoclonal antibody processes (Treier *et al.* [2012]) as well as for different model proteins like BSA and lipolase for breakthrough and elution screenings (Wiendahl *et al.* [2008]). Further approaches for scale-down models employ small-scale chromatography columns for liquid chromatography systems, e.g. MediaScout® MiniColumns® (Atoll, Germany) (Bhambure *et al.* [2011]). In general, miniaturized systems result in a reduction in time and materials and allows for parallelization and automation of the workflow for screening a large number of process parameters. A summary of different commercially available tools for HTPD in fermentation and cell culture science, precipitation and refolding processes as well as chromatography was reviewed by Bhambure *et al.* [2011].

1.4.3 Model-based Approaches

Besides all other approaches for protein purification process development discussed above, modeling and simulations can help reducing the number of experiments during the optimization process by *in silico* predictions (Nfor *et al.* [2009]). Especially in the field of chromatography various accurate and predictive models exist. Generally two types of models must be distinguished, being of empirical or mechanistic nature.

1.4.3.1 Empirical Models

Empirical models are based on process output data and calibration relies on empirical model fitting to respective process data. Such models are only valid within the calibrated range and are less robust towards changes in media composition and variations in the design space (Osberghaus *et al.* [2012]). After cross validations and evaluating the model with external test sets, data predictions in terms of extra- and intrapolations from the model can follow. Examples of empirical models are multi-variate data analysis (MVDA) approaches e.g. used for predictions of protein concentrations based on evaluating protein spectra as shown for in-line quantification and real-time pooling decisions for model proteins (Brestrich *et al.* [2015]), antibodies and serum proteins (Brestrich *et al.* [2014]).

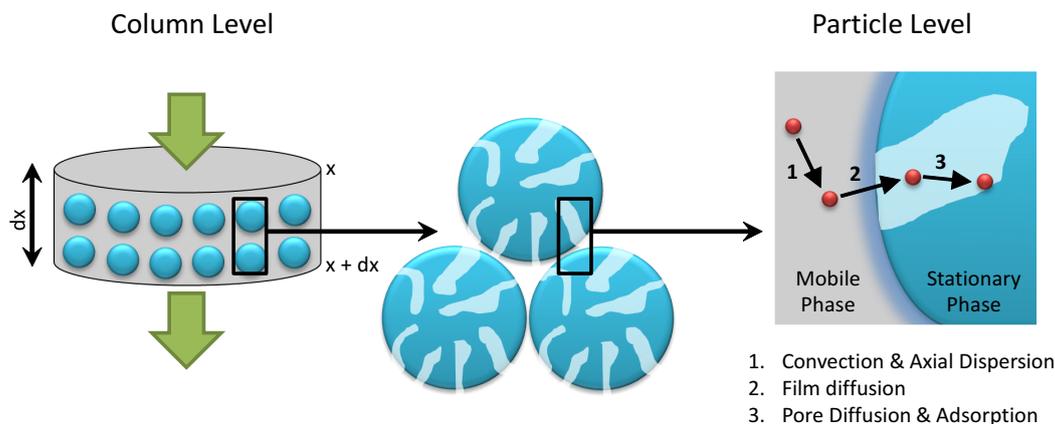


Figure 8: Transport and adsorption phenomena as a basis of mechanistic modeling. Transport on the column level (left) includes convection and dispersion in axial direction. Transport on the particle level (right) includes film and pore diffusion as well as adsorption phenomena according to Michel *et al.* [2005].

1.4.3.2 Mechanistic Models

Mechanistic models, in contrast, exploit knowledge about physico-chemical parameters for accurate predictive models (Osberghaus *et al.* [2012]). Model calibration results in a set of parameters which rely on fundamental understanding of the biochemical process (e.g. protein parameters for size and charge). Examples are molecular dynamics (MD) approaches, where the 3D structure of the protein is investigated and the interactions in a certain environment is assessed as used by Lang *et al.* [2015] in all-atom molecular dynamics simulations on α -lactalbumin, β -lactalbumin, phospholipase A2 and ribonuclease A on a cation-exchanger. Starting from such sophisticated models on the molecular level, certain properties can be correlated to a predicted behavior of the protein as e.g. the overall activity or binding behavior as used for quantitative structure-activity relationship (QSAR) and quantitative structure-property relationship (QSPR) approaches as shown by Ladiwala *et al.* [2005].

Similar to empirical modeling, model validation experiments determine whether the mechanistic model is truly predictive or if a recalibration using different and broader data ranges are needed. In some scenarios the employed model is not suited for the given setup, making need of a reevaluation of the chosen model.

Liquid chromatography can be modeled by an equation of the convection diffusion reaction (CDR) type describing transport phenomena inside the column. Such models include equations of continuity (mass balances etc.) and the reaction term which follows in terms of protein binding processes mostly an isotherm equation. Different models were suggested being suited depending on the application. For chromatography processes of slow mass transfer kinetics the general rate model is commonly used as it includes axial dispersion, as well as resistances in external and internal mass transport (Piatkowski *et al.* [2002]). In this thesis, the transport-dispersive model was used (Michel *et al.* [2005]) describing transport phenomena inside the chromatography column in the interstitial volume of the mobile phase and the mass transfer into the pore volume (Eqs. (1) and (2)). As a simplification compared to the general rate model the film and pore diffusion effects are lumped in an effective mass transfer coefficient $k_{eff,i}$. Eq. (1) models

the change of the concentration $c_i(x, t)$ of component i in the inter-particle phase of a column with length L . Term 1 in Eq. (1) describes the convective transport influenced by the inter-particle velocity of the fluid u . The middle term describes hydrodynamic dispersion in axial direction influenced by the dispersion coefficient D_{ax} . The right term describes the transition of molecules from the inter-particle phase into the particle pores $c_i - c_{p,i}$. The last term is influenced by the bed porosity ε_b , the adsorbent particle radius r_p , and a lumped effective mass transfer coefficient $k_{eff,i}$. Eq. (2) models the mass inside the pores and includes the protein pore concentration in the liquid phase $c_{p,i}$ and the protein concentration bound to the stationary phase q_i . A major influencing factor is the particle porosity ε_p . The transport and adsorption phenomena on the column (left) and pore level (right) are illustrated in Fig. 8.

$$\frac{\partial c_i}{\partial t} = -u \frac{\partial c_i}{\partial x} + D_{ax} \frac{\partial^2 c_i}{\partial x^2} - \frac{1 - \varepsilon_b}{\varepsilon_b} k_{eff,i} \frac{3}{r_p} (c_i - c_{p,i}) \quad (1)$$

$$\frac{\partial c_{p,i}}{\partial t} = k_{eff,i} \frac{3}{r_p \varepsilon_p} (c_i - c_{p,i}) - \frac{1 - \varepsilon_p}{\varepsilon_p} \frac{\partial q_i}{\partial t} \quad (2)$$

A review on different isotherms for defining the binding mechanism is given by Foo and Hameed [2010] including common adsorption models such as Langmuir, Freundlich and BET.

In this thesis the semi-mechanistic steric mass-action (SMA) isotherm (Brooks and Cramer [1992]) was applied for modeling ion-exchange chromatography binding behavior including the influence of varying salt concentrations (Fig. 9). Effects of counter-ions on the retention behavior of proteins are taken into account, as well as the proteins' characteristic charge ν_i . The model also includes different adsorbent properties, such as the total ionic capacity Λ . The steric shielding factor σ_i describes effects caused by proteins covering binding sites without electrostatic interactions. The kinetic form of the SMA isotherm is shown in Eq. (3) for k proteins, with q_i and $c_{p,i}$ being the concentration of the protein i adsorbed and in solution inside the pore, respectively. $c_{p,salt}$ describes the effective pore salt concentration. The constants of adsorption and desorption are described by $k_{ads,i}$ and $k_{des,i}$ and q_{salt} (Eq. (4)) equals the number of salt ions still attached to the adsorbent surface.

$$\frac{\partial q_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) q_j \right)^{\nu_i} c_{p,i} - k_{des,i} c_{p,salt}^{\nu_i} q_i \quad (3)$$

$$q_{salt} = \Lambda - \sum_{j=1}^k \nu_j q_j \quad (4)$$

With the governing equations being defined the model generation step follows in which the calibration data is fitted to the respective model. Here, the inverse method was used by solving a non-linear optimization problem and minimizing the error between measured data points and the employed chromatography model (Soetaert and Petzoldt [2010]). The inverse method was successfully employed for small molecules (Forssén *et al.* [2006]), model proteins (Osberghaus *et al.* [2012]), antibodies (Degerman *et al.* [2006], Karlsson *et al.* [2004]), as well as industrial process applications (Close *et al.* [2014], Mollerup *et al.*

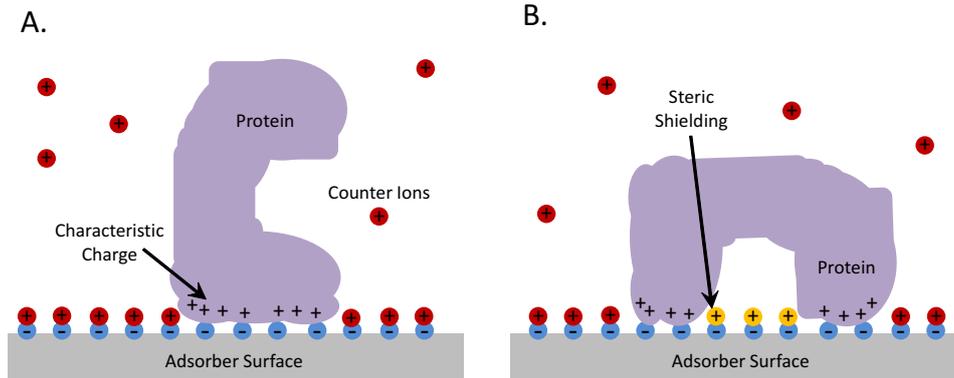


Figure 9: Schematic illustration of the mechanism of the steric mass-action model according to Brooks and Cramer [1992]. Besides adsorption and desorption processes the model includes factors like the steric shielding of binding sites (yellow cations) and the characteristic charge of the protein.

[2007]). Huuk *et al.* [2014] used the inverse method for modeling a multi-step optimization by comparing a consecutive anion- and cation-exchange step in different orders for model proteins.

1.5 Integrated USP - DSP Approach

As shown in diverse case studies, modern tools such as high-throughput process development and *in silico* methods can be used to improve process understanding, development timelines, and ultimately the product quality. Nevertheless, such technology is mostly used for separate optimization of upstream and downstream processes. In other words, cultivation conditions yielding in maximal product titer are selected for downstream processing independently of investigating the impurity feed composition. This might result in an overall suboptimal process as impurity levels are highly influenced by the cultivation conditions (Gawlitzeck *et al.* [1995]).

The level of dissolved oxygen strongly influences HCP levels in terms of stress-response proteins and enzymes for alternative metabolic pathways (e.g. for carbon metabolism and electron transport) (Marino *et al.* [2000], Starck *et al.* [2004], Ye *et al.* [2000]). Temperature changes were found to induce the formation of heat or cold shock proteins, temperature-adapted metabolic enzymes, and proteins involved in transport mechanisms (Budde *et al.* [2006], Smoot *et al.* [2001], Wang *et al.* [2007], Weiner III *et al.* [2003], Woo *et al.* [2009]). Moreover, the chemical environment and media, in which the cultivation takes place (e.g. exposure to different pH levels) result in an up- or downward regulation of certain gene areas (Lee *et al.* [2008], Stancik *et al.* [2002]). In general, alternative pathways and transport mechanisms are used to either metabolize or eliminate the respective chemicals from the cells (Guina *et al.* [2003], Kirkpatrick *et al.* [2001], Pedersen *et al.* [1978]). Additionally, the state of the cell in the cell cycle as well as the induction process change the proteome of cells drastically (Grünenfelder *et al.* [2001]). Especially induction leads to an overburden in the host cell mechanism (Dürschmid *et al.* [2008]). As a consequence of such differences in impurity profiles, conditions of reduced initial product titer might result in an overall superior biopharmaceutical process. It is therefore not only important to develop tools for fast USP and DSP development but there is also the

need for feed stock characterization tools using modern technologies such as HTPD and *in silico* methods.

Another prerequisite for HTPD data of high quality for modeling is the development of a sophisticated analytics which might in terms of e.g. analytical high-performance liquid chromatography (HPLC) and product-specific immunoassays (Maiser *et al.* [2012], Van Lierop *et al.* [2002]) not be compatible with the HTS workflow and suffers from its invasive nature.

2 Research Proposal

This research work is part of the project 'ERA Net Euro Trans Bio - 6: Development of an integrated strategy for a high-throughput process development platform on a micro-scale format - FORECAST'. The overall project structure ($B^n =$ process development for biologics) covers three sub sections as shown in Fig. 10. The S^3 package includes strain construction and clone screening, as well as lab-scale validation experiments. The main focus are besides product titers also downstream relevant aspects such as colloidal and conformational stability of proteins, feed compositions for downstream applications and including genetic elements for analytical purposes. The P^3 package mostly focuses on the integrated upstream-downstream process development platform. It includes feed stock characterization technologies, purification strategies, product performance tests, as well as back coupling to the upstream process development. The final section C^3 covers chromatographic procedures including *in silico* methods, crystallization and precipitation screenings as well as working with highly concentrated protein solutions and protein engineering technologies e.g. by fusion tags.

The development of a biopharmaceutical process is very complex and the final outcome depends on the results of a large number of subprocesses. Each step from choosing the expression host/ system, cloning, clone screening as well as up- and downstream process development poses major challenges. Although each process step can decide on the success of a bioprocess industry is mostly still based on heuristics (expert/ experience-based) and sequential optimization of the separate process steps. In most research facilities the upstream conditions resulting in highest product titers are chosen for downstream process development meaning that the purification site must work with the presented lysates and cannot influence concentration levels of critical impurities. Thus, the major challenge in industry is the fact that the overall success of a process depends on the decisions made during each individual stage which are driven by contradictory goals and motivations. Besides problems caused by the strict process development hierarchy also the 'Time to Market' demands restrict exploring the design space in a sufficient way, resulting in sub-optimal process conditions.

To eliminate these shortcomings and to propose feasible concepts for industrial process development the main focus of this thesis was to develop an integrated platform combining up- and downstream processing features (P^3) and to loosen the barrier set up between up- and downstream processing. By using the latest technologies in high-throughput process development and employing sophisticated models even concerted process optimization strategies can be realized. Synchronously there is need for fast, straightforward, exact and at best non-invasive analytical technologies to prevent an analytical bottleneck created by high-throughput technologies. Such analytics need to be of high standard to produce data of modeling quality to meet the demands of the proposed strategy.

The upstream part focuses on process development based on high-throughput technology. The BioLector[®] from m2p-Labs (Germany) enables 48-well cultivations in a sub-milliliter scale and allows for exploring the design space in terms of cultivation conditions and parameters with low demands on sample volume and process control. The large number of cell lysates then need to be scored in terms of product titers, as well as on impurity profiles. Therefore high-throughput compatible analytical technologies need to be developed including the Cherry-Tag[™] technology by Delphi Genetics (Belgium) as a visual

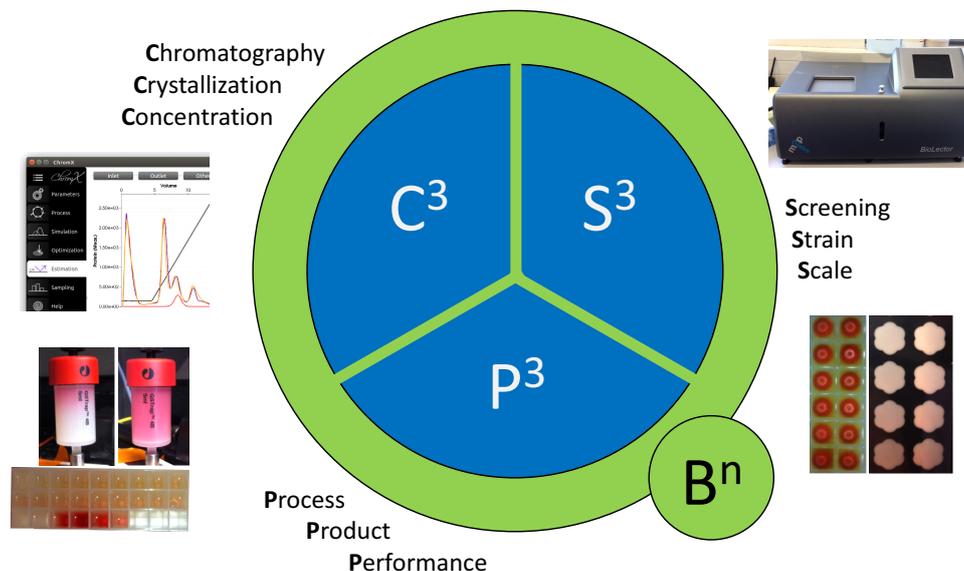


Figure 10: Overview over the B^n ERA Net Euro Trans Bio - 6 project structure. The project is subdivided into the upstream section S^3 , the downstream part C^3 and the platform development P^3 .

protein marker. Scale up experiments shall validate the results obtained in micro-scale. The focus of the downstream part is to develop straightforward tools for designing efficient purification steps in different chromatography modes. Such technologies include systematic strategies (deterministic approach), as well as HTS technologies (experimental approach) and modeling for parameter estimation and *in silico* process optimization. The integrated approach must then be capable of bringing together these technologies for concerted up- and downstream process development. In that context especially tools for lysate characterization and scoring as well as HTS compatible analytical tools need to be developed.

The overall objective of the presented thesis is to create and to link different powerful bioprocess development tools in up- and downstream processing using the latest technology in high-throughput process development and *in silico* tools. The ultimate goal is to meet both, the 'Quality by Design (QbD)' and 'Time to Market' demands and also to increase the overall process efficiency.

3 Publications & Manuscripts

1. **Integrated Development of Up- and Downstream Processes Supported by the Cherry-TagTM for Real-time Tracking of Stability and Solubility of Proteins**

P. Baumann, N. Bluthardt, S. Renner, H. Burghardt, A. Osberghaus, J. Hubbuch
J. Biotechnol. (2015), doi:10.1016/j.jbiotec.2015.02.024

This article presents a protein tag as a quantitative real-time and in-line product tracking tool which allows for a distinction between insoluble/ non-native proteins and soluble/ native species shown in a Glutathione-S-Transferase (GST) case study. In upstream processing conditions of inclusion body formation were identified and in downstream applications it was shown that the colored GST fusion protein was catalytically active and could be bound effectively to a GST affinity column.

2. **Deconvolution of High-throughput Multi-component Isotherms Using Multi-variate Data Analysis of Protein Spectra**

P. Baumann[‡], T. Huuk[‡], T. Hahn, A. Osberghaus, J. Hubbuch ([‡]: Contributed equally)

Eng. Life Sci., accepted manuscript

This case study explores the usability of UV spectra and multi-variate statistics to prevent an analytical bottleneck in high-throughput batch chromatography, incorporating multiple protein species. The presented approach enables integration of the analytical setup in the batch chromatographic workflow, using a standard UV/VIS spectrophotometer. The quality of the analytical data was sufficient to fit steric mass-action isotherms at various pH values and ionic strengths and get a mechanistic insight into the competitive binding behavior.

3. **Establishment of a High-throughput Platform for the Production of Virus-like Particles in *Escherichia coli***

C. Ladd Effio[‡], P. Baumann[‡], P. Vormittag, C. Weigel, A. Middelberg and J. Hubbuch ([‡]: Contributed equally)

Biotechnol. Bioeng., submitted manuscript

This study describes the development of an HTS platform for the production and characterization of virus-like particles (VLPs) produced in *Escherichia coli*. The upstream optimization was based on high-throughput cultivations in 48-well format followed by HTS cell disruption. Capillary gel electrophoresis served as a HTS-compatible analytical tool for scoring the large number of lysates. After a successful scale-up culture and product purification, the assembled VLP product was characterized by transmission electron microscopy (TEM).

4. Systematic Purification of Salt-intolerant Proteins in Ion-exchange Chromatography: The Example of Human α -Galactosidase A

P. Baumann, A. Osberghaus, J. Hubbuch

Eng. Life Sci. (2015), doi:10.1002/elsc.201400210

This paper describes a systematic strategy for purifying salt-intolerant proteins using ion-exchange chromatography. The strategy considers product stability as well as operational conditions and purification steps can be designed by well-chosen prior screening experiments. The efficiency was shown in a case study on capturing human α -Galactosidase A. The applied strategy allowed for a quick establishment of a dedicated IEX capture step at low salt concentrations under stable conditions by well-chosen prior screening experiments.

5. Influence of Binding pH and Protein Solubility on the Dynamic Binding Capacity in Hydrophobic Interaction Chromatography

P. Baumann[‡], K. Baumgartner[‡], J. Hubbuch ([‡]: Contributed equally)

J. Chrom. A (2015), doi:10.1016/j.chroma.2015.04.001

This article discusses the influence of the binding pH on the dynamic binding capacity in hydrophobic interaction chromatography for proteins of acidic, neutral and alkaline isoelectric point. Binding close to the solubility limit increased the binding capacity, whereas an inverse correlation was found for binding kinetics and protein binding. The latter might indicate protein rearrangement on the adsorber surface during the binding process. Despite increased protein binding the protein's integrity after elution was ensured.

6. UV Absorption-based Inverse Modeling of Protein Chromatography

T. Hahn, P. Baumann, T. Huuk, V. Heuveline, J. Hubbuch

Eng. Life Sci. (2015), doi:10.1002/elsc.201400247

This study explores the feasibility of modeling ion-exchange chromatography without knowledge of feed composition in terms of molar or mass concentration. This is especially valuable in early-stage process development when no information is available on the impurities. It was shown that all model parameters can be determined uniquely from single-component elution curves. Here, the concentration of Cherry-tagged glutathione-S-transferase within a crude feedstock could be determined from chromatograms at a particular wavelength.

7. High-throughput Micro-scale Cultivations and Chromatography Modeling: Powerful Tools for Integrated Process Development

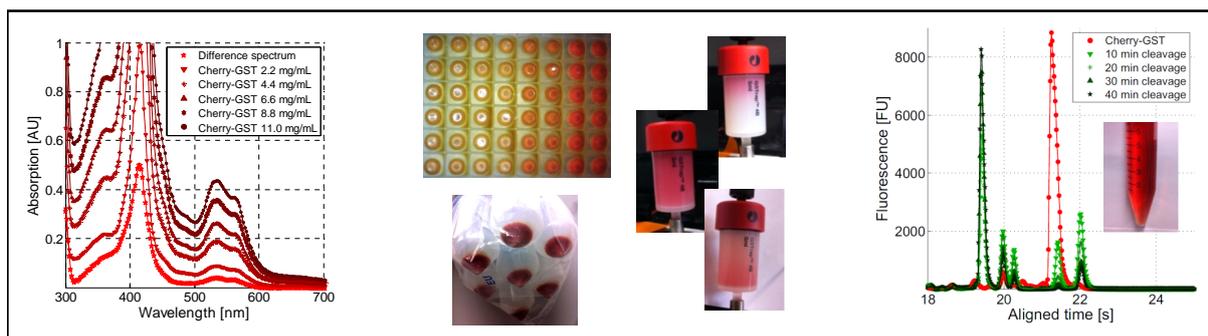
P. Baumann, T. Hahn, J. Hubbuch

Biotechnol. Bioeng. (2015), doi:10.1002/bit.25630

This paper presents a combined up- and downstream optimization approach based on high-throughput micro-scale cultivation experiments and chromatography modeling. The methodology is presented in a case study for the enzyme Glutathione-S-Transferase from *Escherichia coli* SE1. It was shown that the overall optimized system must not necessarily be the one with highest product titers, but the one resulting in an overall superior process performance in up- and downstream.

Integrated Development of Up- and Downstream Processes Supported by the Cherry-TagTM for Real-time Tracking of Stability and Solubility of Proteins

P. Baumann¹, N. Bluthardt¹, S. Renner¹, H. Burghardt¹, A. Osberghaus¹
and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany

* : Corresponding author. *E-mail-address*:juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

Product analytics is the bottleneck of most processes in bioprocess engineering, as it is rather time-consuming. Real-time and in-line product tracing without sample pre-treatment is only possible for few products. The Cherry-TagTM (Delphi Genetics, Belgium) which can be fused to any target protein allows for straightforward product analytics by VIS absorption measurements. When the fused protein becomes unstable or insoluble, the chromophore function of the group is lost, which makes this technology an ideal screening tool for solubility and stability in up- and downstream process development. The Cherry-TagTM technology will be presented for the tagged enzyme Glutathione-S-Transferase (GST) from *E. coli* in a combined up- and downstream process development study. High-throughput cultivations were carried out in a 48-well format in a BioLector[®] system (m2p-Labs, Germany). The best cultivation setup of highest product titer was scaled up to a 2.5 L shake flask culture, followed by a selective affinity chromatography product capturing step. In upstream applications the tag was capable of identifying conditions where insoluble and non-native inclusion bodies were formed. In downstream applications the red-colored product was found to be bound effectively to a GST affinity column. Thus, it was identified to be a native and active protein, as the binding mechanism relies on catalytic activity of the enzyme. The Cherry-TagTM was found to be a reliable and quantitative tool for real-time tracking of stable and soluble proteins in up- and downstream processing applications. Denaturation and aggregation of the product can be detected in-line at any stage of the process. Critical stages can be identified and subsequently changed or replaced.

Keywords: Cherry-TagTM, BioLector[®], Micro-cultivation, Product Analytics, Native Protein

1 Introduction

Product analytics is a major bottleneck during process development in bioprocess engineering. Most methods are rather time- and material-consuming and cannot be implemented in-line and in real-time. Many analytic procedures include additional sample preparation steps (i.e. chemical treatment and denaturation), leading to analytical results which can differ from the untreated sample. Common off-line methods for protein analytics are gel electrophoresis (GX) and capillary gel electrophoresis (GX II) [1, 2]. By means of both approaches, protein samples are separated according to their molecular weight in a denatured and linear state [3, 4]. GX is rather qualitative and is mostly used for fast and qualitative analytics of cell cultivation harvests and cell lysates [5, 6]. In contrast to this, GX II allows for the sequential and quantitative analysis of approximately 400 samples within 2 h of processing time. Sample preparation is rather complex and time-consuming. However, this technology produces highly accurate and reproducible analytical results. It was shown that GX and GX II identify almost identical product and contaminant patterns [1, 7].

The most complex and accurate way of protein analysis is mass spectrometry. In analogy to GX and GX II, protein analysis is performed off-line and in a denatured state. The analyte is ionized and subsequently accelerated in a static or dynamic magnetic or

electrical field [8]. Separation is performed according to the mass to charge ratio which is characteristic of a certain type of molecule. Products of higher molecular weights (e.g. proteins) require mild ionization techniques to prevent fragmentation, such as electrospray ionization (ESI-MS) [9], matrix-assisted laser desorption ionization (MALDI-MS) [10, 11], and desorption/ ionization on silicone (DIOS) [12]. MS protein analytics can also be coupled with enzymatic digestion prior to ionization, leading to smaller molecule fragments. The detected peptide pattern can be characterized and identified by an automated peptide mapping software [13]. MS systems are rather expensive and require vast expertise for data post-processing. They are mostly applied for protein identification and proteome analysis studies [14, 15, 16].

A possible off-line alternative to GX, GX II, and MS for analyzing protein samples in a native state is analytical chromatography [17, 18, 19]. Major disadvantages are sample pre-treatment (e.g. buffer exchanges) as well as buffer- and time-consuming column wash, equilibration, binding, and elution steps for each applied sample. Improvements have been achieved by introducing a parallel interlaced chromatography method [20]. Still, the time per sample could not be reduced to below 2 min.

Off-line analysis including selective protein determination combined with analysis in a native state, can be realized by customized assays. Those methods are widely applied for enzymes, hormones, and growth factors: For the granulocyte colony-stimulating factor (G-CSF), the NFS-60 proliferation assay is used [21, 22]. For enzymes, a selective metabolization of a substrate or the accumulation of a product is evaluated [23, 24]. Other products with specific binding sequences can be analyzed by e.g. antibody-based analytical methods like sandwich ELISA or Western Blotting [25, 15, 26]. However, such technologies are not available for all proteins and assay development and optimization are rather complex.

Selective and real-time product tracking without sample pre-treatment so far has been mostly limited to proteins with special optical properties, which are exceptions rather than the rule. Examples are fluorescent proteins like GFP, YFP, mBanana, tdTomato, mCherry [27], and proteins with chromophore groups like cytochrome c [28]. Most other proteins, however, are hard to distinguish spectrometrically. Few setups were reported for on-line analytics by multi-variate data analysis of protein spectra [29, 30]. However, quantitative protein analysis requires calibration experiments and a secondary analysis for the determination of single protein titers.

An elegant way to make the advantages of fluorescent proteins and chromophores accessible for all proteins is the use of protein tags with special optical properties. Examples are fluorescent GFP- and Halo-Tags[®] [31, 32, 33]. Quantitative analysis for fluorophores, however, is limited to a constant chemical environment. Wild-type GFP is quenched under acidic conditions and several GFP mutants like enhanced GFP (EGFP) show a reduction in fluorescence of approximately 50% starting from pH 5.5 [34]. Also fluorescence is influenced by different ions (e.g. chloride) and is subject to photo bleaching [35, 36]. Moreover, GFP- and Halo-Tag[®] both do not allow for a distinction between the native and soluble or denatured state of the fused protein. This said, a GFP-tag alteration exists, the split GFP-tag system, which can distinguish between soluble and insoluble proteins for upstream applications [33]. A fragment of GFP is fused to the product and is not presented, if the protein is in a misfolded state. After product expression, the second part of GFP is added to the sample and only natively folded proteins show fluorescence.

However, this system is not suited for real-time product tracking of native protein, as the GFP assembly needs to be performed post-processively and is still based on fluorescence measurements.

The presented Cherry-TagTM technology must not be confused with the fluorescent protein mCherry that is e.g. used for high-resolution two-color fluorescence microscopy and exists as a split mCherry-tag system as discussed for GFP [37, 38]. The newly developed Cherry-TagTM (Delphi Genetics, Belgium) can be determined by VIS absorption measurements and is not based on fluorescence [39, 40]. Protein instability and denaturation result in a loss of the chromophore group. Thus, the technology can be used for quantitative online and real-time analysis of proteins without sample pre-treatment and allows conclusions to be drawn with respect to the protein's physiological state in all stages of bioprocess engineering. Though the potential of the Cherry-TagTM in terms of increasing thermo-stability and solubility of proteins was shown for proteins like enterokinase, staphylokinase, and the granulocyte colony-stimulating factor [41, 42], the tag has never been used in quantitative studies and its optical properties have never been applied in up- and downstream process development studies.

Here, the potential of the Cherry-TagTM technology is tested and evaluated for the enzyme glutathione-S-transferase (GST) in up- and downstream applications. The upstream optimization is based on 48-well high-throughput cultivations on micro-scale level in a BioLector[®] system (m2p-Labs, Germany) [43, 44, 45]. Different cultivation conditions, such as the induction setup, shaking speeds, and temperatures, are investigated. The best cultivation setup is picked by the simple VIS absorption analytics of the cleared cell lysates. The cleared lysate of a scaled-up cultivation is subsequently purified by GST affinity chromatography. In all applications the above-mentioned GX II will be applied for validation purposes.

2 Materials & Methods

2.1 Materials

2.1.1 Disposables & Reaction Vessels

The Glutathione-S-Transferase (GST) nucleic acid sequence was cloned into the CherryTM Express T7, as well as into the StabyCodon[®]T7 (T7 promoter without adding Cherry-TagTM) protein expression kit (Delphi Genetics, Belgium). Cells were stored as cryocultures in a CryobankTM strain maintenance kit for microorganisms (Mast Diagnostica, Germany). Micro-scale cultivations were carried out in 48-well FlowerPlates[®] covered with adhesive sterile sealing foil (m2p-Labs, Germany). Lab-scale cultivations were performed in 1 L baffled flasks (Schott, Germany) and 2.5 L baffled TUNAIRTM shake flasks (Sigma Aldrich, Germany). Centrifugations for cell harvest and after cell disruption were carried out in 15, 50, and 225 mL centrifugal tubes (Falcon, Germany). Absorption measurements were carried out in 96-well flat-bottom UV-Star[®] half-area micro-plates (Greiner Bio-One, Germany) and 70 μ L UV-Cuvettes micro (Brand GMBH + CO KG, Germany). For gel electrophoresis (GX), Novex[®]NuPAGE[®]4 - 12% Bis-Tris (1.0 mm, 12 well) protein gels were purchased from Life Technologies, Sweden. Capillary gel electrophoresis (GX II) was performed in an HT Protein Express & Pico

LabChip® (Perkin Elmer, USA). Sample preparation for GX II was conducted in skirted 96-well twin.tec® PCR plates (Eppendorf, Germany). For affinity chromatography runs, 5 mL GStrap™ 4B columns and 50 mL super loops for sample application were purchased from GE Healthcare, Sweden. Fractionations were carried out in 2 mL 96-well square deep-well plates (VWR, Germany). Filtrations were performed using polyether-sulfone 0.2 µm syringe sterile filters (VWR, Germany).

2.1.2 Chemicals & Buffers

Cultivations were carried out in TB (Terrific Broth) medium, consisting of 12 g/L wheat peptone for microbiology (Fluka, Germany), 24 g/L bacteriological yeast extract (Amresco, USA), 5 g/L glycerol bidistilled 99.5%, 17 mM potassium dihydrogen phosphate, and 7 mM di-potassium hydrogen phosphate (VWR, Germany). For induction of the T7 promoter isopropyl-β-D-thiogalactopyranosid (IPTG) from a 0.25 M and for repression, glucose from a 20% (w/w) stock solution (VWR, Germany) was applied. Chemicals for scale-up by determination of the oxygen transfer rate were applied as follows: 12 mM di-sodium hydrogen phosphate (VWR, Germany), 500 mM sodium sulfite BioUltra anhydrous as substrate, $2 \cdot 10^{-7}$ M cobalt(II) sulfate heptahydrate ReagentPlus as reaction catalyst, and Bromothymol Blue as indicator (Sigma Aldrich, Germany). The initial pH value was adjusted to pH 8 by titration with sulfuric acid (Sigma Aldrich, Germany).

The lysis buffer for cell disruption was composed of 50 mM Tris-HCl (VWR, Germany) and 1X SigmaFAST™ Protease Inhibitor (Sigma Aldrich, Germany) adjusted to pH 8 with hydrochloric acid (Merck, Germany). For reduction of the nucleic acid content, 1 µL/10 mL lysate of benzonase purity level II ($\geq 90\%$) for biotechnology (Merck, Germany) and 10 mM magnesium chloride anhydrous (Sigma Life Sciences, Germany) from a 1 M stock solution were applied. Cell disruption under denaturing conditions was realized using trichloroacetic acid BioChemica (AppliChem, Germany) and acetone for liquid chromatography (Merck, Germany).

For gel electrophoresis (GX), 20X Bolt®MES SDS running buffer, NuPAGE®4X LDS sample buffer, and the Mark 12™ unstained standard (Life Technologies, Sweden) were used. Dithiothreitol (DTT) was purchased from Sigma-Aldrich, Germany, and prepared as a 1 M stock solution. Staining of the gel protein bands was realized with 'blue silver' staining solution, consisting of 10% (v/v) phosphoric acid of 85% purity (Roth, Germany), 100 g/L ammonium sulfate BioChemica (AppliChem, Germany), 1.2 g/L Coomassie Brilliant Blue G-250 (Merck, Germany), and 20% (v/v) Methanol (Sigma Aldrich, Germany). For the capillary gel electrophoresis (GX II) experiments, an HT protein express reagent kit (Perkin Elmer, USA) was purchased. Lysozyme served as an internal standard of known concentration (Hampton research, USA).

For the enzymatic tag removal reaction, a recombinant bovine enterokinase (43 kDa catalytic fragment from *P. pastoris*) from BioIngenium (Spain) was used. The reaction buffer consisted of 50 mM Tris-HCl (VWR, Germany), 1 mM calcium chloride, and 0.1% of Tween-20 (Sigma-Aldrich, Germany).

The binding buffer for the 5 mL GStrap™ 4B affinity column consisted of 50 mM Tris-HCl pH 8. Elution was performed in 50 mM Tris-HCl pH 8 with 10 mM of reduced 98% purity L-glutathione (Sigma-Aldrich, Germany). For column regeneration, a 50 mM Tris-HCl pH 8 CIP buffer was applied, including 3 M guanidine hydrochloride for molec-

ular biology (Sigma-Aldrich, Germany). The pH adjustment of all buffers was realized by titration using sodium hydroxide (Merck, Germany).

2.1.3 Instrumentation & Software

Micro-scale cultivations were carried out in a BioLector[®]MB micro-fermentation system using the BioLection[®]HMI & analysis software (m2p-Labs, Germany). Lab-scale fermentations were realized in a MaxQ[™] 6000 incubator (Thermo scientific, USA). For time documentation of the indicator's change of color during scale-up, a Spotlight Webcam Pro (Trust, Netherlands) coupled with the ContaCam Software (Contaware, USA) was used. Solid-liquid separation for cell harvest and removal of cell debris was performed with a 5810 R centrifuge (Eppendorf, Germany). For cell disruption, a Digital Sonifier[®] 450 (Branson Ultrasonic Corporation, USA) was used. Absorption measurements were conducted in an Infinite M200 Reader controlled with I-control 1.9 (Tecan, Germany). The enzymatic tag removal reaction and the sample denaturation for the GX system was realized in a MUR 13 thermo-shaker with additional lid heating (HLC BioTech, Germany). The GX was performed in an Xcell SureLock[™] Novex[®]MiniCell gel chamber operated with a PowerEase[®] 500 power supply (Life Technologies, Sweden). The gels were scanned in a Bio-5000 Gel Scanner (Serva Electrophoresis GmbH, Germany). Capillary gel electrophoresis (GX II) was carried out in a Caliper LabChip[®]GX II (Perkin Elmer, USA). For data processing and analysis, the LabChip[®]GX 3.1 software (PerkinElmer) was used. The GST affinity chromatography purification runs were carried out using an ÄKTA[™] purifier system (GE Healthcare Life Sciences, Sweden). The system was equipped with a pump P-900, mixer M-925, UV detector UV-900, motor valve INV-907, pH and conductivity monitoring pH/C-900, and a fraction collector Frac-950 unit. The FPLC system was controlled using Unicorn 5.2 (GE Healthcare Life Sciences, Sweden). For pH determination during pH adjustment of all buffers, an HI-3220 pH meter (Hanna Instruments, USA) was used. Data processing and creation of figures were performed in Matlab[®] R2011a (MathWorks, USA).

2.2 Experimental Setup

2.2.1 Establishment of the Cherry-Tag[™] Analytics

Establishment of the Cherry-Tag[™] analytics was performed using lysates from 250 mL cultivations. Two 1 L baffled flasks were filled with 250 mL sterile TB medium each. 12.5 mL of TB medium were removed from one flask (negative control) and replaced by 20% (w/v) sterile filtered glucose solution, resulting in a final glucose concentration of 1% (w/v). Addition of glucose results in an inhibition of the lac-operon (cAMP positive control). Both shake flasks (positive and negative control) were inoculated with an *E. coli* SE1 Cherry-GST cryo-culture from the Cryobank[™] strain maintenance kit and cultivated for 20 h in the Max Q 6000 incubator at 170 rpm and 37 °C. For cell harvest, 180 mL of cell broth were transferred to 225 mL centrifugal tubes and centrifuged at 4000 rpm and 5 °C for 30 min in the 5810 R centrifuge. The supernatant was discarded and the cell pellets were frozen for 24 h at -20 °C as pre-treatment by physical cell damage. Each thawed cell pellet was resuspended in 20 mL lysis buffer and sonified in the

Branson Digital Sonifier[®] 450 using a 1/2” extension cylindrical sonication probe. Sonication was performed at 70% maximal power output and a total treatment time of 150 s (10 · 15 s pulse on and 30 s pulse off for liquid cooling on ice). After the last sonication cycle, 0.1 μ L/mL of benzonase and 10 mM MgCl₂ were added for reduction of the nucleic acid content. The lysates were transferred to 50 mL centrifugal tubes and centrifuged twice at 12000 rpm and 10 °C for 60 min.

Lysate spectra were monitored to verify the Cherry-Tag[™]-specific wavelength. Spectra of the positive and negative control lysates were determined from 300 to 700 nm in 1 nm steps in an Infinite M200 Reader. The spectra were scaled to an equal lysate protein concentration by using a 300 nm scaling factor. The Cherry-Tag[™]-specific wavelength was then determined from the difference spectrum of the Cherry-GST and negative glucose control sample. Purified Cherry-GST samples in a concentration of 2.2, 4.4, 6.6, 8.8, and 11 mg/mL, respectively, served as a reference and to investigate samples of higher protein concentrations. The purified Cherry-GST samples were produced as described in the ‘Purification of Cherry-GST’ section.

The Cherry-GST extinction coefficient was determined to calculate product concentrations directly from VIS absorption measurements. The Cherry-GST lysate was diluted to a final absorption at the Cherry-Tag[™]-specific wavelength of approximately 1.2 AU. 100, 75, 50, 25, and 0% dilution steps of the lysate using 50 mM Tris-HCl pH 8 were measured as triplicates in 96-well flat-bottom UV-Star[®] micro-plates in the spectrophotometer. For determination of the corresponding concentrations, the samples were also measured in the LabChip[®] GX II device as quadruplicates using the HT Protein Express LabChip[®] kit. The sample and chip preparation procedures were performed as described in the manufacturer’s protocol for the HT Protein Express Assay [46]. The samples were analyzed using the HT Protein Express 200 assay and the LabChip[®] GX 3.1 software. The resulting electropherograms were aligned to the size of Cherry-GST (approx. 39 kDa). Quantification of the proteins in each sample was realized by using peak-baseline integration and scaling to an external lysozyme protein standard of 1 mg/mL.

2.2.2 HTE Upstream Screening

Cultures for the upstream screening were prepared in 250 mL standard TB medium as described in the product analytics section. After 17 h (end of the exponential growth), the cells were diluted in sterile TB medium to a starting OD_{600 nm} of 0.1 AU. 1 mL of this inoculum culture was then transferred to each well of the FlowerPlate[®] which was sealed with an adhesive gas-permeable sterile membrane for micro-scale cultivations. Thus, an equally treated pre-culture was ensured for all upstream experiments.

The cultivation setups of the HTE upstream screening were performed as quadruplicates, leading to a total of 12 different cultivation conditions per 48-well FlowerPlate[®]. Four different induction times (OD_{600 nm} of 1, 2, 4, and 8) and three different inducer concentrations (IPTG concentrations of 0.1, 0.5, and 1 mM) were investigated. The induction times were calculated based on scattered light values in the BioLector[®] system using previously determined OD_{600 nm}/ scattered light correlation functions. For 600 rpm, a linear correlation slope of 17.26 SU/AU was used. For 800 and 1200 rpm, slightly higher linear correlation slopes of 18.69 and 19.8 SU/AU were applied. When the scattered light signals exceeded the threshold value of the pre-defined induction time, the fermentation

process was interrupted for the addition of IPTG. The experiments were performed at three different temperatures (27, 32, and 37 °C) and three different shaking speeds in the BioLector[®] system (600, 800, and 1200 rpm). Including the quadruplicates of three non-induced 600 rpm cultivations, a total of $48 \cdot 9 + 12 = 444$ fermentations (111 cultivation conditions) were investigated as a full factorial experimental design. The fermentation was stopped after 20 h. 1 mL fermentation broth of each well was transferred to a 2 mL 96-well square deep-well plate and centrifuged at 4000 rpm and 10 °C for 30 min. The supernatants were discarded and the cell pellets were frozen at -20 °C for physical pre-treatment until cell disruption.

For Cherry-GST quantification from the pellets of the HTE upstream screening, a standardized cell disruption was developed. The frozen cell pellets from 1 mL cultivations were resuspended in 1.25 mL lysis buffer. 4 pellets were pooled in a 15 mL centrifugal tube, leading to 5 mL suspension. The sonification process was performed in the Branson Digital Sonifier[®] 450 using a 1/8" tapered micro-tip sonication probe. For a standardization of the cell disruption process, the immersion depth of the probe was kept constant at 0.5 cm above the centrifugal tube bottom for all cell disruptions. The sonication was performed at 30% maximal power output and a total treatment time of 90 s (6 · 15 s pulse on and 30 s pulse off for cooling on ice). Following the cell disruption, 0.1 μL/mL of benzonase and 10 mM MgCl₂ were added for nucleic acid reduction. The lysates were centrifuged at 12000 rpm and 10 °C for 60 min. Disruption of 12 identical cell samples resulted in a relative standard deviation below 5%. As a result, this method was found suitable for quantitative comparison of the different conditions of the HTE upstream screening.

2.2.3 Detection of Inclusion Bodies

As a proof, that proteins with Cherry-Tag[™] lose their color, as soon as they become unstable or are in a denatured state a methodology for detection of inclusion bodies was used. Cultivations with inclusion body formation are to show no red color, but to reveal product bands in an SDS PAGE after total protein extraction under denaturing conditions.

Protein extraction under denaturing conditions was used to extract all soluble and insoluble proteins from cell pellets. A cell pellet from a 1 mL cultivation broth was resuspended in 1 mL of ultrapure water and 50 μL of 100% (w/v) cold trichloroacetic acid (TCA) were added. After mixing for 20 s, the solution was stored on ice for 10 min and centrifuged (12000 rpm, 4 °C for 10 min). The supernatant was discarded and the pellet was washed twice with 500 μL of cold, ultrapure acetone to remove residual TCA. After each washing step, another centrifugation cycle followed. The final pellet was left for two hours under a hood at room temperature for removal of residual acetone. The OD_{600nm} value of the original 1 mL cell suspension sample served as a scaling factor for redissolution. The dried pellet from the denatured total protein extraction process was dissolved in OD_{600nm} · 200 μL of 1X NuPAGE[®]LDS sample buffer and 12.5 mM DTT for gel electrophoresis. Redissolution was performed for at least 1 h in an overhead shaker. The dissolved samples were denatured at 100 °C and 300 rpm for 15 min on a plate shaker with lid heating. The further GX procedure was performed as described in the Novex[®]NuPAGE[®]4 - 12% Bis-Tris Protein Gel manual [47]. The gels were stained with

'blue silver' staining solution and analyzed in a Bio-5000 gel scanner.

2.2.4 Scale-up Culture

The scale-up from the 1 mL BioLector[®] culture of highest product titers to the 800 mL shake flask was based on a constant oxygen transfer rate. OTR and kL_a values were determined by the sulfite oxidation method [48]. 800 mL of the initial solution (as described in the materials section) were transferred to a 2.5 L baffled TUNAIR[™] shake flask and the initial pH value was set to pH 8. The solution was agitated in the Max Q 6000 incubator. The exact time of the change of color was determined using a Spotlight Webcam Pro coupled with the ContaCam Software. The kL_a value can be calculated as follows:

$$kL_a = \frac{c_{Na_2SO_3} \cdot k_1}{\nu \cdot t \cdot k_1 \cdot c_{O_2}^* - c_{Na_2SO_3}} \quad (1)$$

with $c_{Na_2SO_3}$ being the sulfite initial concentration, k_1 the first-order reaction constant, ν the stoichiometric coefficient of oxygen, t the time until the indicator's change of color, and $c_{O_2}^*$ the oxygen solubility. The determined shaking speed was then used for the scale-up cultivation. The adapted cultivation and the cell disruption were performed as described in the product analytics section.

2.2.5 Purification of Cherry-GST

The purification of Cherry-GST was performed using GST affinity chromatography on a 5 mL GSTrap[™] 4B affinity column. 10 mL of lysate from the scale-up experiment were mixed with 30 mL of 50 mM Tris-HCl pH 8 binding buffer. For removal of aggregates and particulates from the thawing process, the mixture was centrifuged at 12000 rpm, 20 °C for 15 min and filtered using a sterile 0.2 μ m PES syringe filter. 40 mL of sample were injected into the affinity column. For sample loading and the washing step, a flow of 1 mL/min was applied. For elution and regeneration, the flow was increased to 5 mL/min. The GSTrap[™] 4B affinity column was equilibrated with 2 CV (10 mL) of binding buffer until 8 CV (40 mL) flushed sample injection, followed by 2 CV (10 mL) of additional binding buffer wash. The product elution was performed in a 100% step to the elution buffer which included 10 mM of reduced L-glutathione. Fractions of 2 mL were taken throughout the purification process. The fractions and the applied sample were then analyzed by Cherry-Tag[™]-specific absorption measurements and compared to capillary gel electrophoresis analysis.

2.2.6 Production of Untagged GST

To investigate whether the productivity trends derived from the HTE upstream screening for Cherry-GST are identical for GST alone, two experimental extremes (low and high natively produced product titers) were investigated for an *E. coli* SE1 GST strain constructed in a StabyCodon T7 protein expression kit. The scale up cultures were produced as described above. As the Cherry-Tag[™] was not present in these experiments, the amount of natively produced GST was determined by binding to a 5 mL GSTrap[™] 4B affinity column, as described above.

As an alternative for producing untagged GST, an enzymatic tag removal reaction was

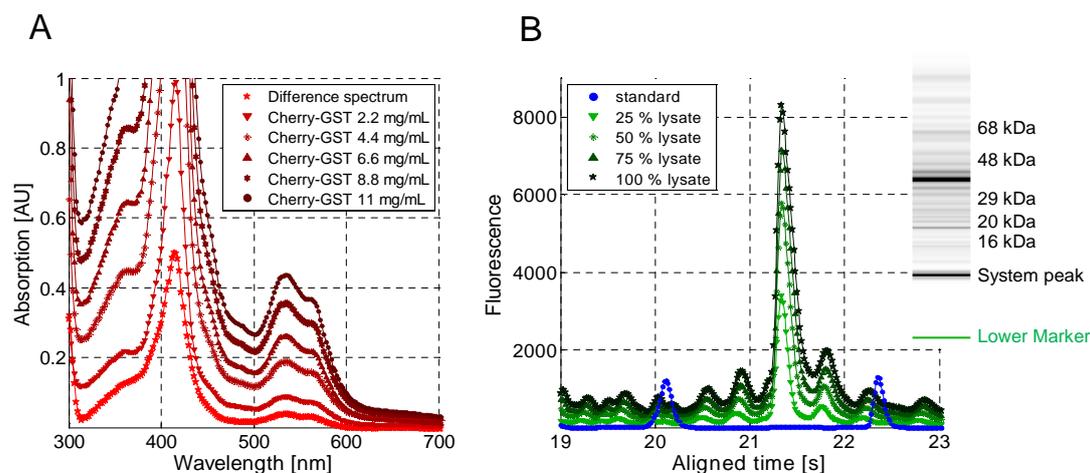


Figure 1: Development of a Cherry-TagTM analytics - A: Lysate spectra for the determination of a specific detection wavelength. Purified Cherry-GST samples in a concentration range of 2.2 to 11 mg/mL are compared to a difference spectrum of a lysate with and without Cherry-GST. B: Fluorescence signal of four dilution steps of a Cherry-GST lysate determined by capillary gel electrophoresis and a virtual gel of the applied lysate. The 29 and 48 kDa standards are illustrated by blue circles.

investigated using a recombinant bovine enterokinase. Purified Cherry-GST was stored in enterokinase reaction buffer. The cleavage reaction was performed in 200 μ L batches using a 2 mL 96-well square deep-well plate at 37 $^{\circ}$ C on a MUR 13 thermo-shaker with 5 μ L enterokinase per 1 mg of fusion protein. The reaction was stopped after 10, 20, 30, and 40 min for investigating the reaction kinetics. All samples were then analyzed in the GX II as duplicates, as described above.

3 Results & Discussion

3.1 Establishment of the Cherry-TagTM Analytics

Lysate spectra were taken to determine a specific wavelength for the quantification of Cherry-GST. The blanked and 300 nm scaled spectra of a Cherry-GST (positive control) and a glucose-repressed lysate (negative control) were used to calculate a difference spectrum (Fig. 1A - pentagrams). The spectrum shows one major maximum at 414 nm and one minor maximum at 536 nm with a shoulder at 562 nm. The first maximum agrees with the Cherry-TagTM-specific wavelength from the Cherry Express kit manual of 413 nm [39, 40]. The second wavelength has not yet been reported and shows a lower sensitivity.

The Cherry-GST extinction coefficient was determined to find a linear correlation between 414 nm/ 536 nm absorption measurements and the product concentration in solution. The Cherry-GST concentrations of four lysate dilutions (25, 50, 75, 100% lysate) were measured in the capillary gel electrophoresis system as quadruplicates. The denatured samples were stained with a fluorescent dye. Small molecules have a higher mobility in the highly viscous capillary gel matrix and are detected at an earlier aligned time. Fig. 1B shows the mean fluorescence signal plotted versus the aligned time for the lysate dilution quadruplicates and for the system internal 29 kDa (20.1 s) and 48 kDa (22.3 s) protein

Shaking speed/Temperature	27 °C	32 °C	37 °C
600 rpm	18.7 ± 1.3 SU/h	19.9 ± 0.7 SU/h	22.9 ± 1.1 SU/h
800 rpm	29.9 ± 0.3 SU/h	34.2 ± 0.5 SU/h	39.7 ± 0.4 SU/h
1200 rpm	37.0 ± 0.5 SU/h	59.4 ± 2.0 SU/h	65.4 ± 2.8 SU/h

Table 1: Maximal growth rates of *E. coli* SE1 at cultivation temperatures of 27, 32, and 37 °C and rotational shaking speeds of 600, 800, and 1200 rpm in the exponential growth phase in SU/h.

standards of known concentration that were injected after each 12 assayed samples for internal calibration. The Cherry-GST product peak was correlated to a molecular weight of 39 kDa at an aligned time of 21.3 s. As GST has a molecular weight of approximately 27 kDa and the Cherry-TagTM alone of 11 kDa, this determined size agrees well with the size of the protein fusion. The concentrations were calculated by peak-baseline integration and compared to the external lysozyme standard of 1 mg/mL to compensate lot-to-lot variations of GX II chips and chemicals. As expected, the fluorescence signals intensify with increasing lysate concentration. In Fig. 1B a virtual gel of the measured sample (calculated from the integrated fluorescence signals) is shown on the right. The lysate dilutions were also measured at 414 nm in a spectrophotometer as triplicates. Comparison of Cherry-GST concentrations with the absorption signal at 414 nm leads to a linear correlation up to an absorption of 1.2 AU with a coefficient of determination (R^2) of 99.4%. The slope of this linear correlation of 2.213 mg/(mL·AU_{414 nm}) was used as the extinction coefficient of Cherry-GST at 414 nm. The GX II measurements only served for method validation and can be neglected in future work with Cherry-TagTM proteins. For the 536 nm extinction coefficient, spectra of purified Cherry-GST samples (see purification of Cherry-GST section) with concentrations of 2.2 to 11 mg/mL were compared (Fig. 1A). As expected for Cherry-GST concentrations above 3 mg/mL, the 414 nm signal was outside of the linear measurement range. The 536 nm signal showed a 10 times lower sensitivity. The slope of the linear correlation was determined to be 24.290 mg/(mL·AU_{536 nm}) and a coefficient of determination (R^2) of 99.8% was obtained. Hence, the 536 nm signal can be a useful alternative for the determination of more concentrated Cherry-TagTM solutions.

3.2 HTE Upstream Screening

On each 48-well FlowerPlate[®] used for cultivation screening, four different induction times at cell densities of OD_{600 nm} 1, 2, 4, and 8 AU and three inducer concentrations of 0.1, 0.5, and 1 mM IPTG were investigated as quadruplicates. Coupled with three cultivation temperatures of 27, 32, and 37 °C and three shaking speeds of 600, 800, and 1200 rpm, 108 different cultivation setups resulted. Only cells cultivated at a shaking speed of 600 rpm qualitatively showed colored pellets (Fig. 2B).

Quantitative evaluation of all setups was realized with the standardized cell disruption method by sonication and lysate absorption measurements at 414 nm. The spectral data were converted to concentrations via the extinction coefficient. The product titers of all cultivation setups were determined within less than 5 min in a standard plate reader. Analyzing that amount of samples in a GX II, including system preparation, would take

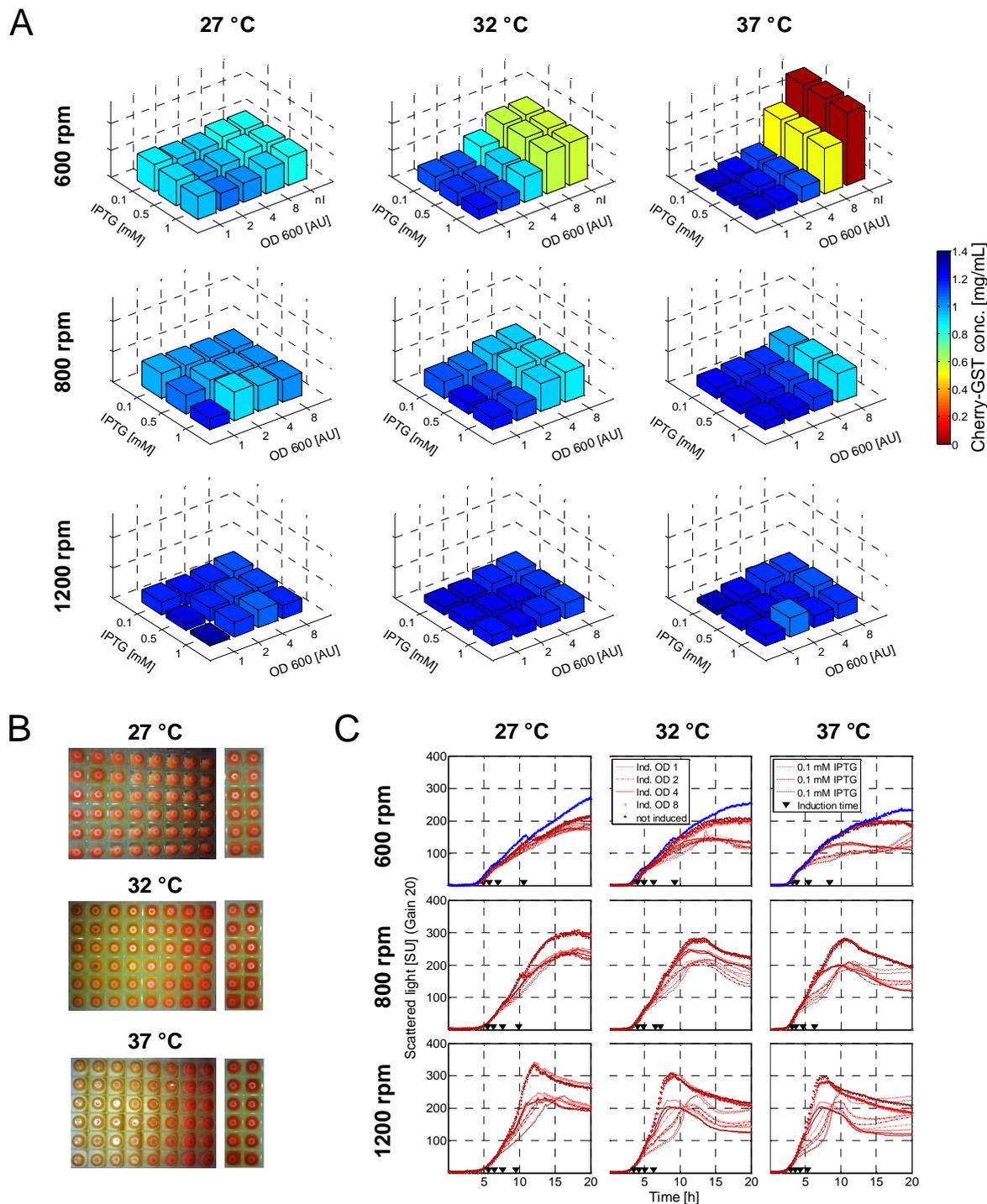


Figure 2: BioLector[®] growth curves and determined product titers after cell disruption of *E. coli* SE1 producing Cherry-GST under different cultivation conditions: Temperatures of 27, 32, and 37 °C and three different shaking speeds of 600, 800, and 1200 rpm coupled with different induction times (OD 1, 2, 4, and 8) and inducer concentrations (0.1, 0.5, and 1 mM IPTG) are shown. A: Product titers of soluble Cherry-GST [mg/mL] under the above-mentioned cultivation conditions. B: Photographs of the centrifuged cell pellets cultivated at 600 rpm shaking speed. C: BioLector[®] growth curves under the above-mentioned cultivation conditions. The scattered light signal is plotted versus the time starting from inoculation. The black triangles indicate the respective induction times.

approximately 3 - 4 h.

The titers of soluble and native Cherry-GST of all cultivation setups are illustrated in 3D bar plots (Fig. 2A). For 600 rpm experiments, the non-induced samples are shown as well. The best product titers were achieved for cultivations at 600 rpm shaking frequency, yielding a maximum of 1.34 mg/mL Cherry-GST for the not-induced setup at 600 rpm and 37 °C. Shaking frequencies higher than 600 rpm resulted in maximum product concentrations of 0.49 mg/mL (at 800 rpm, 32 °C, induction at OD_{600 nm} of 8). The increased oxygen transfer rates at higher shaking speed (15.0 mmol/L/h at 600 rpm, 24.5 mmol/L/h at 800 rpm, and 65.0 mmol/L/h at 1200 rpm as known from the FlowerPlate[®] product sheet by m2p-labs) were found to decrease native Cherry-GST production.

For the 600 rpm setups at 37 and 32 °C, a late or not-induced cultivation led to highest product titers. Early induced cells at OD_{600 nm} of 1 yielded only 0.09 to 0.27 mg/mL native Cherry-GST, whereas late or not-induced setups resulted in titers of 0.77 to 1.34 mg/mL. The T7 promoter seems to be too strongly induced to produce Cherry-GST in a native state, resulting in inclusion bodies for early induced cells. This assumption was investigated by a total protein extraction under denaturing conditions, as discussed in the experimental section. For the 600 rpm setups at 27 °C, an almost constant product level was determined in a range of 0.32 to 0.53 mg/mL. The variation of the IPTG concentration, showed no pronounced influence on the bacterial growth and product formation in all investigated setups. The T7 promoter in *E. coli SE1* is bound very strongly by the T7 RNA polymerase and results in a switch-on or -off mechanism which was concentration-independent in the investigated setups.

The growth curves of all experiments are illustrated in Fig. 2C. The slope of the exponential growth phase determines the maximal growth rates and, thus, the proliferation frequency and activity (Table 1). The slopes decrease with a reduction of shaking speed and also with decreasing temperatures. The highest growth rates of 65.4 ± 2.8 SU/h were determined at 37 °C and 1200 rpm shaking speed of the BioLector[®] and the lowest growth rates of 18.7 ± 1.3 SU/h were observed at 27 °C and 600 rpm shaking speed. The shaking speed was found to be the more influencing parameter on bacterial growth. For all 600 rpm experiments, the slopes for all temperatures were found to be in a similar range of 18.7 ± 1.3 SU/h and 22.9 ± 1.1 SU/h. Those comparably low proliferation rates had been expected, as *E. coli* is an aerobic and mesophilic microorganism and exhibits decreased growth under oxygen-limited conditions and temperatures other than 37 °C. Growth under conditions other than the optimum were beneficial for slow and correctly folded proteins, however. The lower stationary phase plateaus of the early induced cells imply a strong change in metabolism due to product formation. This again indicates formation of inclusion bodies which were not detectable in the 414 nm signals.

3.3 Detection of Inclusion Bodies

The HTE upstream screening resulted in low product titers for early induced cells. As proteins with Cherry-Tag[™] are said to lose color in a denatured and insoluble state, those findings suggested the formation of inclusion bodies. A protein extraction under denaturing conditions was applied, where all soluble and insoluble proteins were extracted and unfolded. The stained gel from a gel electrophoresis run under different cultivation conditions is illustrated in Fig. 3. Lane 1 shows the protein standard with markers

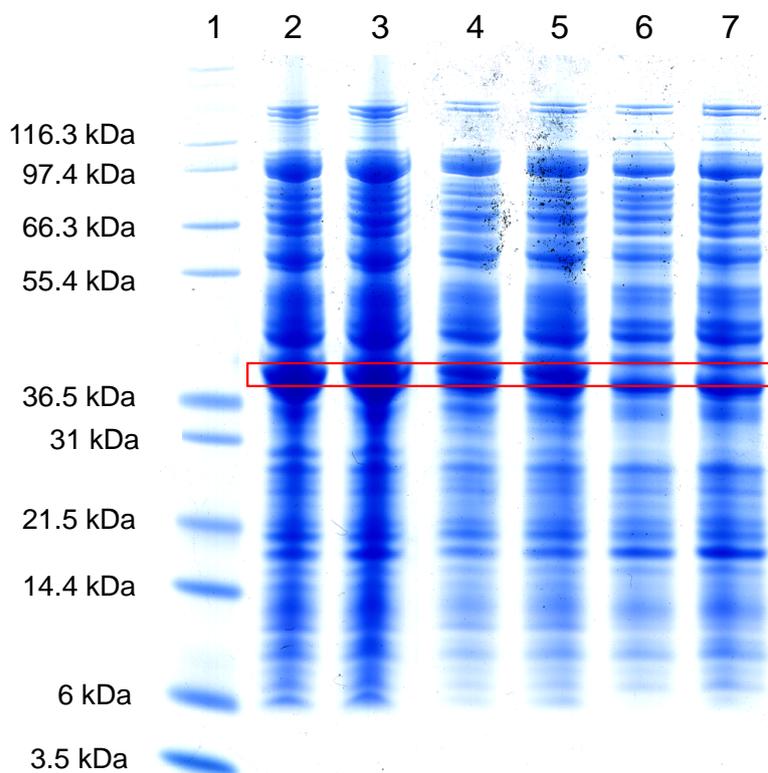


Figure 3: SDS-PAGE of the denaturing protein extraction for determination of the total protein content and identification of inclusion bodies of different cultivation setups at 600 rpm and 37 °C. Lane 1: Protein standard of molecular weight in a range from 3.5 to 116.3 kDa. Lanes 2 + 3: Induction at $OD_{600\ nm}$ 1 with 1 mM IPTG. Lanes 4 + 5: Induction at $OD_{600\ nm}$ 8 with 1 mM IPTG. Lanes 6 + 7: Not induced.

of known molecular weight in a range of 3.5 to 116.3 kDa. Lanes 2 and 3 are total protein extracts from an early induced ($OD_{600\ nm}$ of 1) BioLector[®] cultivation at 37 °C and 600 rpm. Lanes 4 and 5 show extracts of a not induced sample under the same cultivation conditions and lanes 6 and 7 of a glucose-repressed sample. All samples exhibit an identical contaminant pattern with protein bands over the entire molecular weight range. As expected, the Cherry-GST band at 39 kDa is missing in lanes 6 and 7 due to the inhibition of the lac-operon (cAMP positive control). Note that the close band at 37 kDa represents a contaminant and, hence, is present in all investigated samples. Lanes 2 to 5 show an intense product band at 39 kDa. As gel electrophoresis is a qualitative analytical method, product titers cannot be specified. However, the induced sample shows an at least equal amount of Cherry-GST than the not induced sample. The lysate measurements of those samples (compare Fig. 2A - 37 °C, 600 rpm), by contrast, showed 10 times higher absorption values for not induced cells. Hence, the intense product band for the induced cultivation can be correlated to misfolded inclusion bodies. The *E. coli* SE1 T7 promoter was proven to be too strong for the native production of Cherry-GST under induced conditions. The assertion of uncolored protein in the unfolded state was verified for upstream applications.

3.4 Scale-up Culture

The optimized 1 mL BioLector[®] culture at 600 rpm shaking frequency and 37 °C was scaled up by keeping the oxygen transfer rate constant. From the FlowerPlate[®] product sheet by m2p-labs, it was known that for 1 mL filling volume at 600 rpm, a kL_a value of 88.9 h⁻¹ was determined by the sulfite oxidation method [48]. The incubator shaking speed of equivalent oxygen uptake for the 2.5 L baffled TUNAIR[™] shake flask with 800 mL filling volume was found to be 180 rpm with a kL_a value of 87.1 h⁻¹.

In analogy to the BioLector[®] setup, cell harvest was performed after 20 h, followed by the cell disruption as described in the experimental section. Capillary gel electrophoresis runs of the lysate resulted in a Cherry-GST concentration of 12.96 mg/mL. Note that the concentration is in 20 mL lysis buffer from 180 mL cell broth (concentration factor 9). The 600 rpm and 37 °C setup in the BioLector[®] resulted in a Cherry-GST concentration of 1.34 mg/mL in 5 mL lysis buffer from 4 mL cell broth (concentration factor 0.8). To compare the product titers of those two scales, the determined concentrations were divided by the concentration factors, resulting in 1.68 mg/mL in micro-scale and 1.44 mg/mL in the scaled-up cultivation. Although the oxygen transfer rate was kept constant, a deviation in productivity of 0.24 mg/mL was observed. Garcia Ochoa et al. proposed that apart from physical and biochemical system properties also geometrical parameters of the bioreactor, such as the size and design, influence the scale up process [49].

3.5 Purification of Cherry-GST

Capturing of GST using an affinity resin (GSTrap[™] 4B) is based on the catalytic activity of the enzyme. Only soluble and enzymatically active GSTs bind to the chosen resin, as was outlined in [50, 51]. The chromatogram of the purification process is shown in Fig. 4A. The 280 nm signal (solid line) is the non-specific protein signal of aromatic amino acids. The 414 nm and 536 nm signals (dash-dotted and dotted lines) were used as the Cherry-Tag[™]-specific signals. The glutathione concentration is represented by the dashed line.

After sample injection, the 414 nm signal increases to approximately 300 mAU and it continues to increase with time up to 720 mAU after 40 mL of injection. The first sharp increase of the 414 nm signal can be correlated to colored media components from the cultivation procedure. In contrast to this, the steady increase with time indicates the advanced loading of the affinity column. After the column wash at 50 mL, the product is eluted by shifting the glutathione concentration to 10 mM. The 280 nm and the 414 nm signals increase above the linear range of the UV measurement cell. The 536 nm signal can still be recorded with a maximum of 1080 mAU. The less sensitive 536 nm signal was proven to be useful to evaluate high product titers of different purification runs. This ensures measurement signals in the linear range even for product elution, which is a prerequisite for e.g. modeling.

Fractionation times for the 2 mL samples are marked by downward triangles. Those fractions were 10 times diluted and analyzed in the Caliper GX II capillary gel electrophoresis system (reference analytics). Fig. 4A shows virtual gels of the applied sample (left of the dashed line) and all fractions of the affinity chromatography run, calculated from the integrated fluorescence signals. The 39 kDa Cherry-GST band intensifies with sample

Sample	Purity [%]	Concentration [mg/mL]
Sample	43.65 ± 0.42	3.24 ± 0.10
FT ₂ mL	51.30 ± 0.71	0.04 ± 0.01
FT ₄ mL	25.21 ± 1.92	0.30 ± 0.03
FT ₆ mL	13.46 ± 0.53	0.66 ± 0.05
FT ₈ mL	13.70 ± 0.19	0.78 ± 0.06
FT ₁₀ mL	14.30 ± 0.23	0.76 ± 0.06
FT ₁₂ mL	14.34 ± 0.44	0.76 ± 0.08
FT ₁₄ mL	15.01 ± 0.10	0.81 ± 0.06
FT ₁₆ mL	15.56 ± 0.27	0.82 ± 0.04
FT ₁₈ mL	16.09 ± 0.90	0.85 ± 0.09
FT ₂₀ mL	16.11 ± 0.08	0.86 ± 0.03
FT ₂₂ mL	16.39 ± 0.02	0.88 ± 0.05
FT ₂₄ mL	16.57 ± 0.21	0.91 ± 0.07
FT ₂₆ mL	17.15 ± 0.01	0.88 ± 0.00
FT ₂₈ mL	18.38 ± 0.13	0.89 ± 0.02
FT ₃₀ mL	18.02 ± 0.36	0.96 ± 0.01
FT ₃₂ mL	19.08 ± 0.95	0.96 ± 0.05
FT ₃₄ mL	20.05 ± 0.68	1.06 ± 0.14
FT ₃₆ mL	21.56 ± 0.01	0.98 ± 0.03
FT ₃₈ mL	20.91 ± 1.45	1.06 ± 0.05
FT ₄₀ mL	24.44 ± 3.69	0.74 ± 0.49
W ₂ mL	24.63 ± 0.69	1.04 ± 0.10
W ₄ mL	38.00 ± 0.08	0.35 ± 0.00
W ₆ mL	100.00 ± 0.00	0.04 ± 0.00
W ₈ mL	100.00 ± 0.00	0.02 ± 0.00
W ₁₀ mL	-	0.00 ± 0.00
E ₂ mL	100.00 ± 0.00	0.03 ± 0.01
E ₄ mL	89.86 ± 0.11	11.89 ± 0.24
E ₆ mL	90.10 ± 0.12	20.14 ± 0.00
E ₈ mL	89.62 ± 0.18	7.70 ± 0.16
E ₁₀ mL	88.98 ± 0.35	3.08 ± 0.15

Table 2: GX II fraction analysis of the GST affinity chromatography procedure: Purity and concentration of all flow-through (FT), wash (W) and elution (E) fractions are compared to the performance of the applied sample.

loading, while the contaminant levels remain constant. Fluorescence signals of the GX II for selected flow-through (FT) and elution (E) fractions are illustrated in Fig. 4B + C. The system internal 29 kDa (20.1 s) and 48 kDa (22.3 s) protein standards are indicated as circles. The Cherry-GST product peak was correlated to a molecular weight of 39 kDa (21.3 s). In Fig. 4B the sample and three flow-through fractions after 6, 24, and 40 mL are compared. The contaminant peaks for all illustrated samples are identical, while the product fluorescence peak (21.3 s) of the flow-through fractions is far below the Cherry-GST signal of the applied sample. Still, a minor increase of the product peak can be observed over the sample loading volume from 6 to 40 mL agreeing well with the findings derived in 414 and 536 nm real-time measurements on the ÄKTA™ purifier as discussed above. Fig. 4C shows the fluorescence signals of all elution fractions. There is only one major peak corresponding to Cherry-GST (21.3 s) with a concentration ranging from 3.08 to 20.14 mg/mL (Table 2).

Cherry-GST concentrations and purities of the sample and all fractions are listed in Table 2. The concentration of the applied sample was determined to be 3.24 ± 0.10 mg/mL. In the flow-through fractions the product concentration increased from 0.30 ± 0.03 mg/mL (9.3% of the applied sample) to 1.06 ± 0.05 mg/mL (32.7% of the applied sample). The yield of the purification was determined to be 71.1% and the product was three times concentrated to 10.7 mg/mL in the elution pool E_{4 mL} to E_{10 mL}. As over 90% of Cherry-GST were binding after sample application, Cherry-GST was proven to be in a catalytically active state.

Comparison of the 536 nm signal from the integrated ÄKTA™ purifier system to the concentrations from GX II resulted in a linear correlation of $11.521 \text{ mg}/(\text{mL}^2 \cdot \text{AU}_{536 \text{ nm}})$ with an R^2 of 99.1%. Hence, concentrations can be determined directly from online measured signals, also for downstream applications making sampling obsolete. Another major advantage of using the Cherry-Tag™ in DSP applications is the possibility of making decisions in real-time, e.g. making product fractionation decisions without needing post-processive analytics and having to repeat the experiment.

3.6 Production of Untagged GST

To prove that the productivity trends derived from the HTE upstream screening for the tagged product hold for GST alone for the *E. coli* SE1 GST strain, two experimental extremes of low and high natively produced product titers (induction at OD_{600 nm} of 1 and 8 using 0.5 mM IPTG) were investigated at 37 °C and 180 rpm as scale-up cultures. The titer of natively produced GST was determined by binding to a 5 mL GSTrap™ 4B affinity column and evaluation of the eluate fractions. The chromatography runs of the two investigated cell lysates are shown in Fig. 5A. The red curves represent the UV signals of the early induced cultures (OD_{600 nm} of 1) and the blue curves indicate the late induced cultures (OD_{600 nm} of 8). The 300 nm signal serves as a less sensitive protein signal in addition to the standard 280 nm signal. The flow-through fraction of 40 mL is similar for both lysates with both UV signals within the measurement noise. After the wash, however, differences were observed for the elution peaks. The early induced cell lysate shows a much smaller 300 nm peak with an integrated area of 1230 mAU*mL compared to the late induced cell lysate with an integrated peak area of 21026 mAU*mL. GX II analysis of the collected elution fractions indicated a total yield of 30.0 mg GST

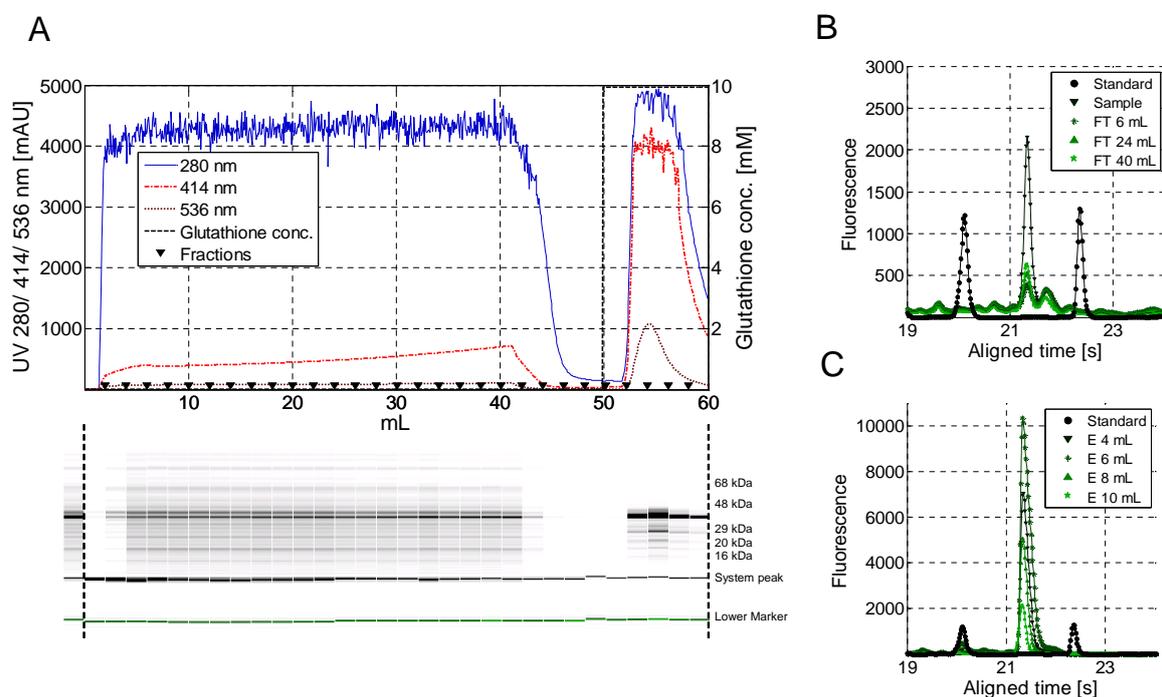


Figure 4: GST affinity chromatography procedure as a capture step for Cherry-GST. A: Chromatogram (280, 414, and 536 nm signals) from the Cherry-GST capturing procedure. Fractions are indicated by downward triangles. Fraction analysis by capillary gel electrophoresis is shown below as virtual gels. B: Capillary gel electrophoresis fluorescence signals of the applied sample and three different flow-through fractions after 6, 24, and 40 mL loading. C: Capillary gel electrophoresis fluorescence signals of the four elution fractions.

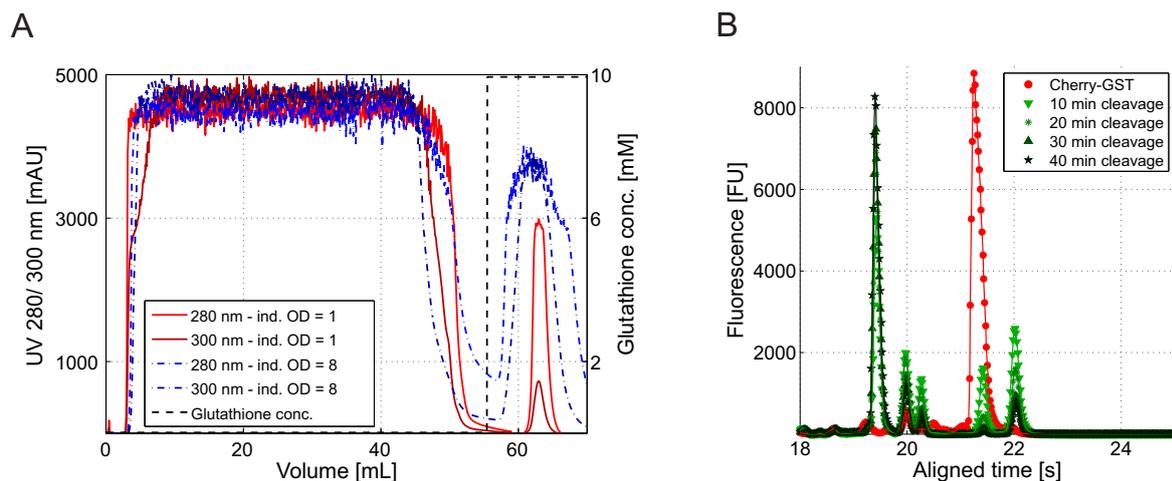


Figure 5: Experiments for the production of untagged GST. A: GST affinity chromatography procedure as a capture step for untagged GST produced at 37 °C and 180 rpm as scale up cultures. 280 nm and 300 nm signals of the chromatography runs of lysates from early induced cells (induction at $OD_{600\text{ nm}}$ of 1) are indicated in red, late induced cells (induction at $OD_{600\text{ nm}}$ of 8) in blue. B: Capillary gel electrophoresis fluorescence signals of a purified Cherry-GST sample (red) and treated samples 10, 20, 30, and 40 min after starting the enzymatic cleavage reaction using enterokinase.

for OD_{600 nm} of 1 induced cells and 511.2 mg GST for OD_{600 nm} of 8 induced cells. The trends of a benefit of a late induction, as seen for Cherry-GST (approximately 7 times increase in productivity compared to early induction), were thus confirmed for GST alone, leading to an over 17-fold increase in productivity for late induced cells. Whereas relative trends were conserved, absolute values differed due to the influence of the fused tag. Nevertheless, the results from the HTE upstream screening using the Cherry-TagTM were directly transferable to a production process for the protein of interest alone.

When using the Cherry-TagTM as an in-production assay, a final removal of the tag after production is mandatory. The enzyme enterokinase recognizes the Asp-Asp-Asp-Asp-Lys linker sequence situated between the protein of interest and the Cherry-TagTM and cleaves at the Lys-residue [41]. Fig. 5B shows fluorescence signals for the applied Cherry-GST sample (red curve) and the reaction mixture after 10, 20, 30, and 40 min of reaction time determined in the GX II (green curves). The 43 kDa enterokinase is represented by the fluorescence peak at 22 s and the enzyme-related impurities are found at 20.0 s and 20.3 s. The peak at 19.5 s indicates untagged GST, whereas Cherry-GST is detected at 21.3 s. The Cherry-GST concentration decreases over time, yielding 14.0% residual substrate after 10 min and 3.9% after 20 min reaction time, as can be seen by the decreasing fluorescence peak at 21.3 s in Fig. 5B. This indicates an efficient catalytic process within the first 20 to 30 min after starting the enzyme reaction. The GST peak at 19.5 s, in contrast, maximizes steadily, whereas the highest increase is detected again within the first 20 to 30 min of catalytic time. It was shown that the tag was effectively cleaved using enterokinase making the Cherry-TagTM a valuable tool for in-production analytics. Also when co-expressing enterokinase with the tagged protein of interest, an economic autocatalysis can be realized, as was shown by Azhar et al. [41].

4 Conclusions

The Cherry-TagTM was found to be a fast and reliable tool for product tracking in diverse fields of bioprocess engineering by simple VIS absorption measurements. The presented technology revealed various advantages when compared to traditional analytical technologies used for proteins.

One major aspect is that no sample pre-treatment is required and the protein can be investigated in the natively folded state. The absence of such pre-processing steps as needed for GX, GX II, and MS leads to a reduction in time and material costs. Nevertheless, those analytical tools are accurate and perform well, as was shown for GX II as the reference analytical method in the presented study for Cherry-GST. However, the total time per sample to be analyzed is much higher for GX II. Preparation of the analytical microfluidics chip as well as sample pre-treatment can easily take up to 2 h. Additionally, approximately 40 s per analyzed sample follow, adding up to another 64 min for analyzing a set of 96 samples. In contrast, analyzing such a number of samples using the Cherry-TagTM analytics takes approximately 1 min in total with a standard 96 well plate reader. Thus, the proposed technology can be implemented easily into HTS systems like robotic workstations.

The possibility to perform in-line and real-time protein analysis, as was shown in the affinity purification setup, even makes sampling obsolete and enables continuous process

control. Such setups are hardly implementable for most other analytical tools, namely, GX, GX II, MS, analytical chromatography, and customized assays. In those setups, fractionated samples have to be analyzed post-processively using e.g. auto-sampling units to perform systematic analysis.

Finally, the Cherry-TagTM was even capable of distinguishing between insoluble - unstable and soluble - active products, as shown in up- and downstream applications. Although strongly differing chemical environments were encountered in up- and downstream experiments, the optical properties of the Cherry-TagTM were not altered in contrast to fluorescent proteins, as was discussed previously. The Cherry-TagTM analytics was applied for the evaluation of the HTE induction screening in the BioLector[®] system. It was shown that only the lowest shaking frequency of 600 rpm investigated resulted in soluble and native Cherry-GST. Early induced cells were found to result in inclusion body formation, leading to uncolored cell pellets. This agrees with the assumptions that only native and soluble proteins show red color. The best upstream setup for soluble Cherry-GST was scaled up successfully. Cherry-GST was purified from the lysate in an affinity chromatography step binding over 90% of the product after injection. As only catalytically active enzymes are retained in GST affinity chromatography columns, the colored proteins were in a native and soluble state. It was possible to calculate product concentrations directly from online 536 nm data in the ÄKTATM purifier system.

The established Cherry-TagTM analytics resulted in two possible wavelengths for the quantitative detection of tagged proteins. 414 nm signals can be used in a concentration range of approximately 0 to 3 mg/mL, whereas 536 nm measurements can be applied even up to 30 mg/mL in a standardized cuvette of 10 mm path length.

This technology has a big potential for simplifying product analytics during bioprocess engineering. One possible application is for pre-production HTS process development by fast and straightforward plate reader analytics on automated robotic workstations. In upstream processing it can be used as a fast tool for clone screening experiments to find suitable promoters for product expression and for the prevention of inclusion bodies. In the study presented the optimized upstream conditions were verified for the product alone, allowing for a scale up of the cultivation process for the untagged protein based on the screenings including the tag. Additionally, the technology is useful for in-production analytics, as was shown in the GST affinity chromatography experiments. The tag is a powerful instrument for real-time determination of protein instabilities during different unit operations. Buffer compositions and unit operations can be changed to make those processes more efficient with less product losses. Electrostatic and hydrophobic protein interaction in downstream processing can, however, strongly differ for the tagged and the pure product species due to the nature of the tag. Despite those obvious drawbacks for DSP, this fact can be used for protein engineering by shifting the elution conditions of proteins that are difficult to separate.

The Cherry-TagTM product sheet indicates its additional potential for stabilizing proteins compared to the untagged species. This might be useful for highly concentrated protein solutions which might be examined in a next step. However, it also remains to be investigated how the tag influences the biological activity of the fused protein.

References

- [1] S. G. Anema, The use of 'lab-on-a-chip' microfluidic SDS electrophoresis technology for the separation and quantification of milk proteins, *Int Dairy J* 19 (2009) 198–204.
- [2] S. Pandey, C. M. Lu, D. A. Herold, Measurement of microalbuminuria using protein chip electrophoresis., *Am J Clin Pathol* 129 (2008) 432–438.
- [3] L. Bousse, S. Mouradian, A. Minalla, H. Yee, K. Williams, R. Dubrow, Protein sizing on a microchip., *Anal chem* 73 (2001) 1207–1212.
- [4] F. Kröner, D. Elsässer, J. Hubbuch, A high-throughput 2D-analytical technique to obtain single protein parameters from complex cell lysates for in silico process development of ion exchange chromatography., *J Chrom A* 1318 (2013) 84–91.
- [5] A. K. Patra, R. Mukhopadhyay, R. Mukhija, A. Krishnan, L. C. Garg, A. K. Panda, Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from *Escherichia coli*., *Protein Express Purif* 18 (2000) 182–192.
- [6] T. I. Potgieter, M. I. Cukan, J. E. Drummond, N. R. Houston-Cummings, Y. Jiang, F. Li, H. Lynaugh, M. Mallem, T. W. McKelvey, T. Mitchell, A. Nysten, A. Rittenhour, D. Stadheim, T. A. and Zha, M. D'Anjou, Production of monoclonal antibodies by glycoengineered *Pichia pastoris*., *J Biotech* 139 (2009) 318–325.
- [7] M. Kuschel, T. Neumann, P. Barthmaier, M. Kratzmeier, Use of lab-on-a-chip technology for protein sizing and quantitation., *J Biomol Tech* 13 (2002) 172–178.
- [8] E. de Hoffmann, V. Stroobant, *Mass spectrometry: Principles and applications*, Wiley VCH, 2007, pp. 306–342.
- [9] J. B. Fenn, M. Mann, C. K. A. I. Meng, S. F. Wong, C. M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules., *Science* 246 (1989) 64–71.
- [10] S. A. Trauger, W. Webb, G. Siuzdak, Peptide and protein analysis with mass spectrometry., *Spectroscopy* 16 (2002) 15–28.
- [11] B. Domon, R. Aebersold, Mass spectrometry and protein analysis., *Science* 312 (2006) 212–217.
- [12] W. G. Lewis, Z. Shen, M. G. Finn, G. Siuzdak, Desorption/ ionization on silicon (DIOS) mass spectrometry: background and applications., *Int J Mass Spectrom* 226 (2003) 107–116.
- [13] D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data., *Electrophoresis* 20 (1999) 3551–3567.
- [14] A. Shevchenko, O. N. Jensen, A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, M. Mann, Linking genome and proteome by mass spectrometry: Large-scale., *Proc Natl Acad Sci USA* 93 (1996) 14440–14445.

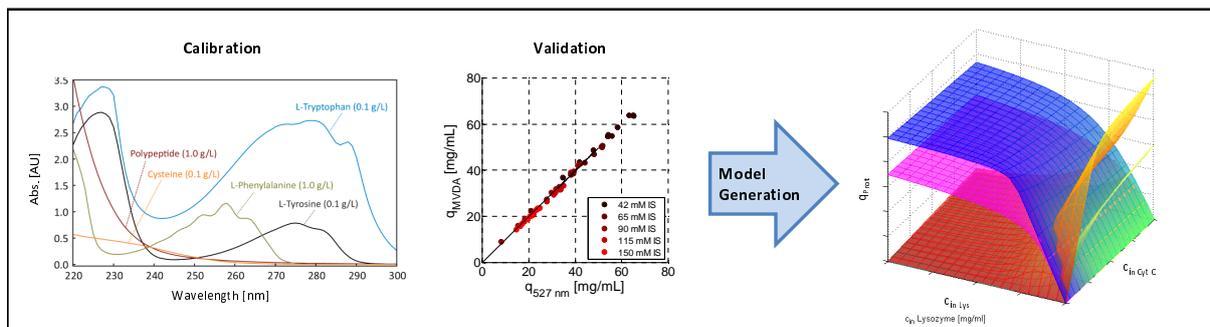
- [15] F. Kröner, A. T. Hanke, B. K. Nfor, M. W. H. Pinkse, P. D. E. M. Verhaert, M. Ottens, J. Hubbuch, Analytical characterization of complex , biotechnological feedstocks by pH gradient ion exchange chromatography for purification process development., *J Chrom A* 1311 (2013) 55–64.
- [16] B. Maiser, F. Kröner, F. Dimer, G. Brenner-Weiss, J. Hubbuch, Isoform separation and binding site determination of mono-PEGylated lysozyme with pH gradient chromatography., *J Chrom A* 1268 (2012) 102–108.
- [17] S. Flatman, I. Alam, J. Gerard, N. Mussa, Process analytics for purification of monoclonal antibodies., *J Chrom B* 848 (2007) 79–87.
- [18] B. Maiser, F. Dimer, J. Hubbuch, Optimization of random PEGylation reactions by means of high throughput screening., *Biotechnol Bioeng* 111 (2014) 104–14.
- [19] H. Goetz, M. Kuschel, T. Wulff, C. Sauber, C. Miller, S. Fisher, C. Woodward, Comparison of selected analytical techniques for protein sizing, quantitation and molecular weight determination., *J Biochem Bioph Meth* 60 (2004) 281–293.
- [20] P. Diederich, S. K. Hansen, S. A. Oelmeier, B. Stolzenberger, J. Hubbuch, A sub-two minutes method for monoclonal antibody-aggregate quantification using parallel interlaced size exclusion high performance liquid chromatography., *J Chrom A* 1218 (2011) 9010–9018.
- [21] A. Apte-Deshpande, S. Somani, G. Mandal, S. Soorapaneni, S. Padmanabhan, Over expression and analysis of O-glycosylated recombinant human granulocyte colony stimulating factor in *Pichia pastoris* using Agilent 2100 Bioanalyzer., *J Biotech* 143 (2009) 44–50.
- [22] H. Jin, G. T. Cantin, S. Maki, L. C. Chew, S. M. Resnick, J. Ngai, D. M. Retallack, Soluble periplasmic production of human granulocyte colony-stimulating factor (G-CSF) in *Pseudomonas fluorescens*., *Protein Express Purif* 78 (2011) 69–77.
- [23] H. Forootanfar, M. A. Faramarzi, A. R. Shahverdi, M. T. Yazdi, Purification and biochemical characterization of extracellular laccase from the ascomycete *Paraconiothyrium variabile*., *Bioresource technol* 102 (2011) 1808–1814.
- [24] C. Su, B. H. Chiang, Partitioning and purification of lysozyme from chicken egg white using aqueous two-phase system, *Process Biochem* 41 (2006) 257–263.
- [25] M. J. C. Van Lierop, F. Wijnands, Y. W. Loke, P. M. Emmer, H. G. M. Lukassen, D. D. M. Braat, A. van der Meer, I. Mosselman, S. and Joosten, Detection of HLA-G by a specific sandwich ELISA using monoclonal antibodies G233 and 56B., *Mol hum reprod* 8 (2002) 776–784.
- [26] B. Gasser, M. Maurer, J. Gach, R. Kunert, D. Mattanovich, Engineering of *Pichia pastoris* for improved production of antibody fragments., *Biotechnol Bioeng* 94 (2006) 353–361.

-
- [27] N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer, R. Y. Tsien, Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein., *Nature biotechnol* 22 (2004) 1567–1572.
- [28] A. Osberghaus, K. Drechsel, S. Hansen, S. K. Hepbildikler, S. Nath, M. Haindl, E. Von Lieres, J. Hubbuch, Model-integrated process development demonstrated on the optimization of a robotic cation exchange step., *Chem Eng Sci* 76 (2012) 129–139.
- [29] N. Brestrich, T. Briskot, A. Osberghaus, J. Hubbuch, A tool for selective inline quantification of co-eluting proteins in chromatography using spectral analysis and partial least squares regression., *Biotechnol Bioeng* 111 (2014) 1–9.
- [30] F. Dimer, S. Hansen, S. A. Oelmeier, J. Hubbuch, Accurate retention time determination of co-eluting proteins in analytical chromatography by means of spectral data., *Biotechnol Bioeng* 110 (2013) 683–693.
- [31] Promega Corporation, HaloTag Mammalian Protein Detection and Purification Systems., 2013.
- [32] R. F. Ohana, L. P. Encell, K. Zhao, D. Simpson, M. R. Slater, M. Urh, K. V. Wood, HaloTag7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification., *Protein Express Purif* 68 (2009) 110–120.
- [33] S. Cabantous, T. C. Terwilliger, G. S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein., *Nature biotechnol* 23 (2005) 102–107.
- [34] R. Y. Tsien, The Green Fluorescent Protein, *Annu Rev Biochem* 67 (1998) 509–544.
- [35] N. C. Shaner, P. A. Steinbach, R. Y. Tsien, A guide to choosing fluorescent proteins, *Nature Methods* 2 (2005) 905–909.
- [36] O. Griesbeck, G. S. Baird, R. E. Campbell, D. A. Zacharias, R. Y. Tsien, Reducing the Environmental Sensitivity of Yellow Fluorescent Protein, *J Biol Chem* 276 (2001) 29188–29194.
- [37] F. V. Subach, G. H. Patterson, S. Manley, J. M. Gillette, J. Lippincott-Schwartz, V. V. Verkhusha, Photoactivatable mCherry for high-resolution two-color fluorescence microscopy, *Nature Methods* 6 (2008) 153–159.
- [38] J. Fan, Z. Cui, H. Wei, Z. Zhang, Y. Zhou, Y. Wang, X. Zhang, Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells, *Biochem Biophys Res Com* 367 (2007) 47–53.
- [39] Delphi Genetics SA, Cherry Codon Kit., 2009.
- [40] Delphi Genetics SA, Cherry Express Kit., 2009.
- [41] M. Azhar, R. Somashekhar, Cloning, expression and purification of human and bovine Enterokinase light chain with Cherry tag and their activity comparison, *Indian J Applied and Pure Bio* 29 (2014) 125–132.

- [42] K. M. Padmanabha Das, S. Barve, S. Banerjee, S. Bandyopadhyay, S. Padmanabhan, A Novel Thermostability Conferring Property of Cherry Tag and its Application in Purification of Fusion Proteins, *J Microbial Biochem Technol* 1 (2009) 59–63.
- [43] M. Funke, S. Diederichs, F. Kensy, C. Mueller, J. Buechs, The Baffled Microtiter Plate : Increased Oxygen Transfer and Improved Online Monitoring in Small Scale Fermentations., *Biotechnol Bioeng* 103 (2009) 1118–1128.
- [44] F. Kensy, E. Zang, C. Faulhammer, R. Tan, J. Buechs, Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates., *Microb Cell Fact* 8 (2009) 1–17.
- [45] K. Kottmeier, J. Weber, C. Mueller, T. Bley, J. Buechs, Asymmetric Division of *Hansenula polymorpha* Reflected by a Drop of Light Scatter Intensity Measured in Batch Microtiter Plate Cultivations at Phosphate Limitation., *Biotechnol Bioeng* 104 (2009) 554–561.
- [46] Perkin Elmer, HT Protein Express LabChip Kit, Version 2 LabChip GXII User Guide., 2012.
- [47] Life Technologies, NuPAGE Technical Guide., 2010.
- [48] R. Hermann, M. Lehmann, J. Büchs, Characterization of gas-liquid mass transfer phenomena in microtiter plates., *Biotechnol Bioeng* 81 (2003) 178–186.
- [49] F. Garcia-Ochoa, E. Gomez, Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview, *Biotech Adv* 27 (2009) 153–176.
- [50] Promega Corporation, MagneGST Protein Purification System Product Sheet., 2014.
- [51] Clontech Laboratories, Inc., Glutathione Resin User Manual, 2014.

Deconvolution of High-throughput Multi-component Isotherms Using Multi-variate Data Analysis of Protein Spectra

P. Baumann^{1,‡}, T. Huuk^{1,‡}, T. Hahn¹, A. Osberghaus¹ and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany

[‡] : Contributed equally to this work

* : Corresponding author. *E-mail-address*: juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

Gaining a more profound understanding of biopharmaceutical downstream processes is a key demand of the 'Quality by Design (QbD)' guideline. One of the most dominant approaches to gain process understanding is the extensive use of experimental high-throughput formats, such as batch chromatography on robotic liquid handling stations. Using these high-throughput experimental formats, the generation of numerous samples poses an enormous problem to subsequent analytical techniques.

Here, a high-throughput case study for batch chromatographic multi-component isotherms is presented. To debottleneck the subsequent analytics, a non-invasive technique using UV spectra and multi-variate statistics was adapted to a batch chromatographic format. Using this approach, it was possible to integrate the entire analytical setup into the robotic workflow.

As a case study, batch isotherms for SP Sepharose FF and the model proteins cytochrome c and lysozyme at various pH values and ionic strengths were recorded. A successful examination of the quality of the analytical procedure as compared to classical single wavelength photometry was carried out. To address the growing demand for a more profound process understanding, the experimental data were fitted to the steric mass-action isotherm, getting a more detailed insight into the competitive binding behavior at various pH values and ionic strengths.

Keywords: Ion-exchange Chromatography, Multi-component Batch Isotherms, Steric Mass-action Isotherm, High-throughput Process Development, Multi-variate Data Analysis

1 Introduction

Nowadays, biopharmaceutical downstream process (DSP) development is mainly based on chromatographic separation techniques. Chromatography, as many other techniques for biopharmaceuticals, is influenced by numerous operational parameters that affect process performance. The growing need for a deeper mechanistic understanding of the technical process on a molecular level and the impact of the process on the product quality, demanded by the 'Quality by Design (QbD)' approach, pose a great challenge to biopharmaceutical DSP [1, 2]. The requested process understanding can be generated by several approaches. One very obvious approach is the systematic experimental evaluation of the impact of different operational parameters on the downstream process performance.

A possibility to support the exploration of the design space, as basis for many statistical DoE approaches, is the use of high-throughput techniques [3]. These approaches enable miniaturization of the experimental systems and parallelization and automation of the experimental workflow allowing for full factorial experimental designs [4, 5, 6]. Therefore, high-throughput techniques enable a reduction of time and material consumption. In the context of chromatographic process development, there are several experimental systems adapted to robotic high-throughput experiments.

Batch chromatography can be implemented very easily on a liquid handling station using a defined adsorbent volume provided in a 96-well plate. The batch chromatographic systems can be used for resin screening [3], batch bind-elute studies or the measurement

of adsorption isotherms [3, 7, 8, 9] and kinetic data [10].

One major disadvantage of these experimental high-throughput formats is that the above-mentioned experimental bottleneck is not prevented but often only shifted to subsequent analytics. In the simplest case, when only single protein data have to be recorded, a photometric measurement within the robotic workflow is possible. This simple analytical approach is restricted to the quantification of a single protein species or several protein species with different exclusive absorption maxima beside 280 *nm* [11]. The major advantage of photometric assays is the high sample throughput, which can easily cope with the amount of samples generated in robotic high-throughput experiments and the non-invasive nature of the technique.

In pharmaceutical high-throughput process development, a simple photometric measurement will be unable to differentiate between several protein species. In this case, more sophisticated analytics have to be performed. Such advanced analytics might be e.g., analytical high-performance liquid chromatography (HPLC) and product-specific immunoassays [12, 13]. The time requirement of these immunoassays is still high in comparison to that of the chromatographic experiments. HPLC assays are predominant in the characterization of the size and charge heterogeneity of a target protein. One major drawback of processing high-throughput experimental samples with HPLC assays is the low sample throughput of HPLC techniques. Despite the mentioned drawbacks, HPLC and other analytical techniques such as immunoassays bear the major disadvantage of an invasive nature.

An ideal analytical technique in the context of high-throughput experimental workflows would be a photometric assay that can be carried out in standard UV/VIS plate readers typically installed on robotic workstations. This assay has to be protein species-specific and quantitative. Such an assay, based on the measurement of UV spectra and subsequent multi-variate statistics, was introduced by Hansen et al. in 2011 [14]. When using protein mid-UV (200 - 300 nm) spectra, the absorption of different protein species is strongly influenced by properties of peptide bonds and amino acid residues. A partial least squares - projection to latent structures (PLS) regression enables subsequent quantitative evaluation of mixtures of different protein species. Here, we present a case study on the high-throughput generation of multi-component isotherm data. We combine a high-throughput method for the collection of multi-component isotherms in an automated batch format with a non-invasive protein-specific quantification method based on the measurement of mid-UV spectra. The collected data are subsequently fitted using the steric mass-action isotherm including competitive protein binding introduced by Brooks and Cramer [15].

2 Materials & Methods

2.1 Materials

2.1.1 Disposables & Reaction Vessels

As a strong cation-exchange adsorbent, sulfopropyl (SP) sepharose fast flow (FF) provided by GE Healthcare Life Sciences, Sweden, was used. Binding experiments were carried out in 2 mL 96-well square deep well plates (VWR, Germany). Absorption and protein spectra

measurements were carried out in 96-well flat bottom UV-Star[®] microplates (Greiner Bio-One, Germany). Buffers were filtered using 0.2 μm cellulose acetate filters supplied by Sartorius, Germany.

2.1.2 Chemicals & Buffers

Binding experiments at pH 5 were carried out in 20 mM acetate buffer consisting of acetic acid (Merck, Germany) and sodium acetate (Sigma-Aldrich, USA). For experiments at pH 7, a 20 mM phosphate buffer consisting of di-sodium hydrogen phosphate and sodium di-hydrogen phosphate (Sigma-Aldrich, USA) was used. The ionic strength of the applied buffers was adjusted to 42, 65, 90, 115, and 150 mM using sodium chloride (AppliChem, Germany). The model proteins lysozyme from chicken egg white and cytochrome c from equine heart were purchased from Sigma-Aldrich, USA. Calibrations for the multi-variate data analysis model of the two-component system were carried out using a 2.4 mg/mL protein solution. For binding experiments, a 10 mg/mL stock solution with 70% cytochrome c and 30% lysozyme was used.

2.1.3 Instrumentation & Software

For pH adjustment of all buffers, a HI-3220 pH meter (Hanna Instruments, USA) was used. The instrument was calibrated using high-precision standards from Hanna Instruments (USA). For generation of equal amounts of adsorbent per well, a MediaScout[®] Resi-Quot System (Atoll, Germany) was applied. For pressure adjustment, a vacuum pump with pressure regulation was used. For the batch isotherms, a Freedom EVO[®]200 liquid handling station (Tecan, Germany) was used, operated with Freedom EVOware[®]2.1. The system is equipped with eight fixed tips, a plate-moving arm and an orbital shaker (Tecan, Germany). A Rotanta 46RSC centrifuge (Hettich, Germany) and an Infinite M200 UV plate spectrophotometer (Tecan, Germany) are integrated in the system. The spectrometer was controlled by i-control 1.9 (Tecan, Germany). Data processing and creation of figures was performed in Matlab[®]R2011a (MathWorks, USA).

2.2 Experimental Setup

2.2.1 Model Calibration & Validation

The multi-variate data analysis calibration [14] was based on a four-level D-optimal onion design generated with MODDE (Umetrics, Sweden) with additional data points added at low concentration levels. The model included 7 mixing ratios (1:0; 2.5:1; 2:1; 1:1; 1:2; 1:2.5; 0:1) and 15 concentration levels (concentration levels correspond to the protein concentration of cytochrome c and lysozyme in total) from 0 to 1.2 g/L (each 5 concentrations in a range of 0 to 0.1, 0.1 to 0.5, and 0.5 to 1.2 g/L). E.g. a mixing ratio of 2:1 means that the mixture contains 2/3 of cytochrome c and 1/3 of lysozyme. For the example of a total concentration level of 1.2 mg/mL at a mixing ratio of 2:1, the solution contains 0.8 mg/mL cytochrome c and 0.4 mg/mL lysozyme. In total, 32 samples were used for model calibration.

8 samples were added as a test set for model validation. The cytochrome c and lysozyme stock solutions were pipetted in the desired mixing ratios and diluted with the respective

buffers on the liquid handling station. All 40 samples were prepared as 1.8 mL solutions in 96 deep-well plates to avoid small pipetted volumes. 300 μL of each sample were transferred to 96-well flat bottom UV-Star[®] microplates. Sample absorption spectra were measured in a range of 240 - 300 nm in 2 nm steps. The spectral data of the 32 samples mentioned above were used for model calibration in Matlab[®]R2011a, using the PLS toolbox (Eigenvector Research, USA). 5 samples out of 32 were used as an internal cross validation. The regression model was then validated on the external test set of additional 8 samples and used for the concentration determination of unknown samples.

2.2.2 Generation of Equal Adsorbent Volumes

Generation of equal 20.8 μL adsorbent amounts was achieved using the MediaScout[®] Resi-Quot system (Atoll, Germany) described by Herrmann et al. [16]. The system was equipped with a pressure-controlled vacuum pump and the working pressure was set to 800 mbar. To remove the adsorbent storage solution, the adsorbent plaques were washed twice with deionized water and the applied binding buffer of the respective experiment. The equilibrated plaques were transferred into a 2 mL 96-well square deep well plate and suspended in 100 μL binding buffer. The plate was then stored until use on the liquid handling station. The outer wells were not used for isotherm experiments due to the largest variance in adsorbent volume on the plate [17].

2.2.3 Isotherm Experiments

Isotherms covering 10 different starting concentrations c_{in} were generated on the liquid handling station. As only the inner 60 wells were used for the experiments, 6 isotherms could be created per sample plate. Each adsorbent plate was used for one pH level including 5 different salt level isotherms (42, 65, 90, 115, and 150 mM) and one isotherm as duplicate for investigating the repeatability of isotherm data.

The investigated pH levels were pH 5 and 7. 10 mg/mL stock solutions of 70% cytochrome c and 30% lysozyme in the respective buffers were applied. The starting concentrations c_{in} of the isotherm experiments were set to 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, and 8 mg/mL with a final volume of 800 μl per well (including 100 μL adsorbent storage buffer from plaque generation). The adsorbent plate was closed with a lid by the robotic plate-moving arm and placed in the orbital shaker for 2 h. Kinetic studies for lysozyme [18] and a monoclonal antibody [4] have shown that an incubation time of 20 to 40 min is sufficient for reaching the binding equilibrium on SP Sepharose FF. Afterwards, the plate was centrifuged for 10 min at 1000 rpm in the Rotanta 46RSC centrifuge. 300 μL of the resulting supernatant was transferred to a 96-well flat bottom UV-Star[®] microplate. As for model calibration, sample absorption spectra were measured in a range of 240 - 300 nm in 2 nm steps. The spectral data were processed using the previously calibrated regression model. Lysozyme and cytochrome c concentrations could be determined selectively. The validity of cytochrome c levels was additionally assured by comparison to 527 nm absorption measurements.

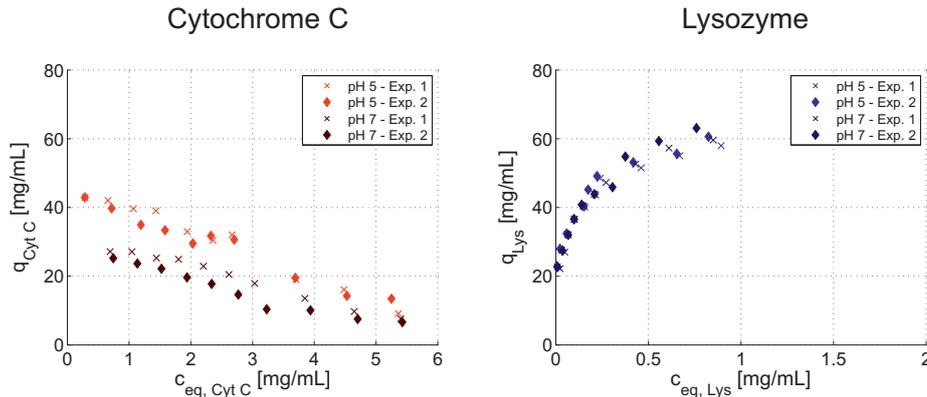


Figure 1: Comparison of isotherm data for determination of the experimental robustness and model performance. Duplicates of the isotherm experiments performed at 90 mM ionic strength for pH 5 and pH 7. The duplicates are indicated as diamonds and crosses, respectively. The results for cytochrome c are shown in red (left) and for lysozyme in blue (right).

2.2.4 Isotherm Fitting

The purpose of isotherm fitting is to validate, whether the observed measurements follow the theoretical framework for protein adsorption in ion-exchange chromatography. The adsorption model applied is the semi-mechanistic steric mass-action isotherm (SMA) introduced by Brooks and Cramer [15]. It incorporates effects of counter-ions on the retention behavior of proteins and includes characteristic charges of proteins ν_i . Additionally, steric shielding effects of proteins are considered as a parameter σ_i , representing sterically hindered binding sites without electrostatic interactions. As a final factor, the total ionic capacity Λ of the applied adsorbent represents the total number of electrostatic binding sites. The isothermal form of the SMA isotherm for a mixture of two proteins is shown in Eq. (1) and Eq. (2), with q_i and $c_{eq,i}$ being the concentration of the protein i adsorbed and in solution, respectively. The effective pore salt concentration is described by c_{salt} . $k_{eq,i}$ is the equilibrium constant of adsorption and desorption.

$$q_1 = k_{eq,1} \left(\frac{\Lambda - (\nu_1 + \sigma_1)q_1 - (\nu_2 + \sigma_2)q_2}{c_{salt}} \right)^{\nu_1} c_{eq,1} \quad (1)$$

$$q_2 = k_{eq,2} \left(\frac{\Lambda - (\nu_2 + \sigma_2)q_2 - (\nu_1 + \sigma_1)q_1}{c_{salt}} \right)^{\nu_2} c_{eq,2} \quad (2)$$

The SMA model allows for fitting isotherm data, including concurring binding behavior and varying salt concentrations. The fitting procedure was carried out in Matlab[®]R2011a using the least squares data fitting function `lsqcurvefit` (trust-region-reflective algorithm). The estimated parameters include the equilibrium constant $k_{eq,i}$, the characteristic charge ν_i , and the steric shielding σ_i .

3 Results & Discussion

3.1 Model Generation & Experimental Performance

Partial least squares regression (PLS) is used to reduce data sets and finding significant variance for correlating several input variables (e.g. wavelengths) with output variables

(e.g. concentrations). Input variables of similar information content are pooled as so-called latent variables (LVs) leading to a data reduction. The first LV carries the highest information content, whereas each additional LV added leads to less and less improvement of the model. At a certain number of LVs, further addition of LVs then leads to incorporation of measurement noise into the system which needs to be avoided. For generation of the MVDA model, the optimal number of latent variables was found to be 5, yielding normally distributed residues for the cross validation of all 32 samples. This setup was found identical for both pH 5 and pH 7. The generated MVDA models were then applied to the external test set consisting of 8 samples. The maximal relative deviation in concentration was determined as 3.9% for pH 5 and 5.3% for pH 7. Hence, the model performed well for a set of samples that were not used for the calibration experiments. Besides testing the performance of the MVDA model, also the repeatability of the experimental data was investigated. The duplicates of the isotherms in the presence of 90 mM ionic strength at pH 5 and pH 7 are shown in Fig. 1 for cytochrome c (left) and lysozyme (right). The protein bound to the adsorbent is plotted versus the residual protein in free solution for equilibrium conditions. The trends of the duplicates (cross/ diamond) agree well for both investigated pH values and proteins. Also, the absolute values of the duplicates are in good accordance considering the experimental difficulties when working with low adsorbent and liquid volumes. Consequently, both the obtained MVDA model and the experimental data were of high quality and were used in the presented study.

3.2 Multi-component Isotherms

To prove the applicability of the MVDA model for real isotherm data in a mixture of proteins, the concentrations of cytochrome c derived from the MVDA are compared to the values of a selective 527 nm analytical wavelength (see Fig. 2). The data points from the MVDA model plotted over the selective 527 nm wavelength measurements are shown in a parity plot. It has to be noted that the MVDA model was calibrated in the UV range (240 - 300 nm) whereas the selective wavelength for cytochrome c is at a much higher wavelength in the VIS region. For pH 5 (Fig. 2A) as well as for pH 7 (Fig. 2B), the isotherm data agree well with coefficients of determination of the data points for all investigated ionic strengths (ISs) of 99.74% for pH 5 and 99.34% for pH 7. Thus, the MVDA model was applicable to this large data set for straightforward HTS applications. The isotherm data for both lysozyme and cytochrome c are shown in Fig. 3 as 2D (left) and 3D (right) scatter plots. The experimental data are illustrated as markers and the fitted SMA model is shown as curves and planes. For all data sets, the binding of proteins decreases with increasing ionic strength (from 42 mM to 150 mM IS). For cytochrome c (red), both pH conditions (pH 5 - Fig. 3A and pH 7 - Fig. 3B) show a similar trend for all ionic strengths: The isotherm data increase for low starting concentrations c_{in} and decrease for higher values of c_{in} . However, the overall binding of cytochrome c at pH 5 (Fig. 3A) is slightly increased when compared to pH 7 (Fig. 3B). E.g., in the setups of 42 mM IS, the maximum binding capacity q_{cytc} was determined to be 63.98 mg/mL for pH 5 and 50.13 mg/mL for pH 7. Again, for lysozyme (blue), both pH conditions behave similarly for all ISs. In contrast to cytochrome c, the isotherm data increase consistently with the starting concentrations c_{in} . The overall binding behavior of lysozyme at pH 5 (Fig. 3A) is slightly decreased when compared to pH 7 (Fig. 3B). The maximum binding

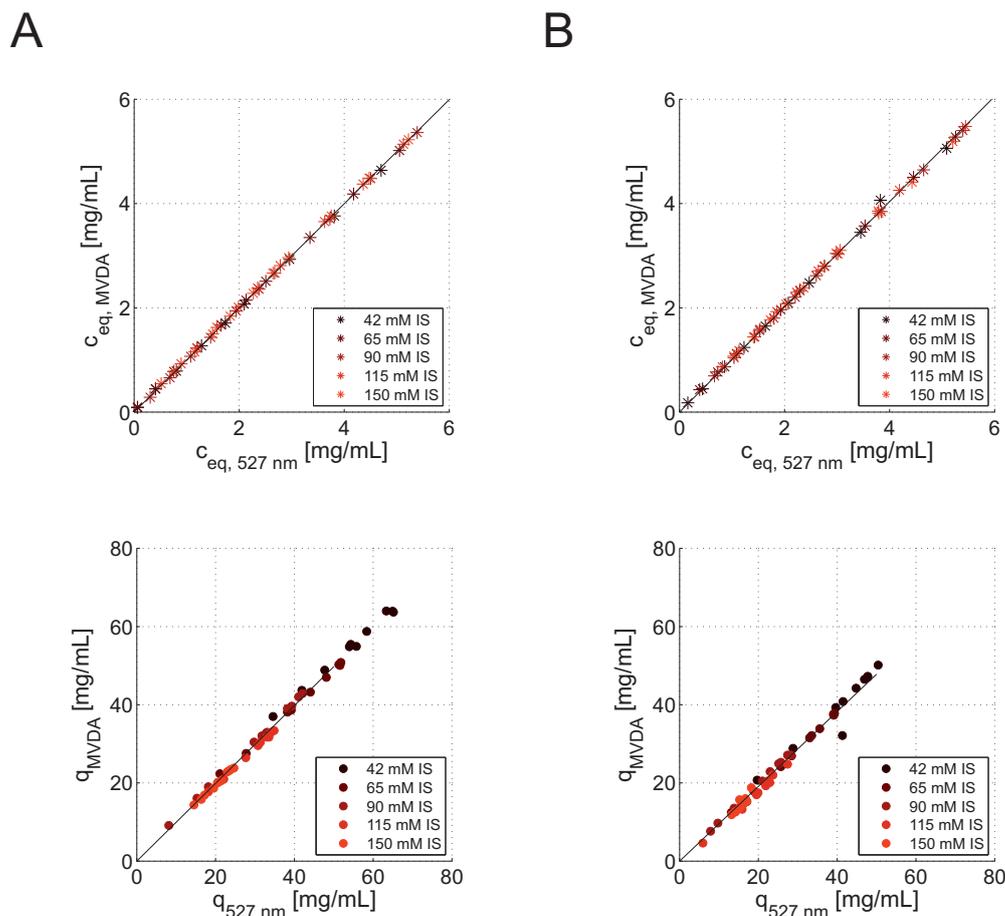


Figure 2: Comparison of cytochrome c isotherm data points at pH 5 (A) and pH 7 (B) derived from the MVDA model and the selective 527 nm wavelength as a secondary analytics for validation. The agreement of the different data point is shown in parity plots for the equilibrium concentration of cytochrome in solution and bound to the adsorbent.

capacity q_{lys} at 42 mM IS was determined to be 67.48 mg/mL for pH 5 and 74.77 mg/mL for pH 7. Although the initial amount of lysozyme in the mixture was much lower (30%) compared to cytochrome c (70%), the maxima of q_{lys} exceed those of q_{cytc} .

The experimental results encountered follow the expected trends. As observed for all investigated setups, an increase in ionic strength causes a weakening of the electrostatic binding in ion-exchange chromatography. Cytochrome c showed a strong increase in protein binding for low starting concentrations c_{in} , starting to decrease with higher values of c_{in} . This indicates a displacement of cytochrome c by lysozyme when the binding process approaches the maximal binding capacity of the adsorbent. This assumption was confirmed by the trends encountered for lysozyme. Here, a continuous increase in protein binding was observed towards a maximum for the highest starting concentrations c_{in} . This displacement under all investigated conditions agrees with the isoelectric points of the two proteins being 10 to 10.5 for cytochrome c and 11.4 for lysozyme (compare data sheet Sigma-Aldrich) yielding a higher net charge for lysozyme and thus a stronger binding to the cation-exchange resin. This also explains the enhancement of the displacement reaction for pH 7 compared to pH 5 due to cytochrome c being closer to its nominal pI.

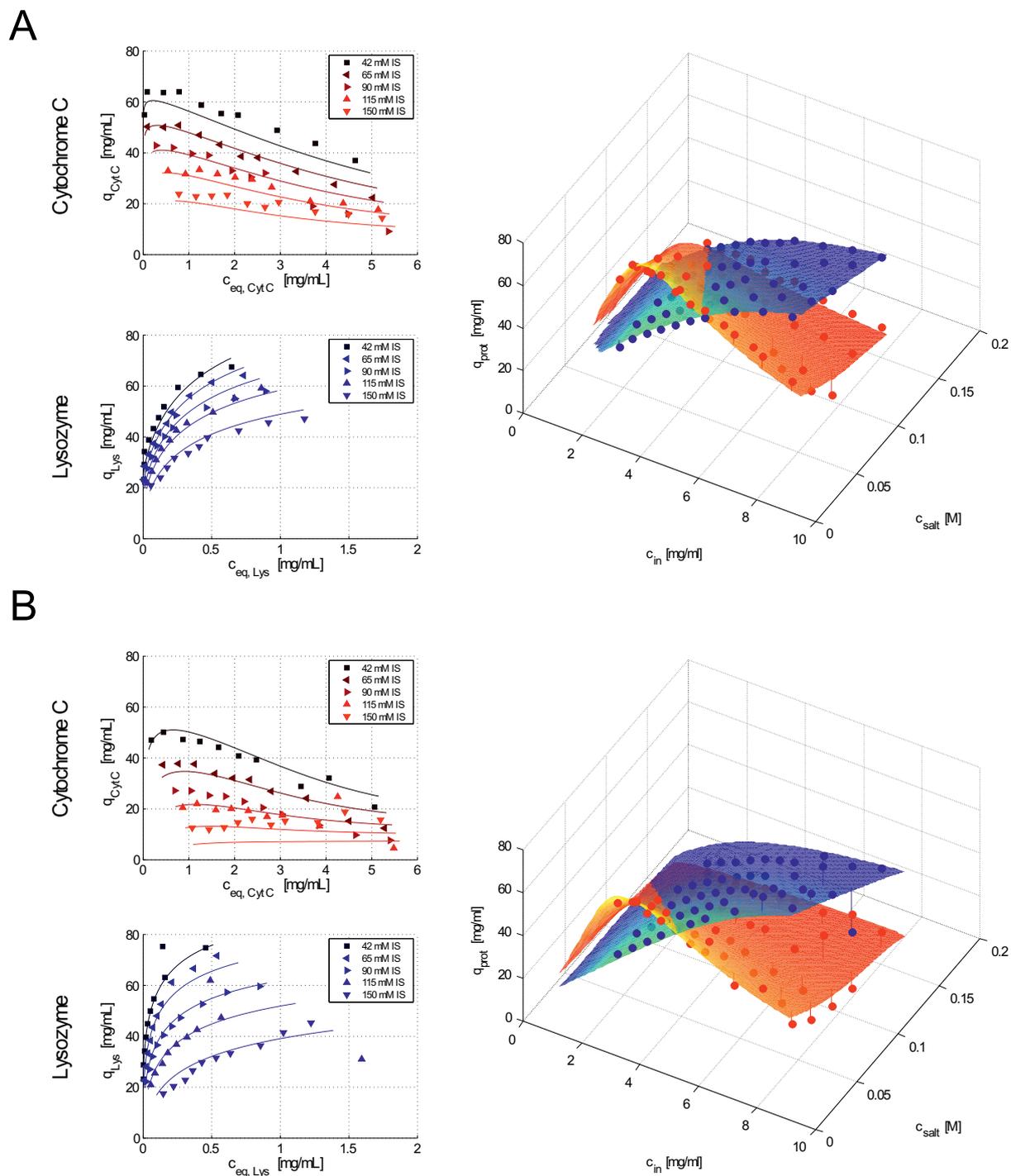


Figure 3: Isotherm data derived from the MVDA model-based equilibrium concentrations for both cytochrome c (red) and lysozyme (blue) and the corresponding SMA isotherm fitting curves. The experimental results at pH 5 (A) and pH 7 (B) are illustrated for all investigated ionic strengths in a range of 42 mM to 150 mM. The experimental data points are highlighted as markers and the SMA isotherm fitting curves are illustrated as solid lines for 2D plots and as planes for 3D plots.

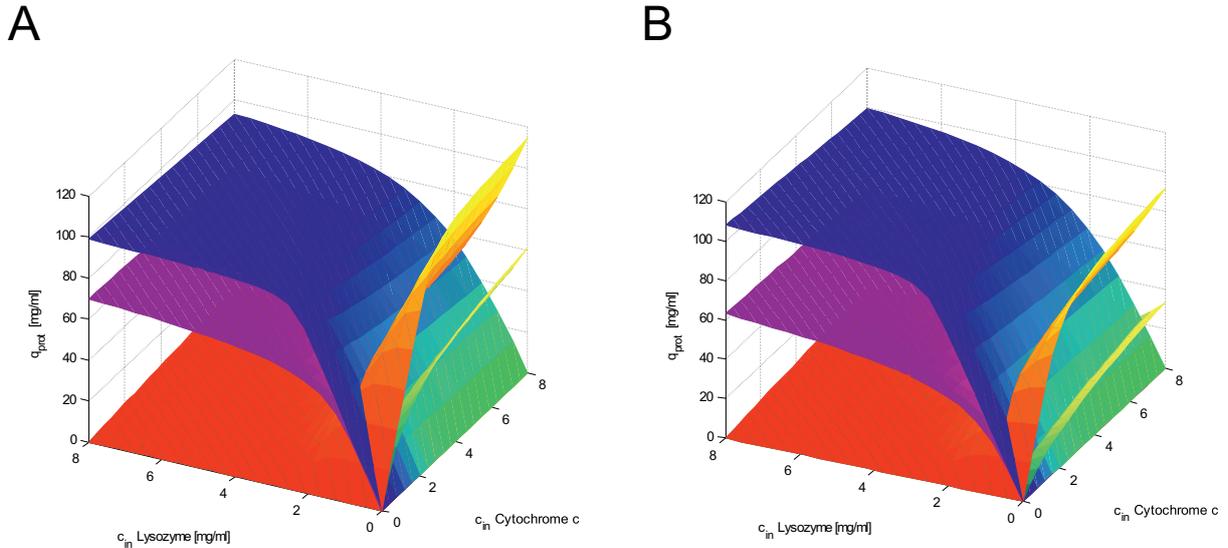


Figure 4: Predicted SMA model for different concentrations and mixing ratios of cytochrome c and lysozyme at pH 5 (A) and pH 7 (B). Lysozyme is shown in blue (42 mM IS) and pink (150 mM IS), cytochrome c is illustrated in red (42 mM IS) and yellow (150 mM).

3.3 SMA Data Fitting

For SMA data fitting, the total ionic capacity Λ for the adsorbent plaques was estimated first. A packed 1 mL of SP Sepharose FF has a total ionic capacity of 800 mM [19]. The calculation of the equivalent Λ for the adsorbent plaques being 0.504 mM followed the description by Hermann et al. using the packed-bed porosity factor [16]. The data fitting was performed in Matlab[®]R2011a using q_i and $c_{eq,i}$ being the concentration of the protein i adsorbed and in solution inside the pore as input variables according to Eqs. (1) - (2). The corresponding SMA fitting functions are illustrated in Fig. 2 by solid lines plotted combined with the respective MVDA data points. The fits were found to be of good agreement as the experimental isotherm points are matched well and the competitive binding trends were conserved.

The resulting SMA parameter sets for both proteins under the investigated conditions are listed in Table 1. The lysozyme equilibrium constants $k_{eq,i}$ are much higher (ranging from 0.454 to 0.632) compared to cytochrome c (0.010 to 0.040) for both pH 5 and pH 7. Comparable values and trends were reported by Gallant et al. for both proteins at pH 6 ($k_{eq,lys} = 0.124$, $k_{eq,cytc} = 0.006$) [20]. As shown in Table 1, the characteristic charge ν_i of cytochrome c was determined to be 20% higher than for lysozyme at pH 5 whereas for pH 7, the same trend is shown inversely. ν_i values given by Gallant et al. at pH 6 ($\nu_{lys} = 5.95$, $\nu_{cytc} = 6.15$) as the center point between pH 5 and pH 7 indicated an identical charge characteristic for both proteins agreeing well with this inversion of ν_i from pH 5 to pH 7 shown in this study. Also, the absolute values for ν_i match well. The steric shielding parameters σ_i shown in Table 1 are similar for all proteins under all conditions ranging from 36.877 to 51.871. Those values are in accordance with Gallant et al. for cytochrome c (53.4) and Osberghaus et al. for lysozyme (29.7 - 36.8) and cytochrome c (28.7 - 40.8) [19].

The overall model quality depends on the identifiability of isotherm parameters from the chosen experiments. To determine the range of the design space used for model

calibration, the total concentration of occupied binding sites in equilibrium Λ_{bound} is calculated by:

$$\Lambda_{bound} = \Lambda - \sum_i (\nu_i + \sigma_i) q_i \quad (3)$$

Pairs (q_{cytc}, q_{lys}) can be taken from the experimental data in [mg/mL] and divided by the respective molecular weights. $MW_{cytc} = 12.38$ kDa and $MW_{lys} = 14.30$ kDa were taken from the Sigma Aldrich data sheet and the SMA parameters for pH 5 and pH 7 are given in Table 1. For the lowest salt concentration and highest initial concentration at pH 5, Λ_{bound} adds up to 432.5 mM (166.7 mM cytochrome c, 265.8 mM lysozyme) and for pH 7 to 371.5 mM (69.7 mM cytochrome c, 301.8 mM lysozyme). For the lowest initial concentration, we obtain $\Lambda_{bound,pH5} = 338.7$ mM and $\Lambda_{bound,pH7} = 302.7$ mM. Comparing these values to the total ionic capacity Λ of 504 mM, the loading is in a range of 60% to 86%.

This could be an explanation for the deviation of the equilibrium constants from literature values. As the linear range of the isotherms is not covered by the experiments, the 95% confidence intervals for k_{eq} are the largest ($k_{eq,cytc,pH5}$ 94%, $k_{eq,cytc,pH7}$ 233%, $k_{eq,lys,pH5}$ 68%, $k_{eq,lys,pH7}$ 187%). However, it should be noted that Gallant et al. were investigating cytochrome c and lysozyme at pH 6 and not at pH 5 and 7. Absolute values were thus not comparable but relative trends were conserved. The confidence intervals for the steric shielding parameter were the smallest ($\sigma_{cytc,pH5}$ 13%, $\sigma_{cytc,pH7}$ 34%, $\sigma_{lys,pH5}$ 11%, $\sigma_{lys,pH7}$ 28%) as they can be read from the concentration in the saturated state. For the characteristic charge we determined slightly larger intervals ($\nu_{cytc,pH5}$ 20%, $\nu_{cytc,pH7}$ 42%, $\nu_{lys,pH5}$ 21%, $\nu_{lys,pH7}$ 41%). As the parameters are still correlated, additional experiments could be done to minimize the uncertainty, e.g. following [21]. Nevertheless, it can be concluded that the SMA model is able to reproduce the observed binding behavior.

Finally, the fitted SMA model was plotted for different protein concentration levels c_{in} of both proteins for pH 5 (Fig. 4A) and pH 7 (Fig. 4B). Lysozyme is shown in blue (42 mM IS) and pink (150 mM IS), cytochrome c is illustrated in red (42 mM IS) and yellow (150 mM IS). Data points along the x- and y-axes (0 mg/mL of lysozyme or cytochrome c) are the single component isotherms showing a linear slope for low concentrations and resulting in a maximum (q_{max}) at the adsorbent saturation. The model predicts an almost constant maximal binding capacity at pH 5 and pH 7 at both ISs for lysozyme whereas the binding of cytochrome c is enhanced at pH 5 especially for 150 mM IS compared to pH 7. Furthermore the model in Fig. 4 predicts that already 2 mg/mL of lysozyme present in the mixture suppresses the binding of cytochrome c on the adsorbent for both pH setups. These findings agree with the experimental data which indicated a diminished binding of cytochrome c (Fig. 3) though being the species of excess concentration in the mixture (70% cytochrome c/ 30% lysozyme).

3.4 SMA Parameters vs. Experimental Data

The equilibrium constant $k_{eq,i}$ as a factor describing the binding affinity of proteins under given conditions being over 10-fold higher for lysozyme compared to cytochrome c (compare Table 1) is in accordance with the isotherms given in Fig. 3. The displacement of cytochrome c by lysozyme at higher starting concentrations c_{in} was discussed above.

Protein	pH	$k_{eq,i}$	ν_i	σ_i
Cytochrome c	5	0.010	9.955	45.775
Cytochrome c	7	0.040	4.827	36.877
Lysozyme	5	0.632	7.945	48.278
Lysozyme	7	0.454	5.834	51.871

Table 1: SMA parameters of cytochrome c and lysozyme at pH 5 and pH 7 determined by least squares fitting of the MVDA isotherm data.

Surprisingly, the corresponding characteristic charge ν_i is not the crucial factor driving this phenomenon. E.g., for pH 5, ν_{lys} of 7.945 is lower than ν_{cytc} of 9.955 but still the equilibrium is strongly shifted towards lysozyme. However, the above-discussed effects of increased displacement of cytochrome c at pH 7 (towards the pI of both proteins) agree with the change in characteristic charge from pH 5 to pH 7 which decreases strongly for cytochrome c from 9.955 to 4.827 whereas lysozyme is much less affected. In summary, the equilibrium constant was found to be a factor for indicating displacement phenomena whereas the characteristic charge shows the extent. The shielding factors σ_i being similar for both proteins were expected as lysozyme and cytochrome c are similar in size. Lysozyme as the slightly larger molecule ($MW_{lys} = 14.30$ kDa, $MW_{cytc} = 12.38$ kDa) resulted in marginally higher σ_i values.

4 Conclusions

It was demonstrated that multi-variate data analysis (MVDA) of protein spectra is a straightforward method of determining accurate concentration levels for protein mixtures for HTS applications such as batch isotherms. The data points were of high quality and were thus usable for modeling purposes. Based on the MVDA data, accurate SMA fits were found that coincide with literature values.

Analyzing the presented amount of samples using a standard analytical method, such as analytical chromatography takes approximately 5 min per sample or above. Using the presented MVDA method, in contrast, 10 - 15 s per sample are sufficient. The time requirement per sample is one of the key aspects in the concept of QbD, as the bottleneck of HTS and exploring the design space is shifted to the analytics. Besides this, the major advantage over traditional analytical tools, its non-invasive nature, is very convenient in the context of performing experiments on robotic workstations, making sampling obsolete. The proposed methodology can be used for any multi-component mixtures that do show differences in single-component absorption spectra. Brestrich et al. used the methodology for a distinction between antibody monomers, aggregates and lower molecular weight species though being spectrometrically similar [22]. Such species can be distinguished due to an increase or decrease of exposed amino acid residues and by included scattering effects for larger aggregates. The analytical technology based on MVDA of protein spectra might become a standard tool for robotic workstations and in-line analytics [23]. The technique of model calibration might be extended to more complex protein compositions and proteins which are not available as pure components.

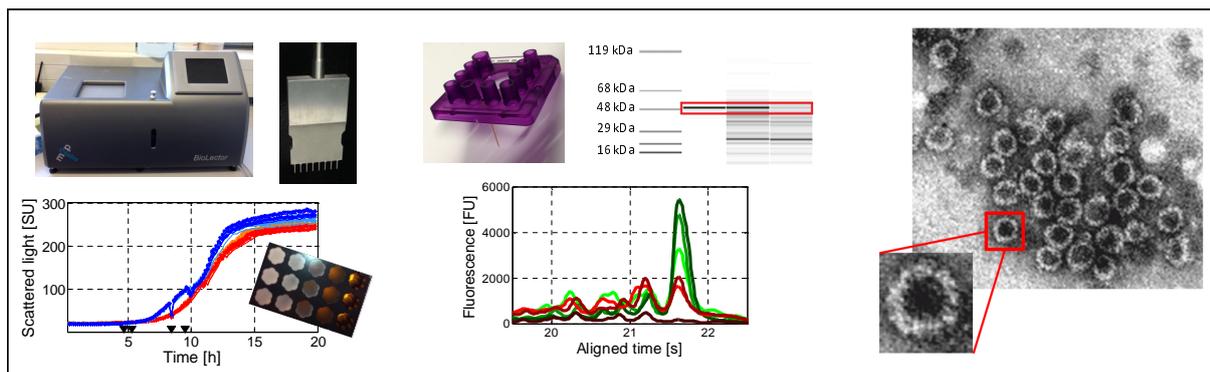
References

- [1] ICH, Ich-endorsed guide for ich q8/q9/q10 implementation, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- [2] S. Chhatre, S. S. Farid, J. Coffman, P. Bird, A. R. Newcombe, N. J. Titchener-Hooker, How implementation of quality by design and advances in biochemical engineering are enabling efficient bioprocess development and manufacture, *J Chem Technol Biotechnol* 86 (2011) 1125–1129.
- [3] B. D. Kelley, M. Switzer, P. Bastek, J. F. Kramarczyk, K. Molnar, T. Yu, J. Coffman, High-throughput screening of chromatographic separations: Iv. ion-exchange, *Biotechnol Bioeng* 100 (2008) 950–963.
- [4] M. Bensch, P. S. Wierling, E. von Lieres, J. Hubbuch, High throughput screening of chromatographic phases for rapid process development, *Chem Eng Technol* 28 (11) (2005) 1274–1284.
- [5] M. Wiendahl, P. Schulze Wierling, J. Nielsen, D. Fomsgaard Christensen, J. Krarup, A. Staby, J. Hubbuch, High throughput screening for the design and optimization of chromatographic processes and miniaturization, automation and parallelization of breakthrough and elution studies, *Chem Eng Technol* 21 (2008) 893–903.
- [6] S. Oelmeier, F. Dimer, J. Hubbuch, Application of an aqueous two-phase systems high-throughput screening method to evaluate mab hcp separation, *Biotechnol Bioeng* 108 (2011) 69–81.
- [7] B. K. Nfor, M. Noverraz, S. Chilamkurthi, P. D. E. M. Verhaert, L. A. M. van der Wielen, M. Ottens, High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents, *J Chrom A* 1217 (2010) 6829–6850.
- [8] B. K. Nfor, J. Ripic, A. von der Padt, M. Jacobs, M. Ottens, Model-based high-throughput process development for chromatographic whey proteins separation, *Biotechnol J* 7 (2012) 1221–1232.
- [9] E. J. Close, J. R. Salm, D. G. Bracewell, E. Sorensen, Modelling of industrial biopharmaceutical multicomponent chromatography, *Chem Eng Res Des* 92 (2014) 1304–1314.
- [10] J. L. Coffman, J. F. Kramarczyk, B. D. Kelley, High-throughput screening of chromatographic separations: I. method development and column modeling, *Biotechnol Bioeng* 100 (2008) 605–618.
- [11] P. Baumann, N. Bluthardt, S. Renner, A. Osberhaus, J. Hubbuch, Integrated Development of Up- and Downstream Processes Supported by the Cherry-Tag for Real-time Tracking of Stability and Solubility of Proteins, *J Biotech* 200 (2015) 27–37.

- [12] B. Maiser, F. Kröner, F. Dimer, G. Brenner-Weiss, J. Hubbuch, Isoform separation and binding site determination of mono-PEGylated lysozyme with pH gradient chromatography., *J Chrom A* 1268 (2012) 102–108.
- [13] M. J. C. Van Lierop, F. Wijnands, Y. W. Loke, P. M. Emmer, H. G. M. Lukassen, D. D. M. Braat, A. van der Meer, S. Mosselman, I. Joosten, Detection of HLA-G by a specific sandwich ELISA using monoclonal antibodies G233 and 56B., *Mol hum reprod* 8 (2002) 776–84.
- [14] S. K. Hansen, E. Skibsted, A. Staby, J. Hubbuch, A label-free methodology for selective protein quantification by means of absorption measurements, *Biotechnol Bioeng* 108 (2011) 2661–2669.
- [15] C. A. Brooks, S. M. Cramer, Steric mass-action ion exchange: Displacement profiles and induced salt gradients, *AIChE Journal* 38 (1992) 1969–1978.
- [16] T. Herrmann, M. Schroeder, J. Hubbuch, Generation of equally sized particle plaques using solid-liquid suspensions, *Biotechnol Prog* 22 (2006) 914–918.
- [17] A. Osberghaus, P. Baumann, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch, Detection, quantification, and propagation of uncertainty in high-throughput experimentation by monte carlo methods, *Chem Eng Technol* 35 (2012) 1456–1464.
- [18] S. R. Dziennik, E. B. Belcher, G. A. Barker, A. M. Lenhoff, Effects of ionic strength on lysozyme uptake rates in cation exchangers. i: Uptake in sp sepharose ff., *Biotechnol Bioeng* 91 (2005) 139–153.
- [19] A. Osberghaus, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch, Determination of parameters for the steric mass action model - a comparison between two approaches, *J Chrom A* 1233 (2012) 54–65.
- [20] S. R. Gallant, A. Kundu, S. M. Cramer, Optimization of step gradient separations: Consideration of noniinear adsorption, *Biotechnol Bioeng* 47 (1995) 355–372.
- [21] T. Barz, V. Loeffler, H. Arellano-Garcia, G. Wozny, Optimal determination of steric mass action model parameters for beta-lactoglobulin using static batch experiments., *J Chromatogr A* 1217 (2010) 4267–4277.
- [22] N. Brestrich, T. Briskot, A. Osberhaus, J. Hubbuch, Advances in inline quantification of co-eluting proteins in chromatography: Process-data-based model calibration and application towards real-life separation issues, *Biotechnol Bioeng* 111 (2015) 1365–1373.
- [23] N. Brestrich, A. Sanden, A. Kraft, K. McCann, J. Bertolini, J. Hubbuch, A tool for selective inline quantification of co-eluting proteins in chromatography using spectral analysis and partial least squares regression, *Biotechnol Bioeng* doi:10.1002/bit.25546.

Establishment of a High-throughput Platform for the Production of Virus-like Particles in *Escherichia coli*

C. Ladd Effio^{1,‡}, P. Baumann^{1,‡}, P. Vormittag¹, C. Weigel¹, A. Middelberg²
and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany

² : Australian Institute for Bioengineering and Nanotechnology, University of Queensland, Brisbane, Australia

[‡] : Contributed equally to this work

^{*} : Corresponding author. *E-mail-address*:juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

The production of safe vaccines against untreatable or new diseases has pushed the research in the field of virus-like particles (VLPs). Currently, a large number of commercial VLP-based human vaccines and vaccine candidates are available or under development. A promising VLP production route is the controlled in vitro assembly of virus proteins into capsids with antigen epitopes on their surface. A rapid and economic production host for industrial processes is *Escherichia coli*.

In the study reported here a high-throughput (HTS) platform for the production and characterization of VLPs in bacterial cell systems was established. High-throughput cultivations were carried out in 48-well format in the BioLector[®] system (m2p-Labs, Germany) using an *Escherichia coli* strain with a tac promoter producing the murine polyomavirus capsid protein (VP1). The screening covered different cultivation temperatures and shaking speeds, as well as varying induction conditions and additives. An effective cell disruption process was established using a HTS sonication device. Capillary gel electrophoresis served as an HTS-compatible analytical tool for scoring the large number of lysates under the cultivation conditions investigated. The most efficient system was selected based on an evaluation of soluble and insoluble product concentrations as well as on the percentage of product in the total soluble protein fraction. The optimized system was scaled up to cultivation 2.5 L shaker flask scale and purified using an anion-exchange chromatography membrane adsorber, followed by a size exclusion chromatography polishing procedure. Purified VP1 capsomeres were assembled under defined buffer conditions into empty capsids and characterized using transmission electron microscopy (TEM). The presented HTS platform allowed for a fast development of an efficient production process of VLPs in *E. coli*. Under optimized cultivation conditions, the VP1 product totaled up to 43% of the total soluble protein fraction, yielding 1.63 mg VP1 per mL of applied cultivation medium. The developed production process strongly promotes the murine polyoma-VLP platform, moving towards an industrially feasible technology for new chimeric vaccines.

Keywords: Virus-like Particles, *Escherichia coli*, High-throughput Screening, Micro-scale Cultivation, BioLector[®]

1 Introduction

Virus-like particles (VLPs) represent novel molecular means for the containment of infectious diseases and the immunotherapeutic treatment of cancer, Alzheimer's disease, and autoimmune diseases ([1, 2, 3, 4]. A steadily increasing number of VLP-based vaccines is currently undergoing clinical phase studies and several recombinant VLPs have been licensed for prophylactic vaccination against cervical cancer (Gardasil[®], Cervarix[®]), hepatitis B (Recombivax[®], Engerix[®]), and hepatitis E (Hecolin[®]) [5, 6]. Protein-based VLPs are produced by recombinant expression of virus proteins in yeast, insect, or bacteria cells [7]. For the production of homogeneous particles, currently licensed VLP-based vaccines are assembled in vitro under defined and controllable buffer conditions. Separate disassembly process steps are needed for in vivo assembled VLPs. A promising approach to the rapid and scalable production of VLPs is the expression of virus proteins in *Es-*

cherichia coli. *E. coli* is the recombinant system of choice for multiple biopharmaceutical products, such as human insulin [8], human growth hormones [9], antibody fragments ([10], or hepatitis vaccines [11], providing high product titers and total soluble protein values of up to 50% [12].

In recent years, several microbial VLP production systems were developed and tailored to epitope presentations in vaccine formulations, such as the bacteriophage Q β system [4, 13], the human hepatitis B core protein [14, 15, 16], the woodchuck hepadnavirus core protein (WHcAg) [17] or the papaya mosaic virus (PapMV) system [18]. Another well-characterized VLP-based nanocarrier is the murine polyoma virus protein VP1 that is capable of spontaneously self-assembling into capsids and lacking the risk of human pre-existing carrier-specific immunity. *E. coli*-derived VP1 builds pentamer structures that allow processing small protein complexes instead of whole particles and, thus, do not require any additional disassembly steps. Chimeric murine polyoma-VLPs were used in the past for the presentation of influenza antigen epitopes (H190, M2e) [19, 20, 21], group A streptococcus epitopes (J8i) [22, 23], tumor cell epitopes [24, 25], and packaging of small molecules and DNA [26, 27]. However, the current state-of-the-art platform for *E. coli*-derived VP1 is limited by the expression and processing of glutathion-S-transferase (GST)-VP1 fusion tag complexes [28]. These protein tags are widely used for enhancing the solubility of proteins, increasing product titers, and facilitating lab-scale purification by using affinity chromatography for capturing the product of interest with a high selectivity [29, 30, 19]. Major drawbacks of protein tags are due to multimerization effects, slow and high-priced affinity chromatography media, and the demand for additional downstream procedures for the removal of the fusion tags [28, 31]. Boosting the productivity of the murine polyoma-VLP platform requires an optimization, acceleration, and simplification of both upstream and downstream processing.

An optimization of protein expression in *E. coli* often is time-intensive when performing shaker flask cultivations. Rapid process development can be performed *in silico* using molecular dynamics simulations of the product of interest [32] or by high-throughput experimentation with multiple recently developed tools for miniaturized automated cultivations [33, 34, 35]. However, miniaturization may result in a reduction of accessible process information, leading to difficulties in process control and scale-up [33]. To overcome these shortcomings, more sophisticated systems were developed, e.g. the Advanced Micro-scale Bioreactors (ambr[®]) technology (Sartorius Stedim Biotech, Germany) mimicking lab-scale fermenters on a robotic workstation [36] and the BioLector[®] system (m2p-Labs, Germany) used in this study [37, 38]. For the BioLector[®] system, special 48-well cultivation plates with flower-shaped wells exist for improving mixing and oxygen uptake into the medium [39]. Successful scale-up from the micro-scale to lab-scale can be implemented to gain a large pool of cell broth for subsequent downstream process development [40, 41].

In this paper, we present a high-throughput approach to the design of the upstream process of murine polyoma-VLPs. In addition, we suggest an alternative membrane-based downstream process enabling the rapid large-scale production of future chimeric murine polyoma-VLP-based vaccine candidates.

2 Materials & Methods

2.1 Materials

2.1.1 Disposables

Microbial cells were stored in a CryobankTM strain maintenance kit for microorganisms (Mast Diagnostica, Germany). Cell transformation was carried out in 1 mm polycarbonate electroporation cuvettes (Biozym Scientific GmbH, Germany). Micro-scale cultivation experiments were performed in 48-well FlowerPlates[®] covered with adhesive sterile sealing foil (m2p-Labs, Germany). All pre-cultures were cultivated in 250 mL baffled flasks (Schott, Germany). For scale-up, 2.5 L baffled TUNAIRTM shake flasks (Sigma Aldrich, Germany) were used. Cell lysis and centrifugation procedures were carried out in 300 μ L polypropylene 96-well plates with a conical bottom (VWR, Germany). Sterile filtration of cell lysates prior to chromatography runs was conducted with 0.45 μ m and 0.1 μ m cellulose acetate filters (Sartorius AG, Germany). 3 mL Slide-A-LyzerTM G2 dialysis cassettes with 10 K MWCO cellulose membranes (Life Technologies, USA) were used for the in vitro VLP assembly. For gel electrophoresis (GX) experiments, Novex[®]NuPAGE[®]4 - 12% Bis-Tris (1.0 mm, 12 wells) protein gels were purchased from Life Technologies, Sweden. Capillary gel electrophoresis (GX II) was carried out in an HT Protein Express & Pico LabChip[®] (Perkin Elmer, USA). GX II sample preparation was conducted in skirted 96-well twin.tec[®]PCR plates (Eppendorf, Germany). For western blotting, XCellTM blot sponge pads, filter papers, and a nitrocellulose transfer membrane were purchased from Life Technologies, USA.

2.1.2 Chemicals & Buffers

For the electroporation experiments, super optimal broth (SOB) medium composed of 20 g/L wheat peptone for microbiology (Fluka, Germany), 5 g/L bacteriological yeast extract (Amresco, USA), 10 mM sodium chloride, 2.5 mM potassium chloride (VWR, Germany), and super optimal broth with catabolite repression (SOC) medium, including the ingredients of SOB with additional 10 mM magnesium chloride, 10 mM magnesium sulfate (Merck kGaA, Germany), and 20 mM glucose (VWR, Germany) were used.

For cultivations, a terrific broth (TB) medium was applied, composed of 12 g/L wheat peptone, 24 g/L bacteriological yeast extract, 5 g/L glycerol bidistilled 99.5% (VWR, Germany), and 89 mM potassium dihydrogen phosphate (VWR, Germany). For the selection of transformed cells, antibiotics were added to a final concentration of 100 μ g/mL carbenicillin and 34 μ g/mL chloramphenicol (AppliChem GmbH, Germany). The TB medium was adjusted to pH 6 or pH 7 using potassium hydroxide and for experiments including magnesium, magnesium sulfate heptahydrate (Merck kGaA, Germany) was added to a final concentration of 2 mM. The induction procedure of the pTac promoter was realized with isopropyl- β -D-thiogalactopyranosid (IPTG) from a 1 M stock solution (VWR, Germany).

As a lysis buffer for native protein release (compatible with the GX II system), a 20 mM Tris buffer (Merck KGaA, Germany) adjusted to pH 8 with hydrochloric acid (Merck KGaA, Germany), including 1X SigmaFASTTM protease inhibitor and 5 mM dithiothreitol (DTT) (Sigma Aldrich, Germany), was used. The lysis buffer for purification experi-

ments consisted of 20 mM Tris-HCl (VWR, Germany), 1 mM EDTA disodium dihydrate (Fluka Chemie GmbH, Switzerland), 5% (v/v) glycerol, 1X SigmaFAST™ protease inhibitor, and 5 mM DTT (Sigma Aldrich, Germany). Total protein extraction under denaturing conditions was carried out using trichloroacetic acid (TCA) BioChemica (AppliChem, Germany) and acetone for liquid chromatography (Merck KGaA, Germany).

The compositions of the binding and elution buffer for anion-exchange chromatography were 20 mM Tris (pH 8), 5 mM DTT, 1 mM EDTA, and 5% (v/v) glycerol, with an additional 1 M NaCl (Merck KGaA, Germany) in the elution buffer. 20 mM Tris (pH 8), 5 mM DTT, 1 mM EDTA, 5% (v/v) glycerol, and 250 mM NaCl were used as running buffer for purification by size-exclusion chromatography. The VLP assembly buffer developed by [19] was used, which consists of 0.5 M ammonium sulfate, 20 mM Tris (pH 7.4), 5% (v/v) glycerol, and 1 mM calcium chloride (VWR, Germany).

Gel electrophoresis (GX) was carried out using 20X Bolt®MES SDS running buffer, NuPAGE®4X LDS sample buffer, and the Novex®Sharp unstained protein standard (Life Technologies, Sweden). DTT was prepared as a 1 M stock solution. For protein staining, a 'blue silver' staining solution, consisting of 10% (v/v) phosphoric acid of 85% purity (Roth, Germany), 100 g/L ammonium sulfate BioChemica (AppliChem, Germany), 1.2 g/L Coomassie Brilliant Blue G-250 (Merck, Germany), and 20% (v/v) methanol (Sigma Aldrich, Germany), was used. For capillary gel electrophoresis (GX II) experiments, an HT protein express reagent kit was purchased from Perkin Elmer, USA. A 1 mg/mL lysozyme solution (Hampton research, USA) served as an internal concentration standard.

The transfer buffer for western blotting was composed of 192 mM glycine, 20% (v/v) methanol (Sigma Aldrich, Germany), and 25 mM Tris(hydroxymethyl)-aminomethane (Merck KGaA, Germany) and was adjusted to pH 8.3. As further buffers, Tris-buffered saline (TBS), including 500 mM chloride and 20 mM Tris adjusted to pH 7.5, and TBS-T with additional 0.05% (v/v) of Tween 20 were applied. The solution for color development (AP color development reagent kit) was purchased from BioRad Laboratories, USA.

2.1.3 Instrumentation & Software

Transformations were conducted using a MicroPulser™ Electroporator (Bio-Rad Laboratories, USA). High-throughput micro-scale cultivation experiments were carried out in a BioLector®MB micro-scale fermentation system run with the BioLectio®HMI & analysis software (m2p-Labs, Germany). Shake flask cultivations were conducted in a MaxQ™ 6000 incubator (Thermo Fisher Scientific, USA). A 5810 R centrifuge (Eppendorf, Germany) was used for liquid-solid separation procedures, including cell harvest. Release of the intracellular product was realized using a Model 120 Sonic Dismembrator equipped with an eight-tip horn positioner (Thermo Fisher Scientific, USA) for high-throughput micro-scale cell disruption and with a Digital Sonifier® 450 (Branson Ultrasonic Corporation, USA) for the scale-up cell lysis procedure. Product purification experiments were carried out in an ÄKTA™ Purifier system (GE Healthcare Life Sciences, Sweden), equipped with a pump P-900, mixer M-925, UV detector UV-900, motor valve INV-907, pH and conductivity monitoring unit pH/C-900, and a fraction collector Frac-950 unit. The FPLC system was controlled using Unicorn 5.2 (GE Health-

care Life Sciences, Sweden). For pH adjustment of all buffers, an HI-3220 pH meter (Hanna Instruments, USA) was used. Gel electrophoresis was carried out in an Xcell SureLock™ Novex® MiniCell gel chamber equipped with a PowerEase® 500 power supply (Life Technologies, USA). Capillary gel electrophoresis (GX II) was carried out in a Caliper LabChip® GX II run with the LabChip® GX 3.1 software (Perkin Elmer, USA). Sample denaturation for both the GX and GX II system was realized in a MUR 13 thermo-shaker with additional lid heating (HLC BioTech, Germany). Matlab® R2011a (MathWorks, USA) served for data processing and creation of figures. Assembled virus-like particles were inspected by transmission electron microscopy (TEM) on a CM 200 FEG/ST electron microscope (Philips, Netherlands).

2.2 Upstream Process Development

2.2.1 Plasmid Construction & Host Strain

The plasmid pALVP1TAC [42] was generously provided by Prof. Robert Garcea (University of Colorado, USA) and sequenced by JenaGen GmbH, Germany. The plasmid pTacVP1 was constructed by inserting the nucleic acid sequence of murine polyoma virus capsid protein VP1 (sequence M34958) with optimized codon usage (Fig. 1) between the NdeI and HindIII sites of a pTac-MAT-Tag-1 expression vector (Sigma-Aldrich, USA). A stop codon was inserted in front of HindIII to prevent the translation of the N-terminal metal affinity tag. The plasmid was designed and synthesized by Centic Biotec, Germany. *E. coli* Rosetta(DE3)pLysS cells (Merck KGaA, Germany) were prepared for electroporation by the standard procedure: Growth in SOB medium (250 mL shaker flasks, 37°C, 180 rpm), harvest at an OD_{600 nm}-value of 0.3 by centrifugation (4°C, 12000 rpm), followed by several cooling and washing steps with sterile bi-distilled water. Transformation was carried out in a MicroPulser™ Electroporator at 1.8 kV with 100 µL cells and 100 ng DNA. Transformed cells containing pTacVP1 with the ampicillin resistance gene were selected on carbenicillin agar plates.

2.2.2 Micro-scale Cultivations

The pre-cultures for the micro-scale cultivations were prepared as 80 mL cultures in TB medium using *E. coli* Rosetta pTacVP1 from a Cryobank™ strain maintenance kit cryo culture in the MaxQ™ 6000 incubator (180 rpm - 37 °C). 4 different pre-cultures were employed, applying TB medium of variable composition (pH 6 - w or w/o 2 mM magnesium sulfate and pH 7 - w or w/o 2 mM magnesium sulfate). After 16 h (late exponential phase), the cells were diluted to a final OD_{600 nm} of 0.1 AU using the different TB media (equally treated pre-culture for all experiments). The working volume of the 48-well FlowerPlate® was set to 1 mL for all cultivation experiments. The cultivation plate was sealed with an adhesive gas-permeable sterile membrane.

The micro-scale cultivation experiments were performed at shaking speeds of 600 rpm or 1200 rpm, covering a temperature set of 27, 32, and 37 °C. Per 48-well FlowerPlate®, each cultivation experiment was performed in triplicate, allowing for a total of 16 conditions to be screened (combined with all shaking speeds and temperatures, this adds up to 16·6 = 96 conditions). These 16 conditions covered 2 different pH values (pH 6 and pH 7), 2 different additive setups (w and w/o 2 mM magnesium sulfate), 2 different

induction times ($OD_{600\text{ nm}}$ of 0.5 AU and 4 AU), as well as 2 different inducer concentrations (0.1 mM and 0.5 mM of IPTG) as a full factorial experimental design. The respective induction times, being in scattered light units in the BioLector[®] system, were calculated using scattered light - $OD_{600\text{ nm}}$ correlation functions (600 rpm: Scattered light = $11.69 \cdot OD_{600\text{ nm}} + 17.59$; 1200 rpm: Scattered light = $11.36 \cdot OD_{600\text{ nm}} + 16.14$). After 20 h, the cultivation procedure was stopped. 50 μL of each well of the FlowerPlate[®] were transferred to 300 μL polypropylene 96-well plates with a conical bottom for TCA total protein extraction under denaturing conditions. Additionally, 150 μL of cell broth were transferred to another 300 μL polypropylene 96-well plate for native cell disruption by sonication. All plates for protein release were centrifuged at 4000 rpm and 10 °C for 30 min and the supernatant was discarded prior to freezing at -20 °C.

2.2.3 High-throughput Cell Disruption

For total protein extraction under denaturing conditions using a TCA protocol, the cell pellets were resuspended in 200 μL of ultrapure water. Then, 10 μL of 100% (w/v) TCA solution were added to each well and the plate was shaken for 20 s and subsequently stored on ice for 10 min. The plate was then centrifuged at 4000 rpm and 4 °C for 10 min and the supernatant was discarded. The pellet was washed twice with 200 μL of cold pure acetone, followed by another centrifugation cycle after each washing step. In a final step the supernatant was removed, the pellet was dried for 3 - 4 h and stored at -20 °C until analysis.

For HTS sonication experiments using the Model 120 Sonic Dismembrator (Fisher Scientific), the cell pellets were resuspended in 150 μL of lysis buffer for native protein release (compatible with the GX II system). The sonication device was operated with an amplitude of 70%, applying 6 cycles of 20 s pulse duration. Between the pulses, 30 s chilling on ice followed. After the final sonication cycle, the plates were centrifuged twice at 4 °C and 4000 rpm for 30 min. The supernatant was transferred to another 96-well plate and stored at -20 °C until analysis in the GX II system.

2.2.4 Scale-up Culture

A scale-up culture from the 1 mL micro-scale cultivation yielding highest VP1 titers to a 2.5 L shake flask culture was realized by keeping the oxygen transfer rate constant, as shown by [43]. The corresponding shaking frequency for scale-up was taken from [41]. The cultivation was performed at 180 rpm shaking speed and 37 °C. The TB medium at pH 6 included 2 mM magnesium sulfate and the induction was carried out at $OD_{600\text{ nm}}$ of 0.5 AU with 0.5 mM IPTG. 20 h after inoculation, the cells were centrifuged at 4000 rpm and 4 °C for 30 min. The supernatant was discarded and the cell pellet was frozen at -80 °C until cell disruption. Product release under native conditions was performed by sonication using a Digital Sonifier[®] 450 equipped with a 1/2" extension cylindrical sonication probe. The cell pellets from 400 mL culture were suspended in 20 mL native lysis buffer for purification experiments. The sonication procedure was carried out at an amplitude of 70% applying 8 cycles of 15 s pulse duration (each cycle followed by 30 s chilling on ice). The lysates were centrifuged twice at 10 °C and 12000 rpm for 30 min, followed by a 0.45 μm filtration step using PES filters by Sartorius, Germany. The lysates were stored at -20 °C until the purification procedure.

2.3 Downstream Process Development

VP1 expressed in *E. coli* Rosetta cells was purified by anion-exchange (AEX) membrane chromatography and size-exclusion chromatography (SEC). A clarified *E. coli* lysate (1 mL) was loaded onto a 3 mL Sartobind[®]Q Nano Membrane Capsule (Sartorius Stedim Biotech GmbH, Goettingen, Germany) at 3 mL/min and purified by applying a salt step elution with NaCl. Pooled VP1 fractions of the AEX eluate were further processed on a Superose 6[®]Increase 10/300 GL column (Ge Healthcare, Uppsala, Sweden) at a flow rate of 0.5 mL/min. Purified VP1 was finally assembled into VLPs by dialysis into the assembly buffer, followed by dialysis against PBS for 16 h prior to analysis by transmission electron microscopy (TEM).

2.4 Analytical Methods

2.4.1 SDS-PAGE & Western Blot

For gel electrophoresis, the TCA pellet from the total protein extraction protocol was dissolved in 1X NuPAGE[®]LDS sample buffer, including 12.5 mM DTT for gel electrophoresis ($OD_{600nm} \cdot 200 \mu\text{L}$), for 2 h in an overhead shaker. For native samples in solution, the sample preparation was performed according to the Novex[®]NuPAGE[®]4 - 12% Bis-Tris protein gel manual [44]. Purified VP1 derived from the GST-tag process developed by [19] served as reference standard. All samples were denatured at 100 °C and 300 rpm for 15 min on a plate shaker with lid heating. All following steps were carried out as described in the user manual. For each set of samples, two distinct gels were created, one for western blotting without staining and another stained with 'blue silver' staining solution.

Proteins were transferred from SDS-PAGE gels to a nitrocellulose membrane at 30 kV for 1 h using the XCell II[™] blot module from Life Technologies, USA. Washing, blocking with 5% BSA, and incubation with primary and secondary antibodies were performed according to the user manual [45]. Rabbit-derived anti-VP1 antibody was a gift of Prof. Robert Garcea [42] and was used as primary antibody. Alkaline phosphatase affinity-purified donkey anti-rabbit IgG (Jackson ImmunoResearch Laboratories, USA) served as secondary antibody. Product detection was carried out using alkaline phosphatase color reagents (BioRad Laboratories, USA). All gels and membranes were analyzed in a Bio-5000 gel scanner.

2.4.2 Capillary Gel Electrophoresis

For quantitative determination of purities and VP1 concentrations, all samples were analyzed in a LabChip[®]GX II capillary gel electrophoresis device as duplicates with an HT Protein Express LabChip[®] kit. Purified VP1 derived from the GST-VP1 process developed by [19] served as reference standard. The sample and chip preparation procedures for liquid samples were performed as described in the manufacturer's protocol for the HT Protein Express Assay [46]. In contrast to this, the dried samples from the TCA extraction were dissolved in 100 μL HT Protein Express sample buffer, including 34 mM of DTT. All further steps followed the above-mentioned protocol for liquid samples. Sample analysis was carried out using the HT Protein Express 200 assay in the LabChip[®]GX 3.1

software. Product quantification was based on peak-baseline integration of the product peak and comparison to a lysozyme standard of 1 mg/mL.

2.4.3 Transmission Electron Microscopy

Visualization and characterization of assembled virus-like particles were carried out by transmission electron microscopy (TEM) on a CM 200 FEG/ST electron microscope. Sample preparation was performed as described previously by [47].

3 Results & Discussion

3.1 Codon Optimization & Product Detection

The expression of virus proteins in *E. coli* depends on various parameters, such as strain and vector design, promoter, cultivation and induction conditions, as well as on codon usage. First cultures with *E. coli* Rosetta(DE3)pLysS cells and the pALVP1TAC [42] plasmid showed very poor yields below 1% of total soluble protein (TSP) (data not shown) following cultivation protocols by [30]. [30] demonstrated the presence of several rare codons in the pALVP1TAC-derived insert of pGexVP1. Hence, we sequenced pALVP1TAC and looked for under-represented codons in the *E. coli* genome. According to [48], [49], and [50], codons which are considered rare in *E. coli* are AGG, AGA, CGG, and CGA encoding for arginine, GGA encoding for glycine, ATA encoding for isoleucine, CTA encoding for leucine, and CCC encoding for proline. As shown in Fig. 1, seven of these rare codons are represented in the insert of the VP1 sequence.

Aiming at a higher protein expression, the sequence was codon-optimized for *E. coli* by Centic Biotec (Heidelberg, Germany), obtaining a slightly different insert sequence than [30] (85% identity). Subsequently, the codon-optimized DNA sequence was inserted into a vector with a tac promoter and transformed into *E. coli* Rosetta cells (see section 2.2.1). For the proof of concept, expression of the product of interest on the shaker flask scale was evaluated by several analytical methods: SDS-PAGE, western blot, capillary gel electrophoresis, and reversed-phase ultra-high-performance liquid chromatography (data not shown). Fig. 2 shows an overview of the different procedures for the detection of VP1 in *E. coli* cell lysates with and without the constructed pTacVP1 plasmid. A protein ladder and a VP1 standard are included in each of the analytical technologies. Fig. 2A illustrates the VP1 identification using capillary gel electrophoresis, showing virtual gels (left) with VP1 highlighted as a box and as an electropherogram (right). The VP1 standard (red solid line) was detected at 49 kDa (aligned time of 21.7 s). Accordingly, the transformed *E. coli* Rosetta pTacVP1 strain (black solid line) revealed a pronounced peak at the same molecular weight as the VP1 standard, whereas the *E. coli* Rosetta wild type (blue dashed line) shows minor fluorescence at 49 kDa only. In Fig. 2B a stained SDS-PAGE gel is shown with lane 1 representing the VP1 standard highlighted by a red box. Lanes 2 and 3 represent the TCA extracted total protein fraction, lanes 4 and 5 show the soluble protein fraction after cell disruption. Both samples of *E. coli* Rosetta pTacVP1 (lanes 2 and 4) reveal an intense band at the size of the VP1 standard, whereas the *E. coli* Rosetta wild type samples (lanes 3 and 5) show minor bands in this molecular weight region (compare capillary gel electrophoresis results). The product-specific western blot is illustrated in

3 PUBLICATIONS & MANUSCRIPTS

VP1	1	ATG GCC CCC AAA AGA AAA AGC GGC GTC TCT AAA TGC GAG ACA AAA TGT ACA AAG GCC TGT
VP1*	1	ATG GCA CCT AAG CGT AAG AGC GGT GTC TCT AAG TGC GAG ACT AAG TGC ACC AAG GCA TGC
VP1	61	CCA AGA CCC GCA CCC GTT CCC AAA CTG CTT ATT AAA GGG GGT ATG GAG GTG CTG GAC CTT
VP1*	61	CCA CGT CCG GCA CCA GTT CCA AAA CTG CTG ATT AAG GGT GGT ATG GAG GTA CTG GAT CTG
VP1	121	GTG ACA GGG CCA GAC AGT GTG ACA GAA ATA GAA GCT TTT CTG AAC CCC AGA ATG GGG CAG
VP1*	121	GTA ACG GGT CCG GAT TCT GTA ACT GAG ATC GAG GCT TTT CTG AAC CCA CGT ATG GGT CAG
VP1	181	CCA CCC ACC CCT GAA AGC CTA ACA GAG GGA GGG CAA TAC TAT GGT TGG AGC AGA GGG ATT
VP1*	181	CCG CCG ACT CCG GAA TCT CTG ACT GAA GGT GGT CAG TAC TAC GGT TGG TCT CGT GGT ATT
VP1	241	AAT TTG GCT ACA TCA GAT ACA GAG GAT TCC CCA GGA AAT AAT ACA CTT CCC ACA TGG AGT
VP1*	241	AAC CTG GCT ACT AGC GAT ACC GAG GAT TCC CCA GGT AAC AAT ACC CTG CCG ACT TGG TCT
VP1	301	ATG GCA AAG CTC CAG CTT CCC ATG CTC AAT GAG GAC CTC ACC TGT GAC ACC CTA CAA ATG
VP1*	301	ATG GCT AAA CTG CAG CTG CCG ATG CTG AAC GAG GAT CTG ACT TGC GAC ACT CTG CAG ATG
VP1	361	TGG GAG GCA GTC TCA GTG AAA ACC GAG GTG GTG GGC TCT GGC TCA CTG TTA GAT GTG CAT
VP1*	361	TGG GAA GCG GTT TCT GTA AAA ACC GAA GTG GTG GGC TCT GGT TCC CTG CTG GTA CAC
VP1	421	GGG TTC AAC AAA CCC ACA GAT ACA GTA AAC ACA AAA GGA ATT TCC ACT CCA GTG GAA GGC
VP1*	421	GGT TTC AAC AAA CCG ACC GAC ACT GTG AAC ACC AAA GGC ATC TCC ACC CCG GTA GAA GGT
VP1	481	AGC CAA TAT CAT GTG TTT GCT GTG GGC GGG GAA CCG CTT GAC CTC CAG GGA CTT GTG ACA
VP1*	481	AGC CAG TAC CAC GTA TTC GCC GTT GGT GGC GAA CCT CTG GAC CTG CAA GGT CTG GTT ACC
VP1	541	GAT GCC AGA ACA AAA TAC AAG GAA GAA GGG GTA GTA ACA ATC AAA ACA ATC ACA AAG AAG
VP1*	541	GAT GCG CGT ACC AAA TAC AAA GAA GAA GGT GTT GTG ACC ATC AAA ACC ATC ACC AAA AAA
VP1	601	GAC ATG GTC AAC AAA GAC CAA GTC CTG AAT CCA ATT AGC AAG GCC AAG CTG GAT AAG GAC
VP1*	601	GAC ATG GTC AAC AAA GAC CAG GTC CTG AAC CCG ATC AGC AAA GCG AAA CTG GAC AAA GAC
VP1	661	GGA ATG TAT CCA GTT GAA ATC TGG CAT CCA GAT CCA GCA AAA AAT GAG AAC ACA AGG TAC
VP1*	661	GGC ATG TAC CCG GTG GAA ATC TGG CAC CCG GAC CCT GCC AAA AAC GAA AAC ACG CGT TAC
VP1	721	TTT GGC AAT TAC ACT GGA GGC ACA ACA ACT CCA CCC GTC CTG CAG TTC ACA AAC ACC CTG
VP1*	721	TTC GGC AAC TAC ACG GGC GGC ACC ACC ACC CCG CCG GTT CTG CAG TTC ACT AAC ACT CTG
VP1	781	ACA ACT GTG CTC CTA GAT GAA AAT GGA GTT GGG CCC CTC TGT AAA GGA GAG GGC CTA TAC
VP1*	781	ACT ACC GTG CTG CTG GAC GAA AAC GGC GTT GGT CCG CTG TGT AAA GGT GAA GGC CTG TAT
VP1	841	CTC TCC TGT GTA GAT ATA ATG GGC TGG AGA GTT ACA AGA AAC TAT GAT GTC CAT CAC TGG
VP1*	841	CTG TCC TGT GTT GAT ATC ATG GGC TGG CGT GTT ACC CGT AAC TAC GAC GTC CAT CAT TGG
VP1	901	AGA GGG CTT CCC AGA TAT TTC AAA ATC ACC CTG AGA AAA AGA TGG GTC AAA AAT CCC TAT
VP1*	901	CGT GGC CTG CCG CGT TAT TTC AAA ATT ACC CTG CGC AAA CGC TGG GTT AAA AAC CCG TAT
VP1	961	CCC ATG GCC TCC CTC ATA AGT TCC CTT TTC AAC AAC ATG CTC CCC CAA GTG CAG GGC CAA
VP1*	961	CCG ATG GCG TCC CTG ATT AGC TCC CTG TTC AAC AAC ATG CTG CCG CAG GTG CAA GGC CAG
VP1	1021	CCC ATG GAA GGG GAG AAC ACC CAG GTA GAG GAG GTT AGA GTG TAT GAT GGG ACT GAA CCT
VP1*	1021	CCT ATG GAA GGT GAA AAT ACC CAG GTG GAA GAA GTT CGC GTT TAT GAC GGC ACC GAA CCG
VP1	1081	GTA CCG GGG GAC CCT GAT ATG ACG CGC TAT GTT GAC CGC TTT GGA AAA ACA AAG ACT GTA
VP1*	1081	GTG CCG GGC GAT CCG GAT ATG ACG CGC TAT GTT GAC CGC TTT GGC AAA ACC AAA ACG GTT
VP1	1141	TTT CCT GGA AAT TAA 1155
VP1*	1141	TTT CCG GGC AAT TAA 1155

Figure 1: Codon optimization of the VP1 sequence for recombinant protein expression in *E. coli*. The original DNA sequence of VP1 is illustrated above the sequence after codon optimization (VP1*) with rare codons highlighted in red. Unchanged sequences in the VP1 genetic code are indicated by lines.

Fig. 2C with VP1 highlighted. Lanes 1 to 5 correspond to the samples shown in the SDS-PAGE (Fig. 2B), again revealing intense bands like the VP1 standard (lane 1) for denatured and native lysates of *E. coli* Rosetta pTacVP1 (lanes 2 and 4). Lanes 3 and 5 as the *E. coli* Rosetta wild type are insensitive to the product-specific antibody.

Different orthogonal analytical methodologies, namely, capillary gel electrophoresis, SDS-PAGE, western blot, as well as reversed-phase chromatography (data not shown), indicated the presence of VP1 in the codon-optimized *E. coli* Rosetta pTacVP1 strain. As shown in the *E. coli* Rosetta wild type lysate, a minor amount of contaminants was detected with the same molecular weight as VP1. The slight bands of the *E. coli* Rosetta wild type lysate (Fig. 2C, lanes 3 and 5) is explained by cross reactivity of the proteins to the primary antibody for VP1, as reported by [42].

3.2 High-throughput Cultivations

3.2.1 Growth Curves

The high-throughput cultivation experiments covered shaking speeds of 600 rpm and 1200 rpm, temperatures of 27, 32, and 37 °C, two medium pH values of pH 6 and pH 7, two additive setups (w and w/o 2 mM magnesium sulfate), two induction times of $OD_{600\text{ nm}} = 0.5$ AU and 4 AU, as well as two inducer concentrations of 0.1 mM and 0.5 mM IPTG. The growth curves for all conditions investigated are illustrated in Fig. 3 with separated plots for the different combinations of temperature and shaking speeds. In each individual plot, cultivations performed at pH 6 are marked in red and at pH 7 in blue (dark colored conditions include 2 mM magnesium sulfate). Conditions induced at an $OD_{600\text{ nm}}$ of 0.5 AU are illustrated as lines (dashed lines for induction using 0.1 mM IPTG, solid lines for 0.5 mM IPTG), whereas inductions at $OD_{600\text{ nm}}$ of 4 AU are shown as markers (asterisk for induction using 0.1 mM IPTG, circles for 0.5 mM IPTG).

The lag-phase of the *E. coli* Rosetta pTacVP1 cells decreases steadily from 27 °C (6 to 7 h) up to 37 °C (3 h), shifting the induction times (downward triangles) closer to the time of inoculation (0 h). All other factors show almost no impact on the lag-phase, with the medium pH being the only exception. In all setups the lag-phase of cells cultivated at pH 6 is elongated compared to those grown at pH 7, with this effect being most pronounced for the cells cultivated at 27 °C. The increased time until starting the exponential phase for cultivations at lower temperatures is intuitive due to the decrease in metabolic activity of *E. coli* cells at a temperature other than the optimum of 37 °C. The negative effect of slightly acidic pH on the lag-phase is due to *E. coli* having its pH optimum in the neutral region.

The shape of the exponential growth shows highest variations when comparing conditions of different shaking speeds. All cells cultivated at 1200 rpm (right graphs, Fig. 3) display an ideal exponential growth until the stationary phase. In contrast to this, all cultivations performed at 600 rpm result in multiple varying slopes of the exponential growth separated by intermediate plateaus. This effect is especially pronounced for cells cultivated at pH 7. Such behavior was also described by Funke et al. [51] for *E. coli* when accumulating acidic metabolites under fermenting conditions. Lower shaking frequencies (600 rpm) result in poor mixing and low oxygen uptake into the medium inducing such fermenting conditions. Cultivations performed at pH 6 are less influenced as cells grow under 'pseudo-fermenting' conditions starting from inoculation due to the acidic environ-

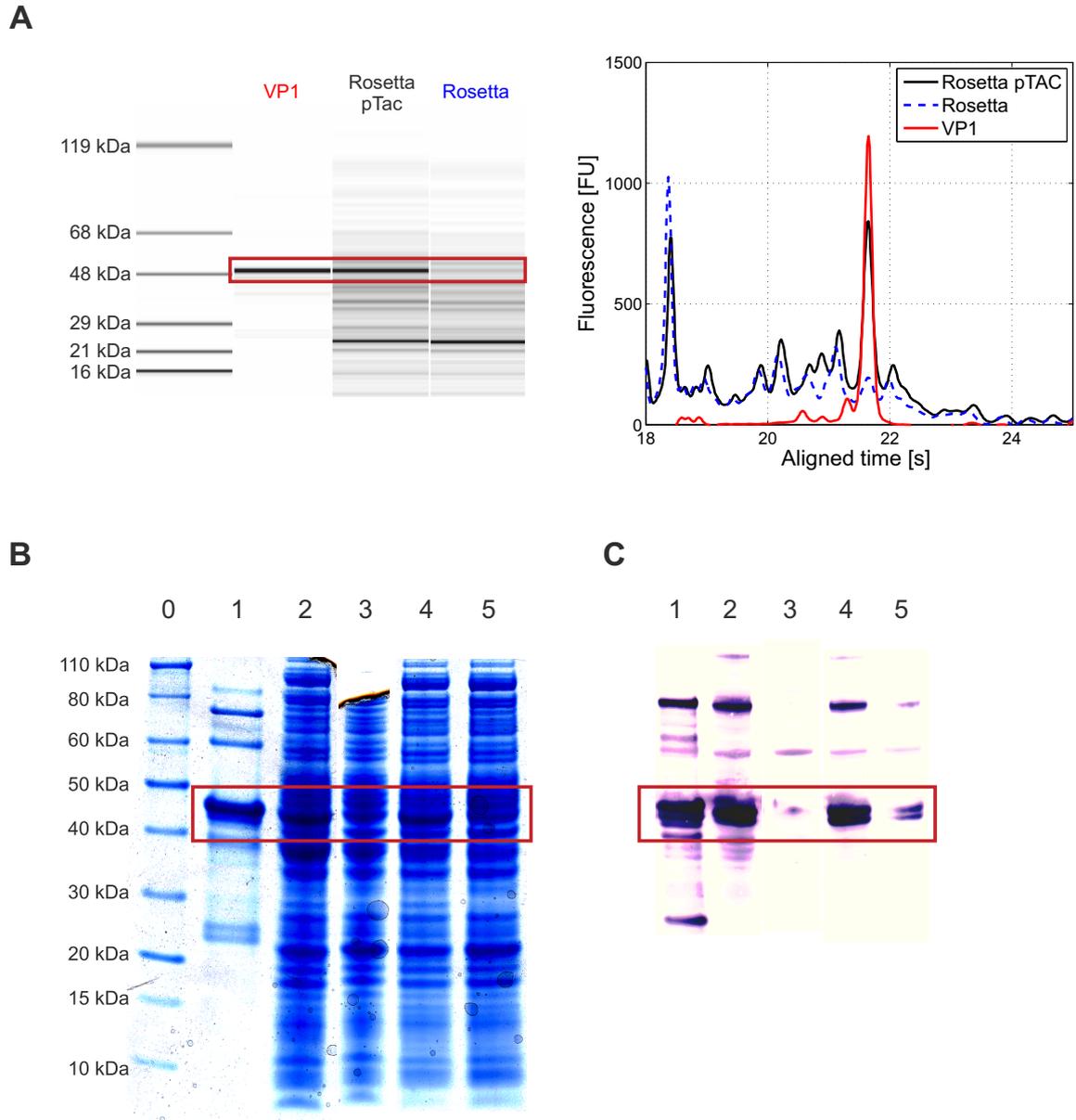


Figure 2: Product (VP1) identification out of a crude feedstock of transformed *E. coli* Rosetta pTacVP1. A: Capillary gel electrophoresis of an *E. coli* Rosetta pTacVP1 lysate (black curve), an *E. coli* Rosetta wild type lysate (blue dashed curve), and a purified VP1 standard (red curve). VP1 is detected at 49 kDa in the virtual gel (left) equivalent to an aligned time of 21.6 s in the electropherogram (right). B: Gel electrophoresis of a purified VP1 standard (lane 1), total protein extracts of *E. coli* Rosetta pTacVP1 and the wild type (lanes 2 + 3), and the soluble protein fraction after sonication of Rosetta pTacVP1 and the wild type (lanes 4 + 5). The protein standard of known molecular weight is shown in lane 0 and the product is highlighted in red at a molecular weight of 46 kDa. C: Western blot of lanes 1 - 5 of the gel electrophoresis run (B) for VP1 identification. The relevant product-related bands are highlighted in red.

ment.

The final plateaus of the stationary phase are elevated for all conditions at a shaking speed of 1200 rpm compared to those at 600 rpm, yielding a maximal scattered light plateau of 285 SU at 27 °C. In comparison, the maximal scattered light signal for the same temperature at 600 rpm is about 40% lower (180 SU). With an increase in temperature up to 37 °C, this effect becomes less pronounced. Whereas the stationary phase plateaus vary significantly for different shaker speeds and temperatures, all other cultivation factors investigated are negligible. As discussed above, higher shaking speeds result in a better oxygen uptake into the medium and, thus, enhance growth under aerobic conditions.

3.2.2 Soluble VP1 Fraction

The soluble protein fraction of cells after sonication and lysate clarification for all conditions investigated is shown in Fig. 4 as 3D bar plots. Again, plots are separated for different combinations of shaking speeds and temperatures. All VP1 concentrations c_{VP1} are normalized to the maximal concentration attained under all conditions $c_{VP1,max}$ of $1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$ (37 °C, 600 rpm, pH 6, including 2 mM magnesium sulfate, induced at $OD_{600 \text{ nm}}$ of 0.5 AU using 0.5 mM IPTG).

All cultivations performed at 1200 rpm (right) resulted in lower VP1 titers compared to those at shaker speeds of 600 rpm (left), with the highest value $38.8\% \pm 4.5\%$ being reached at 27 °C and 1200 rpm compared to the overall maximum $c_{VP1,max}$. A decrease in temperature is slightly beneficial for the formation of soluble VP1 at 1200 rpm, shifting $c_{VP1}/c_{VP1,max}$ from up to $16.4\% \pm 1.9\%$ at 37 °C to $38.8\% \pm 4.5\%$ at 27 °C. Nevertheless, the maximal VP1 concentration investigated at 1200 rpm was below that of most conditions examined for 600 rpm shaking speed, ranging from $c_{VP1}/c_{VP1,max}$ of $30.8\% \pm 4.2\%$ to $100\% \pm 8.2\%$. For experiments performed at 600 rpm, the VP1 concentration in the soluble protein fraction follows an inverse trend towards cultivation temperature, shifting $c_{VP1}/c_{VP1,max}$ from up to $49.2\% \pm 4.9\%$ at 27 °C to $100\% \pm 8.2\%$ at 37 °C. For 32 °C and 37 °C, additional pH effects become apparent, showing a positive effect for cells cultivated under slightly acidic conditions at pH 6 compared to experiments performed in a medium of pH 7. The addition of 2 mM magnesium sulfate did not show pronounced effects in all setups investigated. Also the induction time and inducer concentration were determined to be factors of minor importance.

These findings of high native protein production under fermenting conditions with low oxygen supply and increased temperature agree with the findings of Baumann et al. [41] for glutathione-S-transferase and Losen et al. [52] for benzoylformate decarboxylase. However, they contradict the experiments of Chuan et al. [30], who found an optimal production of 0.18 mg/mL VP1 at a decreased temperature of 26 °C. Note that the maximal VP1 concentration in this study of $1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$ corresponds to an increase of approximately a factor of 10 compared to Chuan et al. [30]. Finally, the pH of the cultivation medium was found to be an important factor for enhancing native VP1 production. The effect of increased soluble protein production under slightly acidic conditions was also reported by Kopetzki et al. [53] for α -glucosidase produced in *E. coli*. As fermenting conditions with acidic by-products have a positive impact on soluble VP1 formation, the decreased pH after inoculation might have enhanced this effect. Whereas the minor influence of the inducer concentration on the product formation agrees with

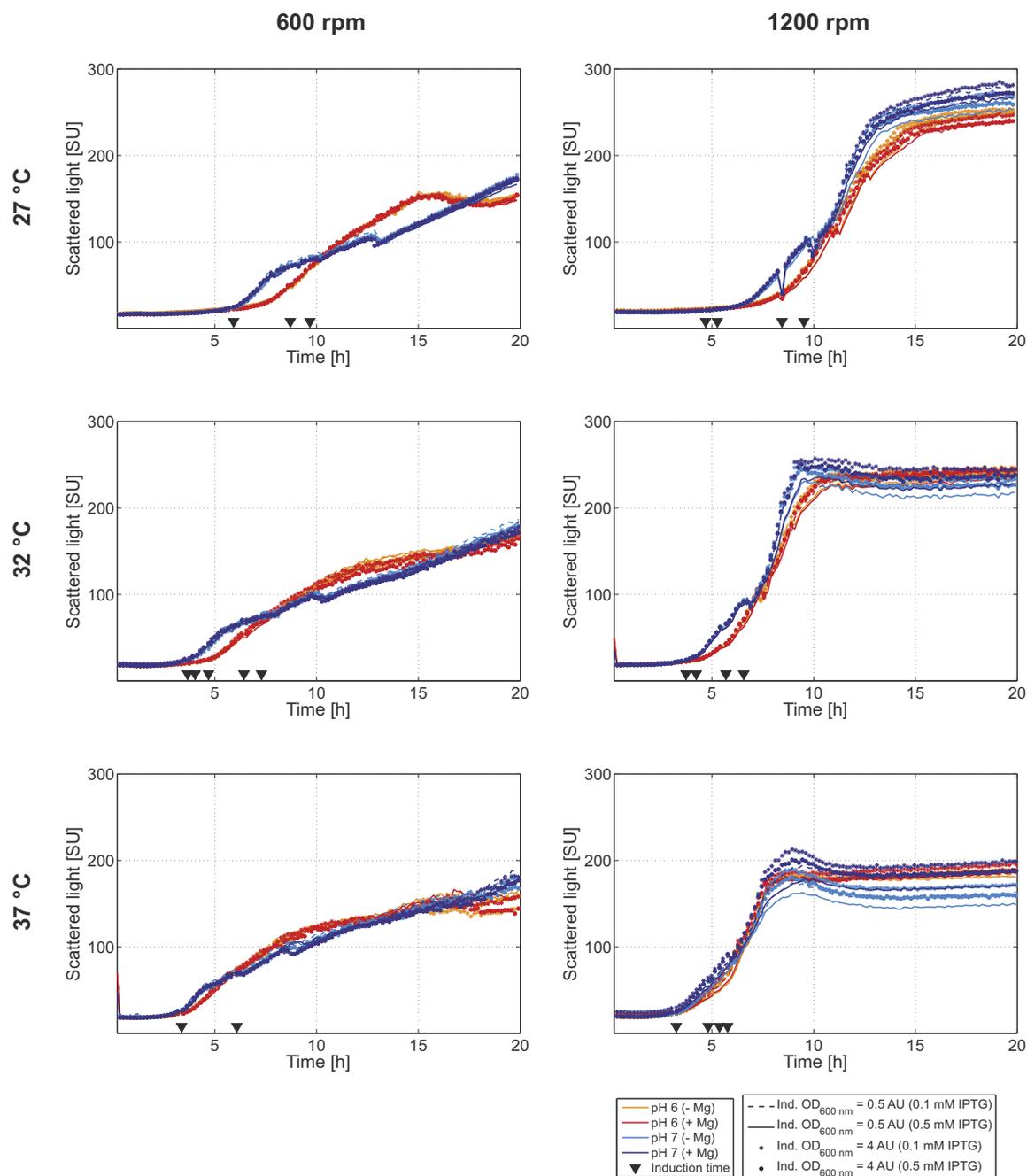


Figure 3: BioLector[®] growth curves of *E. coli* Rosetta pTacVP1 under different cultivation conditions: Temperatures of 27, 32, and 37 °C and two shaking speeds of 600 and 1200 rpm were investigated. For each setup, four distinct media compositions were examined, being two pH values of pH 6 (red curves) and pH 7 (blue curves) with (dark colored) and without (light colored) addition of 2 mM magnesium sulfate. As a final factor, four different induction setups were investigated, being the induction times $OD_{600\text{ nm}}$ of 0.5 (lines) and 4 (markers) using 0.1 mM (dashed line/ asterisk) or 0.5 mM IPTG (solid line/ circle), respectively.

Baumann et al. [41], the minor effect of the induction time investigated in this study is in contradiction, as was also reported by Galloway et al. [54].

When comparing the trends of the growth curves to the soluble VP1 titers, there is a clear correlation of increased productivity of *E. coli* cells under non-optimal growth conditions. All experiments performed at high shaker speed (1200 rpm) resulted in unlimited growth (Fig. 3, right) and, consequently, yielded low titers on soluble VP1 (Fig. 4, right). Fermenting conditions at 600 rpm shaker speed, by contrast, resulted in varying slopes of the exponential growth separated by intermediate plateaus and were found beneficial for soluble VP1 formation with product concentrations of up to $1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$ (Fig. 4, left). Note that excess stress on cell growth, as shown for fermenting conditions (600 rpm) and low cultivation temperatures ($27 \text{ }^\circ\text{C}$), reduced this effect, which might be due to product degradation as a consequence of cell starvation.

3.2.3 Ratio of VP1 to Total Soluble Protein

Besides the overall product titer of soluble VP1, also the initial purity of the sample (ratio of VP1 to total soluble protein) is an important factor for subsequent purification processes. A comparison of soluble VP1 titers and total soluble protein (TSP) is shown in Fig. 5A. The different combinations of temperature and shaker speeds are indicated as different markers, whereas different media compositions are highlighted as colors. All data points were fitted with a linear regression function yielding a coefficient of determination R^2 of 85.6%, indicating that the purity of the samples increased linearly with the amount of produced soluble VP1. All cell lysates derived from cultivations at 1200 rpm are found in the bottom left corner of Fig. 5A (circles, upward triangles, diamonds). 600 rpm cultures, by contrast, are found at higher soluble VP1 and TSP values in Fig. 5A (asterisks, downward triangles, hexagrams) with the overall best result found for the $37 \text{ }^\circ\text{C}$ cultivations (hexagrams), as was discussed earlier. The linear increase of TSP with soluble VP1 indicates a product enrichment without causing additional upward regulation of impurity levels. Fig. 5C illustrates this effect of shaker speed and temperature on the VP1 titer and the TSP for lysates from cultivations equivalent to the optimal system point ($1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$ VP1) by capillary gel electrophoresis runs. The product band in the virtual gel (Fig. 5C, left) at 49 kDa is highlighted (red box). Lanes 1, 3, and 5 show cultivations performed at 600 rpm (27 , 32 , and $37 \text{ }^\circ\text{C}$) whereas lanes 2, 4, and 6 are lysates from 1200 rpm cultivations. In Fig. 5C (right graph) the corresponding electropherograms are shown for the same samples (600 rpm in green and 1200 rpm in red). Both illustrations emphasize that contaminant levels are rather constant, whereas VP1 levels vary strongly, which is most pronounced when comparing $37 \text{ }^\circ\text{C}$ - 600 rpm as the maximum VP1 concentration and $37 \text{ }^\circ\text{C}$ - 1200 rpm as the condition of almost no expressed product.

The overall optimal system point for both soluble VP1 and TSP was found for the optimum identified in the 'Soluble VP1 fraction' section, resulting in a TSP of 42.6% with a concentration of $1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$ soluble VP1, as magnified in Fig. 5B (red leftward triangle). Hence, the system point of $37 \text{ }^\circ\text{C}$, 600 rpm, pH 6, including 2 mM magnesium sulfate, induced at $\text{OD}_{600 \text{ nm}}$ of 0.5 AU using 0.5 mM IPTG was selected for scale-up and downstream process development, as both initial purity and product levels were optimal under the investigated upstream conditions.

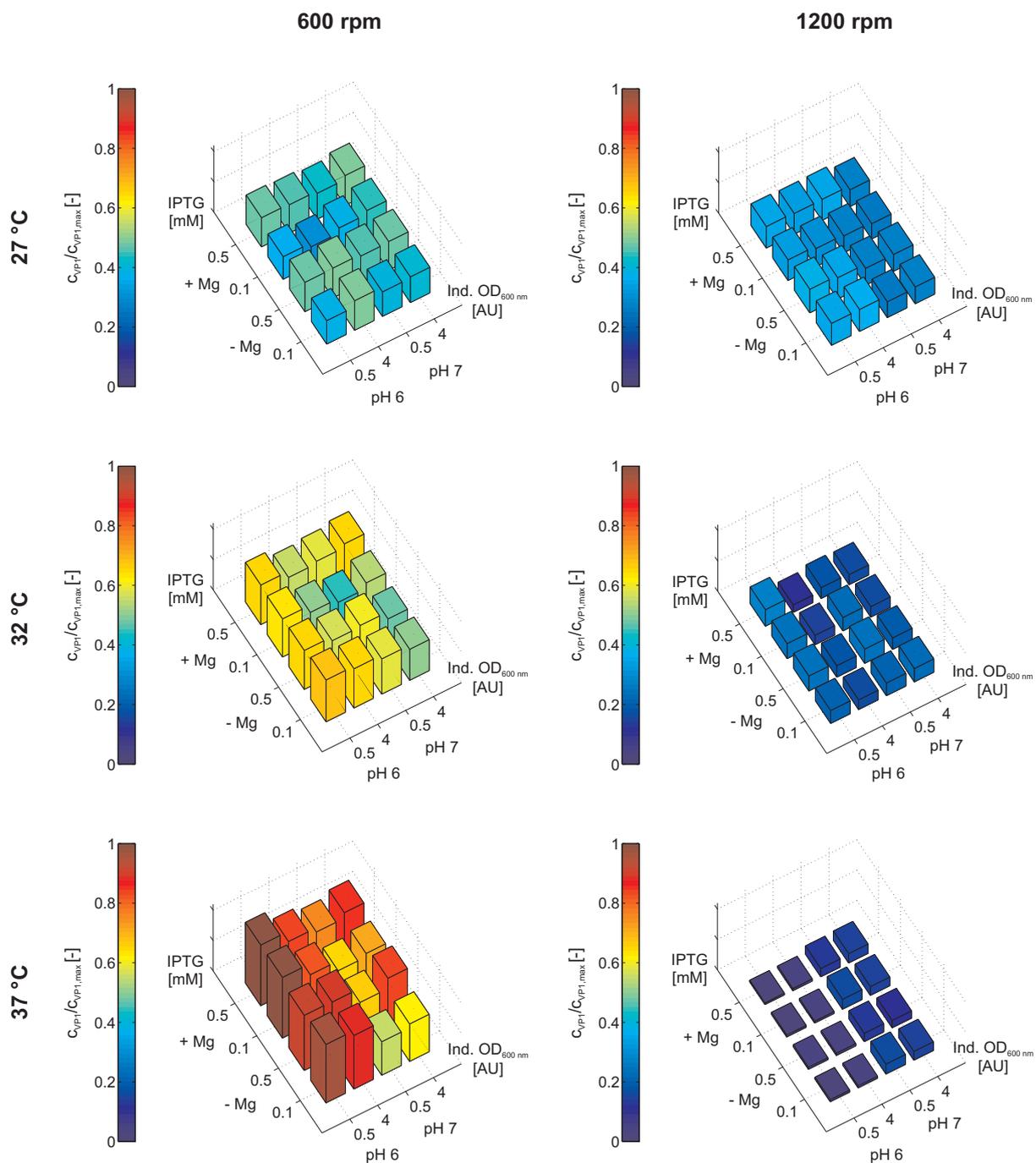


Figure 4: Soluble protein fraction of lysed cells shown in 3D bar plots. Separated plots are shown for different combinations of shaking speeds (600 rpm and 1200 rpm) and temperatures (27, 32, and 37 °C). Results for different media compositions (pH and magnesium sulfate concentrations), induction times, and inducer concentrations are shown for each individual sub-plot. All VP1 concentrations C_{VP1} are normalized to the maximal concentration attained under all conditions $C_{VP1,max}$ being $1.63 \text{ mg/mL} \pm 0.13 \text{ mg/mL}$.

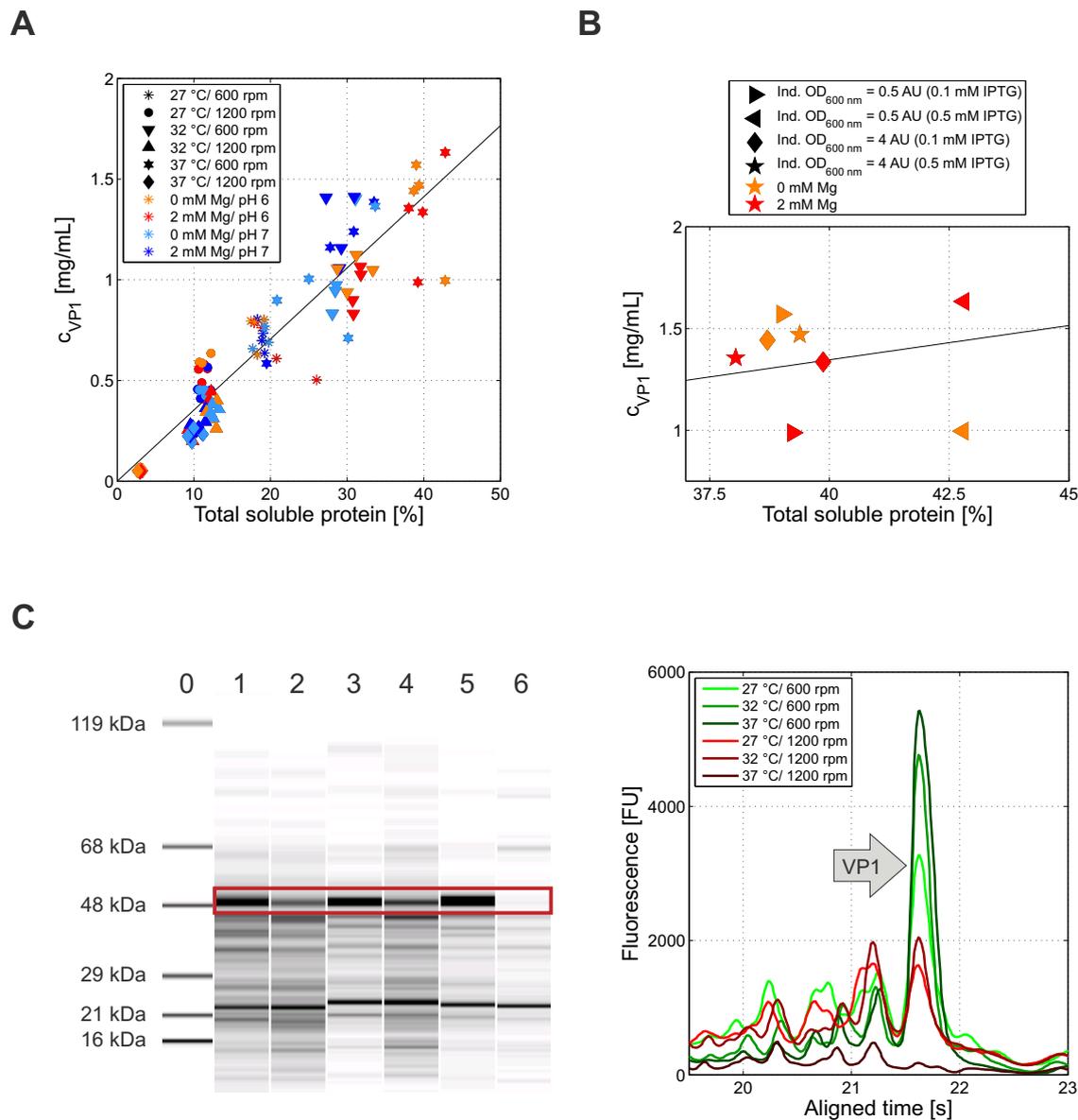


Figure 5: Impurity levels of the different lysates investigated in the BioLector[®] screening. A: Ratio of VP1 to total soluble protein, including a linear regression of data points. The different combinations of temperature and shaker speeds are indicated as different markers and different media compositions are highlighted as colors. B: Ratio of VP1 to total soluble protein, including a linear regression of data points for the optimal system points (cultivation at 37 °C, 600 rpm, and pH 6). The different combinations of induction times and inducer concentrations are indicated as different markers, different levels of magnesium sulfate are highlighted as colors. C: Comparison of the impact of shaker speed and temperature on the VP1 titer and the impurity levels shown for the media conditions of the overall optimal system point (pH 6, including 2 mM magnesium sulfate, induced at OD_{600 nm} of 0.5 AU using 0.5 mM IPTG). The left figure shows virtual gels of these 6 combinations of temperature and shaking speed, with lanes 1, 3, and 5 representing cultivations at 600 rpm (27, 32, and 37 °C) and lanes 2, 4, and 6 for 1200 rpm (27, 32, and 37 °C) and the VP1 band at 49 kDa highlighted in red. The corresponding electropherograms are shown on the right with samples from cultivations at 600 rpm highlighted in green and those of 1200 rpm shown in red.

3.2.4 Ratio of Soluble to Total Expressed VP1

Apart from the soluble VP1 fraction, also the total product fraction, including insoluble species, was determined using a TCA protocol for total protein extraction under denaturing conditions (Fig. 6). Fig. 7 illustrates the ratio of soluble to total expressed VP1 for each setup. For cultivations performed at 1200 rpm, the fraction of soluble VP1 is low, ranging from 4.6% to 39.1% compared to experiments at 600 rpm, where the fraction ranges from 20.9% to 106.1%. The optimal ratio of soluble to total VP1 was obtained for cells grown at 600 rpm and 27 °C, the range being 67.4% to 106.1%. Under these conditions, the neutral medium condition of pH 7 was slightly beneficial compared to pH 6. The same trend was found for cultivations at 600 rpm under 32 °C as well as 37 °C. 32 °C was found to be the condition of the lowest soluble to insoluble VP1 ratio for cells cultivated at 600 rpm.

It was shown that a good mixing and oxygen uptake at 1200 rpm do not only result in poorly soluble VP1 titers as discussed earlier, but also in the formation of aggregates and inclusion bodies. Poor mixing at 600 rpm shaker speed, by contrast, resulted in ratios of soluble VP1 of up to 100%, mostly for conditions of reduced temperature, as was reported in different studies [55, 56, 53]. Note that a factor of 100% does not imply optimal conditions for yielding high overall product titers (compare Fig. 4), but conditions of non-soluble VP1 formation. For the condition of highest productivity of 1.63 mg/mL \pm 0.13 mg/mL VP1, the ratio of soluble product was determined to be 72.8%, indicating an economic system point for scale-up, as was discussed earlier.

3.3 VP1 Purification & Assembly of Virus-like Particles

The development of an alternative production pathway for murine polyoma VLPs without an affinity protein tag entails the establishment of a novel downstream process. Fig. 8 summarizes the outcome of the purification procedure with chromatograms for the capture and polishing step, electropherograms of VP1 process samples, and a TEM micrograph of assembled murine polyoma VLPs.

Capturing of VP1 from clarified *E. coli* cell lysate was performed by anion-exchange membrane chromatography. The chromatogram of the optimized salt step elution is shown in Fig. 8A. UV absorption (blue line), conductivity signals (dashed gray line), and VP1 concentration (green bars) are plotted against the mobile phase volume. Virtual gels for all fractions are plotted below the chromatogram to track the elution of VP1 and HCPs. As shown in the chromatogram of Fig. 8A, VP1 eluted mainly in the second salt step removing weaker and stronger charged molecules in a low salt step at 0.09 M NaCl and a high salt step at 1 M NaCl, respectively. $69 \pm 4\%$ VP1 were recovered in the second salt step with a protein purity of $61 \pm 0.9\%$ in the pooled fractions. Fig. 8B shows the chromatogram of the polishing step by size-exclusion chromatography for VP1 captured by anion-exchange membrane chromatography. The UV signal (blue line) reveals four major peaks. VP1 (green bars) elutes in the major peak at a retention volume of 15 mL, separating smaller host cell impurities (10 to 40 kDa) as displayed in the virtual gels of the FPLC fractions. The final purity of the pooled VP1 fractions was $92 \pm 2\%$ recovering $58 \pm 1.5\%$ of VP1. In Fig. 8C capillary gel electrophoresis analysis of clarified cell lysate (dashed black line), captured VP1 (blue line), and polished VP1 (red line) is compared in an electropherogram. The comparison reveals the high purity achieved at the end of the

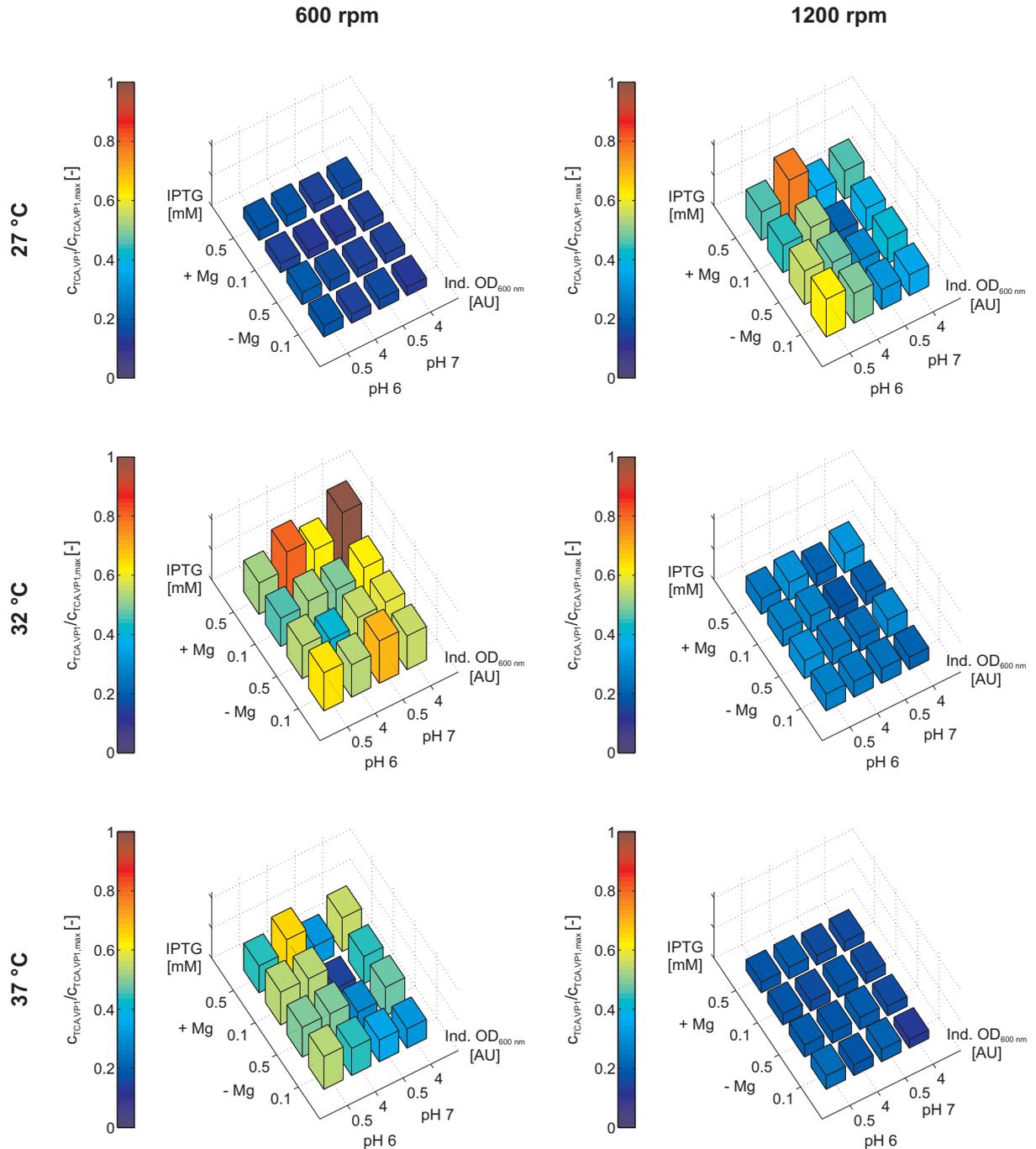


Figure 6: Total (soluble and insoluble) VP1 fraction shown in 3D bar plots. Separated plots are shown for different combinations of shaking speeds (600 rpm and 1200 rpm) and temperatures (27, 32, and 37 °C). Results for different media compositions (pH and magnesium sulfate concentrations), induction times and inducer concentrations are shown for each individual sub plot. All VP1 concentrations $C_{TCA,VP1}$ are normalized to the maximal concentration attained in all conditions $C_{TCA,VP1,max}$.

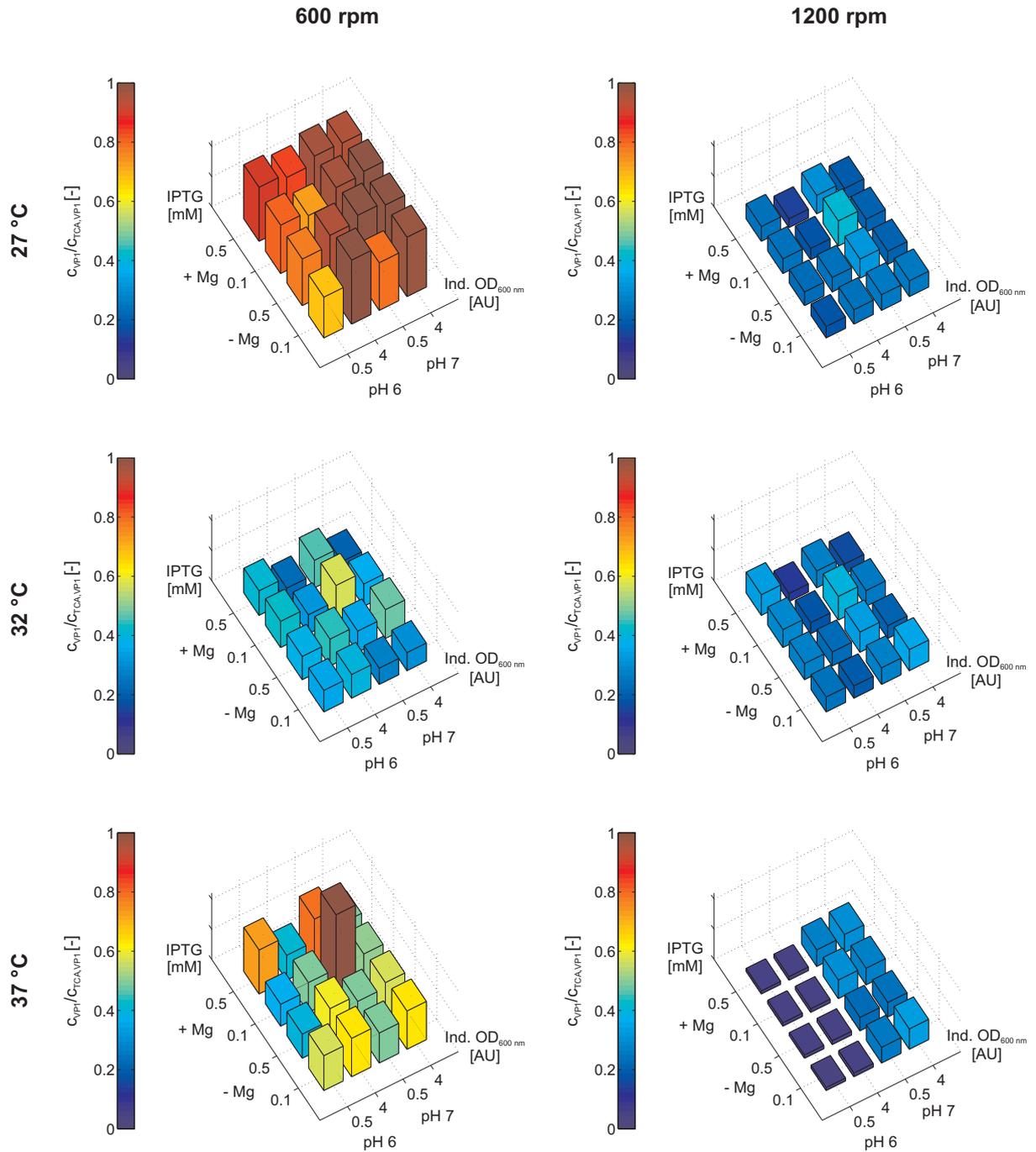


Figure 7: Ratio of soluble to total expressed VP1 of lysed cells shown in 3D bar plots. Separated plots are shown for different combinations of shaking speeds (600 rpm and 1200 rpm) and temperatures (27, 32, and 37 °C). Results for different media compositions (pH and magnesium sulfate concentrations), induction times, and inducer concentrations are shown for each individual sub-plot.

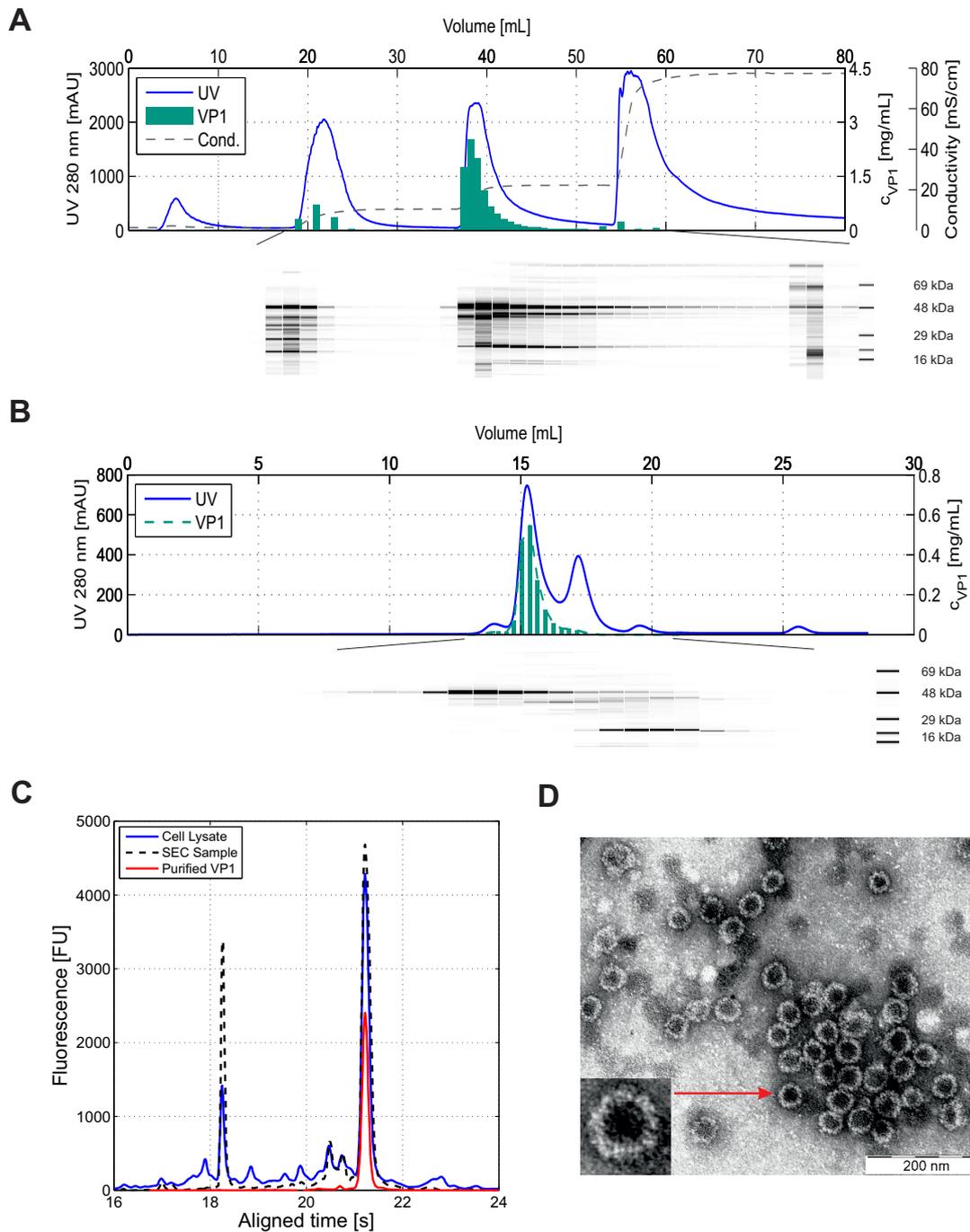


Figure 8: VP1 purification and VLP assembly procedure. A: Chromatogram of the AEX step for VP1 capturing. The UV signal is illustrated as a blue line, the conductivity is shown as a dashed line, and the VP1 concentrations are marked as green bars. The product-containing fractions are magnified in virtual gels of the capillary gel electrophoresis procedure, with the VP1 fraction shown at 49 kDa. B: Chromatogram of the SEC step for VP1 polishing. The UV signal is illustrated as a blue line and the VP1 concentrations are marked as green bars. The product-containing fractions are magnified in virtual gels of the capillary gel electrophoresis procedure with the VP1 fraction shown at 49 kDa. C: Electropherograms of the cell lysate (blue solid line), the pooled elution fraction after AEX (black dashed line), and the purified product after SEC polishing (red solid line). The VP1 peak is detected at 21.6 s of aligned time. D: TEM micrographs of the generated VLP sample after VLP assembly, showing spherical homogeneous particles (diameter of 40 - 50 nm).

two-step downstream process with only one minor impurity left. The cell lysate obtained from the 2.5 L shaker flask-scale fermentation showed almost the same VP1 concentration of 1.5 ± 0.2 mg/mL as in the 1 mL micro-scale cultivation and a similar TSP value of $42 \pm 2\%$.

Purified VP1 at a concentration of 0.3 mg/mL was finally assembled to VLPs, as described earlier by Middelberg et al. [19]. Fig. 8D shows the TEM micrographs of the generated VLP sample, revealing spherical homogeneous particles with diameters of 40 - 50 nm. The empty particles can hardly be distinguished from those murine polyoma VLPs produced in former studies [19, 27, 57]. While the overall recovery of the downstream process can still be optimized, the concept was proven by purifying non GST-tagged VP1 with a rapid and scalable two-step purification procedure, yielding a protein purity close to the desired purity level for recombinant protein-based VLP vaccines of 95 - 99% [6]. Both anion-exchange chromatography and size-exclusion chromatography currently are the methods of choice for the purification of numerous viral or VLP-based vaccines and vaccine candidates [58, 59, 60, 61]. Thus, an industrially common downstream process setup was developed for the processing of VP1 capsomeres.

4 Conclusions

Generating high numbers of vaccine doses in short time and at low cost remains the main challenge for the vaccine industry, especially in the light of pandemic threats and arising pathogens. VLPs represent promising nanocarriers for antigen epitopes of pathogens, but still lack straightforward and easily controllable upstream and downstream procedures.

In this study, we aimed at optimizing both expression and purification methods of the murine polyoma VLP platform by using a high-throughput screening procedure for micro-scale upstream process development and product analysis. A novel insert and expression system was designed and used to produce up to 1.63 mg/mL \pm 0.13 mg/mL VP1. The highest product titers were obtained at a high temperature (37 °C), a low oxygen supply at 600 rpm shaker speed, and a slightly acid pH of 6. To the best of our knowledge, this exceeds all reported VLP yields achieved in shaker flask cultivations so far. In general, the micro-scale cultivations suggested that induced stress during cell growth and product formation (e.g. by oxygen limitation or slightly acidic pH) is beneficial for high titers of soluble product. The main parameter was identified to be the oxygen supply that needed to be low to ensure fermenting conditions.

Following the successful scale-up from micro-titer plates to 2.5 L shaker flasks, VP1 was purified by a simple two-step downstream process and finally assembled into homogeneous, spherical particles. The produced VLPs constitute an ideal potential carrier platform for the presentation of antigen epitopes, packaging of small molecules, DNA or as a model nanoparticle or viral system. The microbial murine polyoma VLP platform has now been advanced further towards an industrially feasible option for tailored vaccines.

References

- [1] R. L. Garcea, L. Gissmann, Virus-like particles as vaccines and vessels for the delivery of small molecules., *Curr Opin Biotechnol* 15 (2004) 513–517.
- [2] G. T. Jennings, M. F. Bachmann, The coming of age of virus-like particle vaccines, *Biol Chem* 389 (2008) 521–536.
- [3] N. Kushnir, S. J. Streatfield, V. Yusibov, Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development., *Vaccine* 31 (2012) 58–83.
- [4] M. F. Bachmann, P. Whitehead, Active immunotherapy for chronic diseases., *Vaccine* 31 (2013) 1777–84.
- [5] World Health Organization, Guidelines on the quality , safety and efficacy of recombinant malaria vaccines targeting the pre-erythrocytic and blood stages of *Plasmodium falciparum* (2014).
- [6] C. Ladd Effio, J. Hubbuch, Next generation vaccines and vectors: Designing downstream processes for recombinant protein based virus-like particles, *Biotechnol J* 10 (2015) 715–727.
- [7] L. H. L. Lua, N. K. Connors, F. Sainsbury, Y. P. Chuan, N. Wibowo, A. P. J. Middelberg, Bioengineering virus-like particles as vaccines., *Biotechnol Bioeng* 111 (2014) 425–440.
- [8] J. R. Swartz, Advances in *Escherichia coli* production of therapeutic proteins, *Curr Opin Biotechnol* 12 (2001) 195–201.
- [9] T. K. Joergensen, L. H. Bagger, J. Christiansen, G. H. Johnsen, J. R. Faarbaek, L. Joergensen, B. S. Welinder, Quantifying biosynthetic human growth hormone in *Escherichia coli* with capillary electrophoresis under hydrophobic conditions, *J Chrom A* 817 (1998) 205–214.
- [10] D. C. Andersen, L. Krummen, Recombinant protein expression for therapeutic applications, *Curr Opin Biotechnol* 13 (2002) 117–123.
- [11] X. Zhang, M. Wei, H. Pan, Z. Lin, K. Wang, Z. Weng, Y. Zhu, L. Xin, J. Zhang, S. Li, N. Xia, Q. Zhao, Energetic changes caused by antigenic module insertion in a virus-like particle revealed by experiment and molecular dynamics simulations., *Vaccine* 32 (2014) 4039–4050.
- [12] V. I. Tishkov, A. G. Galkin, V. V. Fedorchuk, P. A. Savitsky, A. M. Rojkova, H. Gieren, M. R. Kula, Pilot scale production and isolation of recombinant NAD⁺- and NADP⁺-specific formate dehydrogenases., *Biotechnol Bioeng* 64 (1999) 187–193.
- [13] M.-E. Riviere, A. Caputo, N. Laurent, I. Vostiar, N. Andreasen, S. Cohen, R. W. Kressig, J. M. Ryan, A. Graf, Active Ab Immunotherapy Cad106 Phase II Dose-Adjuvant Finding Study: Immune Response, *Alzheimer's and Dementia* 10 (2014) 447–448.

- [14] W. Fiers, M. De Filette, K. El Bakkouri, B. Schepens, K. Roose, M. Schotsaert, A. Birkett, X. Saelens, M2e-based universal influenza A vaccine., *Vaccine* 27 (2009) 6280–6283.
- [15] J. Vekemans, A. Leach, J. Cohen, Development of the RTS,S/AS malaria candidate vaccine., *Vaccine* 27 (2009) 67–71.
- [16] A. Bouchie, GSK plows ahead with EMA malaria vaccine submission., *Nat Biotechnol* 31 (2013) 1066.
- [17] D. C. Whitacre, B. O. Lee, D. R. Milich, Use of hepadnavirus core proteins as vaccine platforms, *Expert Rev Vaccines* 8 (2010) 1565–1573.
- [18] M. Baz, Y. Abed, C. Savard, C. Pare, C. Lopez, G. Boivin, D. Leclerc, Development of a universal influenza A vaccine based on the M2e peptide fused to the papaya mosaic virus (PapMV) vaccine platform, *Vaccine* 26 (2008) 3395–3403.
- [19] A. P. J. Middelberg, T. Rivera-Hernandez, N. Wibowo, L. H. L. Lua, Y. Fan, G. Magor, C. Chang, Y. P. Chuan, M. F. Good, M. R. Batzloff, A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines., *Vaccine* 29 (2011) 7154–62.
- [20] N. Wibowo, Y. P. Chuan, L. H. L. Lua, A. P. J. Middelberg, Modular engineering of a microbially-produced viral capsomere vaccine for influenza, *Chem Eng Sci* 103 (2013) 12–20.
- [21] M. R. Anggraeni, N. K. Connors, Y. Wu, Y. P. Chuan, L. H. L. Lua, A. P. J. Middelberg, Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle., *Vaccine* 31 (2013) 4428–4435.
- [22] T. Rivera-Hernandez, J. Hartas, Y. Wu, Y. P. Chuan, L. H. L. Lua, M. Good, M. R. Batzloff, A. P. J. Middelberg, Self-adjuvanting modular virus-like particles for mucosal vaccination against group A streptococcus (GAS)., *Vaccine* 31 (2013) 1950–1955.
- [23] Y. P. Chuan, T. Rivera-Hernandez, N. Wibowo, N. K. Connors, Y. Wu, F. K. Hughes, L. H. L. Lua, A. P. J. Middelberg, Effects of pre-existing anti-carrier immunity and antigenic element multiplicity on efficacy of a modular virus-like particle vaccine., *Biotechnol Bioeng* 110 (2013) 2343–2351.
- [24] K. Tegerstedt, A. Franzén, T. Ramqvist, T. Dalianis, Dendritic cells loaded with polyomavirus VP1/VP2Her2 virus-like particles efficiently prevent outgrowth of a Her2 / neu expressing tumor, *Cancer Immunol Immunother* 56 (2007) 1335–1344.
- [25] M. Eriksson, K. Andreasson, J. Weidmann, K. Lundberg, K. Tegerstedt, Murine Polyomavirus Virus-Like Particles Carrying Full- Length Human PSA Protect BALB / c Mice from Outgrowth of a PSA Expressing Tumor, *PLoS One* 6 (2011) 1–9.

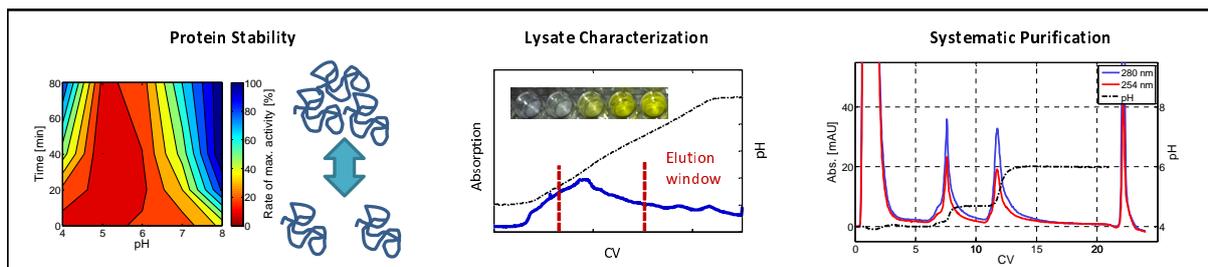
-
- [26] K. Tegerstedt, A. V. Franzén, K. Andreasson, J. Joneberg, S. Heidari, T. Ramqvist, T. Dalianis, Murine polyomavirus virus-like particles (VLPs) as vectors for gene and immune therapy and vaccines against viral infections and cancer., *Anticancer Res* 25 (2005) 2601–2608.
- [27] D. I. Lipin, Y. P. Chuan, L. H. L. Lua, A. P. J. Middelberg, Encapsulation of DNA and non-viral protein changes the structure of murine polyomavirus virus-like particles., *Arch Virol* 153 (2008) 2027–2039.
- [28] D. I. Lipin, L. H. L. Lua, A. P. J. Middelberg, Quaternary size distribution of soluble aggregates of glutathione-S-transferase-purified viral protein as determined by asymmetrical flow field flow fractionation and dynamic light scattering., *J Chrom A* 1190 (2008) 204–214.
- [29] M. E. Kimple, J. Sondek, Overview of affinity tags for protein purification, *Curr Protoc Protein Sci* 9 (2004) 36.
- [30] Y. P. Chuan, L. H. L. Lua, A. P. J. Middelberg, High-level expression of soluble viral structural protein in *Escherichia coli*., *J Biotechnol* 134 (2008) 64–71.
- [31] N. K. Connors, Y. Wu, L. H. L. Lua, A. P. J. Middelberg, Improved fusion tag cleavage strategies in the downstream processing of self-assembling virus-like particle vaccines, *Food Bioprod Process* 92 (2014) 143–151.
- [32] A. Schaller, N. K. Connors, S. A. Oelmeier, J. Hubbuch, A. P. J. Middelberg, Predicting recombinant protein expression Experiments using molecular Dynamics simulation, *Chem Eng Sci* 121 (2015) 340–350.
- [33] J. I. Betts, F. Baganz, Miniature bioreactors: current practices and future opportunities., *Microb Cell Fact* 5 (2006) 1–14.
- [34] J. Altenbach-Rehm, C. Nell, M. Arnold, D. Weuster-Botz, Parallel Bubble Columns with Fed-Batch Technique for Microbial Process Development on a Small Scale, *Chem Eng Technol* 22 (1999) 1051–1058.
- [35] R. Puskeiler, K. Kaufmann, D. Weuster-Botz, Development, parallelization, and automation of a gas-inducing milliliter-scale bioreactor for high-throughput bioprocess design (HTBD)., *Biotechnol Bioeng* 89 (2005) 512–23.
- [36] S. Rameez, S. S. Mostafa, C. Miller, A. A. Shukla, High-Throughput Miniaturized Bioreactors for Cell Culture Process Development: Reproducibility, Scalability, and Control., *Biotechnol Prog* 30 (2011) 497–505.
- [37] K. Kottmeier, J. Weber, C. Mueller, T. Bley, J. Buechs, Asymmetric Division of *Hansenula polymorpha* Reflected by a Drop of Light Scatter Intensity Measured in Batch Microtiter Plate Cultivations at Phosphate Limitation, *Biotechnol Bioeng* 104 (2009) 554–561.
- [38] R. Huber, S. Roth, N. Rahmen, J. Buechs, Utilizing highthroughput experimentation to enhance specific productivity of an *E.coli* T7 expression system by phosphate limitation., *BMC Biotechnol* 22 (2011) 1–11.

- [39] M. Funke, S. Diederichs, F. Kesy, C. Mueller, J. Buechs, The Baffled Microtiter Plate : Increased Oxygen Transfer and Improved Online Monitoring in Small Scale Fermentations, *Biotechnol Bioeng* 103 (2009) 1118–1128.
- [40] F. Kesy, C. Engelbrecht, J. Buechs, 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, *Microb Cell Fact* 68 (2009) 1–15.
- [41] P. Baumann, N. Bluthardt, S. Renner, A. Osberhaus, J. Hubbuch, Integrated Development of Up- and Downstream Processes Supported by the Cherry-Tag for Real-time Tracking of Stability and Solubility of Proteins, *J Biotech* 200 (2015) 27–37.
- [42] A. D. Leavitt, T. M. Roberts, R. L. Garcea, Polyoma Virus Major Capsid Protein, VP1 - purification after high-level expression in *Escherichia coli*, *J Biol Chem* 260 (1985) 12803–12809.
- [43] R. Hermann, M. Lehmann, J. Büchs, Characterization of gas-liquid mass transfer phenomena in microtiter plates., *Biotechnol Bioeng* 81 (2003) 178–86.
- [44] Life Technologies, NuPAGE Technical Guide. (2010).
- [45] Life Technologies, XCell II Blot Module - User Manual (2009).
- [46] Perkin Elmer, HT Protein Express LabChip Kit, Version 2 LabChip GXII User Guide. (2012).
- [47] C. Ladd Effio, L. Wenger, O. Ötes, S. A. Oelmeier, J. Hubbuch, Downstream processing of virus-like particles: Single-stage and multi-stage aqueous two-phase extraction, *J Chrom A* 1383 (2015) 35–46.
- [48] A. Henaut, A. Danchin, *Escherichia coli* *Salmonella typhimurium*. *Cell. Mol. Biol.* Vol. I II., American Society for Microbiology, 1996, pp. 2047–2066.
- [49] D. Chen, D. E. Texada, Low-usage codons and rare codons of *Escherichia coli* Mini Review, *Gene Ther Mol Biol* 10 (2006) 1–12.
- [50] Y. Nakamura, T. Gojobori, T. Ikemura, Codon usage tabulated from the international DNA sequence databases: status for the year 2000, *Nucleic Acids Res* 27 (1999) 292.
- [51] M. Funke, A. Buchenauer, U. Schnakenberg, W. Mokwa, S. Diederichs, A. Mertens, C. Mueller, F. Kesy, J. Buechs, Microfluidic BioLector - Microfluidic Bioprocess Control in Microtiter Plates, *Biotechnol Bioeng* 107 (2010) 497–505.
- [52] M. Losen, B. Frölich, M. Pohl, J. Büchs, Effect of oxygen limitation and medium composition on *Escherichia coli* fermentation in shake-flask cultures., *Biotechnol Prog* 20 (2004) 1062–8.

-
- [53] E. Kopetzki, G. Schumacher, P. Buckel, Control of formation of active soluble or inactive insoluble baker's yeast alpha-glucosidase PI in *Escherichia coli* by induction and growth conditions., *Mol Gen Genet* 216 (1989) 149–155.
- [54] C. A. Galloway, M. P. Sowden, H. C. Smith, Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase., *Biotechniques* 34 (2003) 524–526.
- [55] Y. Shirano, D. Shibata, Low temperature cultivation of *Escherichia coli* carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level, *FEBS Lett* 271 (1990) 128–130.
- [56] C. H. Schein, M. H. M. Noteborn, Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature, *Nat Biotechnol* 6 (1988) 291–294.
- [57] H. Ewers, A. E. Smith, I. F. Sbalzarini, H. Lilie, P. Koumoutsakos, A. Helenius, Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes., *Proc Natl Acad Sci USA* 102 (2005) 15110–15115.
- [58] C. Peixoto, M. F. Q. Sousa, A. C. Silva, M. J. T. Carrondo, P. M. Alves, Downstream processing of triple layered rotavirus like particles., *J Biotechnol* 127 (2007) 452–461.
- [59] B. T. J. Hahn, D. Courbron, M. Hamer, M. Masoud, J. Wong, K. Taylor, J. Hatch, M. Sowers, E. Shane, M. Nathan, H. Jiang, Z. Wei, J. Higgins, K.-H. Roh, J. Burd, D. Chinchilla-Olszar, M. Malou-Williams, D. P. Baskind, G. E. Smith, Rapid Manufacture and Release of a GMP Batch of Avian Influenza A(H7N9) Virus-Like Particle Vaccine Made Using Recombinant Baculovirus-Sf9 Insect Cell Culture Technology, *Bioprocess J* 12 (2013) 1–10.
- [60] T. Kröber, M. W. Wolff, B. Hundt, A. Seidel-Morgenstern, U. Reichl, Continuous purification of influenza virus using simulated moving bed chromatography, *J Chrom A* 1307 (2013) 99–110.
- [61] J. M. Wagner, J. D. Pajeroski, C. L. Daniels, P. M. McHugh, J. A. Flynn, J. W. Balliet, D. R. Casimiro, S. Subramanian, Enhanced production of Chikungunya virus-like particles using a high-pH adapted *spodoptera frugiperda* insect cell line., *PLoS One* 9 (2014) 1–14.

Systematic Purification of Salt-intolerant Proteins in Ion-exchange Chromatography: The Example of Human α -Galactosidase A

P. Baumann¹, A. Osberghaus¹ and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1, 76131 Karlsruhe, Germany

* : Corresponding author. *E-mail-address*: juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

Chromatography is an essential tool for purifying biopharmaceutical products. Many processes are still developed based on traditional routines and empiric procedures. Product losses are mostly due to insufficient optimization of purification setups and product sensitivity to process conditions. In order to eliminate these shortcomings, a systematic strategy for the setup of ion-exchange chromatography is presented, which considers both product stability as well as operational conditions. The stages - a hybrid approach combining high-throughput screening and analytical small-scale chromatography - are as follows: (1) '*pH stability - short-term*'; (2) '*pH stability - long-term*', followed by a screening of additives to enhance protein stability, if required; (3) Analytical '*pH gradient chromatography*' for evaluation of the operational pH window; and (4) '*salt stability - long-term*' in the operational pH window determined. Efficiency and straightforwardness of the strategy were shown in a case study on capturing the human α -Galactosidase A enzyme. Following the above procedure, the enzyme was found to be salt-unstable. The applied strategy allowed for a quick establishment of a dedicated capture step at low salt concentrations under stable conditions by well-chosen prior screening experiments. In this setup a purification factor of 13.2, a concentration factor of 4 and an overall yield of 84.3% were achieved.

Keywords: Ion-exchange Chromatography, High-throughput Screening, Hybrid Approach, Salt-intolerance, α -Galactosidase A

1 Introduction

Chromatographic procedures are commonly used to capture biological products from complex mixtures. The critical process parameters determining purity and yield are among others the pH, salt type and salt concentration used during binding, wash and elution procedures. Ion-exchange chromatography has become the method of choice for capture steps due to the rather mild process conditions and high binding capacities. However, even moderate concentrations of salts being low in chaotropic or cosmotropic nature might affect protein integrity as was reported in a variety of studies on therapeutically relevant proteins. The granulocyte colony-stimulating factor was found to form aggregates in the presence of 150 mM sodium chloride under physiological conditions in phosphate-buffered saline [1, 2]. Analog aggregation behavior was observed for the recombinant human tissue factor pathway inhibitor [3]. A slight increase in sodium chloride concentration from 0.04 mM to 0.11 mM induces major insulin fibrillation under acidic conditions [4]. RNase A was found to be destabilized by moderate sodium chloride concentrations of 250 and 500 mM [5]. Lysozyme also showed a strong decrease in solubility at salt concentrations below 0.5 M [6].

A very simple approach to enhancing separation efficiency of IEX chromatography is the initial determination of an operational pH window exploiting the charge-pH dependency of product and contaminants. In a heuristic approach, the binding pH in ion-exchange chromatography is chosen to be ± 0.5 to 1 pH units above or below the isoelectric point (pI) of the product [7]. Those pI values from literature are either calculated *in silico* based on molecular modeling or determined experimentally using isoelectric focusing.

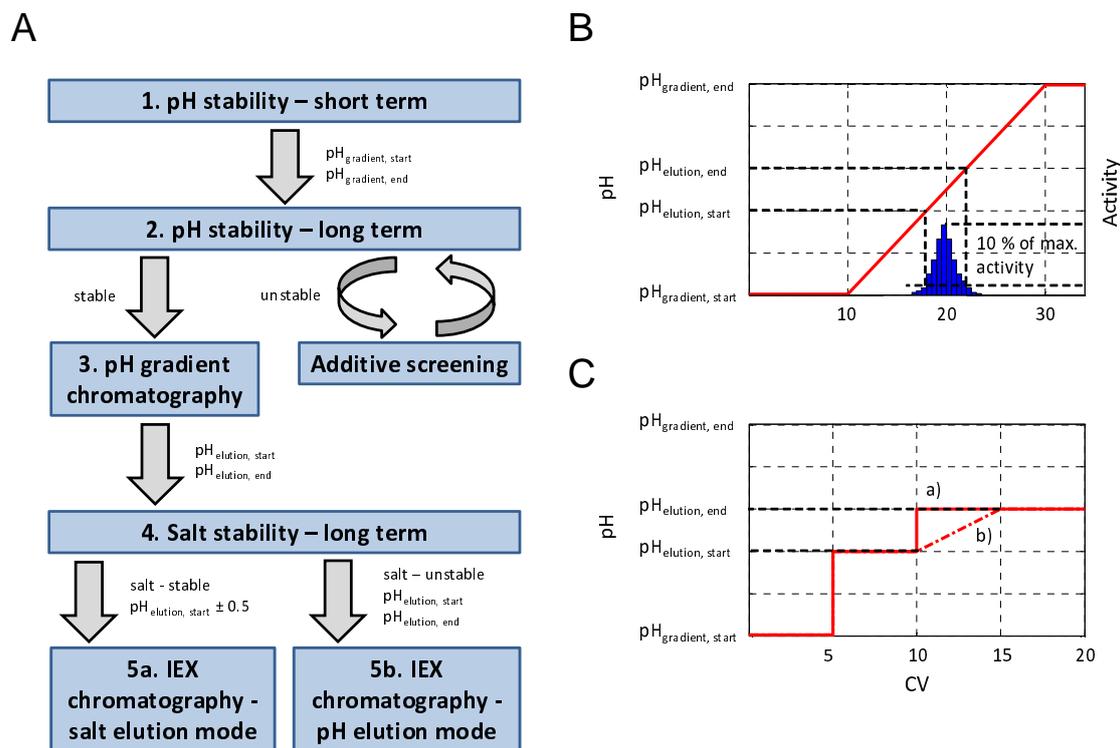


Figure 1: Schematic illustration of the systematic strategy for the purification of salt-intolerant proteins in ion-exchange chromatography. A: Sequential order of screenings for an optimized IEX chromatography process. B: pH gradient chromatography for determination of the product's elution pH. C: IEX chromatography - pH elution mode with two different elution setups of a step (a) or 5 CV gradient (b).

Such values are, however, mostly not identical to the actual product elution pH on a specific column. Charge distributions and charge clusters on protein surfaces may lead to binding to the adsorber even at a molecular net charge of zero [8]. Another factor to be considered is the interference with the resin backbone, where non-electrostatic forces, such as hydrophobic interactions, come into play [9]. Especially proteins with a neutral pI (pI 6 to 8) tend to elute experimentally at much higher pH values than expected [10, 11]. Also posttranslational modifications like glycosylation patterns were shown to affect the elution pH of proteins [12].

A deterministic approach to evaluating an operational pH window applies pH gradient chromatography as an elegant analytical tool in IEX process development. Instead of relying on pI values from literature, the elution pH values of products and contaminants are determined experimentally on a specific column [10]. The salt gradient purification can then be designed for a pH value of ± 0.5 apart from the experimental elution pH. Previously, pH gradients for the determination of the pH operational window in IEX chromatography were used in the purification of a monoclonal antibody from a cell culture supernatant on an AEX and CEX column [13]. Similar approaches were applied for separating Nucleolin from an Sf9 insect cell extract by AEX and CEX chromatography [14] and the anthrax-protective gene (PA) from an *E. coli* cell lysate by AEX chromatography [15].

However, protein stability and use of additives under typical process conditions have not

been considered so far. In fact, colloidal and conformational stability problems, namely, aggregation and unfolding of proteins, will lead to product losses [16]. A pH close to the isoelectric point of a protein might lead to aggregation due to a reduction of repulsive intermolecular forces [17]. In contrast to this, a pH value far from the isoelectric point leads to identical intramolecular charges [17]. This can induce an unfolding of proteins. Another factor to be considered is the amount and type of salt, as discussed above.

So far, protein stability and additive screenings have mostly been applied for protein refolding and crystallization studies. Lysozyme and carbonic anhydrase were stabilized by addition of cyclodextrins and surfactants [18]. The native state of β -Lactamase was maintained by adding sucrose to the buffer. Other examples of additives are co-factors, amino acids (e.g. lysine, proline), several sugars, alcohols (e.g. glycerol), and other detergents like CHAPS [19]. Reducing agents, such as DTT in low concentrations, were applied as well [20]. Lysozyme is shown to be stabilized by low-molecular-weight polyethylene glycols (PEGs), detergents, such as Tween 20, and different sugar molecules like sorbitol and sucrose [6]. For this reason, an integral strategy for optimizing an IEX capture step combining product stability and operational parameters over a wide pH range needs to be considered. In the following chapters the term 'stability' refers to the conservation of enzymatic activity.

A systematic strategy for the purification of salt-intolerant proteins in IEX chromatography will be introduced in this paper. The strategy is illustrated in Fig. 1A and consists of four consecutive stages, including screenings for short- and long-term pH stability, salt tolerance, additives, and optimal pH for a pH-based IEX procedure: In a first stage (1) '*pH stability - short-term*' the general pH stability of the product/ feed stream is evaluated. Conditions of instant denaturation or irreversible aggregation can be excluded in the further process development. Product exposure times to a certain chemical environment can, however, be in the range of several minutes to hours. The identified pH region of stage 1 is further evaluated in a second stage (2) analyzing product '*pH stability - long-term*'. If the product is found to be unstable at this point, an additive screening follows. During this stage, the procedures of stage 2 '*pH stability - long-term*' are repeated with different excipients to find out whether the product can be kept in a native state for a predefined period time. In a third stage (3) the crude feedstock is analyzed by '*pH gradient chromatography*' to identify the operational pH window for IEX chromatography. At this point, the mode of IEX chromatography needs to be chosen. In order to define the optimal elution mode, screenings to determine (4) '*salt stability - long-term*' in the operational pH window determined earlier are used to test the effect of product exposure to different concentrations of salt. If the protein is found to remain in a native state during salt exposure, traditional (5a) '*IEX chromatography - salt elution mode*' can follow. For salt-unstable products, a low-salt purification process based on product elution by pH shift '*IEX chromatography - pH elution mode*' (5b) is carried out. Such elution setups are commonly used in affinity chromatography [21, 22]. In several cases IEX chromatography applying pH shifts for elution were used for separating protein isoforms (e.g. PEGylated Lysozyme, human transferrin) or chemically similar variants (e.g. deaminated protein variants) or as an intermediate purification step of D-amino-acid oxidase from *Candida guilliermondii* [23, 24, 25, 26].

This case study focuses on the use of the systematic strategy proposed for ion-exchange chromatography for capturing human α -Galactosidase A from *Pichia pastoris*. The en-

zyme consists of two 50 kDa subunits and is naturally located in human cellular lysosomes [27] and is involved in the cellular detoxification mechanism by metabolizing harmful glycolipids. The absence of α -Galactosidase A inside the human body leads to an accumulation of glycolipids and major cell damage (Fabry Disease).

2 Materials & Methods

2.1 Materials

Sample conditioning was performed using PD-10 desalting columns (bed volume 8.3 mL) packed with Sephadex[®]G-25 Medium resin (GE Healthcare Life Sciences, Sweden). Fractionations were carried out in 2 mL 96-well square deep well plates (VWR, Germany). The enzyme assay was performed in 96-well flat-bottom polypropylene microplates (Greiner Bio-One, Germany). Absorption measurements were carried out in 96-well flat-bottom UV-Star[®] microplates (Greiner Bio-One, Germany). For chromatography runs, Media-Scout[®]MiniChrom[®]5-50 columns (Atoll, Germany) pre-packed with 1 mL of the following resins were purchased: Source 30 S, Source 30 Q (GE Healthcare Life Sciences, Sweden).

The buffer system used, allowing for the generation of highly linear pH gradients, follows a description by Kroener et al. [11]. For cation-exchange, the buffer system uses the following components: 10 mM sodium chloride, 11 mM MES, 9.9 mM formic acid and 13 mM acetic acid (Merck, Germany), 8.7 mM MOPSO, 9.9 mM HEPPSO, 4.6 mM TAPS, 9.4 mM CHES and 15.6 mM CAPS (AppliChem, Germany). The anion-exchange buffer system is composed of the following components: 7.7 mM hydroxylamine, 9.8 mM methylamine, 6.99 mM 1,2-ethane-diamine, 13.47 mM 1,4-dimethylpiperazine (Merck, Germany), 5.8 mM Bis-Tris (Molekula, UK), and 6.4 mM 1-methylpiperazine (Sigma-Aldrich, USA). The pH of the buffers was adjusted by titration using hydrochloric acid or sodium hydroxide (Merck, Germany), respectively. For the adapted 4-nitrophenole assay, the substrate 4-nitrophenyl α -D-galactopyranoside (4-NPG) and the colorimetric product 4-nitrophenol (4-NP) were purchased from Sigma-Aldrich, USA. As a reaction medium, a 100 mM acetate buffer at pH 5 was prepared from acetic acid (Merck, Germany). For termination of the enzyme reaction, disodium carbonate (Merck, Germany) was applied. A supernatant from a *P. pastoris* fermentation (strain NRRLY - 11430), during which the enzyme human α -Galactosidase A is produced extracellularly, as well as purified human α -Galactosidase A standard were obtained from BioIngenium (Spain).

2.2 Instrumentation & Software

For pH adjustment of all buffers, an HI-3220 pH meter (Hanna Instruments, USA) was used. The instrument was calibrated using high-precision standards from Hanna Instruments (USA). For the enzyme assay, an MUR 13 thermo-shaker with additional lid heating (HLC BioTech, Germany) was used. For absorption measurements, an Infinite M200 Reader (Tecan, Germany) was applied and operated with I-control 1.9 (Tecan, Germany). The '*pH gradient chromatography*' and '*IEX chromatography - pH elution mode*' were carried out in an ÄKTA[™] purifier system operated with Unicorn 5.2 (GE Healthcare Life Sciences, Sweden). The system was equipped with a pump P-900, mixer

M-925, UV detector UV 900, motor valve INV-907, pH and conductivity monitoring pH/C-900, and a fraction collector Frac-950 unit. Data processing and creation of figures were performed in Matlab®R2011a (MathWorks, USA).

2.3 Strategic Process Development

2.3.1 Adapted 4-Nitrophenole Assay

A necessary prerequisite for the application of the proposed systematic strategy to the case study was a pH-insensitive assay determining product concentration or activity.

I. Adapted Assay Procedure

For the determination of product titers, a 4-NPG enzyme activity assay was adapted. The basic principles presented by Rezende et al. [28] and Marin et al. [29] were modified and combined. The adjusted 4-NPG protocol works as follows: 200 μL of 2 mM 4-NPG in 100 mM acetate buffer (pH 5) are added to 50 μL sample and incubated at 37 °C and 400 rpm in a plate shaker with temperature control. The enzyme reaction is stopped by addition of 25 μL of a 1 M sodium carbonate solution. The concentration of accumulated 4-NP is determined in a plate reader by 405 nm absorption measurements.

II. Assay Validation

As fractions collected throughout the pH gradient intrinsically exhibit a large variation of the pH value, special emphasis was put on evaluating the sensitivity of the assay to pH changes of the sample. The enzymatic activity of α -Galactosidase A is optimal in the pH range from about 4.5 to 5 [30, 31, 32]. In order to test whether the buffering strength of the receiving buffering system - 200 μL of 100 mM acetate buffer at pH 5 - is sufficient to buffer samples, the following procedure was applied in triplicate measurements: 2.5 mL of a sample from the *P. pastoris* supernatant containing α -Galactosidase A were conditioned by buffer exchange in a PD-10 desalting column into 100 mM acetate buffer at pH 5. By mixing 100 μL of the conditioned sample and 100 μL of 4 mM 4-NPG solution, a final 2 mM substrate concentration was reached. By adding 50 μL of pH gradient buffer in a pH range of 3 to 11 to this basic mixture, a potentially observed decrease in activity can be correlated directly to a strong pH fluctuation in the reaction mixture.

Next to operating at the optimal pH for the enzyme reaction, care needs to be taken in the pH-dependent readout of the colorimetric reaction. It was pointed out by Werner 2001 [33] that the absorption coefficient of the cleaved colorimetric 4-NP is strongly pH-dependent under acidic and neutral conditions. Hence, terminating the enzyme reaction by adding 1 M sodium carbonate solution is used to shift the pH to basic conditions where 4-NP shows no pH dependence in absorption at 405 nm. 25 μL of 1 M sodium carbonate solution were added and the absorbance was determined at 405 nm. The 405 nm signals obtained were then correlated to the effective enzyme activity by determination of released 4-NP per reaction time. For this purpose, a linear calibration of 1 to 25 $\mu\text{g}/\text{mL}$ 4-NP in AEX buffer, pH 11, was performed, leading to a linear regression Δ_{4-NP} of 0.0726 AU/(μg 4-NP/mL). The molar enzyme activity Act [μmol 4-NP/mL/h]

of samples was then calculated as shown in Eq. (1):

$$Act = \frac{(E_{405, sample} - E_{405, blank}) \cdot 5.5}{t \cdot \Delta_{4-NP} \cdot M_{4-NP}} \quad (1)$$

With E_{405} being the extinction of the sample/ blank measured at 405 nm [AU], 5.5 being the dilution factor of the assay, t denoting the assay duration [h], Δ_{4-NP} denoting the slope of the calibration [AU/(μ g 4-NP/mL)], and M_{4-NP} being the molar mass of 4-nitrophenol (139.11 μ g/ μ mol).

2.3.2 Screening Procedure

The screening methods applied as shown in Fig. 1A are described in detail in the following chapter.

I. pH Stability - Short-term

Short-term product stability was defined to be the stability of α -Galactosidase A during a 5-minutes' period under process conditions. To analyze this stability, the following procedure was applied in triplicate measurements: 5 μ L of purified α -Galactosidase A standard in water was mixed with 45 μ L of the respective pH gradient buffer in the range from pH 3 to pH 11 in steps of one pH unit to cover the acidic and basic pH range. Product activity was measured with the 4-NPG assay. The reduction of enzymatic activity was defined to be a measure of the quantity of denatured protein during 5 min exposure time.

II. pH Stability - Long-term

Long-term product stability was defined to be the stability of α -Galactosidase A during a 120-min period under process conditions. To analyze this stability, the following procedure was applied in triplicate measurements: The buffer composition of the *P. pastoris* supernatants was adjusted to pH 4, 5, 6, 7, and 8 with PD-10 desalting columns. From each of these stock solutions, samples of 50 μ L were taken and the 4-NPG assay was started after 0, 20, 40, and 80 min exposure time, respectively. The assay was stopped after 120 min by adding 1 M sodium carbonate. To counteract potential pH instability of the product, an additive screening was carried out after the pH stability screenings. The experimental setup was identical to that of the 'pH stability - long-term' process, with additives being added to the respective pH gradient buffers. Here, 100 mM sodium chloride and 0.02% Tween 20 were chosen.

III. pH Gradient Chromatography

pH gradient chromatography was performed in an ÄKTATM purifier system using Media-Scout[®]MiniChrom[®]5-50 columns (CV 1 mL) packed with Source 30 S/ Q resin. The pH values - $\text{pH}_{\text{gradient, start}}$ and $\text{pH}_{\text{gradient, end}}$ - defining the pH gradient were determined by the 'pH stability - long-term' experiments described above. The buffer composition of the *P. pastoris* supernatant was consequently adjusted to pH 8 for AEX and pH 4 for CEX chromatography in PD-10 desalting columns. A load of 500 μ L sample was applied and the volumetric flow rate was set to 1 mL/min. The integrated pH electrode

was 2-point calibrated for the start and end conditions of the pH gradient. The general chromatographic procedure is shown in Fig. 1B. For CEX experiments, the pH gradient was increased from pH 4 to 8. For AEX experiments, the pH was decreased from pH 8 to 4 (buffer species and concentrations are described above). Initial equilibration was performed with 5 CV of the respective binding buffer. 10 CV of binding buffer were applied after sample application in the wash step. Subsequently, a 20 CV pH gradient was applied to reach the final pH of the respective gradient. This pH was then kept constant for another 3 CV. Finally, regeneration was performed with a 4 CV 1 M sodium chloride step applied at the end of the chromatographic procedure. Additionally, blank runs were conducted to calculate buffer-independent UV signals. Throughout the chromatography run, 500 μL fractions were taken for product-specific analysis. 50 μL of each fraction were evaluated using the adapted 4-NPG assay.

The operational pH window for product elution was defined by a minimum of 10% enzyme activity found in the respective fractions when compared to the maximum enzyme activity found (see Fig. 1B). The framing pH values are referred to as $\text{pH}_{\text{elution,start}}$ and $\text{pH}_{\text{elution,end}}$.

IV. Salt Stability - Long-term

The *P. pastoris* supernatant was adjusted to $\text{pH}_{\text{elution,start}}$ and the respective buffer system using PD-10 desalting columns. For the CEX setup, pH 4.3 and for the AEX setup, pH 7.5 was selected based on the criteria determined above. The conditioned samples were then mixed (dilution factor 2) with 2-fold concentrated salt solutions of the respective buffers, leading to final sodium chloride concentrations of 0, 50, 100, 200, 300, and 500 mM. The 4-NPG assay was started after 0, 20, 40, and 80 min exposure time and stopped after 120 min by adding 1 M sodium carbonate.

V. IEX Chromatography

As human α -Galactosidase A was found to be salt-unstable, a processing mode applying a pH shift was used. The *P. pastoris* samples were conditioned to $\text{pH}_{\text{gradient,start}}$ (compare Fig. 1C) using the respective buffer components of the analytical pH gradients. All experiments were conducted in an ÄKTATM purifier system using MediaScout[®] MiniChrom[®] 5-50 columns packed with Source 30 S/ Q resin. A load of 1000 μL sample was applied and the volumetric flow rate was set to 1 mL/min. Spectroscopic analysis was conducted using the 280 nm signal for total protein and the 254 nm signal for nucleic acids. The chromatographic procedure was conducted as follows (see Fig. 1C): The IEX column was equilibrated with 5 CV of binding buffer, followed by sample application and wash with another 5 CV of binding buffer. A second wash procedure for 5 CV was run by a step-wise pH shift to $\text{pH}_{\text{elution,start}}$, to elute contaminants close to the product's elution pH. Product elution within the operational pH window was then performed in two different setups for both the AEX and CEX chromatography columns: Step elution and gradient elution over 5 CV. Elution was followed by a 5 CV $\text{pH}_{\text{elution,end}}$ wash and a 4 CV 1 M NaCl high-salt regeneration step. Analogously to the pH gradient screening experiments, blank runs were conducted to calculate the buffer-independent UV signals.

2.3.3 Analysis of Chromatographic Procedures

Throughout the '*IEX chromatography - pH elution mode*' run, fractions of 1000 μL were taken for product-specific analytics.

I. Mass Balances

To investigate denaturation, protein remaining bound to the adsorber or formation of aggregates inside the chromatography column, recovery rates were calculated. The recovery of total protein RTP was performed by dividing the integrated blanked 280 nm signal from the entire chromatography run $A_{280,run}$ (sample injection up to the final salt step) by the integrated signal of a sample injected into the ÄKTATM system $A_{280,sample}$ - Eq. (2). For this purpose, the signals have to be in a linear measurement range. The integration process was performed in Matlab R2011a (MathWorks, USA) using a trapezoidal numerical integration.

$$RTP = \frac{\int(A_{280,run})}{\int(A_{280,sample})} \quad (2)$$

The recovery of activity RA was calculated considering the enzymatic activity of α -Galactosidase A - Eq. (3). The added activities $Act_{fraction}$ of all fractions (sample injection up to the final salt step) were divided by the total activity of the applied sample Act_{sample} taking the sample volumes $V_{fraction}$ and V_{sample} into account.

$$RA = \frac{\sum(Act_{fraction} \cdot V_{fraction})}{\sum(Act_{sample} \cdot V_{sample})} \quad (3)$$

II. Overall Product Yield

The product yields Y were calculated as the ratio between eluted enzyme activity $Act_{elution}$ and applied sample Act_{sample} taking the sample volumes $V_{elution}$ and V_{sample} into account - Eq. (4).

$$Y = \frac{\sum(Act_{elution} \cdot V_{elution})}{\sum(Act_{sample} \cdot V_{sample})} \quad (4)$$

III. Purification Factor

The product purity was calculated using specific activities as defined by Eq. (5). The purification factor PF was defined to be the ratio between the specific activity of the pooled elution fraction and the applied sample.

$$PF = \frac{\sum(Act_{elution})/\int(A_{280,elution})}{\sum(Act_{sample})/\int(A_{280,sample})} \quad (5)$$

IV. Qualitative Nucleic Acid Factor

As a third criterion, the qualitative nucleic acid factor NF was determined by the ratio of the 254 nm signal and the 280 nm signal. In contrast to measurements at 260 nm, DNA and RNA show almost equivalent absorption at these wavelengths, whereas proteins exhibit a lower absorption at 254 nm and, hence, can be distinguished better from nucleic

acids [34]. The factor is defined to be the ratio between the integrated 254 nm signal A_{254} and the 280 nm signal A_{280} - Eq. (6).

$$NF = \frac{\int(A_{254})}{\int(A_{280})} \quad (6)$$

By comparing the NF values of the applied sample and the elution fractions, the qualitative reduction of nucleic acid levels was determined.

V. Concentration Factor

For capturing unit operations, not only the yield, purity, and nucleic acid level reduction are important, but also the achieved product concentration. The product concentration factor CF was determined to be the ratio between the pooled elution activity Act_{pool} and the applied sample Act_{sample} - Eq. (7).

$$CF = \frac{Act_{pool}}{Act_{sample}} \quad (7)$$

2.3.4 Process Simplification

I. Reduction of Buffer Species

Given the optimal system parameters determined according to the procedure described above, the complex mixture of buffer substances was simplified to a two-component system. The buffer components were chosen in a buffering range of the operational pH window defined by the values of $pH_{elution,start}$ and $pH_{elution,end}$ (see Fig. 1C). For the case study, the binding buffer was set to 10 mM acetate (pK_a 4.76) at pH 4.8. The elution buffer (pH 6) consisted of 10 mM MES (pK_a 6.10). Sample binding occurred at $pH_{elution,start}$, thus eliminating the first pH shift depicted in Fig. 1C.

II. Product Capture at the Capacity Limit

Product capture with a simplified two-buffer component system was performed at the column capacity limit, which was determined to be a load of 20 CV at 10% of product breakthrough. 20 mL of sample were applied and the chromatography procedure was evaluated as described in the analysis of the chromatographic procedures section.

3 Results & Discussion

3.1 Adapted 4-nitrophenole Assay

The successful adaptation of the 4-NPG reaction towards a robust, reproducible, and widely pH-insensitive assay for α -Galactosidase A was a necessary prerequisite for the proposed development strategy shown in Fig. 1. To validate the suitability of the chosen assay, several aspects relating to pH insensitivity during reaction and readout as well as reaction kinetics were evaluated. Fig. 2 illustrates the enzyme activities obtained by assaying samples containing identical enzyme concentrations at different pH values. The

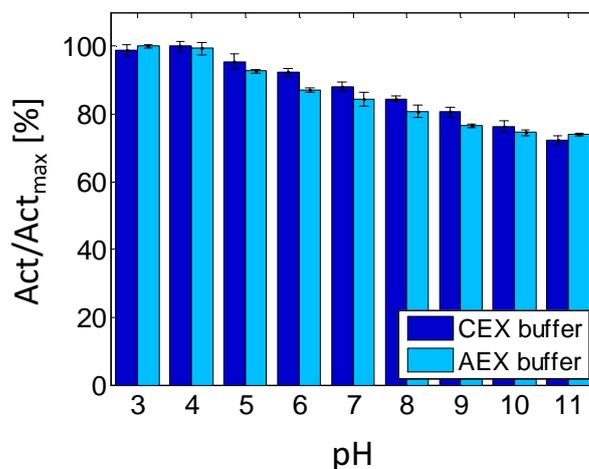


Figure 2: Examination of the adapted 4-NPG assay’s robustness at different pH levels. Evaluation of the reaction buffer’s buffering capacity for the optimal conversion pH of α -Gal A. CEX and AEX buffer samples were assayed in the range from pH 3 to pH 11 with a constant concentration of enzyme. Activities are scaled to the maximal activity determined in the test set. The error bars indicate the data variability of the triplicate measurements.

highest 4-NPG conversion rates were measured for samples in a pH range from pH 3 to pH 5, where the pH shift to the conditions of the reaction buffer is negligible. In general, an increase with sample pH led to a decrease of measured enzyme activity. For the AEX buffer system, an overall decrease in enzyme activity of 26.1% could be observed over the complete pH range from pH 3 to pH 11. In the pH region of pH 3 to pH 8, in which most of the stability and screening experiments were conducted, the decrease was about 19.3%. The use of the CEX buffer system led to a decrease in enzyme activity of 28.8% in the overall range from pH 3 to pH 11. In the region of pH 3 to pH 8 the activity difference was determined to be 15.6%. In summary, the chosen reaction buffer proved to be a good compromise between sufficient buffer capacity and a sufficiently low salt concentration to prevent salt-induced loss of enzyme activity.

The methodology can be used analogously for proteins that can be assayed specifically, including hormones, enzymes, etc. For most other proteins, antibody-based procedures could be applied (e.g. ELISA) for product determination. Another fast method for analyzing proteins is capillary gel electrophoresis using e.g. the Caliper Lab Chip GX II system. Here, samples can be analyzed within approximately 30 - 40 s of processing time.

Regardless of the subsequent analytics, all time-dependent screening steps, such as the ‘pH stability - short-term’ should have a similar setup: After the pre-defined exposure time, a stabilizing buffering solution (analogously to the enzyme reaction buffer) should be added to ‘stop’ the exposure to pH, salt, etc. The remaining soluble and active fraction can then be analyzed.

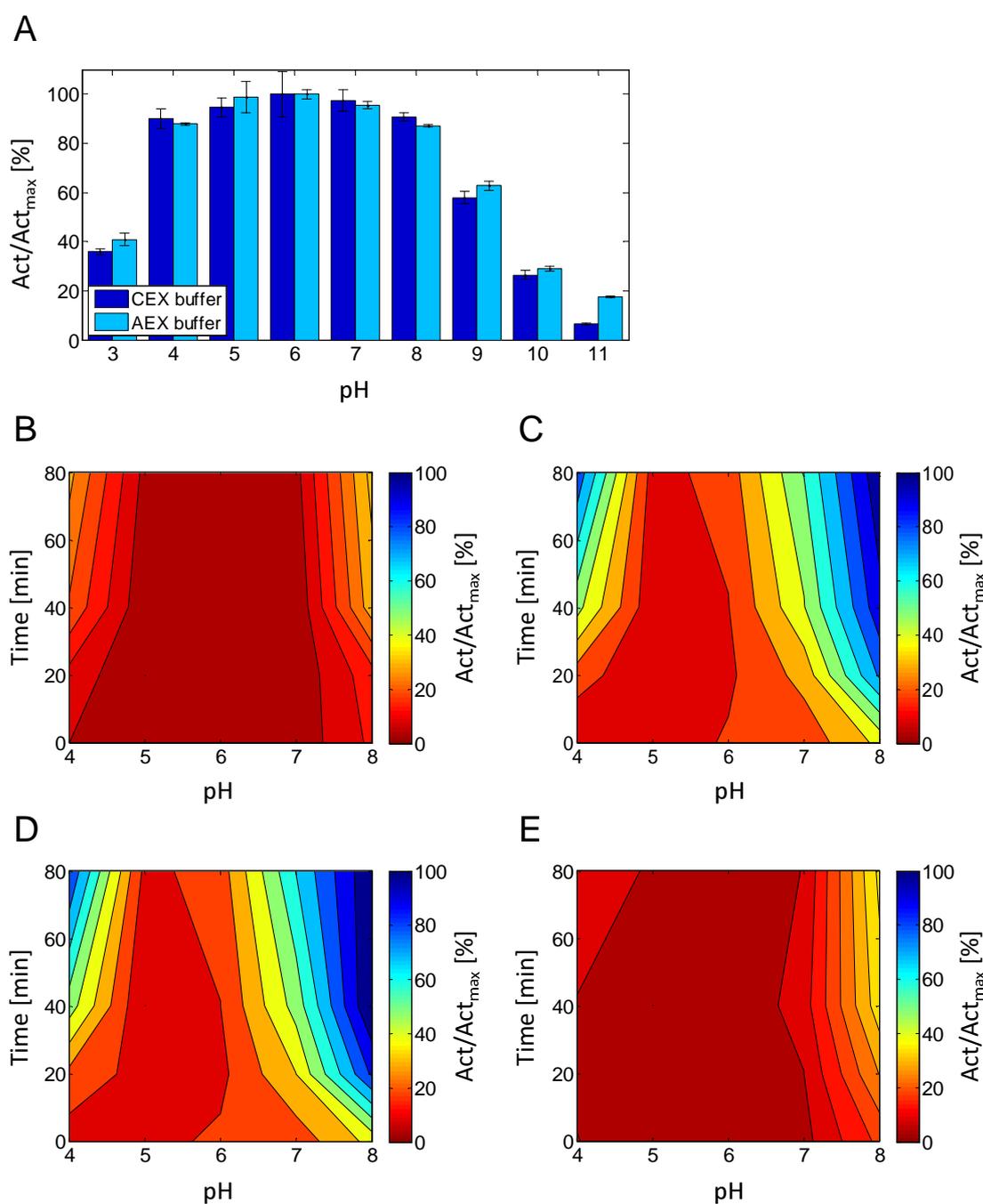


Figure 3: Product pH stability - short and long term in AEX and CEX buffers. Short term experiments were performed in the range from pH 3 to pH 11, long term experiments in the range from pH 4 to pH 8 with product exposure times of 0 to 80 min before stabilization by starting the 4-NPG assay. Activities are scaled to the maximal activity determined in each test set. A: Product pH stability - short term in AEX and CEX buffers. The error bars indicate the data variability of the triplicate measurements. B: Product pH stability - long term for the CEX buffer system without additional excipients. C: Product pH stability - long term for the AEX buffer system without additional excipients. D: Product pH stability - long term for the AEX buffer system with 100 mM sodium chloride added. E: Product pH stability - long term for the AEX buffer system with 0.02% Tween 20 detergent added.

3.2 Screening Procedure

I. pH Stability - Short-term

During the first stage, the '*pH stability - short-term*' of the given target molecule was investigated to identify potential pH limits of instant enzyme denaturation. pH values of 3 to 11 were investigated for AEX and CEX buffer systems. Following 5 min pH exposure to purified α -Galactosidase A standard, enzyme activities were determined using the assay. Fig. 3A shows the activities Act obtained scaled to the maximal activity determined in each test set Act_{max} . Both the AEX and CEX buffer systems show a similar trend: The maximal enzyme activity could be found at pH 6. A stable process window regarding enzyme activity was found in the range from pH 4 to pH 8. In this region, maximal variances of 12.8% (AEX buffers) and 10.1% (CEX buffers) were detected. Buffer systems of pH values higher than 8 led to a significant decrease in converted 4-NPG. 59.1% (AEX) and 64.1% (CEX) of enzyme activity were already lost at pH 3. Similar results were observed for the pH 9 samples. Here, a loss in enzyme activity of 37.2% (AEX) and 42.0% (CEX) was found. At pH 11 (the most basic pH investigated), a maximal degradation of the product was detected, with decreases in activity of 82.4% (AEX) and 93.4% (CEX). Consequently, the appropriate region for all further experiments was fixed to an operational pH window ranging from pH 4 to pH 8.

II. pH Stability - Long-term

During purification processes, product streams are exposed to chemical environments much longer than 5 min. Fig. 3B - E shows relative enzyme activities at pH levels between pH 4 to pH 8 and exposure times of 0 to 80 min. All activities are scaled to the activity maximum determined in the respective setup. For the CEX buffer system (Fig. 3B), a constant activity ratio of ≥ 0.965 was achieved at pH 5 to pH 7, even at 80 min exposure time. Samples exposed to pH 4 and pH 8 showed constant activity ratios ≥ 0.871 at exposure times of up to 20 min. Even at times of up to 80 min, the activities were constantly above 68.3%. For the AEX buffer systems (Fig. 3C), enzyme stability was achieved only at 5 min exposure time (as shown in '*pH stability - short-term*') and at pH 5 to pH 6. At pH 8, an exposure time of 20 min already caused a loss of activity of 71.5%. The CEX buffer systems were found to retain the enzyme activity, while AEX buffers led to a strong enzyme inactivation outside an operational pH window from pH 5 to pH 6.

For the AEX buffer system, an additive screening was carried out to stabilize human α -Galactosidase A. As a mild increase in salt concentration is reported to lead to a stabilizing effect by salting-in for some proteins [17], a first screening was conducted by adding 100 mM of sodium chloride to the AEX buffers (Fig. 3D). The results, however, indicated no improvements in enzymatic stability. A second screening was performed by adding the detergent Tween 20 to the AEX buffer system. In this setup a constant activity of more than 91.1% was reached at pH 4 to pH 7 within the period investigated (Fig. 3E). Exposure to pH 8 revealed activities of more than 74.8% at exposure times of up to 20 min. Up to 80 min of exposure, the activities were constantly above 63.8%. These results obtained for the AEX buffer system containing Tween 20 were then comparable to the enzymatic stability obtained during '*pH stability - long-term*' experiments for the

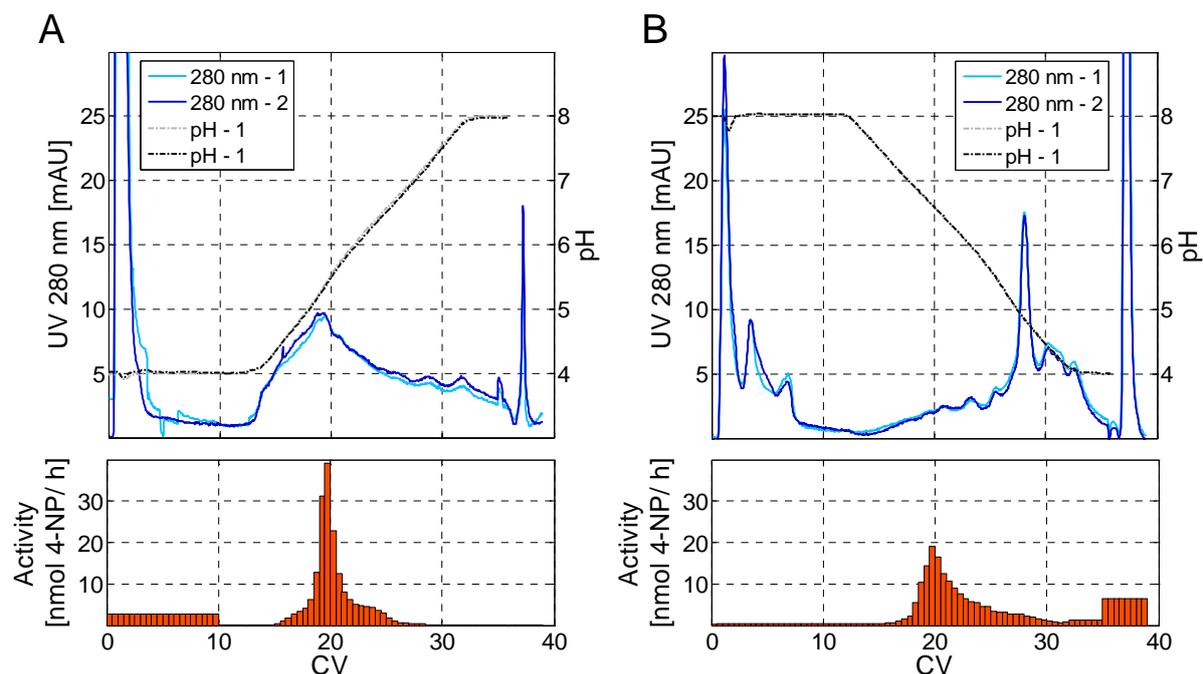


Figure 4: pH gradient chromatography of the *P. pastoris* supernatant in duplicates. UV 280 nm signals and pH values are plotted against column volumes. The corresponding fraction enzyme activities are shown in a bar diagram below. A: CEX pH gradient chromatography on a Source 30 S column (pH 4 to 8) B: AEX pH gradient chromatography on a CEX Source 30 Q column (pH 8 to 4).

CEX buffer system. Tween 20 prevented aggregation of active product molecules, which must have occurred in the presence of AEX buffer substances over time. Hence, the additive screening was capable of stabilizing the enzyme and finding a suitable chemical environment for product purification. In general, suited additives can be found when looking at the final formulation buffer or the cultivation conditions of the investigated product.

III. pH Gradient Chromatography

Analytical '*pH gradient chromatography*' was used for comparing contaminant and product elution profiles of the complex *P. pastoris* feedstock. The 280 nm signals and the enzyme activities of the assayed fractions for both the CEX and AEX (including Tween 20 as a stabilizer) pH gradient runs are shown in Fig. 4. Taking the respective fraction sizes into account, activities are given in nmol 4-NP/h in contrast to the units of nmol 4-NP/mL/h used above.

During the CEX '*pH gradient chromatography*' (Fig. 4A), a flowthrough of negatively charged material is seen during the first 5 CV. When determining the enzymatic activity in fractions covering CV₁ to CV₁₀, a total α -Galactosidase A activity of 51.9 ± 15.5 nmol 4-NP/h was found. This indicates that the dynamic binding capacity was reached. At CV₁₃, the pH gradient was started, leading to the elution of bound species. The maximum elution peak with an UV 280 nm value of 9.5 mAU was found at CV₁₉ and pH 5.3. After this peak, the 280 nm signal decreased until the end of the pH gradient at CV₃₅ and pH 8. Enzymatic activity was found in the region from CV₁₆ to CV₂₆ and

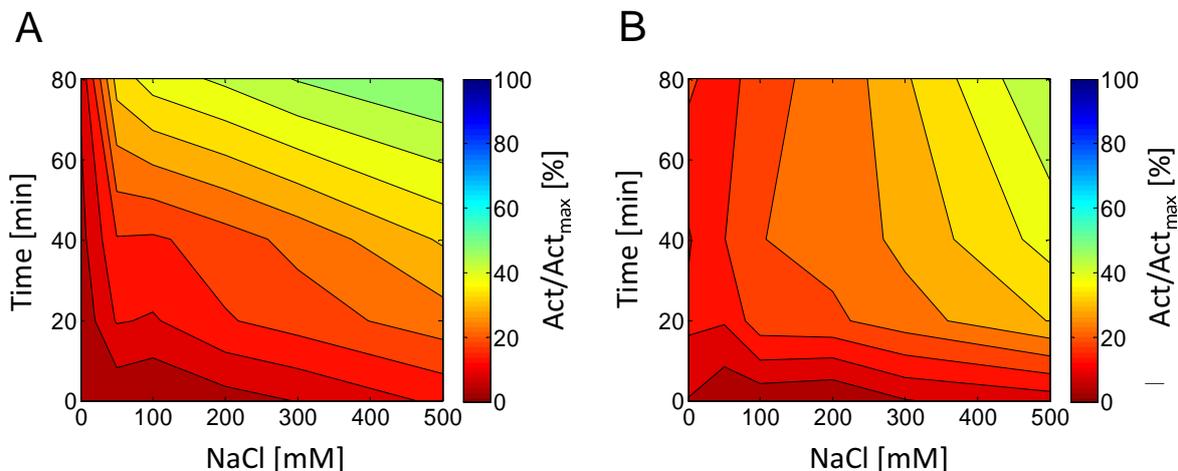


Figure 5: Product salt stability - long-term in AEX and CEX buffers in the sodium chloride concentration range of 0 to 500 mM. Product exposure times of 0 to 80 min were evaluated before starting the 4-NPG assay. Activities are scaled to the maximal activity determined in each test set. A: CEX buffer system at a pH of 4.3 without additional excipients. B: AEX buffer system at a pH of 7.5 with 0.02% Tween 20 detergent added.

at pH values ranging from pH 4.6 to pH 6.7. A maximum of 39.2 ± 2.2 nmol 4-NP/h was reached at CV_{19.5} and pH 5.4. This results in an operational pH window for CEX (declared as 10% of the activity peak maximum) from pH 4.8 to pH 6. During the 1 M NaCl high-salt wash from CV₃₅ to CV₃₉, a final 280 nm peak of 16.6 mAU was observed. Evaluation of the salt fraction showed no major residual enzyme activity.

During the AEX '*pH gradient chromatography*' (Fig. 4B), flowthrough of positively charged material is seen during the first 5 CV. When determining the enzymatic activity in fractions from CV₁ to CV₁₀, a total α -Galactosidase A activity of 7.8 ± 1.2 nmol 4-NP/h was observed. In this setup, less enzyme was detected in the flowthrough compared to the CEX setup. Although the pH gradient started at CV₁₃, the 280 nm signal only started to increase at CV₁₅ and pH 7.5. Beginning at CV₁₅ and pH 7.5, two local maxima at CV₂₈ and pH 4.9 of 17.3 mAU and at CV₃₀ and pH 4.5 of 7.0 mAU were reached. Enzymatic activity was found in the region from CV₁₆ to CV₃₀ and at pH values ranging from pH 7.3 to pH 4.5. the maximum being 19.1 ± 1.6 nmol 4-NP/h at CV_{19.5} and pH 6.7. This results in an operational pH window for AEX (declared as 10% of the activity peak maximum) from pH 7 to pH 4.7. During the 1 M NaCl high-salt wash from CV₃₅ to CV₃₉, a final 280 nm peak of 93.0 mAU was observed. Evaluation of the salt fraction showed a residual summed up enzyme activity of 51.6 ± 5.7 nmol 4-NP/h.

In summary, the operational pH window for CEX was determined from $\text{pH}_{elution,start} = 4.8$ to $\text{pH}_{elution,end} = 6$ and for AEX from $\text{pH}_{elution,start} = 7$ to $\text{pH}_{elution,end} = 4.7$. During both chromatographic procedures using an analytical pH gradient, a broad range of product elution pH values was obtained. This can most probably be explained by glycosylation and charge variants, as *P. pastoris* tends to hyperglycosylate products. Posttranslational modifications of proteins are reported to influence the elution pH of proteins [12].

IV. Salt Stability - Long-term

The traditional '*IEX chromatography - salt elution mode*' requires product stability in the presence of salt at the chosen pH value. The results of the '*salt stability - long-term*' experiments with exposure times of 0 to 80 min are shown in Fig. 5. Operational pH values of 7.5 (AEX) and 4.3 (CEX) were chosen and the salt strength was varied between 0 to 500 mM of sodium chloride. All activities are scaled to the activity maximum determined in the respective setup. After 20 min, enzymatic activity decreased due to enzyme denaturation. This effect increased with increasing salt strength and interaction time. When using the CEX setup (Fig. 5A), an 80 min product exposure to 50 mM sodium chloride already led to a loss in activity of about 32.3%. In contrast to this, the AEX setup (Fig. 5B) was much less influenced by lower added salt concentrations of up to 200 mM sodium chloride over the entire time range. The maximal product inactivation occurred at the highest salt strength of 500 mM tested and an interaction time of 80 min. In this setup, 50.3% (CEX) and 45.2% (AEX) of the product were lost.

V. IEX Chromatography - pH Elution Mode

Based on the results of the '*salt stability - long-term*' experiments, a purification scheme using a salt gradient was no option for the purification of human α -Galactosidase A. As a consequence, a low-salt purification alternative by '*IEX chromatography - pH elution mode*' was applied. Fig. 6A - D shows the results of the pH step and pH gradient elution in the determined operational pH windows (see above). The flowthrough fractions and the high-salt wash fractions will not be discussed any further.

In the CEX setup (Fig. 6A + B), binding of the product occurred at $\text{pH}_{\text{gradient,start}}$ pH 4. Following the binding step, the pH was shifted to pH 4.8 ($\text{pH}_{\text{elution,start}}$). During the stepwise pH shift, the 280 nm signal showed a well-defined peak at CV₇ to CV₈ with an absorption maximum of 35 to 37 mAU. Evaluation of the fractions collected by the 4-NPG assay revealed minor amounts of product. Thus, contaminants close to elution pH of the product were removed. In the next step, elution of the product was performed either by a pH shift in a step (Fig. 6A) or by a 5 CV gradient (Fig. 6B) to pH 6 ($\text{pH}_{\text{elution,end}}$). During the pH shift, the 280 nm signal increased in both setups, resulting in a peak maximum of 33 mAU at CV_{11.5} for the step pH shift and a lower 10 mAU peak maximum at CV_{14.5} for the 5 CV gradient pH shift. The decreased 280 nm maximum at a delayed process time in the 5 CV gradient pH shift setup can be explained by the longer product elution due to an extended time for pH change. The elution fractions were chosen to be CV₁₁ to CV₁₄ with a total activity of 285.2 *pm* 31.9 nmol 4-NP/h for the step and to be CV₁₃ to CV₁₈ with a total activity of 332.4 *pm* 20.7 nmol 4-NP/h for the 5 CV gradient. The binding of the product in the AEX setup (Fig. 6C + D) was performed at pH 8 ($\text{pH}_{\text{gradient,start}}$). A stepwise shift to pH 7 ($\text{pH}_{\text{elution,start}}$) led to a 280 nm peak maximum of 16 to 22 mAU at CV₇ to CV₈. No considerable amounts of product were detected by the 4-NPG assay and only close contaminants were removed. The second pH shift to pH 4.7 ($\text{pH}_{\text{elution,end}}$) for product elution resulted in 280 nm double peak maxima of 47 & 50 mAU at CV₁₃ & CV₁₄ for the step pH shift and lower 25 & 25 mAU peak maxima at CV₁₇ & CV₁₉ for the 5 CV gradient pH shift. The lower 280 nm maxima at a delayed process time in the gradient pH shift setup were discussed previously. The elution fractions were chosen to be CV₁₁ to CV₁₄ with a total activity of 244.0 *pm* 13.4 nmol 4-NP/h for the

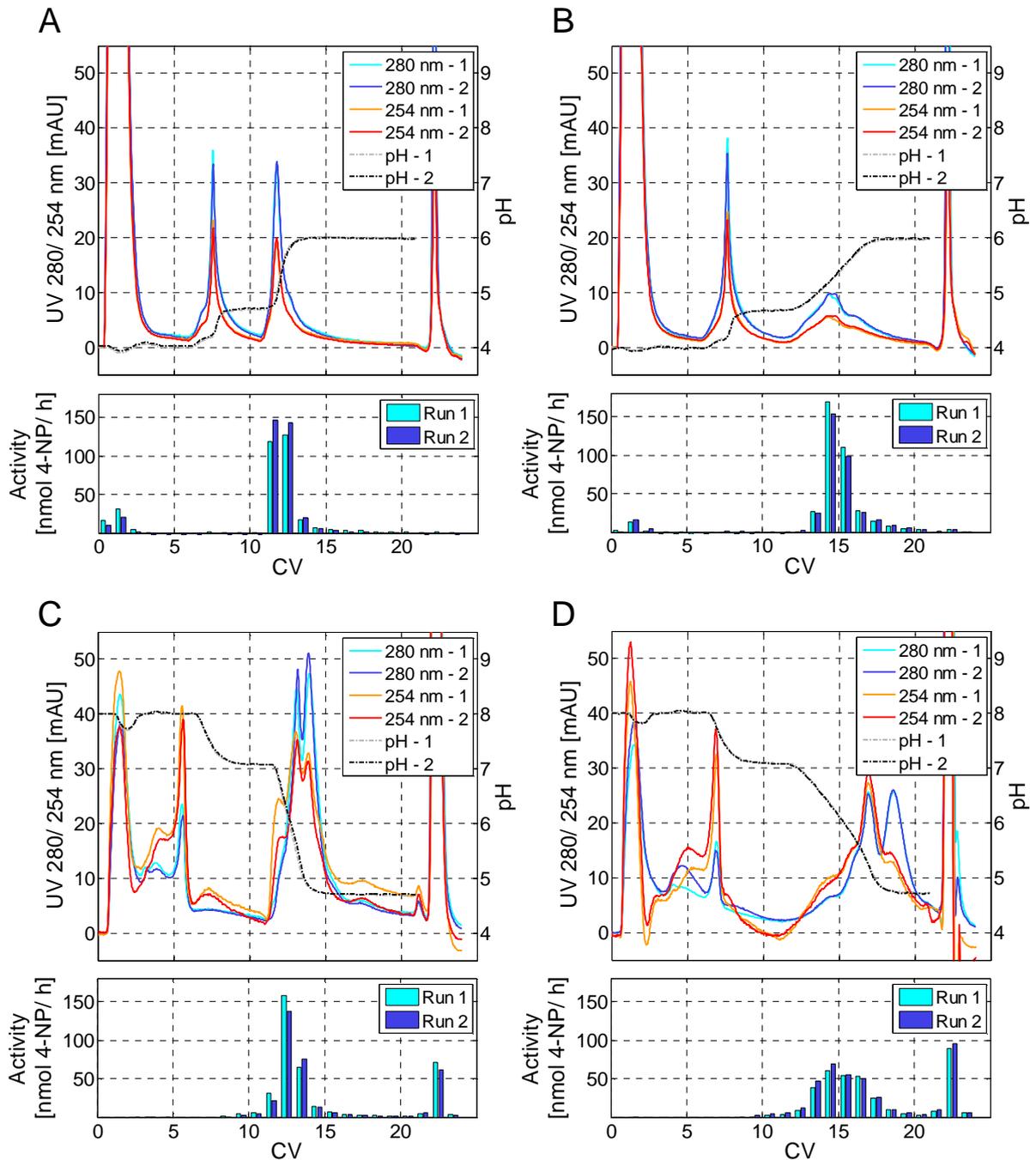


Figure 6: IEX chromatography - pH elution mode runs for the *P. pastoris* supernatant on CEX Source 30 S/ Q columns in duplicates. UV 280/ 254 nm signals and pH values are plotted against column volumes. The corresponding fraction enzyme activities are shown in a bar diagram below. A: Step pH shift for product elution (CEX). B: 5 CV gradient pH shift for product elution (CEX). C: Step pH shift for product elution (AEX). D: 5 CV gradient pH shift for product elution (AEX).

Purification mode	RA [%]	RTP [%]	Y [%]	PF [-]	NF [-]
CEX step	103.9 ± 7.1	94.1 ± 0.3	83.9 ± 10.6	10.9 ± 1.6	0.8 ± 0.0
CEX gradient	107.5 ± 7.2	89.7 ± 1.2	89.9 ± 5.5	12.6 ± 1.4	0.8 ± 0.0
AEX step	98.0 ± 1.1	90.6 ± 3.7	67.8 ± 0.8	4.4 ± 0.0	1.3 ± 0.2
AEX gradient	100.8 ± 0.6	6.9 ± 4.9	65.3 ± 1.1	4.9 ± 0.1	1.5 ± 0.0

Table 1: Recovery rates and quality factors for the different setups in the product capturing by IEX chromatography - pH elution mode. RA and RTP denote the recovery of activity and total protein. Y is the yield of product in the purified elution fractions. PF and NF describe the purification factor and the qualitative nucleic acid reduction, respectively.

step and to be CV_{12} to CV_{18} with a total activity of $249.2 \text{ pm } 13.5 \text{ nmol } 4\text{-NP/h}$ for the 5 CV gradient.

3.3 Analysis of Chromatographic Procedures

I. Mass Balances

Mass balances were determined, as was described in the experimental section. The results are shown in Table 1. For the CEX setups, the recovery of activity RA was found to be $103.9 \pm 7.1\%$ and the recovery of total protein RTP was $94.1 \pm 0.3\%$ for the step elution. For the 5 CV gradient, similar results were observed with an RA of $107.5 \pm 7.2\%$ and RTP of $89.7 \pm 1.2\%$. For the AEX setup, the RA was determined to be $98.0 \pm 1.1\%$ and the RTP was $90.6 \pm 3.7\%$ for the step elution. The 5 CV gradient again produced similar results with an RA of $100.8 \pm 0.6\%$ and RTP of $86.9 \pm 4.9\%$. This indicates that for all processes, no significant product losses were observed. However, minor amounts of impurities remained in the chromatography column or were degraded and not detected.

II. Evaluation of the Purification Setups

For the evaluation of all purification setups, different factors were considered (see experimental section). Yields Y , purification factors PF , and qualitative nucleic acid factors NF are listed in Table 1. All CEX setups resulted in better yields of $83.9 \pm 10.6\%$ for step elution and $89.9 \pm 5.5\%$ for 5 CV gradient elution. The AEX setups produced yields of about 60 to 70%. A similar trend was found for the purification factors. The CEX setup resulted in a PF of 10.9 ± 1.6 for the step and 12.6 ± 1.4 for the 5 CV gradient. For the AEX setup, PF values of 4.4 ± 0.0 and 4.9 ± 0.1 were observed. Hence, they were smaller by factor of 2 to 3. As regards the PF , gradient elution was found to be superior to step elution in all experiments. This can be explained by the broadened elution peak which allows for a more precise and distinct product fractionation. Finally, the qualitative nucleic acid factor NF was evaluated. Whereas the CEX setup was capable of reducing nucleic acid levels by about 20%, AEX chromatography resulted in NF values of above 1. This indicates a qualitative increase in nucleic acid levels. pH dependence of the buffer species or the influence of Tween 20 can be excluded, as blank runs were conducted for all experiments. The increase in NF was thus attributed to nucleic acid fragments attached to the product or co-eluting contaminants. When comparing the pu-

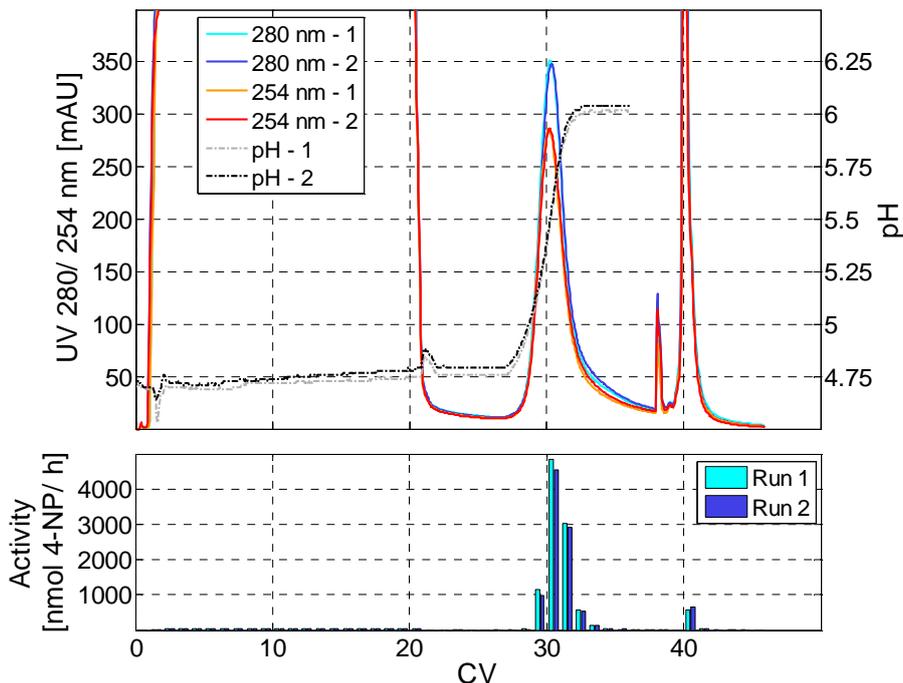


Figure 7: IEX chromatography - pH elution mode runs for the *P. pastoris* supernatant at the capacity limit on a CEX Source 30 S column in duplicates. The complex CEX buffer system was simplified to a two-buffer system using acetate (pK_a 4.76) at pH 4.8 and MES (pK_a 6.10) at pH 6. UV 280/ 254 nm signals and pH values are plotted against column volumes. The corresponding fraction enzyme activities are shown in a bar diagram below.

rification performances of the four purification setups investigated (Table 1), the 'IEX chromatography - pH elution mode' on CEX resins reaches about 20% higher yields as well as purification factors higher by a factor of 3 than the AEX setups. A qualitative reduction of nucleic acid levels was achieved by CEX purification only. The best purification setup for process simplification was found to be the CEX 5 CV gradient elution from pH 4.8 ($pH_{Elution,start}$) to pH 6 ($pH_{Elution,end}$).

3.4 Process Simplification

The complex mixture of buffer substances was reduced to a two-component system using acetate (pK_a 4.76) at pH 4.8 for binding and MES (pK_a 6.10) at pH 6 as the elution buffer. The column capacity limit of 10% product breakthrough was determined to be 20 CV of the loaded sample (data not shown). The 280/ 254 nm signals and the enzyme activities of the assayed fractions of the CEX product capture at the capacity limit are shown in Fig. 7.

The absorption plateau of 490 mAU from CV_1 to CV_{21} describes the flowthrough after sample injection at pH 4.8 ($pH_{Elution,start}$). The evaluation of the fractions by enzymatic assay showed a total activity of 835.3 ± 48.8 nmol 4-NP/h. After the wash step for 8 CV at pH 4.8, the pH was shifted to pH 6 ($pH_{Elution,end}$) using a 5 CV gradient. The 280 nm signal reached a peak maximum of 350 mAU at CV_{30} . The elution fraction was chosen to be CV_{29} to CV_{33} (10% of the maximum enzyme activity as shown in Fig. 1B) with

a total activity of 9284.7 ± 421.5 nmol 4-NP/h. During the 1 M NaCl high-salt wash from CV₃₈ to CV₄₆, a 280 nm peak with a maximum of 2500 mAU was observed. The evaluation of the fractions by enzymatic assay shows a summed up α -Galactosidase A activity of 456.4 ± 59.3 nmol 4-NP/h. The recovery of activity *RA* was determined to be $97.7 \pm 0.5\%$ and the recovery of total protein *RTP* was $92.8 \pm 3.0\%$. These results agree with the previous findings presented in the product capture by pH shift section. A yield of $84.3 \pm 1.2\%$ and a purification factor of 13.2 ± 0.3 were achieved. This agrees with the result obtained for the complex buffer mixture within the standard deviations (Table 1). The qualitative nucleic acid factor *NF* was determined to be 0.9 ± 0.0 . As an additional parameter, the concentration factor *CF* was calculated (see experimental section) to be 4.12 ± 0.08 , indicating an increase in human α -Galactosidase A concentration by more than a factor of 4 compared to the applied sample. It was shown that a process simplification is possible. The pH value is a rather on/ off binding mechanism and is slightly dependent on the buffering species only. However, it is important to keep the total buffering capacity constant.

4 Conclusions

It was demonstrated that optimizing an IEX capture step by product stability screenings and deterministic process development is a powerful tool for protein purification. A systematic strategy for the purification of salt-intolerant proteins in ion-exchange chromatography was evaluated in a case study using human α -Galactosidase A from *P. pastoris*. The 80 min time frame for the determination of long-term stability can be varied. The users of this systematic strategy can extend or shorten the exposure times according to the expected process conditions.

During stages (1) and (2) '*pH stability - short-/ long-term*' of the strategy proposed, pH boundaries of human α -Galactosidase A were determined. For the AEX buffer system, an '*additive screening*' had to be performed to ensure long-term product stability. For this purpose the detergent Tween 20 was used. In a deterministic IEX purification process design the subsequent '*pH gradient chromatography*' (3) resulted in the experimental elution pH of the product. As the enzyme was found to be salt-unstable at its elution pH even at NaCl concentrations of 0.1 M during stage (4) '*salt stability - long-term*', '*IEX chromatography - pH elution mode*' (5b) was the purification method of choice. Different operational modes for pH shift elution were compared, including different IEX modes (CEX vs. AEX) and different setups of elution by pH shift (step vs. 5 CV gradient). The best purification results were obtained in a CEX operational mode with 5 CV gradient elution. Step elution should be chosen especially for proteins that are known to aggregate if eluted at the pH value corresponding to their experimental pI. The critical pH value is then exceeded for a short period of time only.

In a simplified two-buffer component process, human α -Galactosidase A was purified from a *P. pastoris* supernatant based on the results of prior screening experiments. In this setup a purification factor of 13.2 and an overall yield of 84.3% were achieved, which agreed very well with the experimental results for the complex buffer mixture. The product was four times concentrated and the level of nucleic acid content was qualitatively reduced. The results agreed with the experiments using the complex buffer system and

lower sample volumes. The '*IEX chromatography - pH elution mode*' makes IEX purification processes accessible for salt-unstable products and, hence, for a much broader spectrum of products.

The combination of protein stability tests and deterministic screenings should now be extended to cover other purification modes, such as hydrophobic interaction chromatography, affinity chromatography, and aqueous two-phase systems. This will result in a toolbox for process development. For upstream and formulation studies, such an approach may provide valuable information for enhancing the efficiency of these processes.

References

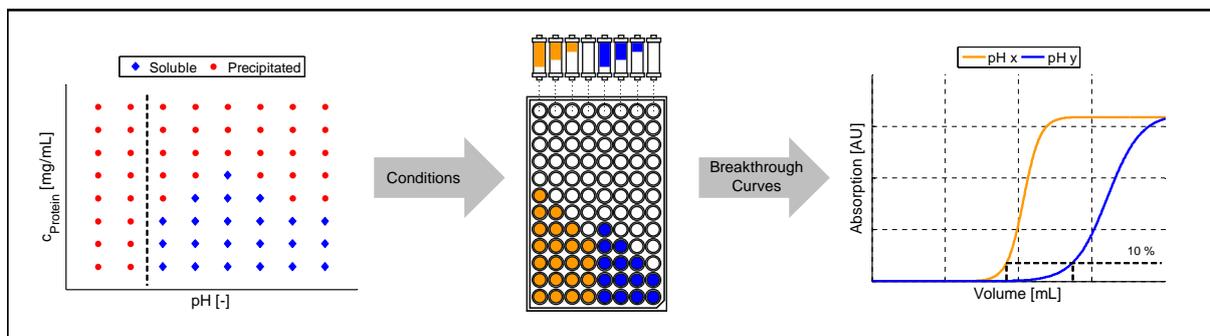
- [1] S. Krishnan, E. Y. Chi, J. N. Webb, B. S. Chang, D. Shan, M. Goldenberg, M. C. Manning, T. W. Randolph, J. F. Carpenter, Aggregation of granulocyte colony stimulating factor under physiological conditions: characterization and thermodynamic inhibition, *Biochemistry* 41 (2002) 6422–6431.
- [2] E. V. A. Y. Chi, S. Krishnan, B. S. Kendrick, B. S. Chang, J. F. Carpenter, T. W. Randolph, Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor, *Protein Sci* 12 (2003) 903–913.
- [3] B. L. Chen, X. Wu, S. J. Babuka, M. Hora, Solubility of recombinant human tissue factor pathway inhibitor, *J Pharm Sci* 88 (1999) 881–888.
- [4] J. Brange, L. Andersen, E. D. Laursen, G. Meyn, E. Rasmussen, Toward understanding insulin fibrillation, *J Pharm Sci* 86 (1997) 517–525.
- [5] A. M. Tsai, J. H. van Zanten, M. J. Betenbaugh, II. Electrostatic effect in the aggregation of heat-denatured RNase A and implications for protein additive design, *Biotechnol Bioeng* 59 (1998) 281–285.
- [6] A. Berg, M. Schuetz, F. Dimer, J. Hubbuch, Automated measurement of apparent protein solubility to rapidly assess complex parameter interactions, *Food Bioproc* 92 (2014) 133–142.
- [7] GE Healthcare, *Ion Exchange Chromatography & Chromatofocusing: Principles and methods*, 2007.
- [8] E. Hallgreen, F. Kalman, D. Farnan, C. Horvath, J. Stahlberg, Protein retention in ion-exchange chromatography: effect of net charge and charge distribution, *J Chrom A* 877 (2000) 13–24.
- [9] P. De Phillips, A. M. Lenhoff, Determinants of protein retention characteristics on cation-exchange adsorbents, *J Chrom A* 933 (2001) 57–72.
- [10] T. Ahamed, B. K. Nfor, P. D. E. M. Verhaert, G. W. K. Van Dedem, L. A. M. Van der Wielen, M. H. M. Eppink, E. J. Van de Sandt, M. Ottens, pH-gradient ion-exchange chromatography: An analytical tool for design and optimization of protein separations, *J Chrom A* 1164 (2007) 181–188.

- [11] F. Kroener, J. Hubbuch, Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application, *J Chrom A* 1285 (2013) 78–87.
- [12] K. Zhu, J. Zhao, D. M. Lubman, F. R. Miller, T. J. Barder, Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins, *Anal Chem* 77 (2005) 2745–2755.
- [13] T. Ahamed, S. Chilamkurth, B. K. Nfor, P. D. E. M. Verhaert, G. W. K. Van Dedem, L. A. M. Van der Wielen, M. H. M. Eppink, E. J. Van de Sandt, M. Ottens, Selection of pH-related parameters in ion-exchange chromatography using pH-gradient operations, *J Chrom A* 1194 (2008) 22–29.
- [14] F. Kroener, A. T. Hanke, B. K. Nfor, M. W. Pinkse, P. D. E. M. Verhaert, M. Ottens, J. Hubbuch, Analytical characterization of complex, biotechnological feedstocks by pH gradient ion exchange chromatography for purification process development, *J Chrom A* 1311 (2013) 55–64.
- [15] B. V. Bhut, K. A. Christensen, S. M. Husson, Membrane chromatography: Protein purification from *E. coli* lysate using newly designed and commercial anion-exchange stationary phases, *J Chrom A* 1217 (2010) 4946–4957.
- [16] M. E. M. Cromwell, E. Hilario, F. Jacobson, Protein Aggregation and Bioprocessing, *AAPS J* 8 (2006) 572–579.
- [17] E. Y. Chi, S. Krishnan, T. W. Randolph, J. F. Carpenter, Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation, *Pharm Res* 20 (2003) 1325–1336.
- [18] W. Wang, Protein aggregation and its inhibition in biopharmaceutics, *Int J Pharm* 289 (2005) 1–30.
- [19] U. B. Ericsson, B. M. Hallberg, G. T. DeTitta, N. Decker, P. Nordlund, Thermofluor-based high-throughput stability optimization of proteins for structural studies, *Anal Biochem* 357 (2006) 289–298.
- [20] J. Jancarik, R. Pufan, C. Hong, S. H. Kima, R. Kim, Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins, *Acta Crystallogr D Biol Crystallogr* 60 (2004) 1670–1673.
- [21] D. Platis, J. Drossard, R. Fischer, J. K. Ma, N. E. Labrou, New downstream processing strategy for the purification of monoclonal antibodies from transgenic tobacco plants, *J Chrom A* 1211 (2008) 80–89.
- [22] Bailon, P. and Ehrlich, G. K. and Fung, W. and Berthold, W., *Affinity Chromatography : Methods and Protocols*, 2000.
- [23] D. Choy, L. Creagh, C. Haynes, Improved isoelectric Focusing Chromatography on Strong Anion Exchange Media via a New Model that Custom Designs Mobile Phases using Simple Buffers, *Biotechnol Bioeng* 111 (2013) 552–564.

-
- [24] X. Kang, J. P. Kutzko, M. L. Hayes, D. D. Frey, Monoclonal antibody heterogeneity analysis and deamidation monitoring with high-performance cation-exchange chromatofocusing using simple , two component buffer systems, *J Chrom A* 1283 (2013) 89–97.
- [25] G. G. K., M. A. Davtyan, A. A. Hambarzumyan, A Simple Purification Procedure of D Amino Acid Oxidase from *Candida guilliermondii*, *Microbiology* 81 (2012) 498–504.
- [26] B. Maiser, F. Kroener, F. Dismar, G. Brenner-Weiss, J. Hubbuch, Isoform separation and binding site determination of mono-PEGylated lysozyme with pH gradient chromatography, *J Chrom A* 1268 (2012) 102–108.
- [27] R. Schiffmann, G. J. Murray, D. Treco, P. Daniel, M. Sellos-Moura, M. Myers, J. M. Quirk, G. C. Zirzow, M. Borowski, K. Loveday, T. Anderson, F. Gillespie, K. L. Oliver, N. O. Jeffries, E. Doo, T. J. Liang, C. Kreps, K. Gunter, K. Frei, K. Crutchfield, R. F. Selden, R. O. Brady, Infusion of alpha-galactosidase A reduces tissue globotriaosylceramide storage in patients with Fabry disease, *Proc Natl Acad Sci USA* 97 (2000) 365–370.
- [28] S. T. De Rezende, V. M. Guimaraes, M. De Castro, C. R. Felix, Purification and Characterization of an alpha-Galactosidase from *Aspergillus fumigatus*, *Brazilian Arch Biol Technol* 48 (2005) 195–202.
- [29] S. Marin, M. E. Guynot, V. Sanchis, A. J. Ramos, Hydrolytic enzyme activities as indicators of fungal spoilage in bakery products, *J Sci Food Agric* 83 (2003) 685–691.
- [30] Y. A. Ioannou, K. M. Zeidner, M. E. Grace, R. J. Desnick, Human Alpha-galactosidase A: glycosylation site 3 is essential for enzyme solubility, *Biochem J* 332 (1998) 789–797.
- [31] S. Zeilinger, D. Kristufek, I. Arisan-Atac, R. Hodits, C. P. Kubicek, Conditions of formation, purification, and characterization of an alpha-galactosidase of *Trichoderma reesei* RUT C-30, *Appl Environ Microbiol* 59 (1993) 1347–1353.
- [32] R. J. Desnick, K. J. Dean, G. Grabowski, D. F. Bishop, C. C. Sweeley, Enzyme therapy in Fabry disease: differential in vivo plasma clearance and metabolic effectiveness of plasma and splenic alpha-galactosidase A isozymes, *Proc Natl Acad Sci USA* 76 (1979) 5326–5330.
- [33] V. A. Werner, Konformerspezifitaet der Proteinphosphatase 2A bei der Dephosphorylierung prolinspezifischer Phosphorylierungsstellen., 2001.
- [34] P. Gagnon, Dissociation of Antibody-Contaminant Complexes With Hydroxyapatite, *Bioprocess J* 9 (2011) 14–24.

Influence of Binding pH and Protein Solubility on the Dynamic Binding Capacity in Hydrophobic Interaction Chromatography

P. Baumann^{1,‡}, K. Baumgartner^{1,‡} and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany

[‡] : Contributed equally to this work

^{*} : Corresponding author. *E-mail-address*: juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

Hydrophobic interaction chromatography (HIC) is one of the most frequently used purification methods in biopharmaceutical industry. A major drawback of HIC, however, is the rather low dynamic binding capacity (DBC) obtained when compared to e.g. ion-exchange chromatography (IEX). The typical purification procedure for HIC includes binding at neutral pH, independently of the protein's nature and isoelectric point. Most approaches to process intensification are based on resin and salt screenings.

In this paper a combination of protein solubility data and varying binding pH leads to a clear enhancement of dynamic binding capacity. This is shown for three proteins of acidic, neutral, and alkaline isoelectric points. High-throughput solubility screenings as well as miniaturized and parallelized breakthrough curves on MediaScout[®]RoboColumns[®] (Atoll, Germany) were conducted at pH 3 to pH 10 on a fully automated robotic workstation. The screening results show a correlation between the DBC and the operational pH, the protein's isoelectric point and the overall solubility. Also, an inverse relationship of DBC in HIC and the binding kinetics was observed. By changing the operational pH, the DBC could be increased up to 30% compared to the standard purification procedure performed at neutral pH. As structural changes of the protein are reported during HIC processes, the applied samples and the elution fractions were proven not to be irreversibly unfolded.

Keywords: Hydrophobic Interaction Chromatography, Dynamic Binding Capacity, Protein Solubility, Binding Kinetics, Multi-variate Data Analysis

1 Introduction

Hydrophobic interaction chromatography (HIC) is a widely applied technique for the intermediate purification and polishing of biomolecules in biopharmaceutical industry. The low dynamic binding capacity compared to e.g. ion-exchange chromatography (IEX) demands for a deeper process understanding and optimization. For IEX, several strategies for process design were reported [1, 2, 3, 4], whereas HIC process development still relies on heuristics and rules of thumb. So far, optimization strategies mostly have been based on the fundamentals of HIC, namely, the salting-out effect of different salts and the stoichiometric displacement of water under entropically favored conditions during protein binding [5, 6]. Considering those two principles, strategic studies mainly focus on screening to identify suitable resins, salt types, and ideal salt concentrations [7].

In [8, 9] the HIC binding mechanism was described in more detail, including additional sub-processes determined by microcalorimetric studies:

- a. Dehydration/ deionization of the protein and adsorber surface.
- b. Van-der-Waals forces between protein and resin.
- c. Structural changes of the protein.
- d. Rearrangement of excluded water molecules in solution.

Especially the structural changes and rearrangements of the protein during the binding process is an important factor to be considered. Various studies have shown, that proteins undergo a partial unfolding during the hydrophobic binding process which can be irreversible especially at higher protein concentrations [10, 11]. In [12, 13], distinct elution peaks were detected which were correlated to the same protein species but of different

conformations. In [14] high protein pore concentrations, as found during the elution step, were found to stabilize the structure of α -lactalbumin. Jones et al. [15] proposed that changes in selectivity for stable proteins can be explained by orientational rearrangement on the adsorber surface, whereas unstable proteins like α -lactalbumin exhibit structural changes during binding to the HIC surface. Both aspects of rearrangement and structural change open up new approaches for HIC optimization strategies. In that context, not only the type of salt or adsorber should be considered, but also the physical state of the protein. Related properties include the protein's compressibility, molecular weight and the solute pH [16].

Although, the pH was identified earlier to be a parameter influencing HIC [17, 18, 19], to the best of our knowledge, the dynamic binding capacity as a function of pH and solubility has never been systematically screened, as was shown for IEX [1, 2, 3]. Still, the gold standard for HIC is using an operational pH in the neutral pH range of pH 6 to 8. Examples are the purification of the cystic fibrosis plasmid vector (operational pH 8), the enrichment of proteins from *Haemophilus influenzae* (pH 7), and the separation of single-, double-stranded, and supercoiled nucleic acids (pH 8) [20, 21, 22]. For antibody purification, varying pH values were considered, resulting in a modified retention behavior. However, the studies covered a range of pH 6 to 8.5 only [18]. Kramarczyk et al. [23] introduced a high-throughput strategy for solubility screenings in dependence of pH, salt type as well as salt concentration and HIC binding experiments for a monoclonal antibody. Nevertheless, all subsequent binding experiments were conducted at the solubility optimum at pH 7, only. A thorough investigation into pH effects as a determining factor for protein solubility and adsorbent interactions under typical HIC conditions is needed. In this study, a high-throughput strategy is developed correlating protein solubility and pH to the dynamic binding capacity (DBC) in HIC. The screening methodology is shown in Fig. 1. Salt type and concentration are kept constant to exclude influences of salt nature and altered isotherm binding effects. High-throughput solubility screenings are performed at pH 3 to pH 10 for proteins of acidic, neutral, and alkaline isoelectric points. pH values of instant protein denaturation are excluded from the following breakthrough curves determined in miniaturized robotic chromatography column experiments. DBCs as well as binding kinetics are subsequently correlated to the protein's pI and solubility data. As protein binding is supposed to vary due to reversible structural changes in the protein, the elution samples of increased DBC need to be investigated for protein integrity. This is done by principal component analysis (PCA) of native and denatured protein spectra of selected setups.

2 Materials & Methods

2.1 Materials

2.1.1 Chemicals & Disposables

The following buffer substances were used: Citric acid monohydrate (Merck, Germany) for pH 3 and 4, sodium acetate trihydrate (Fluka BioChemika, Switzerland) for pH 5, MES monohydrate buffer grade (AppliChem, Germany) for pH 6, sodium dihydrogenphosphate monohydrate (Merck, Germany) for pH 7, TAPS buffer grade (AppliChem,

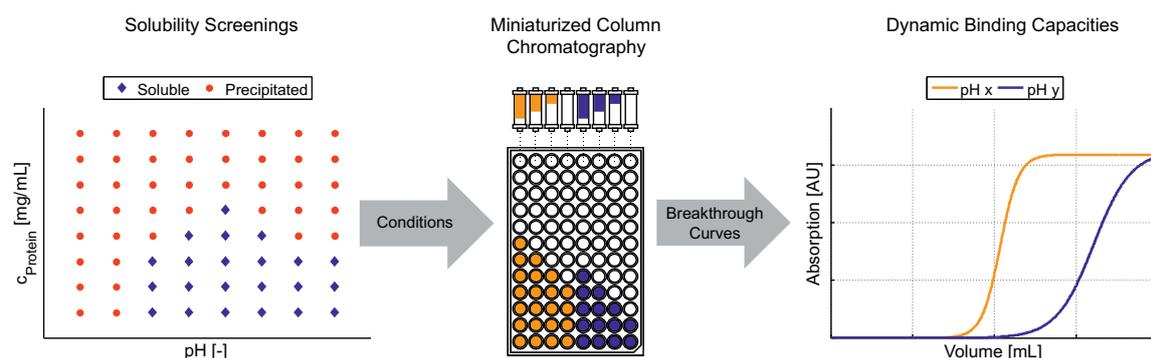


Figure 1: Schematic overview of the HIC pH screening methodology. The solubility screenings determine the feasible pH conditions for the miniaturized column chromatography experiments. Determined breakthrough curves can then be evaluated in terms of dynamic binding capacities, defined as 10% of product breakthrough.

Germany) for pH 8, CHES (AppliChem, Germany) for pH 9, and CAPSO (SantaCruz Biotechnology Inc, USA) for pH 10. All buffers were prepared with a buffer capacity of 40 mM and an additional sodium sulfate (Fluka BioChemika, Switzerland) concentration of 1.875 M. The pH of all buffers was adjusted using hydrochloric acid or sodium hydroxide (Merck, Germany). All buffers were filtered using Supor-450 0.45 μm membrane filters (Pall Life Sciences, Mexico).

Three proteins of different isoelectric points (acidic, neutral, and alkaline) were investigated. All protein solutions were filtered using 0.2 μm syringe filters with PTFE membranes (VWR, Germany). Buffer exchange and protein concentration steps were carried out in Vivaspin[®] centrifugal concentrators (Sartorius, Germany) equipped with PES membranes and molecular weight cutoffs of 5 kDa. Lysozyme from chicken egg white (PDB 1LYZ, HR7-110) was purchased from Hampton Research (USA). A purified single domain antibody was obtained from the industrial partner (BAC, Netherlands) at a concentration of 17.18 mg/mL. Glutathione-S-Transferase (GST) with Cherry-Tag[™] (Delphi genetics, Belgium) was produced in *Escherichia coli* SE1. The cultivation, cell disruption, and purification process was performed according to [24]. Capillary gel electrophoresis (CGE) was performed in a Caliper LabChip[®] GX II system (Perkin Elmer, USA) using an HT Protein Express & Pico LabChip[®] and an HT protein express reagent kit (Perkin Elmer, USA). Sample preparation for the CE was conducted in skirted 96-well twin.tec[®] PCR plates (Eppendorf, Germany). Lysozyme served as an internal standard of known concentration. Protein denaturation as a negative control for protein spectra was accomplished using trichloroacetic acid BioChemika (AppliChem, Germany) combined with acetone for liquid chromatography (Merck, Germany) and urea (Fluka BioChemika, Switzerland).

High-throughput robotic precipitation experiments and UV measurements were carried out in 96-well flat-bottom UV-Star[®] half-area and full-area microplates (Greiner Bio-One, Germany). The plates were sealed with non-sterile Platemax aluminum sealing films (Axygen[®] Scientific, USA). For the high-throughput chromatography experiments, 600 μL MediaScout[®] RoboColumns[®] (Atoll, Germany) filled with Toyopearl[®] Phenyl-650M adsorber (Tosoh Bioscience, Germany) were used. Buffers and protein solutions for the robotic chromatography experiments were prepared in 8-row reservoir plates

(Axygen[®]Scientific, USA).

2.1.2 Instrumentation & Software

For pH adjustment of all buffers, a five-point calibrated HI-3220 pH meter (Hanna Instruments, USA) equipped with a SenTix[®]62 pH electrode (Xylem Inc., USA) was used. The instrument was calibrated using high-precision standards by Hanna Instruments (USA). Concentration and purity measurements of the protein stock solutions were carried out by capillary gel electrophoresis (CGE) in a Caliper LabChip[®]GX II (Perkin Elmer, USA). For data processing and analysis, the LabChip[®]GX 3.1 software (PerkinElmer, USA) was used. A NanoDrop[™] 2000c UV-Vis spectrophotometer operated with the NanoDrop[™] measurement software version 1.4.2 (Thermo Fisher Scientific, USA) served as a fast tool for the determination of protein concentrations. All isoelectric points were calculated using GPMAW (Lighthouse Data, Denmark) for the respective amino acid sequence derived from the protein data bank (PDB).

Solubility screenings and robotic chromatography experiments were conducted on a Freedom EVO[®]200 robotic platform operated with the Freedom EVOware[®]2.5 software (Tecan, Germany). The platform is equipped with a liquid handler with eight fixed pipette tips, a robotic moving arm for transportation of plates, a plate stacker module for storage of plates, a Te-Chrom Bridge for RoboColumns[®], and an orbital shaker. Centrifugations were conducted in an integrated Rotanta 46 RSC centrifuge made by Hettich (Germany). Absorption measurements were performed in an integrated infinite M200 Pro spectrophotometer operated with the software Magellan[™] 7.1 from Tecan (Germany). Principal component analysis (PCA) of protein spectra was performed in Simca (Umetrics, Sweden). Data processing and creation of figures was performed in Matlab[®]R2014a (MathWorks, USA).

2.2 Experimental Setup

As temperature is an influencing factor on protein solubility, pH and the binding behavior in HIC [25, 5] all experiments were performed in a temperature-controlled laboratory (constant temperature of 23 °C). All used buffers were adjusted to respective pH values after equilibration to room temperature. Chromatography columns were stored at room temperature over night for temperature adjustment. Thus, temperature related effects were eliminated.

2.2.1 Protein Solubility Screenings

The protein stock solutions were applied at a concentration of 60 mg/mL in ultra-pure water. Lysozyme was obtained as crystalline powder and was directly dissolved in ultra-pure water. Cherry-GST and the single domain antibody (V_HH), by contrast, were stored in the elution buffer from the purification procedures. Buffer exchanges and concentrating procedures were performed using Vivaspin[®] centrifugal concentrators (Sartorius, Germany) with a cutoff of 5 kDa to the final concentration of 60 mg/mL. With a NanoDrop[™] 2000c by Thermo Fisher Scientific, the protein concentration was determined exactly using absorption coefficients derived from capillary gel electrophoresis (data not shown). Theoretical isoelectric points of the investigated proteins were determined based on the

primary protein structure in the GPMW software. The calculations resulted in $pI = 5.2$ for Cherry-GST, $pI = 7.4$ for the V_HH , and $pI = 10.9$ for lysozyme. Hence, the experiments covered a range of acidic, neutral, and alkaline proteins. As high-salt stock solutions, 40 mM buffers of pH 3 to 10 at a sodium sulfate (NaS) concentration of 1.875 M were prepared.

Protein solubility screenings were carried out on a liquid handling station of the type Freedom EVO[®]200 (Tecan, Germany). A total volume of 300 μL was pipetted into each well of the 96-well UV-Star[®]plate. Each mixture consisted of 200 μL 40 mM high-salt (1.875 M NaS) buffer and 100 μL of protein solution in ultra-pure water (3 times the final protein concentration of the screening), resulting in an effective NaS concentration of 1.25 M. The real concentrations of the applied protein solutions were determined using the NanoDrop[™] system. The microplates were sealed and incubated in an overhead shaker for two hours. The plates were then centrifuged for 10 min at 4000 rpm for removal of potential precipitate. The supernatant was transferred to a 96-well flat-bottom UV-Star[®]full-area plate, measured at 280 nm (414 nm for Cherry-GST) in an Infinite M200 Pro photometer by Tecan and the amount of protein recovery was determined. Data processing to determine the protein phase behavior was performed in Matlab[®]. As a threshold criterion for precipitation conditions, a recovery below 90% was stated. A protein calibration curve in the range from 0 to 1 mg/mL served as a reference.

The initial screenings were performed in a protein concentration range from 0 to 20 mg/mL in 5 mg/mL increments as duplicates. The regions of protein aggregation were subsequently investigated in more detail for each pH in 1 mg/mL steps. For conditions where 20 mg/mL of protein were still soluble, the initial protein concentration was further increased. For those setups, the concentration of the protein stock solution was increased, accordingly. Data points with a relative standard deviation larger than 10% were removed from the evaluation.

2.2.2 Determination of Dynamic Binding Capacities

3 mg/mL of each protein solution were prepared in buffers of the respective pH, including 1.25 M sodium sulfate, as described above and filled into 8-row reservoir plates. The 1.25 M NaS equilibration and wash buffers were applied to the deck of the robotic workstation in the same way.

Dynamic binding capacities were determined using 600 μL Toyopearl[®]Phenyl-650M RoboColumns[®](Atoll, Germany) in an automated liquid handling station (Tecan, Germany). The miniaturized columns were washed with 6 column volumes (CV) of ultra-pure water for removal of the storage solution. Following 12 CV equilibration with 1.25 M high-salt buffer of the respective pH, protein sample loading was performed. During the sample loading procedure, the liquid droplets from the column outlets were fractionated in 96-well UV-Star[®]plates. The fraction volume was 150 μL for experiments with half-area plates and 300 μL for experiments with full-area plates. Each experiment was performed in quadruplicate. For a better curve resolution, sample loading was performed interlaced. Prior to the breakthrough experiment, column 2 was loaded with 0.25 CV, column 3 with 0.5 CV, and column 4 with 0.75 CV of protein solution, resulting in slightly shifted elution pools. A schematic illustration is given in Fig. 1 (center). In total, 15 CV of protein solution were applied to each column for experiments using half-area plates. For

conditions of higher DBCs, a second method for applying 30 CV of protein sample was developed using full-area plates. For both methods, the elution was carried out by applying 3 CV (6 CV for full-area plates) of ultra-pure water after a high salt wash step of 3 CV. After another wash of 6 CV of ultra-pure water, the columns were cleaned with 6 CV of ethanol for storage. As eight robotic columns were used simultaneously, two pH setups could be investigated in one run as an interlaced quadruplicate for a higher resolution.

Data processing and the calculation of the DBCs were performed in Matlab[®]. The variances in sample volume per fractionation due to droplet collection from the RoboColumns[®] were corrected by 990 nm and 900 nm measurements as described in [26]. The real concentrations of the applied protein solutions were determined in the NanoDrop[™] system and were used in the Matlab[®] evaluation software. 10% of protein breakthrough were defined as the DBC threshold. The respective absorption values were determined by a protein calibration curve in the range from 0 to 1 mg/mL. Finally, the DBCs were calculated considering the fraction volumes and column bed volume.

2.2.3 Determination of HIC Binding Kinetics

The measured 280 nm absorption signals of the breakthrough curves were fitted to a logistic function in Matlab[®] as shown in Eq. (1). The fitting was carried out using 'Trust region' optimization in the internal Matlab[®] *curve fitting* toolbox. a , b , and c are the coefficients of the respective fit and V represents the fractionated volume of the droplets.

$$UV_{280nm} = \frac{a}{1 + 10^{b(c-V)}} \quad (1)$$

The maximum of the derivatives of these fits were calculated and used as a measure of the binding kinetics.

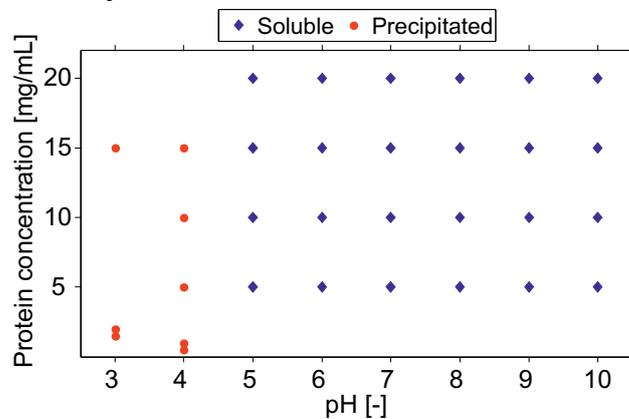
2.2.4 Protein Unfolding Procedures

As HIC protein binding is supposed to be based on reversible partial unfolding, selected samples of increased DBC were investigated for structural integrity after elution by analysis of protein spectra (section 2.2.5). The experiments were performed with lysozyme, as it was obtained with the highest purity (98%) of all investigated proteins to exclude deviations in the spectra due to impurities. Denatured lysozyme spectra were prepared as negative controls. One sample was directly prepared in 6 M urea solution. A second sample was treated with trichloroacetic acid (TCA). For this purpose, 1 mL of 0.3 mg/mL lysozyme solution in ultra-pure water was mixed with 50 μ L of 100% (w/v) TCA solution and stored on ice for 10 min. After a centrifugation cycle at 4000 rpm for 5 min, the supernatant was discarded and the pellet was washed twice with 1 mL of 100% cold acetone. The remaining acetone in the protein pellet was evaporated under a vent for at least 4 h. The dried pellets were then resuspended in 1 mL 6 M urea solution for 2 h at 40 °C.

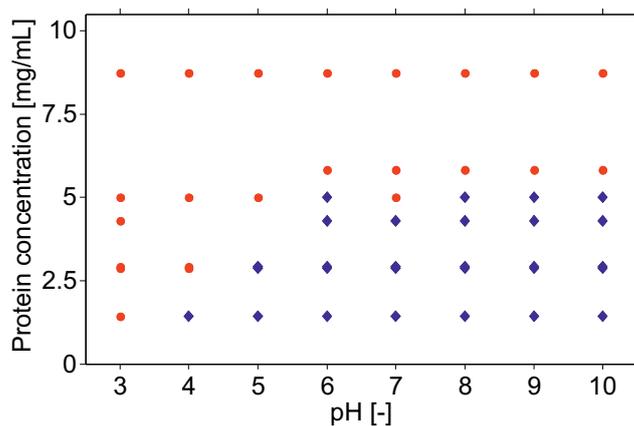
2.2.5 Principal Component Analysis of Protein Spectra

To obtain spectra in a linear range of the spectrophotometer, the protein samples needed to be diluted to a concentration of 0.3 mg/mL. Setups of pH 5, pH 7, and pH 10 were

A: Cherry-GST



B: Single Domain Antibody



C: Lysozyme

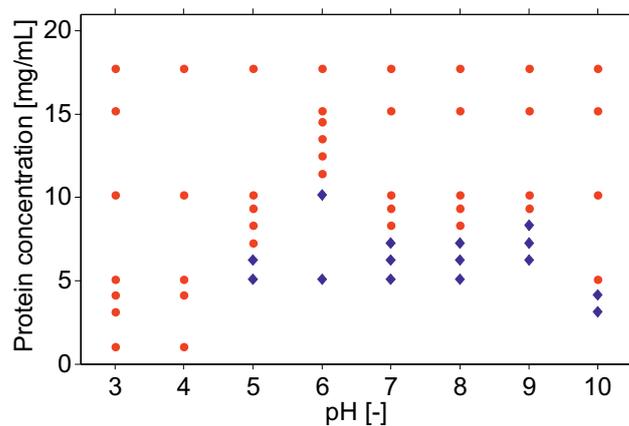


Figure 2: Solubility screenings of three different proteins of acidic, neutral, and alkaline pI in a range from pH 3 to pH 10 in full pH steps, including 1.25 M sodium sulfate. Conditions of soluble proteins are indicated by blue diamonds, precipitated proteins are marked by red circles. A: Cherry-GST (pI = 5.2), B: single domain antibody (V_{HH}) (pI = 7.4), C: Lysozyme (pI = 10.9).

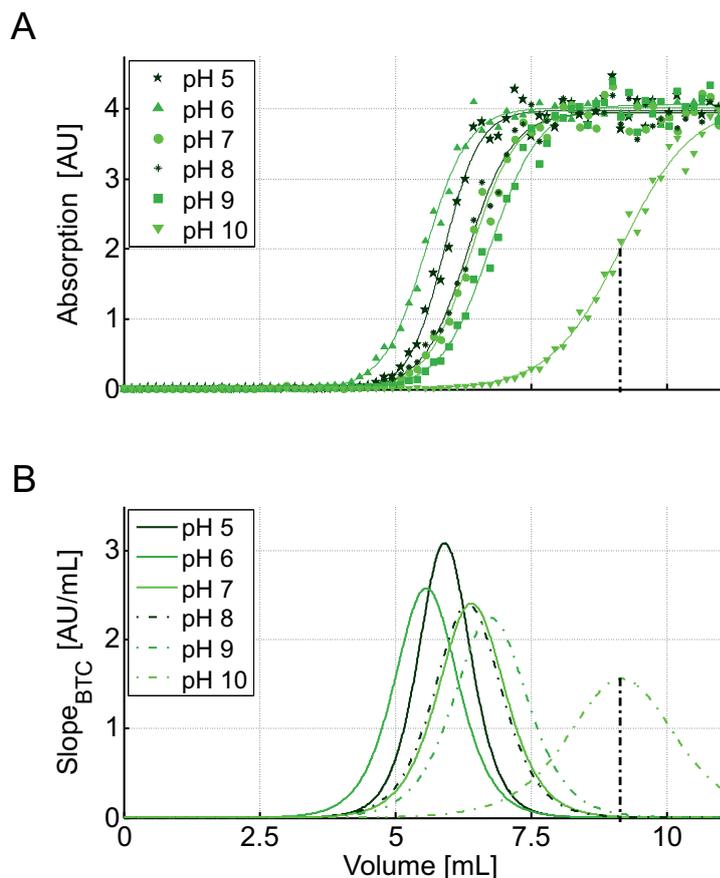


Figure 3: Fitted breakthrough experiments for lysozyme using miniaturized chromatography columns on a robotic workstation (A) and corresponding derivatives of the breakthrough curves (BTCs) (B). The dash-dotted line indicates the maximum of the slope curve, which equals the inflexion point of the BTC shown as an example for pH 10.

investigated. The chosen range covered the boundary conditions of the breakthrough experiments as well as the standard operational neutral pH condition. For each setup, blanked protein spectra were determined in a range from 240 to 300 nm in 2 nm steps. All determined spectra were then analyzed in SIMCA by performing a principal component analysis (PCA). The mean centered data matrix consisted of 31 variables (wavelengths from 240 nm to 300 nm in 2 nm steps) and 9 objects (1 urea and 2 TCA denatured lysozyme samples as well as 6 lysozyme samples before and after chromatography binding at pH 5, 7, and 10). The sample clusters in the score scatter plot were used as a measure of a change in protein integrity.

3 Results

3.1 Protein Solubility Screenings

The results of the protein solubility screenings are illustrated in Fig. 2. All precipitation screenings were carried out in buffers of pH 3 to pH 10 using full pH steps and increasing protein concentrations. The buffers included 1.25 M sodium sulfate. Conditions at which the protein stays soluble are indicated by blue diamonds, whereas conditions of

precipitation are marked by red circles.

For Cherry-GST ($pI = 5.2$), all reaction mixtures of pH 3 and pH 4 resulted in instant product precipitation even at concentrations below 1 mg/mL (Fig. 2A). For all other pH values, the protein was found to be stable even up to concentrations above 80 mg/mL. Starting from pH 11, Cherry-GST showed a behavior analogous to that at pH 3 and pH 4 (data not shown).

The single domain antibody (V_{HH}) ($pI = 7.4$) was not stable above a concentration of 5 mg/mL for all investigated setups (Fig. 2B). pH 3 resulted in an instant precipitation of the product even at a concentration of 1 mg/mL. At pH 4, the V_{HH} was found to be stable up to 1.25 mg/mL. Starting from pH 5, an almost constant solubility limit in a range from 4 to 5 mg/mL was observed. The maximal solubility was achieved for pH 6, 8, 9, and 10 (5 mg/mL), whereas a decline in solubility was determined for pH 7.

Lysozyme (Fig. 2C), as the most alkaline of the investigated proteins ($pI = 10.9$), was not stable at pH 3 and pH 4 analogously to Cherry-GST. The solubility maximum was determined to be up to 10 mg/mL for pH 6 and the protein could be kept soluble up to 6 mg/mL at pH 5. Starting from pH 7, lysozyme was found to be stable in the range from 7 to 8 mg/mL. At pH 10, a sharp decline in solubility was observed, with a maximal soluble protein concentration of 4 mg/mL. For pH 11, lysozyme showed a behavior analogous to that at pH 3 and pH 4 (data not shown).

Due to these results, the DBC experiments were carried out in a range from pH 5 to pH 10 only. The protein concentration for the DBC determinations was set to 3 mg/mL for lysozyme and Cherry-GST. Due to the comparably low solubility of the V_{HH} , the respective experiments were carried out using 2 mg/mL protein solutions.

3.2 Dynamic Binding Capacities (DBCs)

In Fig. 3A the breakthrough curves of lysozyme are exemplarily shown for all investigated pH values. The DBCs were calculated at a breakthrough of 10% of the initial protein concentration, as described above. In Fig. 4A the calculated relative dynamic binding capacities (DBCs) of the three investigated proteins are shown at varying binding pH on the Toyopearl[®]Phenyl-650M adsorber. The DBCs were normalized to the DBC at pH 7 as the standard procedure. For all investigated proteins, the DBCs show a strong pH dependence.

For Cherry-GST (Fig. 4A - red), the minimal DBC was observed for pH 6, with the DBC being $8.2\% \pm 0.8\%$ smaller than under standard pH 7 binding conditions. Towards acidic conditions (pH 5), an increased DBC of $12.8\% \pm 2.4\%$ was observed. Towards the alkaline pH region, higher DBCs were investigated with a maximum increase of $23.4\% \pm 1.3\%$ at pH 9. Overall, the lowest DBCs were found for the neutral pH conditions (pH 6 to 8), while acidic and alkaline buffers increased protein binding by up to more than 23%.

The single domain antibody (Fig. 4A - blue) revealed the highest DBC under the standard operating conditions at pH 7. The lowest DBC was observed for pH 9, the value being $13.8\% \pm 2.4\%$ smaller than under the standard pH 7 binding conditions. Towards acidic as well as alkaline conditions, the DBC decreased compared to pH 7, with the impact being more pronounced in the alkaline region.

For lysozyme from chicken egg white (Fig. 4A - green), the maximal DBC was observed for pH 10 with an increase of $28.3\% \pm 1.0\%$ compared to pH 7. Overall, a trend to-

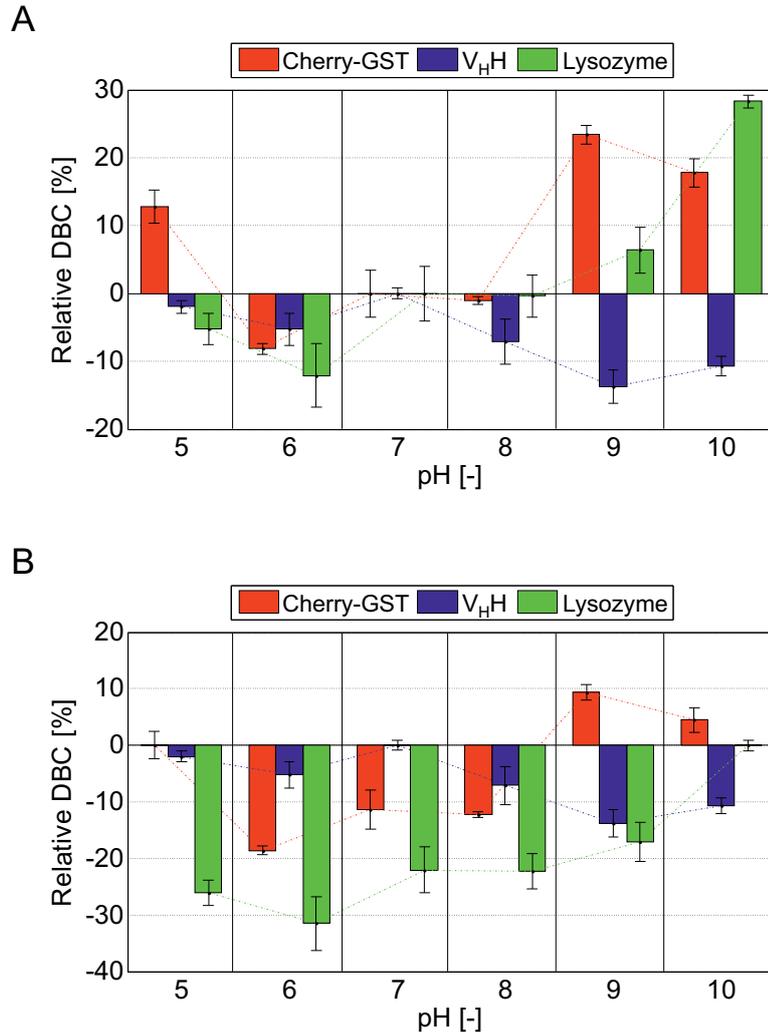


Figure 4: Relative dynamic binding capacities (DBC) of the three investigated proteins: Cherry-GST ($pI = 5.2$) is shown in red, the single domain antibody V_{HH} ($pI = 7.4$) in blue and lysozyme ($pI = 10.9$) in green. A: DBC scaled to the typical operational condition at pH 7. B: DBC scaled to the respective isoelectric points.

wards increased protein binding was found under alkaline operational conditions. This observation is clarified in the chromatograms shown in Fig. 3A. Here, the curves of latest breakthrough correspond to the experiments at pH 9 (squares) and pH 10 (downward triangles). Acidic conditions, by contrast, resulted in a decrease in DBC of up to $12.2\% \pm 4.7\%$ at pH 6. In summary, a continuous increase of DBC was observed for lysozyme from acidic towards alkaline binding conditions on the Toyopearl[®]Phenyl-650M adsorber.

In Fig. 4B the DBCs are scaled to the values of the respective protein pI 's. Except for Cherry-GST at pH 9 and pH 10 all residual conditions yielded in negative relative DBCs. Thus, the isoelectric point was found to be the system point of maximal DBC of almost all investigated conditions.

Fig. 5 shows the correlation between the protein's relative solubility limit (scaled to the solubility maximum) and the corresponding relative DBC (scaled to the maximal investigated DBC) for the single domain antibody (Fig. 5A) and lysozyme (Fig. 5B). Both

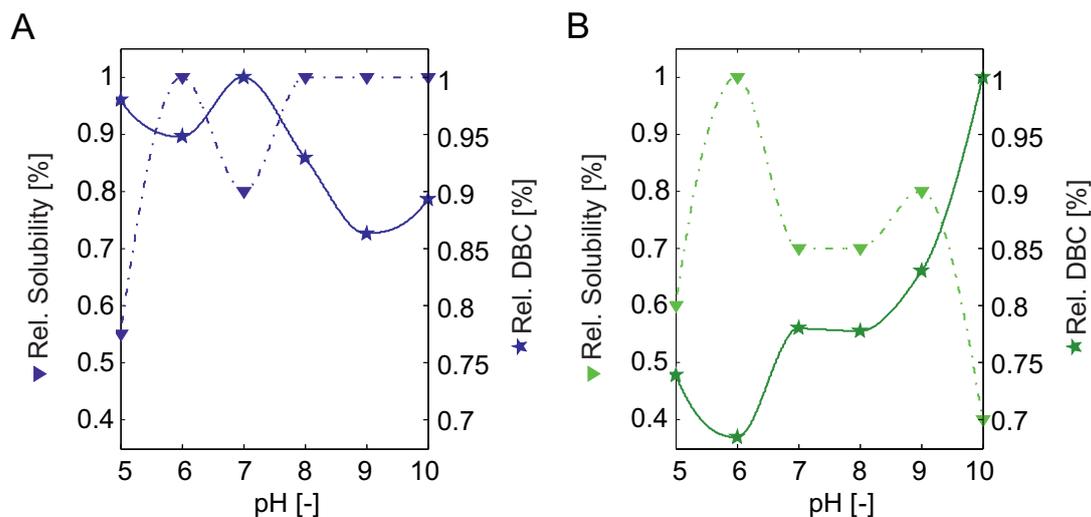


Figure 5: Correlation of relative (scaled to the maximum) protein solubilities (dash-dotted line) and relative (scaled to the maximum) dynamic binding capacities (solid line) of the single domain antibody V_HH (A) and lysozyme (B) in dependence of the operational pH value.

proteins show an inverse trend of relative solubility and relative DBC, most pronounced at pH 6 and pH 10 for lysozyme (Fig. 5B). For Cherry-GST no solubility limit could be determined making such a figure for this protein obsolete.

3.3 Determination of HIC Binding Kinetics

The derivatives of the breakthrough curves (BTCs) are shown as an example for all lysozyme setups in Fig. 3B. The maxima of the derivatives describe the binding kinetics and mark the inflexion points of the BTCs as highlighted by the dash-dotted lines in Fig. 3. In Table 1 the maximal slopes of the fitted breakthrough curves are listed under all investigated conditions. For Cherry-GST, the highest binding kinetics were determined in the range from pH 6 to pH 8 up to 1.26 ± 0.16 AU/mL. At pH 5, the slope was $\approx 30\%$ lower compared to those at neutral pH. At pH 9, the slowest binding kinetics of 0.53 ± 0.04 AU/mL was observed. For pH 10, no fitting to a logistic function was possible, as the kinetics was too slow to reach a final plateau within the defined experimental setup.

For the single domain antibody, the differences in the binding kinetics were negligible between pH 5 and pH 9. In this pH region the minimal slope of 0.61 ± 0.05 AU/mL was obtained for pH 9 and the maximal slope of 0.79 ± 0.06 AU/mL was calculated for pH 8. The only significant variation of the binding kinetics was observed for the pH 10 setup, yielding an overall minimum of 0.43 ± 0.06 AU/mL.

The highest slope for lysozyme of 3.12 ± 0.32 AU/mL was calculated for pH 5 with a continuous decrease towards alkaline conditions. The slowest kinetics was observed for pH 10 with 1.59 ± 0.21 AU/mL.

pH [-]	Kinetics _{Ch-GST} [AU/mL]	Kinetics _{V_HH} [AU/mL]	Kinetics _{Lys} [AU/mL]
5	0.85 ± 0.06	0.76 ± 0.12	3.12 ± 0.32
6	1.24 ± 0.06	0.74 ± 0.08	2.79 ± 0.35
7	1.17 ± 0.24	0.66 ± 0.09	2.46 ± 0.26
8	1.26 ± 0.16	0.79 ± 0.06	2.53 ± 0.19
9	0.53 ± 0.04	0.61 ± 0.05	2.29 ± 0.19
10	n/a	0.43 ± 0.06	1.59 ± 0.21

Table 1: Maximal slopes of the fitted breakthrough curves as a measure of binding kinetics depending on the pH value for the three investigated proteins - Cherry-GST, single domain antibody (V_HH), Lysozyme.

3.4 Investigation of Lysozyme Integrity & Recovery

To investigate samples of increased protein binding for structural integrity after elution, a principal component analysis of protein spectra was conducted. The analysis was performed with lysozyme, as it was the protein of the highest purity (98%) of all investigated biomolecules. Spectra of TCA and, alternatively, urea denatured lysozyme were compared to samples before and after binding to the HIC column. The number of principal components (PCs) describing 99.3% of the system was determined to be 2 (87.3% variance described by PC 1 and 12.0% variance described by PC 2). Creating models of more PCs resulted in no further improvements and measurement noise was found to be included into the model. Fig. 6A illustrates the scatter plot for the score values. The maximal deviance in the score plot of PC 1 was determined for the TCA denatured samples, both located in one isolated cluster. Fig. 6B shows line plots for the loading values of PC 1 and PC 2. Looking at the loading plot for PC 1, those differences were mainly detected in the region from 240 to 248 nm. For PC 2, the score scatter plot revealed a maximal deviation for the urea denatured lysozyme sample. Again, an isolated cluster was observed. The loading plot for PC 2 revealed the wavelength range of 244 to 260 nm to be the most influential one, with a second important region beginning from 284 nm. All remaining samples from the BTC binding experiments merged in one final cluster close to the coordinate origin of the score scatter plot.

In addition to protein integrity studies, lysozyme recoveries were determined in terms of total mass balances for the above-mentioned conditions (pH 5, pH 7, and pH 10). The mass of lysozyme in the elution fraction was divided by the total mass of bound lysozyme until breakthrough. At pH 5 the recovery rate of lysozyme was determined to be 90.7%. For pH 7 94.3% and for pH 10 100.2% of bound lysozyme were recovered.

4 Discussion

4.1 Protein Solubility Screenings

The protein solubility screenings were carried out to identify pH and solubility boundaries suitable to correlate dynamic binding capacities obtained on HIC with systems expressing different protein solubility (Fig. 2). At pH 3 and pH 4, an instant product loss occurred for Cherry-GST, indicating an instability towards acidic conditions induced by protona-

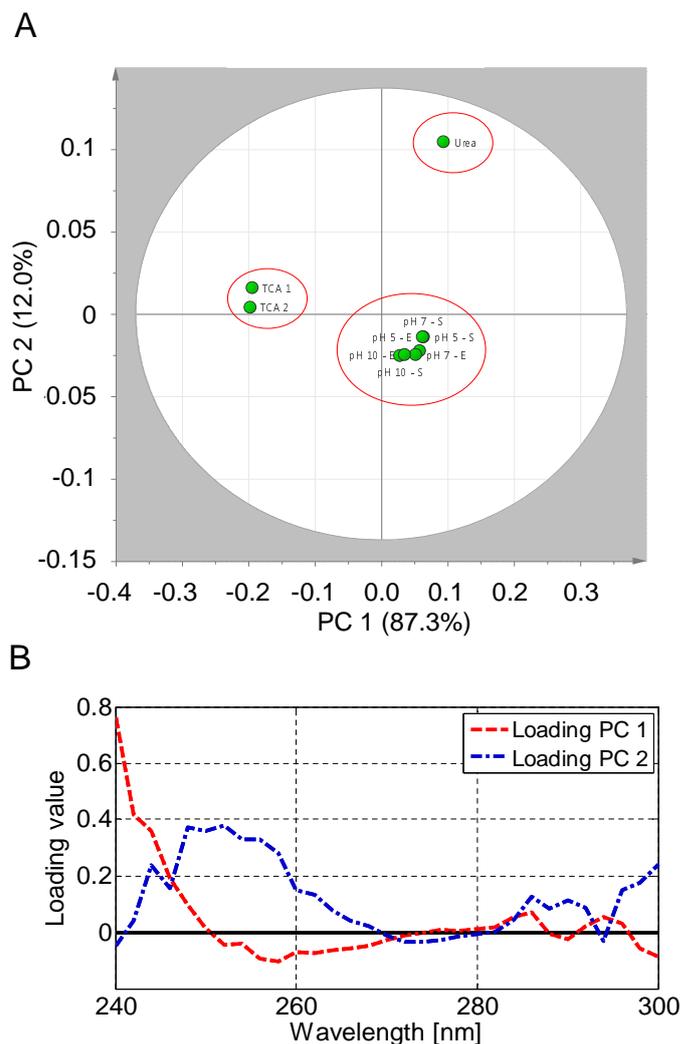


Figure 6: Principal component analysis of lysozyme spectra. Calibration experiments included urea and TCA denatured spectra as well as spectra of samples and elution fractions of the chromatography experiments. A: Scatter plot of the resulting score values for a system of two principal components. B: Line plots of the loading values of the two principal components as a function of the wavelength.

tion of amino acid residues. For all other pH values investigated between pH 5 and pH 10, the protein was stable even at concentrations of up to 60 mg/mL. These findings agree with the reported stabilizing effect of the fused Cherry-TagTM resulting in an increased protein solubility [27, 28].

The single domain antibody (V_HH) showed a much lower solubility over the whole investigated pH range, with a maximum solubility of 5 mg/mL in presence of 1.25 M NaS. Acidic conditions again resulted in instant product loss, as was discussed for Cherry-GST. The reduced protein solubility at pH 7 compared to pH 6 and pH 8 correlates with the isoelectric point of the V_HH of 7.4.

The most alkaline protein investigated, namely lysozyme, again revealed an instant product loss at pH 3 and pH 4. The protonation of amino acid residues showed a strong negative impact on the protein solubility for all investigated protein species. As expected, lysozyme showed a strong decrease in solubility towards alkaline pH conditions

close to its isoelectric point of pH 10.9. The solubility maximum was found in the neutral pH region. Here, lysozyme is stabilized towards amino acid residue protonation, as well as towards aggregation close to the isoelectric point.

When comparing the solubility screenings of all investigated proteins, it can be summarized that strongly acidic conditions (pH 3) result in instant product losses even at protein concentrations below 1 mg/mL. Finally, for all investigated proteins a negative impact of conditions close to the isoelectric point on the protein's solubility was observed.

4.2 Influence of Binding pH Close to the Protein's pI on DBC

In Fig. 4B the DBCs of the different proteins are normalized to the values determined at the respective pI. The acidic protein Cherry-GST (pI = 5.2) revealed an increase in HIC binding towards acidic buffer conditions compared to pH 7 (Fig. 4B - red). However, the maximal DBC was determined for alkaline conditions at pH 9. The single domain antibody (pI = 7.4) as a protein of neutral pI showed the maximal HIC protein binding capacity close to the pI (Fig. 4B - blue). The same results were obtained for lysozyme (pI = 10.9) as the most alkaline protein with a maximal DBC at pH 10 (Fig. 4B - green). In summary, all investigated proteins showed an increase in protein binding towards their pI as illustrated by predominantly negative relative DBCs in Fig. 4B. These findings are in agreement with the pI's definition: As proteins carry a net surface charge of zero at their specified pI, electrostatic repulsions are minimized leading to a domination of hydrophobic forces. This leads to an increased interaction of the protein with hydrophobic ligands of HIC media. As proteins are destabilized close to their respective isoelectric point and structural changes are thus favored, increases in protein binding are most likely based on facilitated partial unfolding in accordance to Jungbauer et al. describing the HIC ligands as a catalyst for partial protein unfolding [10].

4.3 Influence of Protein Solubility on DBC

For Cherry-GST, an increase in DBC was observed under acidic as well as under alkaline conditions. The solubility screenings revealed that Cherry-GST is not stable below pH 5 and above pH 10, indicating an increase in HIC binding towards the solubility boundaries. Those findings match the new insights into the HIC binding mechanism being based on partial protein unfolding upon adsorption. Destabilization of the protein in unfavorable buffers facilitates partial unfolding during the HIC binding process.

Those findings were substantiated by the results obtained for the V_HH , which was most stable under alkaline conditions and revealed the lowest DBCs in this pH region. The HIC binding again increased towards conditions of decreased solubility, namely, acidic pH values and the isoelectric point as illustrated in Fig. 5A.

For lysozyme, as a protein of alkaline pI, the maximal DBCs were determined close to the isoelectric point and at the system point of lowest solubility (Fig. 5B) under the investigated conditions in the miniaturized column chromatography experiments. The lowest protein binding was obtained again at the system point of maximal solubility (pH 6). Starting from pH 6 towards more acidic pH values, protein binding was observed to increase again corresponding to the decline in solubility as illustrated in Fig. 5B.

The maximal DBC was not always found at the respective pI for all investigated proteins.

Although the pI is not generally the optimal system point for HIC protein binding, it may be a good starting point for increasing protein purification productivity. However, there is a correlation between protein solubility and binding behavior in HIC. System points of low protein solubility increase HIC protein binding, whereas conditions of high solubility have the opposite effect (Fig. 5). Those findings again indicate favored protein binding in HIC under destabilized conditions, as was discussed above.

4.4 Correlation of Binding Kinetics and Dynamic Binding Capacities

When comparing the DBCs (Fig. 4A) to the respective binding kinetics (Table 1), a clear inverse trend can be seen for lysozyme and Cherry-GST. Conditions of increased binding in HIC showed a decrease in the binding kinetics. The minimal binding kinetics were observed for the setups resulting in highest capacities and vice versa. For the single domain antibody, those kinetic effects were not that pronounced, as the pH was not a strongly influencing factor in all setups investigated. It can be concluded that the inverse correlation of DBCs and binding kinetics is due to a change of protein orientation during the binding process induced by higher adsorber loadings. This reorientation process slows down protein binding and, thus, decreases the binding kinetics.

4.5 Investigation of Protein Integrity & Recovery

As the determined system points of increased HIC protein binding might be accompanied by structural changes after elution, a spectral protein analysis was conducted. The PCA of the lysozyme spectra resulted in three different score clusters (Fig. 6A). The differently denatured samples formed two distinct score clusters, namely, the TCA cluster (mainly described by PC 1) and the urea cluster (mainly described by PC 2). The two fundamentally different denaturation procedures thus expanded a large area of scoring instabilities of lysozyme. All applied samples and elution pools from the lysozyme DBC experiments merged in one final cluster close to the origin of the score plot, indicating no changes in protein integrity after elution. If increased protein binding as an effect of altered pH was due to partial unfolding of the protein, this would be a reversible process after elution and can be neglected.

As HIC processes are reported to suffer from incomplete product elution [29] the recovery rates of the above-mentioned lysozyme samples were determined. It was shown that all investigated setups yielded in recoveries of 90 - 100% and product losses for conditions of increased DBCs were negligible.

5 Conclusions

The pH and protein solubility are highly influential parameters in HIC and should be considered for process optimization. Choosing an operational pH in the neutral region might be a good condition to choose for some proteins but can also be the reason why HIC can not be employed for some proteins at all until now as shown by [17] for cytochrome c. The screening results show a correlation of the dynamic binding capacity in HIC with

the operational pH value applied. Binding at a pH value close to the protein's isoelectric point as well as conditions near the solubility limit resulted in increased protein binding. These findings indicated a structural change upon binding in HIC, which are favored under destabilizing conditions. However, when comparing lysozyme spectra of the applied samples to the elution fractions in a principal component analysis, native conformation was verified. It can thus be assumed that the induced partial unfolding of lysozyme, if it occurs during binding, is a reversible process under the investigated conditions. To increase the DBCs of HIC adsorbers, choosing a pH close to the protein's pI may be a good starting point for process development. The most influencing factor, however, was shown to be the solubility limit. pH conditions resulting in decreased protein solubility have a positive effect on the binding in HIC.

As an additional factor, the protein binding kinetics were investigated and compared to the protein binding behavior. The kinetics followed an inverse trend compared to the DBCs. The highest DBCs were obtained for the setups of slowest kinetics. Thus, it was concluded that a protein reorientation process took place for high adsorber loadings. This slowed down the binding process.

In summary, different pH-dependent mechanisms like structural changes and protein reorientation upon binding help increase the DBC under varied buffer conditions. As no irreversible structural changes of the protein were investigated without significant product loss during elution for selected setups such an optimization strategy should be exploited for HIC process development.

References

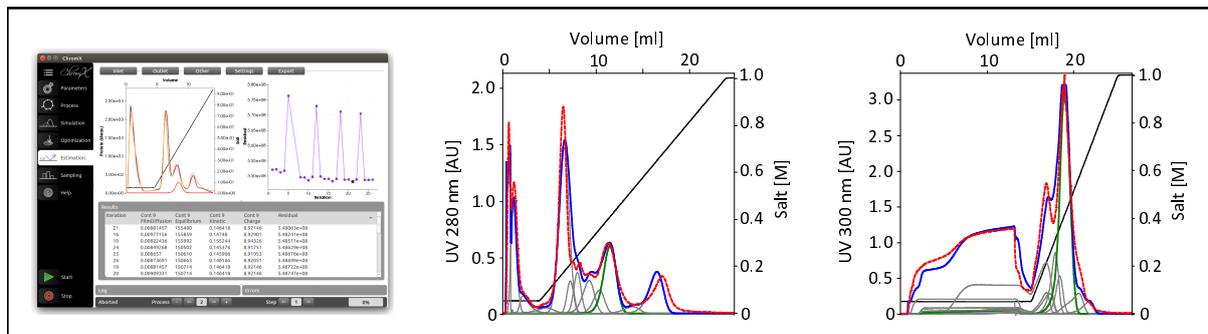
- [1] F. Kröner, D. Elsässer, J. Hubbuch, A high-throughput 2D-analytical technique to obtain single protein parameters from complex cell lysates for in silico process development of ion exchange chromatography, *J Chrom A* 1318 (2013) 84–91.
- [2] F. Kröner, A. T. Hanke, B. K. Nfor, M. W. H. Pinkse, P. D. E. M. Verhaert, M. Ottens, J. Hubbuch, Analytical characterization of complex, biotechnological feedstocks by pH gradient ion exchange chromatography for purification process development, *J Chrom A* 1311 (2013) 55–64.
- [3] T. Ahamed, B. K. Nfor, P. D. E. M. Verhaert, G. W. K. V. Dedem, L. A. M. V. D. Wielen, M. H. M. Eppink, E. J. A. X. V. D. Sandt, M. Ottens, pH-gradient ion-exchange chromatography: An analytical tool for design and optimization of protein separations, *J Chrom A* 1164 (2007) 181–188.
- [4] T. Ahamed, S. Chilamkurthi, B. K. Nfor, P. D. E. M. Verhaert, G. W. K. V. Dedem, L. A. M. V. D. Wielen, M. H. M. Eppink, E. J. A. X. V. D. Sandt, M. Ottens, Selection of pH-related parameters in ion-exchange chromatography using pH-gradient operations, *J Chrom A* 1194 (2008) 22–29.
- [5] J. A. Queiroz, C. T. Tomaz, J. M. S. Cabral, Hydrophobic interaction chromatography of proteins, *J Biotech* 87 (2001) 143–159.
- [6] X. Geng, L. Guo, J. Chang, Study of the retention mechanism of proteins in hydrophobic interaction chromatography, *J Chrom* 507 (1990) 1–23.

- [7] H. P. Jennissen, Hydrophobic interaction chromatography, *Int J Bio Chromatogr* 5 (2000) 131–138.
- [8] H.-M. Huang, F.-Y. Lin, W.-Y. Chen, R.-C. Ruaany, Isothermal Titration Microcalorimetric Studies of the Effect of Temperature on Hydrophobic Interaction between Proteins and Hydrophobic Adsorbents, *J Colloid Interf Sci* 229 (2000) 600–606.
- [9] F.-Y. Lin, W.-Y. Chen, R.-C. Ruaany, H.-M. Huang, Microcalorimetric studies of interactions between proteins and hydrophobic ligands in hydrophobic interaction chromatography: effects of ligand chain length, density and the amount of bound protein, *J Chrom A* 872 (2000) 37–47.
- [10] A. Jungbauer, M. Christine, H. Rainer, Hydrophobic interaction chromatography of proteins III. Unfolding of proteins upon adsorption, *J Chrom A* 1079 (2005) 221–228.
- [11] R. W. Deitcher, J. P. O’Connell, E. J. Fernandez, Changes in solvent exposure reveal the kinetics and equilibria of adsorbed protein unfolding in hydrophobic interaction chromatography, *J Chrom A* 1217 (2010) 5571–5583.
- [12] R. Rosenfeld, K. Benedek, Conformational changes of brain-derived neurotrophic factor during reversed-phase high-performance liquid chromatography, *J Chrom* 632 (1993) 29–36.
- [13] B. L. Karger, B. Rigoberto, The effect of on-column structural changes of proteins on their HPLC behavior, *Talanta* 36 (1989) 243–248.
- [14] J. L. Fogle, J. P. O’Connell, E. J. Fernandez, Loading, stationary phase, and salt effects during hydrophobic interaction chromatography: alpha-Lactalbumin is stabilized at high loadings, *J Chrom A* 1121 (2006) 209–218.
- [15] T. T. Jones, E. J. Fernandez, Hydrophobic Interaction Chromatography Selectivity Changes Among Three Stable Proteins: Conformation Does Not Play a Major Role, *Biotechnol Bioeng* 87 (2003) 388–399.
- [16] B. C. S. To, A. M. Lenhoff, Hydrophobic interaction chromatography of proteins I. The effects of protein and adsorbent properties on retention and recovery, *J Chrom A* 1141 (2007) 191–205.
- [17] S. Hjerten, K. Yao, K.-O. Eriksson, B. Johansson, Gradient and isocratic high-performance hydrophobic interaction chromatography of proteins on agarose columns, *J Chrom* 359 (1986) 99–109.
- [18] J. Chen, J. Tetrault, A. Ley, Comparison of standard and new generation hydrophobic interaction chromatography resins in the monoclonal antibody purification process, *J Chrom A* 1177 (2008) 272–281.
- [19] E. M. Lienqueo, A. Mahn, C. J. Salgado, J. A. Asenjo, Current insights on protein behaviour in hydrophobic interaction chromatography, *J Chrom B* 849 (2007) 53–68.

-
- [20] M. M. Diogo, J. A. Queiroz, M. G. A., G. N. M. Martins, S. A. M. Ferreira, P. D. M. F., Purification of a Cystic Fibrosis Plasmid Vector for Gene Therapy Using Hydrophobic Interaction Chromatography, *Biotechnol Bioeng* 68 (1999) 576–583.
- [21] M. Fountoulakis, M.-F. Takacs, B. Takacs, Enrichment of low-copy-number gene products by hydrophobic interaction chromatography, *J Chrom A* 833 (1999) 157–168.
- [22] M. M. Diogo, J. A. Queiroz, D. M. F. Prazeres, Studies on the retention of plasmid DNA and *Escherichia coli* nucleic acids by hydrophobic interaction chromatography, *Bioseparation* 10 (2001) 211–220.
- [23] J. F. Kramarczyk, B. D. Kelley, C. J. L., High-Throughput Screening of Chromatographic Separations: II. Hydrophobic Interaction, *Biotechnol Bioeng* 100 (2008) 707–720.
- [24] P. Baumann, N. Bluthardt, S. Renner, A. Osberhaus, J. Hubbuch, Integrated Development of Up- and Downstream Processes Supported by the Cherry-Tag for Real-time Tracking of Stability and Solubility of Proteins, *J Biotech* 200 (2015) 27–37.
- [25] A. Werner, E. Hackemann, H. Hasse, Temperature dependence of adsorption of PEGylated lysozyme and pure polyethylene glycol on a hydrophobic resin: Comparison of isothermal titration calorimetry and van't Hoff data, *J Chrom A* 1356 (2014) 188–196.
- [26] E. L. McGown, H. D. G., Multichannel pipettor performance verified by measuring pathlength of reagent dispensed into a microplate, *Anal Biochem* 258.1 (1998) 155–157.
- [27] Delphi Genetics SA, Cherry Codon Kit., 2009.
- [28] Delphi Genetics SA, Cherry Express Kit., 2009.
- [29] R. Hahn, K. Deinhofer, C. Machold, A. Jungbauer, Hydrophobic interaction chromatography of proteins: II. Binding capacity, recovery and mass transfer properties, *J Chrom B* 790 (2003) 99–114.

UV Absorption-based Inverse Modeling of Protein Chromatography

T. Hahn¹, P. Baumann¹, T. Huuk¹, V. Heuveline² and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany

² : Engineering Mathematics and Computing Lab, Heidelberg University, Heidelberg, Germany

* : Corresponding author. *E-mail-address*:juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

UV absorbance measurements play an important role in bioprocess development. Yield and purity are often evaluated in terms of peak percentages in analytical size exclusion chromatography or ion-exchange chromatography. Also, industrial chromatography steps are usually controlled based on UV data with pooling decisions according to absorbance thresholds.

Model-based process development would make elaborate screening experiments redundant, once the model has been calibrated to the specific process step. So far, absorbance measurements could not be used directly for modeling chromatography steps as the commonly applied models rely on mass or molar concentration. This study presents mechanistic modeling of an industrially relevant chromatography setting without any knowledge of the feed composition. The model equations were re-written to employ boundary conditions in UV absorbance units, the absorption coefficients were shifted into the isotherm, and standard parameter estimation procedures could be applied. An anion-exchange chromatography case study of a target protein expressed in *Escherichia coli* and eleven lumped impurity peaks demonstrated practical applicability. The target protein concentration in the feed material was estimated from chromatograms. Using this method, initially unknown feed concentrations can be determined a posteriori for ion-exchange and multi-modal chromatography from single-component absorbance curves.

Keywords: Calibration-free, Inverse Method, Liquid Chromatography, Mechanistic Modeling, UV Absorbance

1 Introduction

Industrial downstream processing (DSP) faces the challenge of efficiently purifying a product out of a very heterogeneous mixture. The purification sequence is commonly based on platform processes that are only slightly adapted to new target components to accelerate process development [1, 2]. Prior to this, high-throughput screening methods are often used to find promising initial conditions for platform processing [3]. Mechanistic modeling is a favorable alternative, provided the model parameters can be determined with less effort, allowing identification of optimal process parameters *in silico*.

The common models in liquid chromatography describe the mass transport in the column by so-called Convection Diffusion Reaction (CDR) equations, where the reaction term models phase transitions and, eventually, the retention of the species. With no a priori knowledge about the components' behavior, the inverse method is a suitable option which alters parameters in a systematic fashion to achieve a match of the recorded chromatogram and the model prediction.

The potential of mathematical modeling and numerical optimization of chromatography has already been demonstrated in academic set-ups for mixtures of small molecules [4, 5], model proteins [6] and antibodies [7, 8], and also for industrial process steps [9, 10]. All applications have in common, that molar concentrations in the feed were known and sensor calibrations existed for all components.

Cornel et al. [11] determined absorption coefficients for a two-component mixture directly from the simulated concentration curves by choosing the best fitting values in each

Parameter	Symbol	Value	Proceeding
Diameter	d	7 mm	From manufacturer
Length	L	25 mm	From manufacturer
Bead radius	r_p	0.045 mm	From manufacturer
System dead volume	V_d	0.07 mL	Acetone injection without column
Ret. volume Acetone	V_{RetAc}	0.96 mL	Acetone peak injection with column
Ret. volume dextran	V_{RetDex}	0.34 mL	Dextran peak injection with column
Std. dev. of dextran	σ_{Dex}	0.029 mL	ÄKTA TM peak integration
Volume of HCl	V_{HCl}	1.48 mL	Acid/base titration
Molarity of HCl	c_{HCl}	0.01 M	Manually controlled
Flow rate	u	0.2 $\frac{mm}{s}$	Manually controlled

Table 1: Measured column parameters.

iteration of the estimation procedure. No sensor calibration was necessary, but mass concentrations in the feed were known a priori.

In the following sections, we describe a mechanistic modeling approach for an industrially relevant chromatography setting that does not require prior knowledge of the feed composition in terms of molar or mass concentrations. We re-write the model equation for boundary conditions in UV absorbance units and aim at determining the unknown feed concentrations a posteriori by taking advantage of the particular structure of stoichiometric exchange models. Standard parameter estimation procedures can be applied if single-component absorption curves are available.

A case study based on an anion-exchange chromatographic (AEX) process step (Q Sepharose FF, GE Healthcare) demonstrates the applicability. The mixture fed into AEX is a crude feed stock of *Escherichia coli* SE1, including the Cherry-tagged enzyme Glutathione-S-Transferase as the product.

2 Materials & Methods

2.1 Column Parameter Determination

To model the mass transport in the chromatography system, the column properties listed in Tables 1 and 2 must be determined by pulse injections of non-interacting tracer molecules [12]. A 1 mL column (effective volume 0.962 mL), with Q Sepharose Fast Flow Resin (GE Healthcare, Germany) was analyzed firstly with an 1% acetone (Merck, Germany) pulse and secondly with a dextran pulse from *Leuconostoc spp.* MW 2,000,000 (Sigma Aldrich, Germany) using an ÄKTATM Purifier system (GE Healthcare, UK) controlled with Unicorn 5.2 (GE Healthcare, Sweden) to determine the essential system parameters [12] presented in Table 1. Acid-base titration was carried out to determine the total ionic capacity: the column was flushed with a 0.5 M NaOH solution (Merck, Germany) until a constant UV and conductivity signal was achieved. Afterwards, the column was washed with ultra pure water until a constant UV and conductivity baseline was reached. Then, the column was titrated at a flow of 0.64 mL/min with a 0.01 M

Parameter	Symbol	Value	Proceeding
Volume	V	0.962 mL	$\frac{1}{4} \cdot \pi \cdot d^2 \cdot L$
Fluid volume	V_f	0.89 mL	$V_{RetAc} - V_d$
Interstitial volume	V_{int}	0.27 mL	$V_{RetDex} - V_d$
Total column porosity	ε_t	0.925	$\frac{V_f}{V}$
Interstitial porosity	ε_b	0.28	$\frac{V_{int}}{V}$
Particle porosity	ε_p	0.896	$\frac{V_f - V_{int}}{V - V_{int}}$
Interstitial flow	u_{int}	0.714 $\frac{mm}{s}$	$\frac{u}{\varepsilon_b}$
Axial dispersion	D_{ax}	0.1 $\frac{mm^2}{s}$	$\frac{1}{2} \cdot L \cdot u_{int} \cdot \left(\frac{\sigma_{Dex}}{V_{int}} \right)^2$
Ionic capacity	Λ	0.22 M	$\frac{c_{HCl} \cdot V_{HCl}}{V(1-\varepsilon_b)(1-\varepsilon_p)}$

Table 2: Calculated column parameters.

HCl solution (Merck, Germany) until an increase in the conductivity signal was recorded. From the Cl^- ion concentration of the titrant and the volume of the applied titrant, the total number of exchangeable ions was calculated. All chemicals used were obtained in highest quality.

With this set of parameters, all system-specific parameters occurring in the mathematical model can be fixed as given in Table 2.

2.2 Sample Production

The applied sample consisted of an *Escherichia coli* SE1 lysate, including Cherry-tagged Glutathione-S-Transferase (GST) as a product. The Cherry-TagTM, which can be fused to any target protein, allows for straightforward product analytics by VIS absorbance measurements [13]. The cultivation was performed for 24 h in 800 mL standard TB (terrific broth) medium at 37 °C and 180 rpm rotational speed in 2.5 l TUNAIRTM flasks (Sigma Aldrich, Germany). Cell disruption was performed by sonication of the cell pellet in 20 mL of 50 mM Tris buffer (pH 8), including 1× SigmaFASTTM protease inhibitor (Sigma Aldrich, Germany), in a Branson Digital Sonifier[®] 450 (70% pulse amplitude, 10 × 15 s pulse duration, 30 s resting on ice between pulses). The lysate was centrifuged at 12000 rpm for 60 min at 10 °C using a 5810 R centrifuge (Eppendorf, Germany) followed by a second clarification step using 0.2 μm sterile PES filters (VWR, Germany). Finally, the permeate was 10 times diluted in 50 mM Tris buffer (pH 8).

2.3 Sample Characterization

For comparison, the product concentration was determined in the Caliper LabChip[®] GX II capillary gel electrophoresis system with LabChip GX 3.1 software (Perkin Elmer, USA). The HT Protein Express and Pico LabChip[®] was run with the HT Protein Express LabChip[®] reagent kit using the HT Protein Express 200 assay. Cherry-tagged GST was identified using the sample ladder from the reagent kit. The product was quantified by peak-baseline integration of the fluorescence signals (Fig. 1) and scaling to an external lysozyme protein standard of 1 mg/mL (Hampton Research, USA).

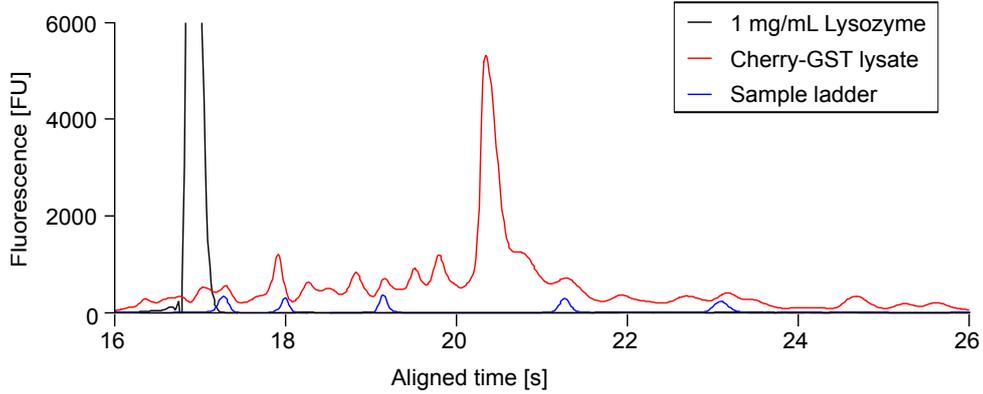


Figure 1: Capillary gel electrophoresis analysis of feed material.

2.4 Bind-elute Experiments

The component-specific isotherm parameters are determined from bind-elute experiments. The general approach is identical to concentration-based parameter estimation [6, 7, 8]. A 50 mM Tris-HCl buffer, pH 8 was employed as the mobile phase during binding and the same buffer supplemented with 1 M NaCl was used for elution. Different salt gradients were generated from these two buffers. After the 0.5 mL sample was injected (12.4 mL for the breakthrough experiment), the column was washed with low-salt buffer for 3 mL of 50 mM Tris-HCl buffer, before initiating linear gradient (0 to 1 M NaCl) elution over 5, 10, 15 and 20 mL. On gradient completion the columns were irrigated with 2 mL of 50 mM Tris-HCl and 1 M NaCl before re-equilibrating with 5 mL of 50 mM Tris-HCl buffer. The linear phase linear velocity employed was 0.2 mm/s throughout.

2.5 Parameter Estimation

In general, estimation of an unknown parameter set \bar{p} solves the least squares optimization problem

$$\min_{\bar{p}} \sum_j \left(m(t_j) - \sum_{i \geq 1} c_i(L, t_j; \bar{p}) \cdot a_i \right)^2, \quad (1)$$

where $m(t_j)$ is the measured chromatogram value at time t_j , typically given in milli absorbance units [mAU]. This measurement might also contain noise, which can be neglected when assuming the noise to be zero-mean Gaussian and isotropic.

$c_i(L, t)$ is the simulated mass or molar concentrations at the outlet of the column with length L . The transformation into absorbance units is performed with a scaling factor a_i . According to Beer's law, the absorption coefficient a_i consists of an extinction coefficient and UV cell path length. It is unknown in this case.

2.6 Chromatography Model

The Transport Dispersive Model (TDM) [12] in Eqs. (2) and (3) is used to model the macroscopic protein transport through the column. For simplicity, the component-specific internal and external diffusion effects are lumped in an effective mass transfer coefficient.

The system is of Convection Diffusion Reaction (CDR) type. Eq. (2) describes the rate of change of the concentration $c_i(x, t)$ of component i in the interstitial volume of a column with length L , which consists of convective mass transport in space with the average interstitial velocity of the fluid u . Peak broadening effects are modeled as dispersion in axial direction with respect to a coefficient D_{ax} . The exchange between the interstitial concentration and the particle pore concentration $c_{p,i}(x, t)$ depends on the porosity of the bed ε_b , the radius of adsorber particles r_p , and a component-specific effective mass transfer coefficient $k_{eff,i}$. The model is one-dimensional in space, such that the concentrations depend on the axial position in the column and time. Eq. (3) models the accumulation of mass in the pore volume $c_{p,i}$ and stationary phase q_i as a function of the particle porosity ε_p . The model is complemented by Danckwert's boundary conditions Eqs. (4) and (5), including the applied inlet concentration $c_{in,i}$, and an isotherm equation modeling the stationary phase concentration q_i .

$$\frac{\partial c_i}{\partial t} = -u(t) \frac{\partial c_i}{\partial x} + D_{ax} \frac{\partial^2 c_i}{\partial x^2} - \frac{1 - \varepsilon_b}{\varepsilon_b} k_{eff,i} \frac{3}{r_p} (c_i - c_{p,i}) \quad (2)$$

$$\frac{\partial c_{p,i}}{\partial t} = -\frac{1 - \varepsilon_p}{\varepsilon_p} \frac{\partial q_i}{\partial t} + k_{eff,i} \frac{3}{\varepsilon_p r_p} (c_i - c_{p,i}) \quad (3)$$

$$\frac{\partial c_i}{\partial x}(0, t) = \frac{u(t)}{D_{ax}} (c_i(0, t) - c_{in,i}(t)) \quad (4)$$

$$\frac{\partial c_i}{\partial x}(L, t) = 0 \quad (5)$$

The steric mass-action isotherm (SMA) [14] is a commonly used semi-mechanistic isotherm in ion-exchange chromatography. It is capable of reproducing the influence of counter ions on the retention behavior of protein species using the proteins' characteristic charges ν_i . In addition, it considers adsorber properties such as the total ionic capacity Λ and steric shielding effects σ_i of the protein covering an amount of binding sites, greater than the actual number of sites it interacts with. The kinetic SMA isotherm is given in Eq. (6) for k proteins, with q_i and $c_{p,i}$ being the concentration of the protein $i \in \{1, \dots, k\}$ adsorbed and in solution, respectively. $c_{p,salt}$ is the salt concentration of the solution. $k_{ads,i}$ and $k_{des,i}$ are the constants of the adsorption and desorption rate.

$$\frac{\partial q_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) q_j \right)^{\nu_i} c_{p,i} - k_{des,i} c_{p,salt}^{\nu_i} q_i \quad (6)$$

$$q_{salt} = \Lambda - \sum_{j=1}^k \nu_j q_j \quad (7)$$

The model is chosen because of its capability of simulating the whole chromatographic process, including elution, by changing the induced salt concentration at the inlet. The model is based on molar concentrations, such that the boundary conditions of the TDM must be set in terms of molarities.

2.7 Transformation

As the SMA model is based on molar concentrations $[M]$, interstitial and pore volume concentrations must be given in $[M]$ as well. This also applies to the boundary conditions. Here, the exact molar concentrations in the feed are unknown, as are the scaling factors for UV absorbance. We will re-write the equations to directly incorporate UV absorbance values.

First, the injected protein concentrations are transformed into absorbance values:

$c'_{in}[mAU] = a[mAU/M] \cdot c_{in}[M]$. These can be determined later from the respective peak area in the chromatogram.

The equations for the interstitial and pore volume as well as the boundary conditions are linear in c . These can be multiplied by a to obtain

$$a_i \frac{\partial c_i}{\partial t} = a_i \left[-u \frac{\partial c_i}{\partial x} + D_{ax} \frac{\partial^2 c_i}{\partial x^2} - \frac{1 - \varepsilon_c}{\varepsilon_c} k_f (c_i - c_{p,i}) \right] \quad (8)$$

$$\iff \frac{\partial c'_i}{\partial t} = -u \frac{\partial c'_i}{\partial x} + D_{ax} \frac{\partial^2 c'_i}{\partial x^2} - \frac{1 - \varepsilon_c}{\varepsilon_c} k_f (c'_i - c'_{p,i}) \quad (9)$$

$$\frac{\partial c'_i}{\partial x}(0, t) = \frac{u}{D_{ax}} (c'_i(0, t) - c'_{in,i}(t)) \quad (10)$$

$$\frac{\partial c'_i}{\partial x}(L, t) = 0 \quad (11)$$

We obtain equations for $c'_i = a_i \cdot c_i$ that use $c'_{p,i} = a_i \cdot c_{p,i}$. This is calculated from the scaled lumped rate model

$$\varepsilon_p \frac{\partial c'_{p,i}}{\partial t} + (1 - \varepsilon_p) \frac{\partial q'_i}{\partial t} = k_{eff,i} \frac{3}{r_p} (c'_i - c'_{p,i}). \quad (12)$$

Again, we require an equation for $q'_i = a_i \cdot q_i$. Scaling the kinetic SMA formulation yields

$$a_i \frac{\partial q_i}{\partial t} = a_i \left[k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) q_j \right)^{\nu_i} c_{p,i} - k_{des,i} c_s^{\nu_i} q_i \right] \quad (13)$$

$$\iff \frac{\partial q'_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) q_j \right)^{\nu_i} c'_{p,i} - k_{des,i} c_s^{\nu_i} q'_i \quad (14)$$

$$\iff \frac{\partial q'_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) \frac{a_j}{a_j} q_j \right)^{\nu_i} c'_{p,i} - k_{des,i} c_s^{\nu_i} q'_i \quad (15)$$

$$\iff \frac{\partial q'_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k \frac{\nu_j + \sigma_j}{a_j} q'_j \right)^{\nu_i} c'_{p,i} - k_{des,i} c_s^{\nu_i} q'_i. \quad (16)$$

Here, we multiplied by $\frac{a_j}{a_j}$ in step Eq. (14)→Eq. (15) to transform the remaining q into q' . Essentially, we shifted the unknown scaling factor from the least-squares problem Eq. (1)

into the isotherm.

We are left with Eq. (7), which can be altered to include q' as above:

$$q_{salt} = \Lambda - \sum_{j=1}^k \nu_j q_j \quad (17)$$

$$= \Lambda - \sum_{j=1}^k \nu_j \frac{a_j}{a_j} q_j \quad (18)$$

$$= \Lambda - \sum_{j=1}^k \frac{\nu_j}{a_j} q'_j. \quad (19)$$

The transformation procedure is also applicable to other isotherms with stoichiometric exchange, e.g. the mixed-mode isotherm in [15].

Binding models of Langmuir type and isotherms without an additional equation for counter-ions, e.g. for hydrophobic interaction chromatography (HIC) [9], can be treated as above. But due to the missing second equation, a will stay hidden in other constants. In case of the kinetic Langmuir isotherm, we obtain

$$\frac{\partial q_i}{\partial t} = k_{ads,i} q_{max,i} \left(1 - \sum_{j=1}^k \frac{q_j}{q_{max,j}} \right) c_{p,i} - k_{des,i} q_i \quad (20)$$

$$\iff \frac{\partial q'_i}{\partial t} = k'_{ads,i} q'_{max,i} \left(1 - \sum_{j=1}^k \frac{q'_j}{q'_{max,j}} \right) c'_{p,i} - k_{des,i} q'_i, \quad (21)$$

with $k'_{ads,i} = k_{ads,i}/a_i$ and $q'_{max,i} = q_{max,i} \cdot a_i$. These parameters can be used for UV-based modeling but not for determining absorption coefficients and molar concentrations.

Other convection-diffusion models, such as the general rate model [12] or models of radial flow chromatography [16], are also linear in the concentration variables and can be treated as the TDM above.

2.8 Uniqueness

In this section, it is shown that the transformation does not affect the parameter determination.

The linear range of the isotherm is uninfluenced by the transformation. For $\Lambda \gg \sum_{j=1}^k \frac{\nu_j + \sigma_j}{a_j} q'_j$, we obtain

$$\frac{\partial q'_i}{\partial t} \approx k_{ads,i} \Lambda^{\nu_i} c'_{p,i} - k_{des,i} c_s^{\nu_i} q'_i. \quad (22)$$

Consequently, all methods to determine the linear SMA parameters k_{ads} , k_{des} , and ν can be employed here. It has been shown that the characteristic charge ν and equilibrium coefficient, defined as $k_{eq} = \frac{k_{ads}}{k_{des}}$, determine the retention time in gradient elution [17]. At least two gradients with different lengths and/ or slopes are necessary to uniquely determine the two values. The kinetic parameter k_{des} can be identified from the peak shape [7]. This method is also applicable to multi-component settings, as no protein-protein interactions are assumed to happen in the linear range.

The non-linear parameter σ is typically determined by a frontal experiment or from batch isotherms. Both methods rely on determining the saturation capacity $q_{max} = \frac{\Lambda}{\nu + \sigma}$ and calculating

$$\sigma = \frac{\Lambda}{q_{max}} - \nu. \quad (23)$$

If the [mAU] equivalent q'_{max} is determined, the equation becomes

$$\sigma = \frac{\Lambda \cdot a}{q'_{max}} - \nu. \quad (24)$$

With known ν and a , σ can be uniquely determined. Alternatively, the steric factor can be identified from the peak shape in non-linear chromatography [18], again with known ν and a . This method allows for including additional steric shielding effects in multi-component settings.

To determine a , one of the original methods for identifying the characteristic charge can be used. It relied on measuring the increase in conductivity caused by freed counter-ions [14]. In the UV-based case, the amount of freed counter-ions can be determined from the second isotherm Eq. (19):

$$\int_0^L \sum_{j=1}^k \frac{\nu_j}{a_j} q'_j dx. \quad (25)$$

If ν has been determined, e.g. from gradient elutions as above, a can be identified uniquely from the increase in conductivity in single-component adsorption. In non-linear multi-component settings, the increase of conductivity will be visible in the chromatogram. The a -dependent locally varying counter-ion concentration in the pores due to adsorbing proteins will lead to a different adsorption behavior.

Summarizing, every parameter plays a distinct role and can be determined from single-component absorbance curves. If only the chromatogram is available in multi-component settings, the parameters might be correlated.

2.9 Numerical Solution

The numerical simulation is performed using the in-house software package ChromX (<http://www.chromx.org>). Following the method of lines, the equation system is first discretized in space using the Finite Element Method (FEM). A Streamline-Upwind-Petrov-Galerkin (SUPG) ansatz was used here with linear basis and test functions. The discretization in time is performed with the fractional step θ -scheme, a semi-implicit procedure providing second-order accuracy [19]. Finally, the non-linearity of the equation system introduced by the isotherm must be treated with an iterative procedure, here, Picard iteration. The resulting linear systems are solved by LU factorization.

A variety of algorithms is available for the solution of the optimization problem in Eq. (1). We employed a heuristic method, the genetic algorithm implementation GAlib, and a deterministic Levenberg-Marquardt implementation CMinpack. Genetic algorithms prevent local minimums by performing random jumps and, hence, explore a larger area of the search space. The result of the genetic algorithm is then refined with the deterministic algorithm. To support this, we divide the kinetic isotherm by k_{des} and use the formulation with equilibrium coefficient k_{eq} . Working with k_{ads} would require to always change

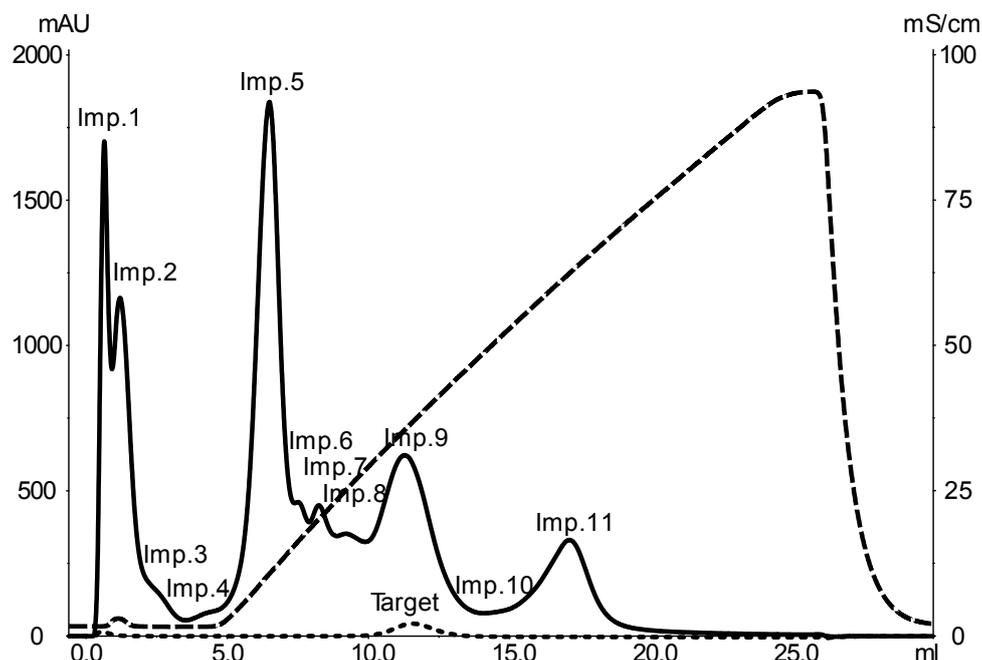


Figure 2: Result of the 20 mL gradient elution: UV 280 nm (solid line), UV 536 nm (dotted line) and conductivity (dashed line). 11 impurities were identified with peak maxima at 0.65 mL (imp. 1), 1.21 mL (imp. 2), 1.96 mL (imp. 3), 5 mL (imp. 4), 6.5 mL (imp. 5), 7.5 mL (imp. 6), 8.23 mL (imp. 7), 9.21 mL (imp. 8), 11.28 mL (imp. 9), 15.52 mL (imp. 10), 17.11 mL (imp. 11). The target component is clearly visible in the 536 nm signal, with the peak maximum at 11.54 mL.

k_{des} at the same time to keep the retention time constant. This is unsuitable for the deterministic algorithm that only uses first derivatives [20].

3 Results & Discussion

3.1 Bind-elute Experiments

Five experiments in bind/ elution mode were performed. Fig. 2 shows the result obtained with a 20 mL gradient. Several impurity peaks could be resolved. The first one is a breakthrough at 0.65 mL. The second peak occurs slightly later at 1.21 mL, followed by two shoulders (impurities 3 and 4) and a high peak at 6.50 mL. The signal continues with three lower peaks (imp. 6, 7, 8), followed by a larger one at 11.28 mL that is also visible at 536 nm. It is identified to be the target component. The fact that this peak's maximum is reached 0.26 mL earlier at 280 nm leads to the assumption that an impurity (imp. 9) is eluting slightly before. A small shoulder (imp. 10) and a final peak at 17.1 mL (imp. 11) complete the elution profile.

3.2 Protein Parameter Estimation

The components' peak areas were determined with Unicorn peak integration from the 280 nm signal of the 20 mL gradient chromatogram. The resulting areas in mAU·mL were divided by the sample volume of 0.5 mL to obtain the inlet absorbance values for modeling c'_{in} in mAU. Hence, the estimated absorption coefficients a refer to 280 nm. The

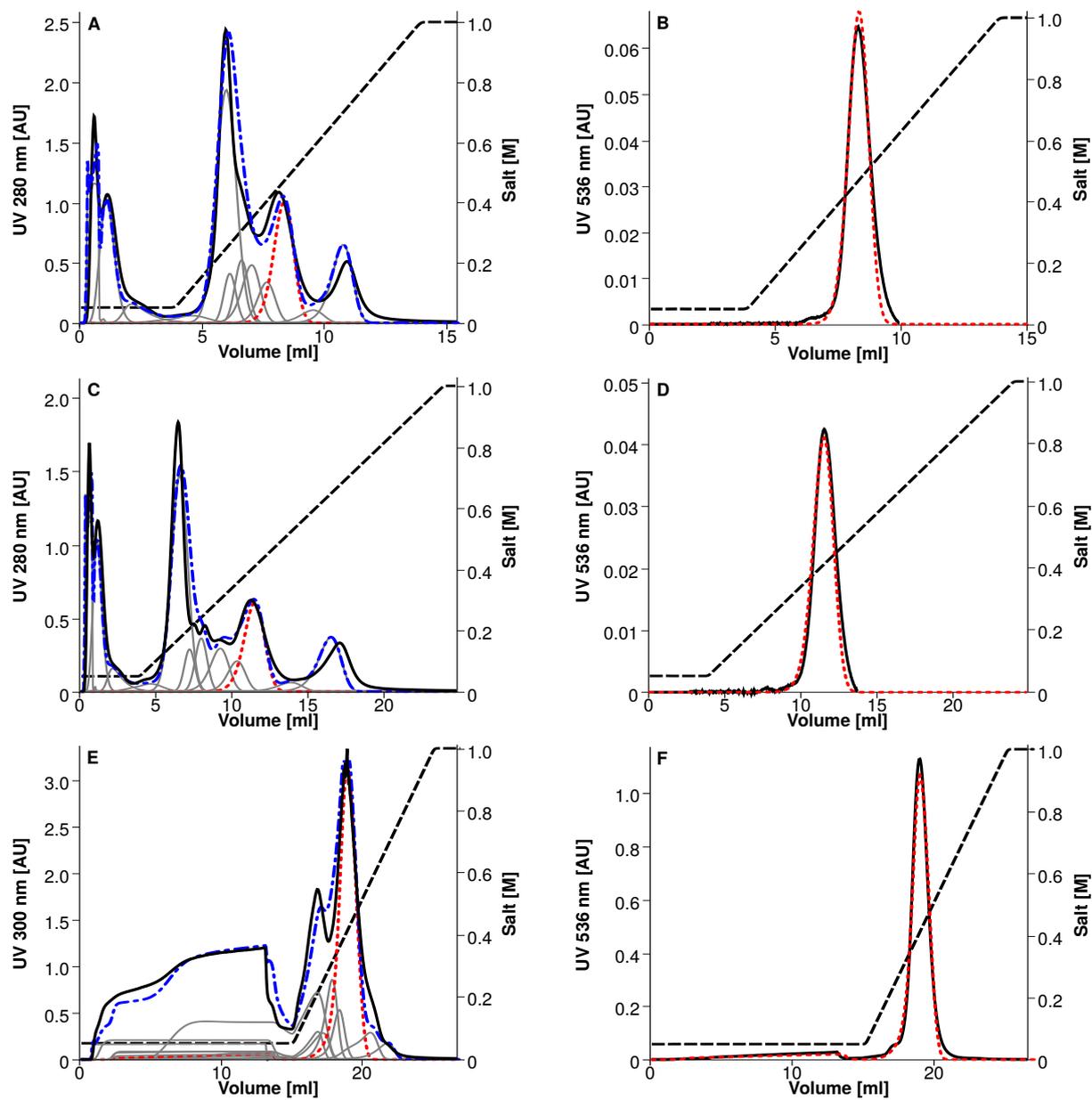


Figure 3: Comparison of measured and simulated chromatograms with UV signals (solid lines), conductivity measurement (dashed lines), simulated Cherry-GST absorbance (dotted lines), impurity traces (light solid lines), and sum of simulated proteins (dot-dashed lines). Plots A and B show the 10 mL gradient elution, C and D the 20 mL gradient elution, and E and F the breakthrough experiment.

same bounds of integration were used for the 300 nm signal. No additional simulation was necessary, the simulated 280 nm peaks were scaled according to the ratio of 300 nm and 280 nm peak areas. The 300 nm signal was used for modeling the breakthrough experiment. Here, the 280 nm signal was incomplete due to sensor saturation.

As the genetic algorithm performs random jumps, admissible parameter ranges have to be set. First estimates of ν and k_{eq} were obtained from the correlation of retention times in gradient elution as in [1] for all binding species. The resulting large equilibrium parameter values in the order of 10^6 fit well to the observed retention. k_{des} has to reside in the natural range $k_{des}^{-1} \in [0, 1]$ and the upper limit for k_{eff} is given by $3 \cdot k_{eff}/r_p = 1$. The steric factor was assumed to be in a range of $\sigma \in [0, 200]$ as it scales approximately with the molecular weight [21] and we expect HCPs of 100 kDa and above. The range of absorption coefficients was chosen large, $a \in [10^5, 10^{10}]$. Eventually, the curve fitting was refined using the Levenberg-Marquardt algorithm.

Selected results are presented in Fig. 3. The left column shows the simulated components, their sum, and the chromatograms at 280 nm for the 10 and 20 mL gradient as well as the 300 nm chromatogram of the breakthrough experiment. The right column shows the same curves at 536 nm. The 15 mL gradient result looks very similar. In the 5 mL gradient chromatogram, highly overlapping impurities lead to only three distinct peaks (data not shown). The corresponding model parameters are given in Table 3. As the first peak is not retained, isotherm parameters could not be determined. The second and third impurity are only slightly retained and the correlation from [1] cannot be used. Furthermore, there might be other parameter combinations of k_{eq} and ν that lead to the same retention volume, and combinations of σ and a with the same amount of occupied ligands. Certainty can be increased by using different low-salt buffer concentrations or including samples with different impurity ratios. But parameter determination for components in the flow-through is not in the focus of this study. The other linear parameters lead to good agreement of simulation and measurement in the first and second row of Fig. 3. As the species do not interact, the four linear parameters can be determined well from the four gradient experiments. The experiment with 12.4 mL sample volume shows good agreement as well, in particular in UV 536 nm. Thanks to this visibility of Cherry-tagged GST in UV 536 nm, a single-component adsorption curve is available that allows for an accurate estimation of the absorption coefficient.

3.3 Capillary Gel Electrophoresis

The target protein can be identified easily in the capillary gel electrophoresis result because of its high fluorescence value. The determined concentration was $3.73 \cdot 10^5$ M, resulting in an absorption coefficient of $7.86 \cdot 10^7$ mAU/M at 280 nm. This is smaller than the estimated value by approximately 9%. The estimate of a is very good, considering the number of interacting species.

Although capillary gel electrophoresis identified even more species, the lumped peaks found by Unicorn peak integration were sufficient to model the elution behavior of the protein components at all investigated wavelengths.

Component	$k_{eff}/10^{-3}$	k_{des}^{-1}	k_{eq}	ν	σ	$a/10^8$
Contaminant 1	0.010	-	-	-	-	-
Contaminant 2	2.090	1	0.004	0.066	157.1	0.600
Contaminant 3	5.006	0	3.939	1.293	102.5	0.056
Contaminant 4	5.586	0.050	22.46	0.566	6.325	0.130
Contaminant 5	15.00	0.151	32.36	0.866	1.895	41.37
Contaminant 6	13.00	0	21.54	1.890	0.001	0.150
Contaminant 7	12.08	0	32.22	1.930	1.702	13.24
Contaminant 8	7.077	0.066	75.83	4.135	0	1.967
Contaminant 9	7.100	0.070	175.8	4.289	0.022	42.91
Cherry-tagged GST	9.271	0.208	334.0	4.000	0.035	0.865
Contaminant 10	14.98	0.424	3084.3	5.350	2.241	2.396
Contaminant 11	7.131	0.235	152227	7.968	0.025	30.00

Table 3: Estimated component-specific model parameters.

4 Conclusions

This study demonstrates that mechanistic modeling can be applied to an anion-exchange step of a crude feed stock, even if the molar concentrations of the feed components are unknown. The model equations were re-written to define injection with respect to the peak areas determined from chromatograms at a chosen wavelengths. The unknown absorption coefficients that scale molar concentration to absorbance units then occur in the isotherm equation. The counter-ion balance of stoichiometric exchange models can be used for estimating these factors using the inverse method. For the steric mass-action model, it was shown theoretically that this additional parameter can be uniquely determined in single-component settings.

In a multi-component case study, the molar concentration of the target protein, Cherry-tagged GST, estimated by chromatogram fitting was only 9% less than the value measured by capillary gel electrophoresis. Here, a single-component absorbance curve was available through the absorbance of the Cherry-TagTM in UV 536 nm. For the other components, a correlation of the steric shielding factor and the absorption coefficient persists. Only the total counter-ion concentration on the adsorber surface is measurable, but not the exact amounts displaced per-component.

Additional reliability can be achieved by including samples with different impurity proportions or fraction analyses that only need to provide peak percentages in one of the observed wavelengths. In preparative chromatography process development, these fraction analyses are performed on a regular basis, e.g. with size-exclusion chromatography [7] or ion-exchange HPLC [10], such that no additional experiments are required.

References

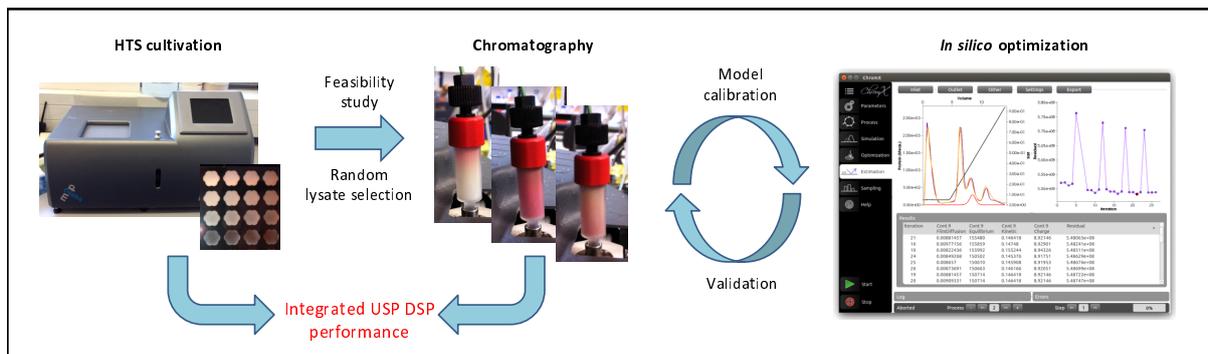
- [1] A. A. Shukla, B. Hubbard, T. Tressel, S. Guhan, D. Low, Downstream processing of monoclonal antibodies - application of platform approaches, J Chrom B 848 (2007)

- 28–39.
- [2] H. F. Liu, J. Ma, C. Winter, R. Bayer, Recovery and purification process development for monoclonal antibody production, *mAbs* 2 (2010) 480–499.
- [3] R. Bhambure, K. Kumar, A. S. Rathore, High-throughput process development for biopharmaceutical drug substances, *Trends Biotechnol* 29 (2011) 127–135.
- [4] S. Abel, G. Erdem, M. Amanullah, M. Morari, M. Mazzotti, M. Morbidelli, Optimizing control of simulated moving beds - experimental implementation, *J Chrom A* 1092 (2005) 2–16.
- [5] P. Forssén, R. Arnell, T. Fornstedt, An improved algorithm for solving inverse problems in liquid chromatography, *Comput Chem Eng* 30 (2006) 1381–1391.
- [6] A. Osberghaus, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch, Optimizing a chromatographic three component separation: a comparison of mechanistic and empiric modeling approaches., *J Chrom A* 1237 (2012) 86–95.
- [7] D. Karlsson, N. Jakobsson, A. Axelsson, B. Nilsson, Model-based optimization of a preparative ion-exchange step for antibody purification, *J Chrom A* 1055 (2004) 29–39.
- [8] M. Degerman, N. Jakobsson, B. Nilsson, Constrained optimization of a preparative ion-exchange step for antibody purification., *J Chrom A* 1113 (2006) 92–100.
- [9] J. M. Mollerup, T. B. Hansen, S. Kidal, L. Sejergaard, A. Staby, Development, modelling, optimisation and scale-up of chromatographic purification of a therapeutic protein, *FFE* 261 (2007) 133–139.
- [10] E. J. Close, J. R. Salm, D. G. Bracewell, E. Sorensen, Modelling of industrial biopharmaceutical multicomponent chromatography, *Chem Eng Res Des* 92 (2014) 1304–1314.
- [11] J. Cornel, A. Tarafder, S. Katsuo, M. Mazzotti, The direct inverse method: A novel approach to estimate adsorption isotherm parameters, *J Chrom A* 1217 (2010) 1934–1941.
- [12] M. Michel, A. Epping, A. Jupke, *Preparative Chromatography: Of Fine Chemicals and Pharmaceutical Agents*, Wiley-VCH, 2005, pp. 215–312.
- [13] P. Baumann, N. Bluthardt, S. Renner, A. Osberhaus, J. Hubbuch, Integrated Development of Up- and Downstream Processes Supported by the Cherry-Tag for Real-time Tracking of Stability and Solubility of Proteins, *J Biotech* 200 (2015) 27–37.
- [14] C. A. Brooks, S. M. Cramer, Steric mass-action ion exchange: Displacement profiles and induced salt gradients, *AIChE Journal* 38 (1992) 1969–1978.
- [15] B. K. Nfor, M. Noverraz, S. Chilamkurthi, P. D. Verhaert, L. A. van der Wielen, M. Ottens, High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents, *J Chrom A* 1217 (2010) 6829–6850.

-
- [16] T. Gu, *Mathematical Modeling and Scale-up of Liquid Chromatography*, Springer Berlin Heidelberg, 1995, pp. 102–110.
- [17] A. A. Shukla, S. S. Bae, J. A. Moore, K. A. Barnthouse, S. M. Cramer, Synthesis and characterization of high-affinity, low molecular weight displacers for cation-exchange chromatography, *Ind Eng Chem Res* 37 (1998) 4090–4098.
- [18] S. R. Gallant, A. Kundu, S. M. Cramer, Modeling non-linear elution of proteins in ion-exchange chromatography, *J Chrom A* 702 (1995) 125–142.
- [19] R. Glowinski, *Handbook of Numerical Analysis*, North-Holland, 2003.
- [20] T. Hahn, A. Sommer, A. Osberghaus, V. Heuveline, J. Hubbuch, Adjoint-based estimation and optimization for column liquid chromatography models, *Comput Chem Eng* 64 (2014) 41–54.
- [21] D. Karlsson, N. Jakobsson, K.-J. Brink, A. Axelsson, B. Nilsson, Methodologies for model calibration to assist the design of a preparative ion-exchange step for antibody purification, *J Chrom A* 1033 (2004) 71–82.

High-throughput Micro-scale Cultivations and Chromatography Modeling: Powerful Tools for Integrated Process Development

P. Baumann¹, T. Hahn¹ and J. Hubbuch^{1,*}



¹ : Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 1,76131 Karlsruhe, Germany
* : Corresponding author. *E-mail-address*:juergen.hubbuch@kit.edu (J. Hubbuch)

Abstract

Upstream processes are rather complex to design and the productivity of cells under suitable cultivation conditions is hard to predict. The method of choice for examining the design space is to execute high-throughput cultivation screenings in micro-scale format. Various predictive *in silico* models have been developed for many downstream processes, leading to a reduction of time and material costs. This paper presents a combined optimization approach based on high-throughput micro-scale cultivation experiments and chromatography modeling. The overall optimized system must not necessarily be the one with highest product titers, but the one resulting in an overall superior process performance in up- and downstream.

The methodology is presented in a case study for the Cherry-tagged enzyme Glutathione-S-Transferase from *Escherichia coli* SE1. The Cherry-TagTM (Delphi Genetics, Belgium) which can be fused to any target protein allows for direct product analytics by simple VIS absorption measurements. High-throughput cultivations were carried out in a 48-well format in a BioLector[®] micro-scale cultivation system (m2p-Labs, Germany). The downstream process optimization for a set of randomly picked upstream conditions producing high yields was performed *in silico* using a chromatography modeling software developed in-house (ChromX). The suggested *in silico*-optimized operational modes for product capturing were validated subsequently. The overall best system was chosen based on a combination of excellent up- and downstream performance.

Keywords: Micro-scale Cultivation, BioLector[®], High-Throughput Screening, UV Absorption Modeling, Ion-exchange Chromatography

1 Introduction

Although the complexity of the protein purification procedure is strongly dependent on the cultivation conditions, up- and downstream processes are mostly optimized separately. When linking those two process parts, the best cultivation must not necessarily be the one leading to highest yields and titers, but the one resulting in an overall superior process performance. For product formation, it is important to have tools to screen large numbers of organisms and production conditions. The subsequent downstream screening needs to be capable of scoring the large number of produced feedstocks in terms of their applicability for downstream process development (DSP).

Cultivation conditions for the effective production of biopharmaceuticals are hard to predict. The chosen cultivation conditions strongly influence product formation and can significantly differ from standard growth conditions. For human interferon- α 2, human interferon- γ , and the interferon-induced murine protein Mx, it was shown that lowered temperatures lead to an increase in soluble protein during product formation in *E. coli* [1]. Similar results were shown for the production of Vitreoscilla globin under oxygen limitations [2]. Different levels of benzoylformate decarboxylase were obtained using various culture media components in *E. coli* [3]. These results, however, cannot be transferred directly when investigating different strains and products, and *in silico* model implementation is limited. Thus, high-throughput process development (HTPD) is the method of choice for screening a large pool of conditions in upstream process development

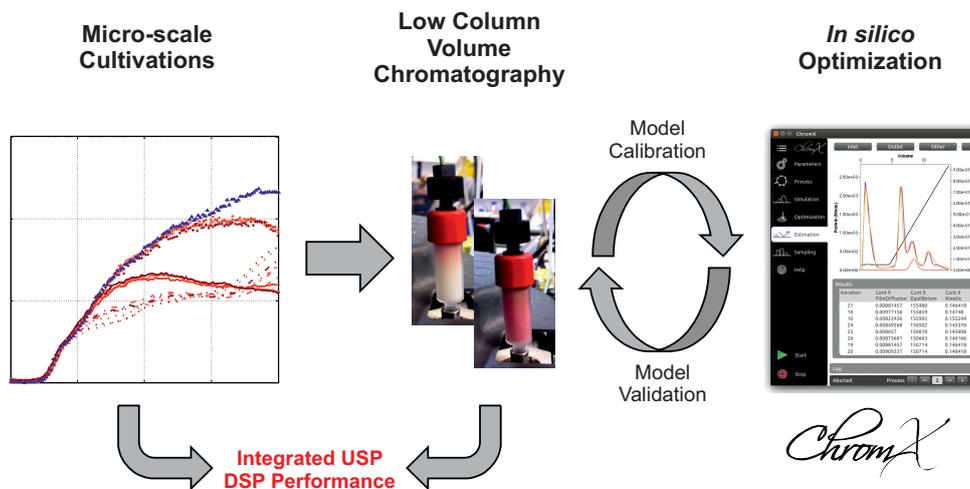


Figure 1: Schematic overview of the integrated optimization approach for bioprocess development based on high-throughput micro-scale cultivations and lysate characterization by chromatography modeling. Different lysates from the micro-scale cultivations are investigated in low column volume chromatography for model calibration. The created models are then used for *in silico* optimization of the purification process. The overall best system is chosen based on the overall optimal performance in up- and down-stream.

(USPD) on the micro-scale. Different miniaturized high-throughput cultivation systems were described, mostly resulting in a strong decrease in accessible process information [4]. One approach uses small-scale vessels which imitate the industrial scale fermenters and can be run in parallel [5, 6, 7]. However, the highest capabilities for high-throughput experiments are reached by cultivations in microtiter plates (MTPs). A major disadvantage is the loss of process information and process control, leading to optimization results that might not be suited for scale up. An approach linking the advantages of high-throughput experimentation in microtiter plates with those of continuous process control is the BioLector[®] technology (m2p-Labs, Germany). Cultivations are carried out in continuously shaken 48-well plates on the sub-milliliter scale. For better mixing and oxygen uptake, microtiter plates with baffled wells (FlowerPlates[®]) were developed [8]. For continuous monitoring of cell growth by scattered light measurements, the plates are equipped with an optical bottom, making sampling obsolete [9, 10].

Besides the effects on product formation, influences of cultivation conditions on the contaminant levels are not considered in most cases. E.g. dissolved oxygen levels [11, 12], temperature changes [13, 14], cultivation media [15, 16], as well as the state of the cell in the cell cycle [17, 18] are highly influential on e.g. host cell protein levels. Consequently the initial purity of the product will differ largely.

The above-mentioned variations of the initial purity and concentrations of critical contaminants play a crucial role in downstream process development. Hence, cultivations should not only be scored by the product titers, but also by the contaminant profiles of respective upstream conditions. Besides HTPD, a large number of predictive *in silico* models have been developed for many downstream unit operations, leading to a reduction of time and material costs. Mathematical modeling and numerical optimization of chromatography were applied successfully for model proteins [19], antibodies [20, 21] as well as for industrial processes [22, 23]. All simulation procedures mentioned rely on known

molar concentrations of the biomolecules involved. However, models can also be based on absorption data using an inverse method [24]. The major advantage of this procedure is that data from chromatograms can be used directly for model development. No sensor calibration is needed, which is why this method appears to be an ideal tool for comparing a large number of feedstocks within a short period of time.

This paper presents an integrated optimization approach based on high-throughput micro-scale cultivations and lysate characterization by chromatography modeling using chromatogram fitting (Fig. 1). The methodology is presented in a case study for the Cherry-tagged enzyme Glutathione-S-Transferase from *Escherichia coli* SE1. The 11 kDa Cherry-TagTM (Delphi Genetics, Belgium) is fused to the target protein and can be detected by 414 nm absorption measurements [25, 26]. The upstream screening for varying cultivation conditions, such as temperature and induction conditions, was performed in a BioLector[®] system (m2p-Labs, Germany). A set of randomly picked upstream conditions producing high yields was then characterized and optimized *in silico* using a chromatography modeling software developed in-house (ChromX). The operational modes for process optimization suggested were validated subsequently. Instead of choosing the cultivation leading to highest product concentrations, the ideal system was chosen based on a combination of overall excellent up- and downstream performance.

2 Materials & Methods

2.1 Mathematical Methods

2.1.1 Chromatogram Fitting

In the following section, the governing equations of the simulations are presented. \bar{p} describes the solution of the least squares optimization problem given in Eq. (1):

$$\arg \min_{\bar{p}} \sum_j \left(m(t_j) - \sum_{i \geq 1} c_i(L, t_j; \bar{p}) \cdot a_i \right)^2, \quad (1)$$

with $m(t_j)$ being the absorption value derived from the chromatogram at point in time t_j , c_i being the protein concentration i at the end of the column of length L , and a_i describing the absorption factor. The algorithms used for solving the above-mentioned problem were a heuristic method based on a genetic algorithm [27] and a deterministic method based on Levenberg-Marquardt implementation [28]. Genetic algorithms prevent local minima by performing random jumps for exploring a larger area in the search space. The deterministic algorithm served for result refinement.

2.1.2 Steric Mass-action (SMA) Model

As an adsorption model for ion-exchange chromatography, the semi-mechanistic steric mass-action isotherm (SMA) [29] was applied allowing for chromatography simulations including varying salt concentrations. It covers effects of counter-ions on the retention behavior of proteins, including the proteins' characteristic charges ν_i . Adsorbent properties, such as the total ionic capacity Λ , are also considered. Steric shielding effects caused by

proteins covering binding sites without electrostatic interactions are included as a factor σ_i . The kinetic form of the SMA isotherm is shown in Eq. (2) for k proteins, with q_i and $c_{p,i}$ being the concentration of the protein i adsorbed and in solution inside the pore, respectively. $c_{p,salt}$ describes the effective pore salt concentration. $k_{ads,i}$ and $k_{des,i}$ are the constants of adsorption and desorption, respectively. q_{salt} in Eq. (3) describes the number of salt ions still attached to the adsorbent surface and is defined as the difference between the total ionic capacity Λ and the number of areas blocked by electrostatic adsorption $\sum_{j=1}^k \nu_j q_j$. The areas blocked due to steric shielding still possess the counter-ions and, hence, are not taken into account.

$$\frac{\partial q_i}{\partial t} = k_{ads,i} \left(\Lambda - \sum_{j=1}^k (\nu_j + \sigma_j) q_j \right)^{\nu_i} c_{p,i} - k_{des,i} c_{p,salt}^{\nu_i} q_i \quad (2)$$

$$q_{salt} = \Lambda - \sum_{j=1}^k \nu_j q_j \quad (3)$$

For small sample loadings ranging from 0 to 5% of the maximum binding capacity of the applied adsorbent resin as in this study the ionic capacity predominates and the occupied adsorption slots can be neglected. This results in a quasi-linear form of the SMA isotherm, meaning that the peak shape is not influenced by the concentration, as shown in Eq. (4). When dividing Eq. (4) by $k_{des,i}$, Eq. (5) results in a kinetic term $\frac{1}{k_{des,i}}$ and an equilibrium term $k_{eq,i}$ which equals $\frac{k_{ads,i}}{k_{des,i}}$.

$$\frac{\partial q_i}{\partial t} \approx k_{ads,i} \Lambda^{\nu_i} c_{p,i} - k_{des,i} c_{p,salt}^{\nu_i} q_i \quad (4)$$

$$\frac{1}{k_{des,i}} \frac{\partial q_i}{\partial t} \approx k_{eq,i} \Lambda^{\nu_i} c_{p,i} - c_{p,salt}^{\nu_i} q_i \quad (5)$$

In contrast, for setups at the column capacity limit besides the 4 gradients presented in this study a frontal experiment (breakthrough experiment) is mandatory. The estimation procedure then includes further parameters to be determined, namely the steric shielding factor σ and the absorption factors. However, considering the limited sample volume derived from high-throughput cultivations such frontal experiments were not possible. Nevertheless, for the determination of critical contaminants the simplified methodology works well for pre-selecting feasible lysates for scale up applications.

2.1.3 Transport Dispersive Model (TDM)

The Transport Dispersive Model (TDM) [30] in Eqs. (6) and (7) describes macroscopic protein transport inside the chromatography column and considers the interstitial volume of the mobile phase, as well as the mass transfer into the pore volume. Film and pore diffusion effects are lumped in an effective mass transfer coefficient $k_{eff,i}$ for simplification. The system is of Convection Diffusion Reaction (CDR) type. Eq. (6) describes the rate of change of the concentration $c_i(x, t)$ of component i in the inter-particle phase of a column with length L . The first term of Eq. (6) refers to the convective transport which is dominated by the inter-particle velocity of the fluid u . The middle term describes hydrodynamic dispersion in axial direction using the dispersion coefficient D_{ax} . The last

term refers to the transition of molecules from the interstitial into the particle pore phase $c_i - c_{p,i}$ which depends on the porosity of the bed ε_b , the radius of adsorbent particles r_p , and a component-specific lumped effective mass transfer coefficient $k_{eff,i}$.

Eq. (7) models the accumulation and distribution of mass inside the pores. It is composed of the protein pore concentration in the liquid phase $c_{p,i}$ and the protein concentration bound to the stationary phase q_i . Particle porosity ε_p is one of the influencing factors.

For the model, Danckwerts boundary conditions in Eqs. (8) and (9) were used describing that the concentration at the inlet is influenced by diffusion and back mixing whereas the concentration at the outlet is unaffected by such effects. The adsorption mechanism was described using the SMA isotherm model shown above in Eqs. (2) and (3).

$$\frac{\partial c_i}{\partial t} = -u \frac{\partial c_i}{\partial x} + D_{ax} \frac{\partial^2 c_i}{\partial x^2} - \frac{1 - \varepsilon_b}{\varepsilon_b} k_{eff,i} \frac{3}{r_p} (c_i - c_{p,i}) \quad (6)$$

$$\frac{\partial c_{p,i}}{\partial t} = k_{eff,i} \frac{3}{r_p \varepsilon_p} (c_i - c_{p,i}) - \frac{1 - \varepsilon_p}{\varepsilon_p} \frac{\partial q_i}{\partial t} \quad (7)$$

$$\frac{\partial c_i}{\partial x}(0, t) = \frac{u}{D_{ax}} (c_i(0, t) - c_{in,i}(t)) \quad (8)$$

$$\frac{\partial c_i}{\partial x}(L, t) = 0 \quad (9)$$

The model was simulated one-dimensionally in space. Consequently, the concentrations depend on the axial position in the column, only. As a quasi-linear form of the SMA isotherm is used, all concentration terms c_i and q_i in Eqs. (3) - (9) can be transformed into absorptions using the absorption factors.

2.1.4 Pareto Optimization

Yields and purities of the elution fractions were investigated for all possible product fractionation setups. Finding a global optimum for a multi-objective problem can be accomplished by Pareto optimization. Single objective problems mostly result in single solutions. For multi-objective problems, by contrast, infinite Pareto optimal solutions exist, as the single objectives interfere. The so-called Pareto front describes the feasible system points of all possible single parameter combinations.

2.2 Materials

2.2.1 Disposables & Reaction Vessels

The nucleic acid sequence of Glutathione-S-Transferase (GST) derived from *S. japonicum* was inserted into the pScherry vector of the CherryTM Express T7 protein expression kit (Delphi Genetics, Belgium). The storage of cells was realized in a CryobankTM strain maintenance kit for microorganisms (Mast Diagnostica, Germany). Micro-scale cultivations were carried out in 48-well FlowerPlates[®] covered with adhesive sterile sealing foil (m2p-Labs, Germany). Pre-cultures were cultivated in 1 L baffled flasks (Schott, Germany). Cell harvest and removal of cell debris after disruption was carried out in 15 mL centrifugal tubes (Falcon, Germany). Absorption measurements were carried out in 96-well flat bottom UV-Star[®] half-area micro-plates (Greiner Bio-One, Germany).

Anion-exchange chromatography experiments were conducted in a '1 mL' HiTrapTM Q FF column with effective volume of 0.962 mL (GE Healthcare Life Sciences, Sweden). For filtration of samples, 0.2 μ m polyethersulfone sterile filters were applied (VWR, Germany). Capillary gel electrophoresis (GX II) was performed in an HT Protein Express & Pico LabChip[®] (Perkin Elmer, USA). Sample preparation for GX II was conducted in skirted 96-well twin.tec[®] PCR plates (Eppendorf, Germany).

2.2.2 Chemicals & Buffers

As a cultivation medium, standard Terrific Broth (TB) medium was used, including 12 g/L wheat peptone for microbiology (Fluka, Germany), 24 g/L bacteriological yeast extract (Amresco, USA), 5 g/L glycerol bidistilled 99.5%, 17 mM potassium dihydrogen phosphate, and 7 mM di-potassium hydrogen phosphate (VWR, Germany). Isopropyl- β -D-thiogalactopyranosid (IPTG) from a 0.25 M stock solution (VWR, Germany) was applied for the induction of the pSCherry's T7 promotor. As a lysis buffer for cell disruption, the 20 mM Tris-HCl AEX running buffer was used (VWR, Germany) adjusted to pH 8 with hydrochloric acid (Merck, Germany). 1X SigmaFASTTM Protease Inhibitor (Sigma Aldrich, Germany) was added for prevention of protease degradation. The binding buffer for the 1 mL HiTrapTM Q FF column was composed of 20 mM Tris-HCl at pH 8, as described above. The elution buffer consisted of the same buffer, including 1 M NaCl (Merck, Germany). For column regeneration, a 0.1 M sodium hydroxide solution (Sigma Aldrich, Germany) was applied. The pH adjustment of all buffers was done by titration using sodium hydroxide (Merck, Germany). A 1% acetone solution (Merck, Germany) and a 10 g/L MW 2,000,000 dextran solution from *leuconostoc spp.* (Sigma Aldrich) were applied to determine porosities and dispersion properties. The total ionic capacity was determined by acid-base titration using 0.01 M HCl solution and 0.5 M NaOH solution (Merck, Germany). For the capillary gel electrophoresis (GX II) experiments, an HT protein express reagent kit (Perkin Elmer, USA) was used. 1 mg/mL Lysozyme solution served as an internal concentration standard (Hampton research, USA).

2.2.3 Instrumentation & Software

Micro-scale cultivations were performed in a BioLector[®]MB micro-scale fermentation system equipped with the BioLection[®]HMI & analysis software (m2p-Labs, Germany). Lab-scale fermentations were carried out in a MaxQTM 6000 incubator (Thermo scientific, USA). Cell harvest and clarification after disruption were carried out in a 5810 R centrifuge (Eppendorf, Germany). The cells were disrupted using a Digital Sonifier[®] 450 (Branson Ultrasonic Corporation, USA). Absorption measurements were conducted in an Infinite M200 Reader controlled with I-control 1.9 (Tecan, Germany). For the chromatography procedures, an ÄKTATM Purifier system (GE Healthcare Life Sciences, Sweden) was used, which was equipped with a pump P-900, mixer M-925, UV-detector UV-900, motor valve INV-907, pH and conductivity monitoring unit pH/C-900, and a fraction collector Frac-950 unit. The FPLC system was controlled using Unicorn 5.2 (GE Healthcare Life Sciences, Sweden). The pH of all buffers was adjusted with an HI-3220 pH meter (Hanna Instruments, USA). Random sampling, data processing, and creation of figures were performed in Matlab[®] R2011a (MathWorks, USA). Simulations and modeling were performed using the software developed in-house, called ChromX (Karlsruhe Institute of

Technology, Germany). Capillary gel electrophoresis (GX II) was carried out in a Caliper LabChip[®] GX II using the LabChip[®] GX 3.1 software (Perkin Elmer, USA).

2.3 Experimental Setup

2.3.1 Micro-scale Cultivations

For the micro-scale cultivations, an equally treated pre-culture was used. A 250 mL TB medium cultivation was prepared by inoculating with an *E. coli* SE1 Cherry-GST cryo culture from the Cryobank[™] strain maintenance kit in a 1 L baffled shake flask. The cultivation was performed in the MaxQ[™] 6000 incubator at 170 rpm and 37 °C. After 17 h (end of the exponential growth), the cells were diluted in sterile TB medium to a starting OD_{600 nm} for the BioLector[®] experiments of 0.1 AU. Each well of the 48-well FlowerPlate[®] was then filled with 1 mL of diluted pre-culture. The micro-scale cultivation plate was subsequently sealed with an adhesive gas-permeable sterile membrane. All cultivations were performed at a BioLector[®] shaking speed of 600 rpm. As the different cultivation processes were performed in quadruplicates, 12 different conditions could be screened per plate in total. Per 48-well FlowerPlate[®], three inducer concentrations (IPTG concentrations of 0.1, 0.5, and 1 mM) and four induction times (OD_{600 nm} of 1, 2, 4, and 8) were investigated. As the BioLector[®] system measures scattered light signals instead of optical densities at 600 nm OD_{600 nm}, a correlation function was used. For 600 rpm shaking frequency, the linear correlation slope of 17.26 SU/AU was determined. As soon as the scattered light signal exceeded the pre-defined induction threshold value, the respective well was induced by adding IPTG. The above-mentioned procedure was repeated for three different cultivation temperatures of 27, 32, and 37 °C as a full factorial experimental design. The cultivation procedure was stopped at 12 h after induction by transferring the cell broth to a 2 mL 96-well square deep-well plate and by centrifuging at 4000 rpm and 10 °C for 30 min. After discarding the supernatants, the cell pellets were frozen at -20 °C. The frozen cells from 1 mL cultivations were resuspended in 1.25 mL of lysis buffer and the quadruplicates were pooled. The cell disruption for the 5 mL suspensions was performed in 15 mL centrifugal tubes in the Branson Digital Sonifier[®] 450 using a 1/8" tapered micro-tip sonication probe at 30% maximal power output and a total treatment time of 90 s (6 X 15 s pulse on and 30 s pulse off for prevention of heat denaturation). The sonication probe was fixed 0.5 cm above the centrifugal tube bottom. Lysate clarification was accomplished by centrifugation at 12000 rpm and 10 °C for 60 min in a 5810 R centrifuge. The supernatants of all cultivations were then analyzed using the Cherry-Tag[™] analytics at 414 nm. 414 nm Cherry-GST signals could be correlated directly to concentrations via the previously determined absorption coefficient of 2.213 mg/(mL·AU_{414 nm}) (data not shown).

2.3.2 Random Sampling Procedure

As a feasibility study, 5 upstream conditions were chosen randomly for ChromX lysate characterization, instead of choosing the cultivations of highest titers. First, the '*Threshold Criterion*' was defined: Conditions resulting in Cherry-GST concentrations below 0.4 mg/mL were deleted for economic reasons and replaced by a concentration of 0 mg/mL.

Parameter	Symbol	Value
Diameter	d	7 mm
Length	L	25 mm
Bead radius	r_p	0.045 mm
System dead volume	V_d	0.07 mL
Retention volume acetone	V_{RetAc}	0.96 mL
Retention volume dextran	V_{RetDex}	0.34 mL
Standard deviation of dextran	σ_{Dex}	0.029 mL
Volume of HCl	V_{HCl}	1.48 mL
Molarity of HCl	c_{HCl}	0.01 M
Flow rate	u	0.2 $\frac{mm}{s}$

Table 1: Column- and system-specific parameters

Parameter	Symbol	Value	Proceeding
Volume	V	0.962 mL	$\frac{1}{4} \cdot \pi \cdot d^2 \cdot L$
Fluid volume	V_f	0.89 mL	$V_{RetAc} - V_d$
Interstitial volume	V_{int}	0.27 mL	$V_{RetDex} - V_d$
Total column porosity	ε_t	0.925	$\frac{V_f}{V}$
Interstitial porosity	ε_b	0.28	$\frac{V_{int}}{V}$
Particle porosity	ε_p	0.896	$\frac{V_f - V_{int}}{V - V_{int}}$
Interstitial flow	u_{int}	0.714 $\frac{mm}{s}$	$\frac{u}{\varepsilon_b}$
Axial dispersion	D_{ax}	0.1 $\frac{mm^2}{s}$	$\frac{1}{2} \cdot L \cdot u_{int} \cdot \left(\frac{\sigma_{Dex}}{V_{int}} \right)^2$
Ionic capacity	Λ	0.22 M	$\frac{c_{HCl} \cdot V_{HCl}}{V(1-\varepsilon_b)(1-\varepsilon_p)}$

Table 2: Model parameters derived from the column- and system-specific parameters

All conditions were then scaled between 0 (= 0 mg/mL) and 1 (= maximal titer investigated), which is referred to as the '*Probability Criterion*'. The random sampling procedure of 5 cultivation conditions was then performed in Matlab[®] R2011a. The probabilities of being chosen were based on the '*Probability Criterion*'. This means that cultivations of lower product concentrations have lower probabilities of being picked, but can still be chosen for downstream investigation.

2.3.3 Model Parameter Determination

For the ChromX model, different chromatography column- and system- specific parameters were determined (compare '*Mathematical Methods*' section). The procedure followed the description of [30]. The bed porosity ε_b and the axial dispersion D_{ax} of the '1 mL' HiTrap[™] Q FF column were determined with a 20 μ L injection of 10 g/L MW 2,000,000 dextran solution from *leuconostoc spp.*. The large dextran molecules are excluded from the pores and, thus, only access the interstitial area of the chromatography column. For the particle porosity ε_p , a 20 μ L injection of 1% acetone solution was performed. In

contrast to this, the small acetone molecules have access to all pores and are a measure for the overall porosity of the system. The system's dead volume V_d was determined by acetone injection without connecting the column to the FPLC system.

For the total ionic capacity Λ , an acid-base titration of the ion-exchange column was carried out. For this purpose, the column was flushed with a 0.5 M NaOH solution until the UV and conductivity signal was constant and all functional groups were blocked with OH^- ions. Following a wash with ultrapure water, a titration with 0.01 M HCl solution followed at a flow of 0.2 mm/s. The amount of acid needed until the conductivity signal increased was recorded. From the number of Cl^- ions replacing the OH^- ions, the total number of binding sites was determined. All other column properties were taken from the manufacturer. All given and determined parameters are listed in Table 1. The calculated model parameters and the underlying proceedings are shown in Table 2.

2.3.4 ChromX Model Calibration

For ChromX model calibration, 4 different gradient experiments were conducted for each cell lysate selected in the '*Random Sampling Procedure*'. The number of gradients should ideally at least be equal to the number of estimated parameters (here: k_{kin} , k_{eq} , k_{eff} , ν). Each sample was 4 times diluted in 20 mM Tris binding buffer pH 8, with a 1 mL injection being ensured for each gradient. After sample application, the column was washed for 2 CVs with binding buffer. The elution was conducted by increasing the NaCl concentration up to 1 M applying 5, 10, 15, and 20 CV gradients. Including steep and shallow slopes helps the created model better predict steps and shallow gradients during the *in silico* optimization procedure. The high-salt stage was kept constant for another 2 CVs, followed by a re-equilibration step for 5 CVs. The resulting chromatograms from Unicorn 5.2 were then exported and implemented in the ChromX model. The areas of the contaminant peaks in each 20 CV run (the longest gradient shows the most '*visible*' peaks) were used as the '*absorption factors*' in the ChromX model.

Finally, the boundary conditions of all estimated parameters were set (namely for k_{kin} , k_{eq} , k_{eff} , ν) to accelerate the estimation process and to exclude non-logical solutions as e.g. negative values for the equilibrium binding coefficient k_{eq} . For initial simulations, a genetic algorithm was used to prevent local minima in the search space, followed by a deterministic Levenberg-Marquardt algorithm for result refinement. To determine the amount of protein loaded onto the chromatography column the fractions taken from the 20 CV gradients were measured in the LabChip[®]GX II device as duplicates using the HT Protein Express LabChip[®]kit. The sample and chip preparation procedures were performed as described in the manufacturer's protocol for the HT Protein Express Assay (www.bioneer.co.kr). The results were analyzed by peak baseline integration using the HT Protein Express 200 assay in the LabChip[®]GX 3.1 software.

2.3.5 Pareto Optimization

The ChromX model of each cell lysate selected in the '*Random Sampling Procedure*' was used for an *in silico* optimization procedure. The operational setup proposed for product purification was subsequently validated using the ÄKTA[™] Purifier system for all 5 selected samples. The best system of all was chosen based on a combination of excellent up- and downstream performance using a Pareto optimization method. The

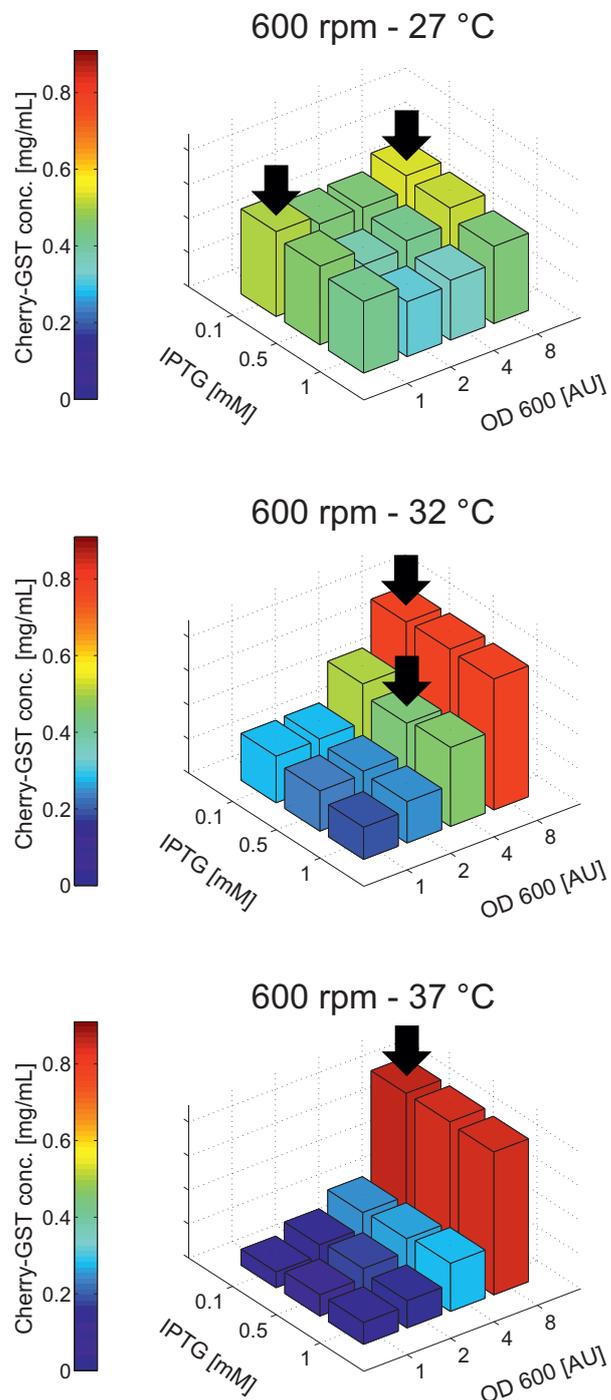


Figure 2: Product titers of soluble Cherry-GST [mg/mL] in *E. coli* SE1 cultivated at 600 rpm shaker frequency under different cultivation conditions. The investigated temperatures were 27, 32, and 37 °C. Additionally, three inducer concentrations (IPTG concentrations of 0.1, 0.5, and 1 mM) and four induction times (OD_{600 nm} of 1, 2, 4, and 8) were investigated. The sample selection of the 'Random Sampling Procedure' is indicated by black arrows.

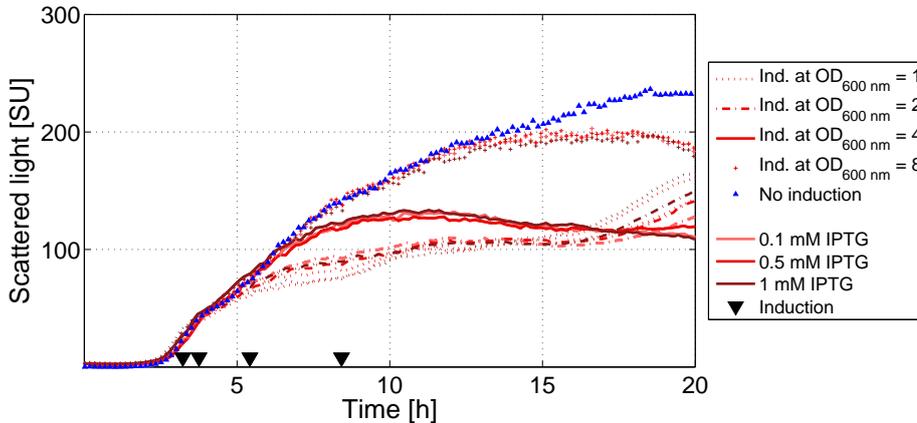


Figure 3: BioLector® growth curves of *E. coli* SE1 cultivated at 600 rpm shaker frequency under different cultivation conditions. Bacterial growth is shown for the 37 °C cultivation at three inducer concentrations (IPTG concentrations of 0.1, 0.5, and 1 mM) and four induction times ($OD_{600\text{ nm}}$ of 1, 2, 4, and 8).

overall recovery Rec_i was defined as the initial titer from the cultivation $c_{init,i}$ multiplied by the yield Y_i in the downstream process of a sample cultivated under condition i , as shown in Eq. (10). The product recoveries were determined *in silico* by integration of the Cherry-GST- specific 414 nm signal. The purity Pur_i was based on integrating the 280 nm signals in ChromX. The 280 nm signal of a purified Cherry-GST sample served as a 100% purity reference $A_{280nm,Cherry-GST,i}$ and was compared to the actual 280 nm signal of the elution fraction $A_{280nm,Elu,i}$, as shown in Eq. (11).

$$Rec_i = c_{init,i} Y_i \quad (10)$$

$$Pur_i = \frac{A_{280nm,Cherry-GST,i}}{A_{280nm,Elu,i}} \quad (11)$$

An *in silico* fractionation was performed to scan all product fractionation possibilities. The resulting pooled recoveries and purities were then plotted in a Pareto optimization diagram, resulting in a Pareto front of optimal operational conditions for each upstream setup. The comparison of those fronts results in a global Pareto front of all selected cultivation conditions.

3 Results & Discussion

3.1 Micro-scale Cultivations & Random Sampling

All cultivations were performed at a BioLector® shaking speed of 600 rpm. Incubation temperatures of 27, 32, and 37 °C were investigated. On each 48-well FlowerPlate®, four different induction times at cell densities of $OD_{600\text{ nm}}$ 1, 2, 4, and 8 AU and three inducer concentrations of 0.1, 0.5, and 1 mM IPTG were investigated as quadruplicates. The product titers were determined after cell disruption by absorption measurements of the cleared lysates at 414 nm. Those data were converted into concentrations using the Cherry-GST extinction coefficient. The results of the micro-scale cultivation screening is

shown in Fig. 2.

The maximal Cherry-GST titer of 0.85 mg/mL was determined for the 37 °C setup applying a late induction at OD_{600 nm} of 8 with 0.1 mg/mL of IPTG. Generally, a pronounced trend of increased soluble product formation towards late induction was observed for the setups of 32 and 37 °C. Here, cells induced at OD_{600 nm} of 1 resulted in product titers of only 0.09 to 0.27 mg/mL. In contrast to this, product titers during induction at OD_{600 nm} of 8 were much higher, ranging from 0.77 to 0.85 mg/mL. Total protein extraction under denaturing conditions revealed that early induced cells resulted in inclusion bodies (data not shown). This implies that an early induction of the *E. coli* SE1 T7-promoter leads to an overproduction of Cherry-GST and, thus, to high amounts of insoluble and incorrectly folded proteins at 32 and 37 °C. The 27 °C setups showed a different trend. Here, almost constant Cherry-GST levels were obtained for all different induction setups in the range from 0.32 to 0.54 mg/mL. As *E. coli* is a mesophilic organism, lowered temperatures result in a downward regulation of the metabolism and, thus, in slowly and correctly folded Cherry-GST. However, the expression levels are far below those determined at higher cultivation temperatures. For all investigated setups, the IPTG concentration did not show any major influence on soluble Cherry-GST formation (Fig. 2). Also the growth curves of different inducer concentrations at 37°C (Fig. 3) do not deviate largely. In contrast to this, growth curves for different induction times and cultivation temperatures (data not shown) differed significantly, indicating large variances in contaminant levels of the respective cell lysates. Those variations in contaminant concentrations can be related to adaption of *E. coli* to changes in temperature, cell cycle and metabolism. Changes in cultivation temperature induce e.g. the formation of heat or cold shock proteins, temperature-adapted enzymes, and alternative transport mechanism proteins [13, 14, 31, 32, 33]. Additionally, different induction conditions resulted in a change of the cell cycle and overburden metabolism [17, 18]. It was therefore suggested that the overall system optimized for up- and downstream processing must not necessarily be the one with highest product titers. As a feasibility study, 5 upstream conditions were chosen randomly in Matlab[®] according to the description of the 'Random Sampling Procedure' in the experimental section. The selected setups (Fig. 2 - arrows) included:

- 27 °C - OD_{600 nm} of 1 AU - 0.1 mM IPTG
- 27 °C - OD_{600 nm} of 8 AU - 0.1 mM IPTG
- 32 °C - OD_{600 nm} of 4 AU - 0.5 mM IPTG
- 32 °C - OD_{600 nm} of 8 AU - 0.1 mM IPTG
- 37 °C - OD_{600 nm} of 8 AU - 0.1 mM IPTG

3.2 ChromX Model Calibration

For the selected lysates from the micro-scale cultivations 5, 10, 15, and 20 CV gradients were performed in the '1 mL' HiTrap[™] Q FF column for ChromX model calibration. To evaluate whether the assumption of low sample loading is valid in the presented setup the collected fractions derived from the 20 CV gradients of all setups were analyzed using quantitative capillary gel electrophoresis. It was shown that the amount of protein bound to the '1 mL' HiTrap[™] Q FF column ranged from 0.2 to 1% of the nominal maximum

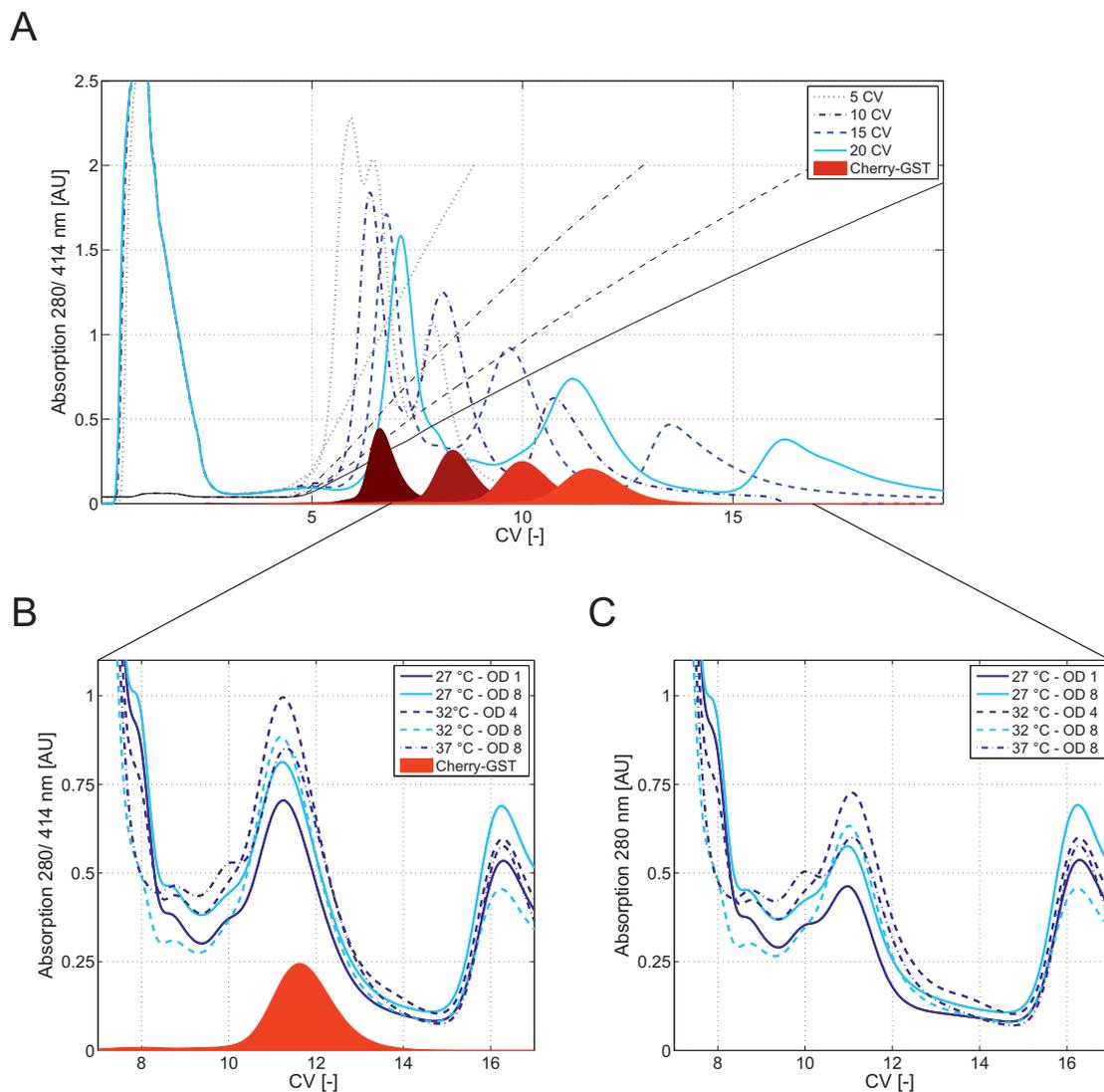


Figure 4: Chromatograms of the ChromX model calibration experiments. The UV_{280 nm} signals are shown as blue lines. Product peaks are indicated as red areas below. The salt concentrations are illustrated as black lines. A: 5, 10, 15, and 20 CV gradients of the lysate derived from the 32 °C cultivation, induced at OD_{600 nm} = 8 with 0.1 mM IPTG. B: Magnification of the product elution window for the 20 CV gradient for all selected cultivation conditions. C: Magnification of the 280 nm signal of the product elution window for the 20 CV gradient for all selected cultivation conditions after subtraction of the product-related 280 nm signal.

	27 °C OD _{600 nm} 1	27 °C OD _{600 nm} 8	32 °C OD _{600 nm} 4	32 °C OD _{600 nm} 8	37 °C OD _{600 nm} 8
	Peak area [mAU · mL]				
Peak 1	1369	711	1455	1049	441
Peak 2	1149	1569	1041	1374	2071
Peak 3	178	210	150	183	309
Peak 4	1661	1651	1042	1374	1594
Peak 5	278	309	189	230	274
Peak 6	198	240	166	205	373
Peak 7	211	256	204	287	376
Peak 8	890	1047	935	1425	1798
Peak 9	574	768	480	688	918
Peak 10	399	548	373	486	653
Peak_{414 nm}	259	260	199	356	459

Table 3: Areas of all 'visible' peaks of the 20 CV gradients determined by zero baseline integration in Unicorn. The 414 nm peak was used for product-specific analysis.

binding capacity (120 mg HSA per mL chromatography resin - Product information sheet www.sigmaaldrich.com). Thus, the simplification of the SMA model was justified.

As an example, the chromatograms of the calibration experiments are shown for the 32 °C cultivation induced at OD_{600 nm} = 8 with 0.1 mM IPTG in Fig. 4A. Flowthrough of non-binding and weakly binding contaminants can be observed during the first 3 CVs. The number of detectable peaks during the salt gradient elution starting after 5 CVs increases with the length of the gradient. Also the elution of Cherry-GST, which is shown as a red area, is shifted to higher column volumes and results in an increased product peak broadening. The 20 CV gradient served as a measure of the number of modeled contaminant peaks, which was identified to be 10 for all setups. The zero baseline peak integration in Unicorn resulted in the peak areas shown in Table 3. Those areas were used as scaling factors for the ChromX model. Obviously, there is no systematic trend of one cultivation condition towards higher or lower peak areas. This means that the previously determined differences in the growth curves due to altered metabolism were confirmed by ion-exchange chromatograms.

A magnification of the product elution window of the 20 CV gradient for all selected cultivation conditions is shown in Fig. 4B. Note that for comparison, the signals are scaled to an equal product peak area. The UV_{280 nm} peak maximum after 11.3 CVs is not equivalent to the Cherry-GST VIS_{414 nm} peak maximum after 11.7 CVs, indicating a closely eluting contaminant. The levels of this critical contaminant (CC) strongly differ within the 5 investigated lysates as shown in Fig. 4C illustrating the overall 280 nm signal subtracted by the 280 nm signal of Cherry-GST. The 27 °C cultivation induced at OD_{600 nm} = 1 with 0.1 mM IPTG reveals the lowest CC levels, resulting in an UV_{280 nm} peak maximum of 0.45 AU. In contrast to this, the 32 °C cultivation induced at OD_{600 nm} = 4 with 0.5 mM IPTG reveals the highest CC levels with an UV_{280 nm} peak maximum of 0.72 AU. The quantitative variances in CC levels can also be seen in the areas of peak

Component	k_{eff}	k_{kin}	k_{eq}	ν
Contaminant 1	0.00049	0.000	0.0	0.000
Contaminant 2	0.00559	0.000	22.5	0.566
Contaminant 3	0.01200	0.000	24.7	1.156
Contaminant 4	0.01300	0.001	21.5	1.890
Contaminant 5	0.01208	0.050	26.2	3.930
Contaminant 6	0.01208	0.020	70.2	4.300
Contaminant 7	0.00808	0.090	243.2	5.670
Contaminant 8	0.00153	0.100	1623.4	6.159
Target	0.00128	0.000	757.1	6.589
Contaminant 9	0.01548	0.125	152227.0	8.968
Contaminant 10	0.00053	0.700	94223.4	7.259

Table 4: Converging SMA parameters of the target and the 10 contaminants for the 5 selected lysates.

8 in Table 3.

The chromatograms were exported from Unicorn 5.2 and implemented in the ChromX simulation. Including the scaling factors and the model parameters (Table 2), the model parameter estimates were obtained using first a heuristic and then a deterministic solver as described in the experimental section. The experimental and simulated data for the 32 °C cultivation induced at $OD_{600\ nm} = 8$ with 0.1 mM IPTG are shown for the 10 CV and 20 CV gradients in Fig. 5. The chromatogram derived from Unicorn (blue dashed line) agrees very well with the $UV_{280\ nm}$ sum signal in ChromX for all cultivation setups and gradients. Hence, the calibration was successful and the models could be used for *in silico* optimization studies. The single contaminant peaks 1 to 10 (grey lines) were resolved decoupled from the $UV_{280\ nm}$ sum signal, and single SMA parameters were estimated. All USP setups consisted of the identical number of 10 contaminant peaks, as was discussed before. As the contaminant levels of all investigated setups differed strongly, regions of overlapping peaks resulted in shifted peak maxima when comparing different cell lysates. For that reason it was first assumed that those shifted peaks incorporated highly distinct protein species and separate models were created for all 5 cell lysates, incorporating the 4 calibration gradients each. However, the SMA parameter sets of all different lysates converged towards similar values as shown in Table 4.

Thus, ChromX revealed that all USP setups consisted of 10 resemblant contaminant peaks which differed in their concentration levels only. To investigate this assumption, the same modeling experiments were carried out in one batch for all lysates, forcing the model to result in identical parameters for all characterized feed stocks. This simulation was found to have a performance equivalent to that of the above-mentioned setup to model all 5 setups in a decoupled way in a much shorter time frame. This even makes an investigation all conditions of micro-scale cultivations feasible as a full factorial design instead of performing a random sampling as shown in this study. For the selected cell systems and cultivation setups, however, it might also be possible to investigate new or different contaminant peaks. Then, the decoupled simulation methodology must be chosen.

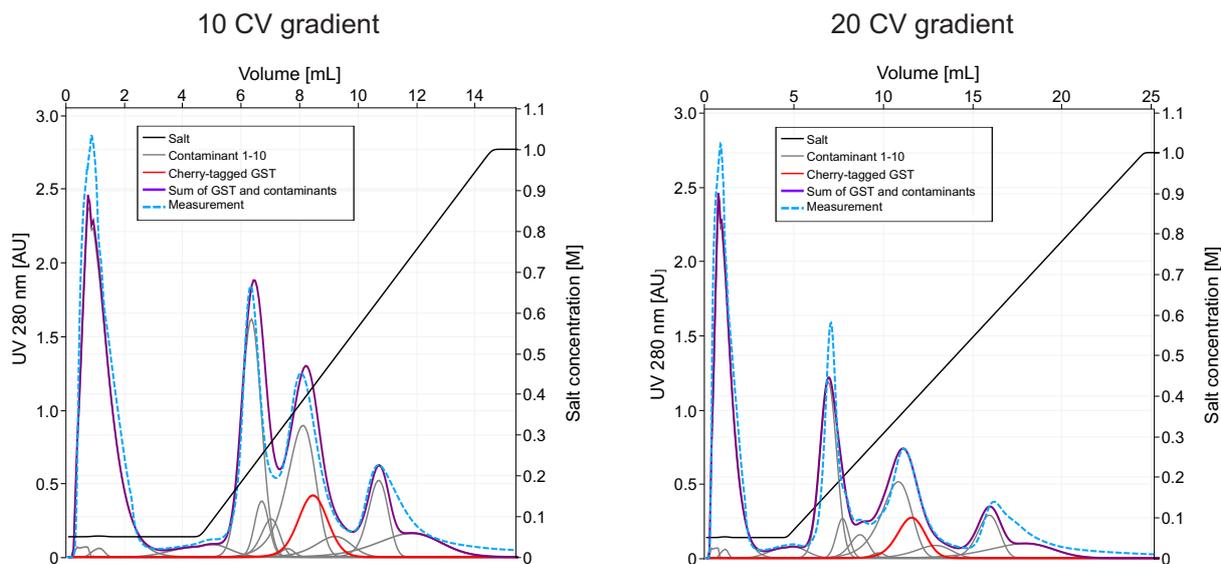


Figure 5: 10 CV and 20 CV gradients of the experimental (blue dashed lines) and ChromX-simulated (purple solid line) chromatograms are shown for the 32 °C cultivation induced at $OD_{600\text{ nm}} = 8$ with 0.1 mM IPTG. The single contaminant peaks (gray lines) and the product peak (red line) are indicated below the $UV_{280\text{ nm}}$ sum signal.

3.3 ChromX Optimization

Using the created models for the 5 lysates, an *in silico* optimization was performed with the help of the ChromX software. The emphasis of the optimization function was laid on yield and purity, resulting in identical purification setups for all samples investigated: After sample application, a 1 CV low-salt wash and a step gradient to 0.21 M NaCl followed, which was kept constant for 5 CVs. Afterwards, the salt concentration was shifted in a step to 0.37 M, followed by a 5 CV gradient to 0.47 M NaCl. In a final high-salt wash and re-equilibration step, the column was prepared for the next chromatography run.

The proposed optimal purification setups were subsequently validated as shown in Fig. 6 for the experimental and simulated data of the 32 °C cultivation, induced at $OD_{600\text{ nm}} = 8$ with 0.1 mM IPTG. Again, the chromatogram derived from Unicorn (blue dashed line) agrees very well with the the $UV_{280\text{ nm}}$ sum signal in ChromX for all cultivation setups. Thus, the chromatogram was correctly predicted by performing 4 different shallow and steep calibration gradients only.

3.4 Integrated Up- & Downstream Performance

The integrated up- and downstream performance was investigated using a Pareto optimization method. The single contaminant peaks 1 to 10 (gray lines) and the product peak (red line) from the optimized purification model were used to calculate the purities Pur_i and recoveries Rec_i for all possible peak fractionation setups, as was described in the experimental section. The product recoveries and purities strongly differed depending on the fractionation boundaries. The *in silico* procedure of scanning all product fractionation possibilities resulted in the Pareto fronts for the 5 lysates shown in Fig. 7. The overlaps of those optimal system points resulted in a global Pareto front of all lysates,

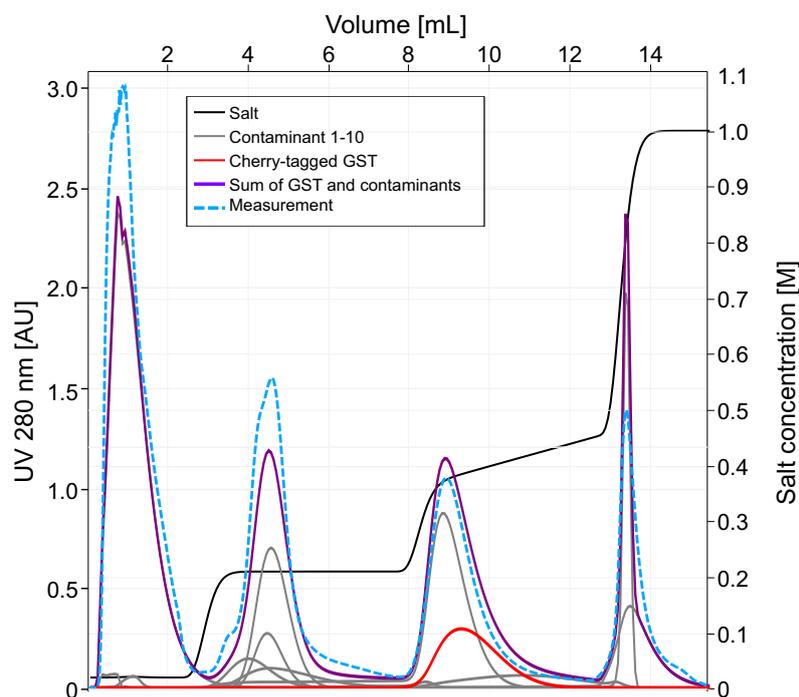


Figure 6: Optimized purification setup of the experimental (blue dashed lines) and ChromX-simulated (purple solid line) chromatograms shown for the 32 °C cultivation induced at $OD_{600\text{ nm}} = 8$ with 0.1 mM IPTG. The single contaminant peaks (gray lines) and the product peak (red line) are indicated below the $UV_{280\text{ nm}}$ sum signal.

which describes the feasible conditions. The optimum recovery was found for the cultivation at 37 °C induced at $OD_{600\text{ nm}} = 8$ with 0.1 mM IPTG (green triangles in Fig. 7). This result is quite intuitive, as most parts of the product peak are fractionated for this setup and the cultivation of highest initial upstream product titers performs best. However, the optimum purity was found for the cultivation at 27 °C induced at $OD_{600\text{ nm}} = 1$ with 0.1 mM IPTG resulting in a purity of up to 65% (dark blue circles in Fig. 7). For all other cultivation setups, a purity above 58% could not be reached. The lowest process performance was identified for the cultivation at 32 °C induced at $OD_{600\text{ nm}} = 4$ with 0.5 mM IPTG (dark red stars in Fig. 7). As discussed, the global Pareto front describes a set of optimal system points which have to be selected based on the value of the product.

It becomes obvious that the cultivation conditions are of crucial importance to downstream processing and the overall process performance due to different contaminant levels and species. The critical contaminants (CCs) which elute close to the product are the factors having the highest influence.

4 Conclusions

The optimization approach based on high-throughput micro-scale cultivation experiments, chromatography modeling, and Pareto optimization was found to be an elegant tool for integrated up- and downstream process optimization. Micro-scale cultivations in the BioLector[®] system helped find feasible cultivation conditions. Besides variations of

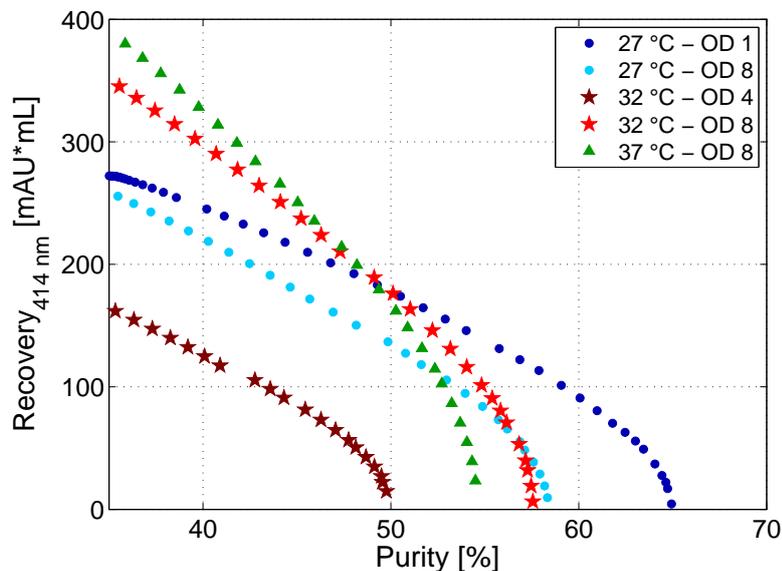


Figure 7: Pareto optimization based on different fractionation setups based on the purities and recoveries of the 5 selected lysates in the pooled elution fractions. The Pareto fronts of the 27 °C setups are shown as blue circles, the 32 °C setups as red and orange stars, and the 37 °C setup as a green triangle.

product titers, large differences in the growth curves were observed, indicating distinct contaminant levels. Instead of selecting cultivations of highest Cherry-GST titers, a feasibility study applying random sampling was conducted. Two criteria were specified: To eliminate all conditions of marginal product concentrations, a *'Threshold Criterion'* was given. Sampling from the resulting conditions was then performed randomly, with the selection probability being based on the product titers (*'Probability Criterion'*). For the 5 chosen lysates, chromatography models were successfully created using 4 calibration gradients in an AEX chromatography column. ChromX revealed that all investigated USP setups consisted of 10 contaminant peaks at different concentration levels only. One coupled model with scaled contaminant levels and identical SMA parameter sets for all investigated lysates was found to perform similarly to decoupled models allowing for distinct SMA parameter sets for the five selected lysates. This even allows for an investigation of all conditions of the micro-scale cultivations as a full factorial design instead of performing a random sampling as shown in this study. However, for cell systems where new contaminant peaks are investigated, induced by different cultivation conditions, the decoupled simulation methodology is the method of choice. An *in silico* optimization of the AEX purification process based on the created models was performed and successfully validated. It was shown that the shallow and steep gradients from the calibration runs could correctly predict steps as well as gradients. As a final step, the integrated up- and downstream performance was determined by a Pareto optimization based on product recovery (initial titer · yield) and purity. The Pareto fronts of the selected lysates resulted in a global set of optimal system points which need to be set based on the value of the product.

In summary, the integrated approach showed large potential for concerted up- and downstream optimization and the upstream conditions were found to result in strong variations in the levels of critical contaminants (CC). The procedure now needs to be tested for fur-

ther systems, e.g. for eucaryotic organisms like *P. pastoris*.

References

- [1] C. H. Schein, M. H. M. Noteborn, Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature, *Nat Biotechnol* 6 (1988) 291–294.
- [2] K. L. Dikshit, R. P. Dikshit, D. A. Webster, Study of *Vitreoscilla globin* (vgb) gene expression and promoter activity in *E. coli* through transcriptional fusion., *Nucleic Acids Res* 18 (1990) 4149–55.
- [3] M. Losen, B. Frölich, M. Pohl, J. Büchs, Effect of oxygen limitation and medium composition on *Escherichia coli* fermentation in shake-flask cultures., *Biotechnol Prog* 20 (2004) 1062–8.
- [4] J. I. Betts, F. Baganz, Miniature bioreactors: current practices and future opportunities., *Microb Cell Fact* 5 (2006) 1–14.
- [5] R. Puskeiler, K. Kaufmann, D. Weuster-Botz, Development, parallelization, and automation of a gas-inducing milliliter-scale bioreactor for high-throughput bioprocess design (HTBD)., *Biotechnol Bioeng* 89 (2005) 512–23.
- [6] S. R. Lamping, H. Zhang, B. Allen, P. Ayazi Shamlou, Design of a prototype miniature bioreactor for high throughput automated bioprocessing, *Chem Eng Sci* 58 (2003) 747–758.
- [7] J. Altenbach-Rehm, C. Nell, M. Arnold, D. Weuster-Botz, Parallel Bubble Columns with Fed-Batch Technique for Microbial Process Development on a Small Scale, *Chem Eng Technol* 22 (1999) 1051–1058.
- [8] M. Funke, S. Diederichs, F. Kensy, C. Mueller, J. Buechs, The Baffled Microtiter Plate : Increased Oxygen Transfer and Improved Online Monitoring in Small Scale Fermentations, *Biotechnol Bioeng* 103 (2009) 1118–1128.
- [9] F. Kensy, E. Zang, C. Faulhammer, R. Tan, J. Buechs, Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates, *Microb Cell Fact* 8 (2009) 1–17.
- [10] K. Kottmeier, J. Weber, C. Mueller, T. Bley, J. Buechs, Asymmetric Division of *Hansenula polymorpha* Reflected by a Drop of Light Scatter Intensity Measured in Batch Microtiter Plate Cultivations at Phosphate Limitation, *Biotechnol Bioeng* 104 (2009) 554–561.
- [11] M. Marino, T. Hoffmann, R. Schmid, H. Möbitz, D. Jahn, Changes in protein synthesis during the adaptation of *Bacillus subtilis* to anaerobic growth conditions., *Microbiology* 146 (2000) 97–105.

-
- [12] J. Starck, G. Källenius, B. Marklund, D. I. Andersson, T. Akerlund, Comparative proteome analysis of *Mycobacterium tuberculosis* grown under aerobic and anaerobic conditions., *Microbiology* 150 (2004) 3821–3829.
- [13] I. Budde, L. Steil, C. Scharf, U. Völker, E. Bremer, Adaptation of *Bacillus subtilis* to growth at low temperature: a combined transcriptomic and proteomic appraisal., *Microbiology* 152 (2006) 831–53.
- [14] P. C. Y. Woo, S. K. P. Lau, H. Tse, J. L. L. Teng, S. O. T. Curreem, A. K. L. Tsang, R. Y. Y. Fan, G. K. M. Wong, Y. Huang, N. J. Loman, L. A. S. Snyder, J. J. Cai, J. Huang, W. Mak, M. J. Pallen, S. Lok, K. Yuen, The complete genome and proteome of *Laribacter hongkongensis* reveal potential mechanisms for adaptations to different temperatures and habitats., *PLoS genetics* 5 (2009) 1–12.
- [15] K. Lee, H. Lee, K. Pi, Y. Choi, The effect of low pH on protein expression by the probiotic bacterium *Lactobacillus reuteri*., *Proteomics* 8 (2008) 1624–1630.
- [16] L. M. Stancik, D. M. Stancik, B. Schmidt, M. Barnhart, Y. N. Yoncheva, D. M. Barnhart, J. L. Slonczewski, pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia coli*, *J Bacteriol* 184 (2002) 4246–4258.
- [17] B. Grünenfelder, G. Rummel, J. Vohradsky, D. Röder, H. Langen, U. Jenal, Proteomic analysis of the bacterial cell cycle., *Proc Natl Acad Sci USA* 98 (2001) 4681–4686.
- [18] K. Dürschmid, H. Reischer, W. Schmidt-Heck, T. Hrebicek, R. Guthke, A. Rizzi, K. Bayer, Monitoring of transcriptome and proteome profiles to investigate the cellular response of *E. coli* towards recombinant protein expression under defined chemostat conditions., *J Biotech* 135 (2008) 34–44.
- [19] A. Osberghaus, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch, Optimizing a chromatographic three component separation: a comparison of mechanistic and empiric modeling approaches., *J Chrom A* 1237 (2012) 86–95.
- [20] D. Karlsson, N. Jakobsson, A. Axelsson, B. Nilsson, Model-based optimization of a preparative ion-exchange step for antibody purification, *J Chrom A* 1055 (2004) 29–39.
- [21] M. Degerman, N. Jakobsson, B. Nilsson, Constrained optimization of a preparative ion-exchange step for antibody purification., *J Chrom A* 1113 (2006) 92–100.
- [22] J. M. Mollerup, T. B. Hansen, S. Kidal, L. Sejergaard, A. Staby, Development, modelling, optimisation and scale-up of chromatographic purification of a therapeutic protein, *FFE* 261 (2007) 133–139.
- [23] E. J. Close, J. R. Salm, D. G. Bracewell, E. Sorensen, Modelling of industrial biopharmaceutical multicomponent chromatography, *Chem Eng Res Des* 92 (2014) 1304–1314.

- [24] J. Cornel, A. Tarafder, S. Katsuo, M. Mazzotti, The direct inverse method: A novel approach to estimate adsorption isotherm parameters, *J Chrom A* 1217 (2010) 1934–1941.
- [25] Delphi Genetics SA, Cherry Codon Kit., 2009.
- [26] Delphi Genetics SA, Cherry Express Kit., 2009.
- [27] M. Wall, Galib: A c++ library of genetic algorithm components, ver. 2.4.7, <http://lancet.mit.edu/ga/> (1996-2005).
- [28] F. Devernay, C/c++ minpack, ver. 1.3.2, <http://devernay.free.fr/hacks/cminpack/> (2007-2013).
- [29] C. A. Brooks, S. M. Cramer, Steric mass-action ion exchange: Displacement profiles and induced salt gradients, *AIChE Journal* 38 (1992) 1969–1978.
- [30] M. Michel, A. Epping, A. Jupke, *Preparative Chromatography: Of Fine Chemicals and Pharmaceutical Agents*, Wiley-VCH Verlag GmbH & Co. KGaA, 2005, pp. 215–312.
- [31] J. Weiner III, C. U. Zimmermann, H. W. Goehlmann, R. Herrmann, Transcription profiles of the bacterium *Mycoplasma pneumoniae* grown at different temperatures, *Nucleic Acids Res* 31 (2003) 6306–6320.
- [32] L. M. Smoot, J. C. Smoot, M. R. Graham, G. A. Somerville, D. E. Sturdevant, C. A. Migliaccio, G. L. Sylva, J. M. Musser, Global differential gene expression in response to growth temperature alteration in group A *Streptococcus*., *Proc Natl Acad Sci USA* 98 (2001) 10416–10421.
- [33] J. Wang, C. Zhao, B. Meng, J. Xie, C. Zhou, X. Chen, K. Zhao, J. Shao, Y. Xue, N. Xu, Y. Ma, S. Liu, The proteomic alterations of *Thermoanaerobacter tengcongensis* cultured at different temperatures., *Proteomics* 7 (2007) 1409–1419.

4 Conclusion & Outlook

This PhD thesis contributes to solving issues that arise in the biopharmaceutical industry during process development and characterization. One main focus was to develop strategies complying with the 'Quality by Design' and 'Time to Market' guidelines in terms of a fast design of robust and well characterized processes. Another challenge was to loosen the strict separation between up- and downstream process development ('over the wall' workflow) and to facilitate concerted up- and downstream optimization strategies. In that context the following topics were covered:

- Advanced analytical technologies for HTS and modeling applications
- Systematic HTS strategies for downstream intensification
- Mechanistic chromatography modeling based on process UV data
- Integrated up- and downstream process development

In the first section different approaches for HTS compatible analytical technologies were evaluated and applied. The Cherry-TagTM as a protein fusion tag of special optical properties proved to be a highly sophisticated tool for real-time and in-line tracking of target molecules. The tag even allowed for a distinction between the soluble and active state of the protein and the insoluble or denatured state in different applications of up- and downstream process development. This property was exploited during screenings on feasible cultivation conditions to prevent the formation of misfolded inclusion bodies and the results were directly transferable to a production process of the untagged protein species. Also the native state of the product can be assured during downstream applications so that critical process steps can be changed or replaced to result in an effective overall process. Despite changing the protein structure, additional straight-to-use analytical technologies were investigated. Multi-variate data analysis (MVDA) of protein spectra was used as a fast and non-invasive method which was implemented into the HTS workflow of a robotic workstation and helped preventing analytical bottlenecks created during HTS experiments. Such empirical partial least squares - projection to latent structures (PLS) models were successfully used for deconvolution of HTS multi-component isotherm data and resulted in modeling quality data. Additionally, MVDA facilitated the fast and straightforward distinction of native and denatured protein species during optimization studies on hydrophobic interaction chromatography. As a final HTS compatible analytical tool, high-throughput capillary gel electrophoresis (HT-GCE) was evaluated. In contrast to traditional analytical methods being based on e.g. chromatography (processing time per sample of several minutes even in optimized setups), HT-CGE (processing time per sample approximately 40 s) was found to be highly effective for large data sets. In HT upstream applications HT-CGE was used for the characterization of the soluble and insoluble protein fractions of crude feed stocks and for providing a ranking on the performance of different cultivation conditions. In addition, a high-resolution characterization of the subsequent downstream process was possible as a large number of collected fractions could be analyzed by HT-CGE. Thus, three very fast, reliable and straightforward analytical tools for HTS and modeling applications were investigated and ready-to-use for diverse applications.

The second section of this PhD thesis proposed systematic HTS strategies for downstream intensification (deterministic experimental approach). For ion-exchange chromatography a methodology was introduced being suited for both salt-stable and salt-intolerant proteins. An optimized product capturing step under stabilized conditions was ensured by a combination of stability and additive screenings in high-throughput as well as a deterministic chromatofocusing strategy for finding optimal operational conditions. For salt-intolerant proteins a straightforward pH shift purification protocol was introduced as an alternative to the traditional salt elution mode. A similar deterministic experimental approach was developed for hydrophobic interaction chromatography (HIC). HTS solubility screenings and miniaturized robotic column chromatography experiments resulted in different correlations of the protein's isoelectric point, overall solubility and the HIC binding behavior. Conditions of decreased solubility as e.g. close to the isoelectric point enhanced protein binding without affecting the protein's structural integrity. Effective protein capture steps were again designed based on well-chosen prior screening experiments. Such systematic strategies should now be extended to other downstream unit operations like aqueous two-phase systems (ATPS) for resulting in a global toolbox for process design.

The third section of the thesis focused on mechanistic models for chromatography using process UV data. The idea was to move away from modeling based on known concentrations and molar fractions to standard industrial process UV data. Changes in the downstream workflow can be detected directly, which opens up new possibilities for real-time intervention and increased effectivity of processes. The UV-based approach for mechanistic modeling of a complex feed stock was realized by re-writing all model equations from molar concentrations to absorption based units by introducing absorption coefficients. As a result standard parameter estimating procedures using chromatogram fitting (inverse modeling) were applicable for standard process UV data.

Finally, high-throughput technologies and the straightforward modeling toolbox were merged to an integrated up- and downstream process optimization strategy. The pre-selection for suited upstream conditions was based on HT cultivation screenings, excluding settings of unfeasible product levels. However, the subsequent lysate characterization using the above-mentioned procedure of mechanistic UV-based inverse modeling of chromatography data was not only performed for the systems of highest product titers as in industry but also on conditions of moderate productivity. The created highly predictive ion-exchange chromatography models for each investigated lysate were then optimized *in silico* for a dedicated product capturing step and analyzed by *in silico* sampling of all possible elution pooling setups. A multi-objective optimization finally resulted in a global set of optimal system points of the integrated process which suggested different cultivation strategies depending on the value of the product. In contradiction to the general practice in industry, this study emphasized that the downstream process is highly influenced by the cultivation conditions and therefore provided a basis for analysis of different other cell systems, vectors and products. With the basic tools given for such investigations fixed conditions for upstream processes in industry (e.g. antibody processes) should be questioned and re-evaluated on the basis of such new insights.

In summary, this PhD thesis covers diverse strategies in process development and characterization as demanded by the biopharmaceutical industry. It resulted in a toolbox of HTS compatible analytical technologies for up- and downstream applications preventing

a bottleneck for large sets of samples. The proposed deterministic experimental strategies might become a routine in process development as an alternative to current practice like heuristic approaches. Finally, the UV-based inverse modeling method helps identifying process inconsistencies and, moreover, realizing sophisticated approaches of concerted up- and downstream optimization.

5 Abbreviations

Abbreviation	Definition
4-NP	4-nitrophenole
4-NPG	4-nitrophenyl- α -D-galactopyranoside
AEX	Anion-exchange
ambr	Advanced micro-scale bioreactor
Amp ^R	Ampicillin resistance
araBAD	Arabinose metabolic operon
ATPS	Aqueous two-phase system
BET	BrunauerEmmettTeller
BIS-TRIS	2,2-Bis(hydroxymethyl)-2,2',2''-nitrilotriethanol
BSA	Bovine serum albumine
BTC	Breakthrough curve
cAMP	Cyclic adenosine monophosphate
CAPS	3-(Cyclohexylamino)-1-propanesulfonic acid
CAPSO	3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic acid
CC	Critical contaminant
ccdA	Bacterial antidote
ccdB	Bacterial toxin
cDNA	Complementary deoxyribonucleic acid
CDR	Convection diffusion reaction
CEX	Cation-exchange
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHES	2-(Cyclohexylamino)ethanesulfonic acid
CHO	Chinese hamster ovary cell
CIP	Cleaning in place
CV	Column volume
Cyt c	Cytochrome c
DBC	Dynamic binding capacity
DIOS	Desorption/ ionization on silicone
DNA	Deoxyribonucleic acid
DoE	Design of experiments
DSP	Downstream process
DTT	Dithiothreitol
EcFbFP	Fluorescent protein
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immuno sorbent assay
EMA	European medicines agency
ESI-MS	Electrospray ionization mass spectrometry
F _c	Crystallisable fragment
Fab	Fragment antigen binding
FEM	Finite element method
FPLC	Fast protein liquid chromatography
G-CSF	Granulocyte colony-stimulating factor

5 ABBREVIATIONS

GA	Genetic algorithm
GFP	Green fluorescent protein
GST	Glutathione-S-Transferase
GX	Gel electrophoresis
HCl	Hydrochloric acid
HCP	Host cell protein
HEPPSO	N-(2-Hydroxyethyl)piperazine-N'-(2-hydroxypropanesulfonic acid)
HIC	Hydrophobic interaction chromatography
His	Histidine
HPLC	High-performance liquid chromatography
HSA	Human serum albumine
HT	High-throughput
HT-CGE	High-throughput capillary gel electrophoresis
HTE	High-throughput experiment
HTPD	High-throughput process development
HTS	High-throughput screening
IB	Inclusion body
IEX	Ion-exchange
IgG	Immunoglobuline G
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
Kan ^R	Kanamycin resistance
kb	kilo base pare
kDa	kilo Dalton
lac	lactose-operon
LB	Lysogeny broth
LDS	Lauryl dodecyl sulfate
LHS	Liquid handling system
LU	Lower upper
LV	Latent variable
Lys	Lysine/ Lysozyme
mAB	Monoclonal antibody
MALDI-MS	Matrix-assisted laser desorption ionization mass spectrometry
MAT	Metal affinity tag
MCS	Multiple cloning site
MD	Molecular dynamics
MES	2-(N-morpholino)ethanesulfonic acid
MOPSO	3-Morpholino-2-hydroxypropanesulfonic acid
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MTP	Microtiter plate
MVDA	Multi-variate data analysis
MW	Molecular weight
MWCO	Molecular weight cutoff
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NaS	Sodium sulfate

OD	Optical density
OED	Optimal experimental design
ori	Origin of replication
OTR	Oxygen transfer rate
PA	anthrax-protective gene
PapMV	Papaya mosaic virus
pBR322	Prokaryotic expression vector (Plasmid)
PBS	Phosphate buffered saline
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PES	Polyether sulfone
pH	Pondus Hydrogenii
pI	Isoelectric point
pK_a	Acid dissociation constant
pL/ pR	Bacteriophage λ promoter
PLS	Partial least squares/ projection to latent structures
PTFE	Polytetrafluoroethylene
pUC	Prokaryotic expression vector (Plasmid)
Q	Quaternary ammonium
QbD	Quality by Design
QSAR	Quantitative structure-activity relationship
QSPR	Quantitative structure-property relationship
RAMOS	Respiration activity monitoring system
RNA	Ribonucleic acid
rPTH	Recombinant parathyroid hormone
S	Sulfite
scFv	Single chain variable fragment
SDR	Short-chain dehydrogenase/ reductase
SDS	Sodium dodecyl sulfate
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Sf9	Spodoptera frugiperda cell line
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SP FF	Sulfopropyl fast flow
SMA	Steric mass-action
SUPG	Streamline-Upwind-Petrov-Galerkin
tac	Hybrid of lactose- and tryptophane-operon
TAPS	N-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TB	Terrific broth
TBS(T)	Tris buffered saline (including Tween 20)
TCA	Trichloro acetic acid
TDM	Transport dispersive model

5 ABBREVIATIONS

TEM	Transmission electron microscopy
Tris	Tris(hydroxymethyl)-aminomethan
tRNA	Transfer ribonucleic acid
trp	tryptophane-operon
TSP	Total soluble protein
USP	Upstream process
USPD	Upstream process development
UV	Ultraviolet
VLP	Virus-like particle
V _H H	Single domain antibody
VIS	Visible
VP1	Murine polyomavirus capsid protein
WHcAg	Woodchuck hepadnavirus core protein
YFP	Yellow fluorescent protein

References

- AHUJA, S. (2000). *Handbook of Bioseparations*, pages 9–21. Academic Press.
- AKKOC, N., GHAMAT, A. and AKCELIK, M. (2011). *Optimisation of bacteriocin production of Lactococcus lactis subsp. lactis MA23, a strain isolated from Boza*. Int J Dairy Technol, 64:425.
- ALTENBACH-REHM, J., NELL, C., ARNOLD, M. and WEUSTER-BOTZ, D. (1999). *Parallel Bubble Columns with Fed-Batch Technique for Microbial Process Development on a Small Scale*. Chem Eng Technol, 22:1051.
- ANDERLEI, T., ZANG, W., PAPASPYROU, M. and BUECHS, J. (2004). *Online respiration activity measurement (OTR, CTR, RQ) in shake flasks*. Biochem Eng J, 17:187.
- AZHAR, M. and SOMASHEKHAR, R. (2014). *Cloning, expression and purification of human and bovine Enterokinase light chain with Cherry tag and their activity comparison*. Indian J Applied and Pure Bio, 29:125.
- BANEYX, F. (2004). *Protein Expression Technologies: Current Status and Future Trends*, pages 40–44. Horizon Bioscience.
- BARTON, L. (2005). *Structural and Functional Relationships in Prokaryotes*, pages 302–307. Springer Verlag.
- BAUMGARTNER, K., GALM, L., NOETZOLD, J., SIGLOCH, H., MORGENSTERN, J., SCHLEINING, K., SUHM, S., OELMEIER, S. and HUBBUCH, J. (2015). *Determination of protein phase diagrams by microbatch experiments: Exploring the influence of precipitants and pH*. Int J Pharm, 479:28.
- BENSCH, M., SELBACH, B. and HUBBUCH, J. (2007). *High-throughput screening techniques in downstream processing: Preparation, characterization and optimization of aqueous two-phase systems*. Chem Eng Sci, 62:2011.
- BERG, A., OELMEIER, S. A., KITTELMANN, J., DISMER, F. and HUBBUCH, J. (2012). *Development and characterization of an automated high throughput screening method for optimization of protein refolding processes*. J Sep Sci, 35:3149.
- BETTS, J. I. and BAGANZ, F. (2006). *Miniature bioreactors: current practices and future opportunities*. Microb Cell Fact, 5:1.
- BHAMBURE, R., KUMAR, K. and RATHORE, A. S. (2011). *High-throughput process development for biopharmaceutical drug substances*. Trends Biotechnol, 29:127.
- BRESTRICH, N., BRISKOT, T., OSBERHAUS, A. and HUBBUCH, J. (2015). *Advances in inline quantification of co-eluting proteins in chromatography: Process-data-based model calibration and application towards real-life separation issues*. Biotechnol Bioeng, 111:1365.

REFERENCES

- BRESTRICH, N., SANDEN, A., KRAFT, A., MCCANN, K., BERTOLINI, J. and HUBBUCH, J. (2014). *A tool for selective inline quantification of co-eluting proteins in chromatography using spectral analysis and partial least squares regression*. Biotechnol Bioeng doi:10.1002/bit.25546.
- BROOKS, C. A. and CRAMER, S. M. (1992). *Steric mass-action ion exchange: Displacement profiles and induced salt gradients*. AIChE Journal, 38:1969.
- BUDDE, I., STEIL, L., SCHARF, C., VÖLKER, U. and BREMER, E. (2006). *Adaptation of Bacillus subtilis to growth at low temperature: a combined transcriptomic and proteomic appraisal*. Microbiology, 152:831.
- BURGESS-BROWN, N. A., SHARMA, S., SOBOTT, F., LOENARZ, C., OPPERMAN, U. and GILEADI, O. (2008). *Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study*. Protein Express Purif, 59:94.
- CABANTOUS, S., TERWILLIGER, T. C. and WALDO, G. S. (2005). *Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein*. Nature Biotechnol, 23:102.
- CARTA, G. and JUNGBAUER, A. (2010). *Protein Chromatography*, pages 85–106. Wiley VCH.
- CHHATRE, S., FARID, S. S., COFFMAN, J., BIRD, P., NEWCOMBE, A. R. and TITCHENER-HOOKER, N. J. (2011). *How implementation of quality by design and advances in biochemical engineering are enabling efficient bioprocess development and manufacture*. J Chem Technol Biotechnol, 86:1125.
- CHMIEL, H. (2011). *Bioprozesstechnik*, pages 373–427. Spektrum Akademischer Verlag.
- CLOSE, E. J., SALM, J. R., BRACEWELL, D. G. and SORENSEN, E. (2014). *Modelling of industrial biopharmaceutical multicomponent chromatography*. Chem Eng Res Des, 92:1304.
- COFFMAN, J. L., KRAMARCZYK, J. F. and KELLEY, B. D. (2008). *High-throughput screening of chromatographic separations: I. method development and column modeling*. Biotechnol Bioeng, 100:605.
- DE BOER, H. A., COMSTOCK, L. J. and VASSER, M. (1983). *The tac promoter: A functional hybrid derived from the trp and lac promoters*. Proc Natl Acad Sci USA, 80:21.
- DE MARCO, A., VIGH, L., DIAMANT, S. and GOLOUBINOFF, P. (2005). *Native folding of aggregation-prone recombinant proteins in escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones*. Cell Stress Chaperones, 10:329.
- DEGERMAN, M., JAKOBSSON, N. and NILSSON, B. (2006). *Constrained optimization of a preparative ion-exchange step for antibody purification*. J Chrom A, 1113:92.
- DEMAIN, A. L. and VAISHNAV, P. (2009). *Production of recombinant proteins by microbes and higher organisms*. Biotech Adv, 27:297.

- DIKSHIT, K. L., DIKSHIT, R. P. and WEBSTER, D. A. (1990). *Study of Vitreoscilla globin (vgb) gene expression and promoter activity in E. coli through transcriptional fusion*. Nucleic Acids Res, 18:4149.
- DUETZ, W. A. (2007). *Microtiter plates as mini-bioreactors: miniaturization of fermentation methods*. Trends Microbiol, 15:469.
- DÜRRSCHMID, K., REISCHER, H., SCHMIDT-HECK, W., HREBICEK, T., GUTHKE, R., RIZZI, A. and BAYER, K. (2008). *Monitoring of transcriptome and proteome profiles to investigate the cellular response of E. coli towards recombinant protein expression under defined chemostat conditions*. J Biotech, 135:34.
- ELVIN, C. M., THOMPSON, P. R., ARGALL, M. E., HENDRY, P., STAMFORD, N. P. J., LILLEY, P. E. and DIXON, N. E. (1990). *Modified bacteriophage lambda promoter vectors for overproduction of proteins in escherichia coli*. Gene, 87:123.
- FOO, K. Y. and HAMEED, B. H. (2010). *Insights into the modeling of adsorption isotherm systems*. Chem Eng J, 156:2.
- FORSSÉN, P., ARNELL, R. and FORNSTEDT, T. (2006). *An improved algorithm for solving inverse problems in liquid chromatography*. Comput Chem Eng, 30:1381.
- FUNKE, M., BUCHENAUER, A., SCHNAKENBERG, U., MOKWA, W., DIEDERICHS, S., MERTENS, A., MUELLER, C., KENSY, F. and BUECHS, J. (2010). *Microfluidic BioLector - Microfluidic Bioprocess Control in Microtiter Plates*. Biotechnol Bioeng, 107:497.
- FUNKE, M., DIEDERICHS, S., KENSY, F., MUELLER, C. and BUECHS, J. (2009). *The Baffled Microtiter Plate : Increased Oxygen Transfer and Improved Online Monitoring in Small Scale Fermentations*. Biotechnol Bioeng, 103:1118.
- GALLOWAY, C. A., SOWDEN, M. P. and SMITH, H. C. (2003). *Increasing the yield of soluble recombinant protein expressed in E. coli by induction during late log phase*. Biotechniques, 34:524.
- GAWLITZEK, M., VALLEY, U., NIMTZ, M., WAGNER, R. and CONRADT, H. S. (1995). *Characterization of changes in the glycosylation pattern of recombinant proteins from bhk-21 cells due to different culture conditions*. J Biotech, 42:117.
- GILL, R. T., DE LISA, M. P., VALDES, J. J. and BENTLEY, W. E. (2000). *Genomic analysis of high-cell-density recombinant Escherichia coli fermentation and 'cell conditioning' for improved recombinant protein yield*. Biotechnol Bioeng, 72:85.
- GRGACIC, E. V. L. and ANDERSON, D. A. (2006). *Virus-like particles: Passport to immune recognition*. Methods, 40:60.
- GRÜNENFELDER, B., RUMMEL, G., VOHRADSKY, J., RÖDER, D., LANGEN, H. and JENAL, U. (2001). *Proteomic analysis of the bacterial cell cycle*. Proc Natl Acad Sci USA, 98:4681.

REFERENCES

- GUINA, T., WU, M., MILLER, S. I., PURVINE, S. O., YI, E. C., ENG, J., GOODLETT, D. R., AEBERSOLD, R., ERNST, R. K. and LEE, K. A. (2003). *Proteomic analysis of Pseudomonas aeruginosa grown under magnesium limitation*. J Am Soc Mass Spectrom Chem, 14:742.
- GUIOCHON, G. and BEAVER, L. A. (2011). *Separation science is the key to successful biopharmaceuticals*. J Chrom A, 1218:8836.
- GUZMAN, L., BELIN, D., CARSON, M. J. and BECKWITH, J. (1995). *Tight regulation, modulation, and high-level expression by vectors containing the arabinose pbad promoter*. J Bacteriol, 177:4121.
- HARRISON, S. T. L. (1991). *Bacterial cell disruption: A key unit operation in the recovery of intracellular products*. Biotech Adv, 9:217.
- HERMANN, R., LEHMANN, M. and BÜCHS, J. (2003). *Characterization of gas-liquid mass transfer phenomena in microtiter plates*. Biotechnol Bioeng, 81:178.
- HERRMANN, T., SCHROEDER, M. and HUBBUCH, J. (2006). *Generation of equally sized particle plaques using solid-liquid suspensions*. Biotechnol Prog, 22:914.
- HOHNADDEL, M., FELDEN, L., FIJULJANIN, D., JOUETTE, S. and CHOLLET, R. (2014). *A new ultrasonic high-throughput instrument for rapid DNA release from microorganisms*. J Microbiol Meth, 99:71.
- HSU, W. and AULAKH, R. P. S. (2012). *Advanced microscale bioreactor system: a representative scale-down model for bench-top bioreactors*. Cytotechnology, 64:667.
- HUBER, R., RITTER, D., HERING, T., HILLMER, A., KENSY, F., MUELLER, C., WANG, L. and BUECHS, J. (2009). *Robo-Lector - a novel platform for automated high-throughput cultivations in microtiter plates with high information content*. Microb Cell Fact, 8:1.
- HUBER, R., ROTH, S., RAHMEN, N. and BUECHS, J. (2011). *Utilizing highthroughput experimentation to enhance specific productivity of an E.coli T7 expression system by phosphate limitation*. BMC Biotechnol, 22:1.
- HUUK, T. C., HAHN, T., OSBERGHAUS, A. and HUBBUCH, J. (2014). *Model-based integrated optimization and evaluation of a multi-step ion exchange chromatography*. Sep Purif Technol, 136:207.
- ICH (2011). *Ich-endorsed guide for ich q8/q9/q10 implementation*. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- JAIN, K. K. (2009). *Textbook of Personalized Medicine*, pages 1–17. Springer.
- JANSON, J. (2011). *Principles, High Resolution Methods, and Applications*, pages 7–14. Wiley VCH.

- KANG, C. and WU, C. (1987). *Studies on sp6 promoter using a new plasmid vector that allows gene insertion at the transcription initiation site*. Nucleic Acids Res, 15:2279.
- KARLSSON, D., JAKOBSSON, N., AXELSSON, A. and NILSSON, B. (2004). *Model-based optimization of a preparative ion-exchange step for antibody purification*. J Chrom A, 1055:29.
- KELLEY, B. D., SWITZER, M., BASTEK, P., KRAMARCZYK, J. F., MOLNAR, K., YU, T. and COFFMAN, J. (2008). *High-throughput screening of chromatographic separations: Iv. ion-exchange*. Biotechnol Bioeng, 100:950.
- KENSY, F., ENGELBRECHT, C. and BUECHS, J. (2009a). *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*. Microb Cell Fact, 8:1.
- KENSY, F., WENK, P. and WANG, L. (2012). *Debottlenecking Fermentation Experiments: New Robot Automates Media Preparation and Facilitates DoE Implementation*. Genet Eng News, 32:38.
- KENSY, F., ZANG, E., FAULHAMMER, C., TAN, R. and BUECHS, J. (2009b). *Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates*. Microb Cell Fact, 8:1.
- KINAST, K., LITSANOV, B., BOTT, M., WIECHERT, W. and OLDIGES, M. (2012). *Biobasierte Produktion von organischen Saeuren in Corynebacterium glutamicum: Stamm- und Prozesscharakterisierung im mL-Massstab*. Chem Eng Technol, 84:1346.
- KIRKPATRICK, C., MAURER, L. M., OYELAKIN, N. E., YONCHEVA, Y. N., MAURER, R. and SLONCZEWSKI, J. L. (2001). *Acetate and formate stress: opposite responses in the proteome of Escherichia coli*. J Bacteriol, 183:6466.
- KOPETZKI, E., SCHUMACHER, G. and BUCKEL, P. (1989). *Control of formation of active soluble or inactive insoluble baker's yeast alpha-glucosidase PI in Escherichia coli by induction and growth conditions*. Mol Gen Genet, 216:149.
- KOTTMEIER, K., WEBER, J., MUELLER, C., BLEY, T. and BUECHS, J. (2009). *Asymmetric Division of Hansenula polymorpha Reflected by a Drop of Light Scatter Intensity Measured in Batch Microtiter Plate Cultivations at Phosphate Limitation*. Biotechnol Bioeng, 104:554.
- LADIWALA, A., REGE, K., BRENNEMAN, C. M. and CRAMER, S. M. (2005). *A priori prediction of adsorption isotherm parameters and chromatographic behavior in ion-exchange systems*. Proc Natl Acad Sci USA, 102:11710.
- LAMPING, S. R., ZHANG, H., ALLEN, B. and AYAZI SHAMLOU, P. (2003). *Design of a prototype miniature bioreactor for high throughput automated bioprocessing*. Chem Eng Sci, 58:747.

REFERENCES

- LANG, K. M. H., KITTELMANN, J., DUERR, C., OSBERHAUS, A. and HUBBUCH, J. (2015). *A comprehensive molecular dynamics approach to protein retention modeling in ion exchange chromatography*. *J Chrom A*, 1381:184.
- LAZAREVIC, V., GAIA, N., GIRARD, M., FRANCOIS, P. and SCHRENZEL, J. (2013). *Comparison of DNA Extraction Methods in Analysis of Salivary Bacterial Communities*. *PLoS ONE*, 8:1.
- LEE, K., LEE, H., PI, K. and CHOI, Y. (2008). *The effect of low pH on protein expression by the probiotic bacterium Lactobacillus reuteri*. *Proteomics*, 8:1624.
- LESER, E. W. and ASENJO, J. A. (1992). *Rational design of purification processes for recombinant proteins*. *J Chrom A*, 584:43.
- LIENQUEO, M. E. and ASENJO, J. A. (2000). *Use of expert systems for the synthesis of downstream protein processes*. *Comput Chem Eng*, 24:2339.
- LIPIN, D. I., LUA, L. H. L. and MIDDELBERG, A. P. J. (2008). *Quaternary size distribution of soluble aggregates of glutathione-S-transferase-purified viral protein as determined by asymmetrical flow field flow fractionation and dynamic light scattering*. *J Chromatogr A*, 1190:204.
- LOSEN, M., FRÖLICH, B., POHL, M. and BÜCHS, J. (2004). *Effect of oxygen limitation and medium composition on Escherichia coli fermentation in shake-flask cultures*. *Biotechnol Prog*, 20:1062.
- MAISER, B., KRÖNER, F., DISMER, F., BRENNER-WEISS, G. and HUBBUCH, J. (2012). *Isoform separation and binding site determination of mono-PEGylated lysozyme with pH gradient chromatography*. *J Chrom A*, 1268:102.
- MAKRIDES, S. C. (1996). *Strategies for Achieving High-Level Expression of Genes in Escherichia coli*. *Microbiol Mol Biol Rev*, 60:512.
- MARINO, M., HOFFMANN, T., SCHMID, R., MÖBITZ, H. and JAHN, D. (2000). *Changes in protein synthesis during the adaptation of Bacillus subtilis to anaerobic growth conditions*. *Microbiology*, 146:97.
- MICHEL, M., EPPING, A. and JUPKE, A. (2005). *Preparative Chromatography: Of Fine Chemicals and Pharmaceutical Agents*, pages 215–312. Wiley-VCH Verlag GmbH & Co. KGaA.
- MOLLERUP, J. M., HANSEN, T. B., KIDAL, S., SEJERGAARD, L. and STABY, A. (2007). *Development, modelling, optimisation and scale-up of chromatographic purification of a therapeutic protein*. *FFE*, 261:133.
- MUELLER, F., MOUSSA, M., EL GHAZALY, M., ROHDE, J., BARTSCH, N., PARTHIER, A. and KENSY, F. (2013). *Efficient production of recombinant parathyroid hormone fragment (rPTH)1-34 in the methylotrophic yeast Hansenula polymorpha*. *GaBI Journal*, 2:114.

- NFOR, B. K., VERHAERT, P. D. E. M., VAN DER WIELEN, L. A. M., HUBBUCH, J. and OTTENS, M. (2009). *Rational and systematic protein purification process development: the next generation*. Trends Biotechnol, 27:673.
- OELMEIER, S., DISMER, F. and HUBBUCH, J. (2011). *Application of an aqueous two-phase systems high-throughput screening method to evaluate mab hcp separation*. Biotechnol Bioeng, 108:69.
- OHANA, R. F., ENCELL, L. P., ZHAO, K., SIMPSON, D., SLATER, M. R., URH, M. and WOOD, K. V. (2009). *HaloTag7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification*. Protein Express Purif, 68:110.
- OSBERGHAUS, A., HEPBILDIKLER, S., NATH, S., HAINDL, M., VON LIERES, E. and HUBBUCH, J. (2012). *Optimizing a chromatographic three component separation: a comparison of mechanistic and empiric modeling approaches*. J Chrom A, 1237:86.
- OU, J., WANG, L., DING, X., DU, J., ZHANG, Y., CHEN, H. and XU, A. (2003). *Stationary phase protein overproduction is a fundamental capability of Escherichia coli*. Biochem Biophys Res Commun, 314:174.
- PADMANABHA DAS, K. M., BARVE, S., BANERJEE, S., BANDYOPADHYAY, S. and PADMANABHAN, S. (2009). *A Novel Thermostability Conferring Property of Cherry Tag and its Application in Purification of Fusion Proteins*. J Microbial Biochem Technol, 1:59.
- PEDERSEN, S., BLOCH, P. L., REEH, S. and NEIDHARDT, F. C. (1978). *Patterns of protein synthesis in E. coli: a catalog of the amount of 140 individual proteins at different growth rates*. Cell, 14:179.
- PEZZINI, J., JOUCLA, G., GANTIER, R., TOUEILLE, M., LOMENTECH, A., LE SENECHAL, C., GARBAY, B., SANTARELLI, X. and CABANNE, C. (2011). *Antibody capture by mixed-mode chromatography: A comprehensive study from determination of optimal purification conditions to identification of contaminating host cell proteins*. J Chrom A, 1218:8197.
- PIATKOWSKI, W., ANTOS, D. and KACZMARSKI, K. (2002). *Modeling of preparative chromatography processes with slow intraparticle mass transport kinetics*. J Chrom A, 988:219.
- POUDINEH, M., MOHAMADI, R. M., SAGE, A., MAHMOUDIAN, L., SARGENT, E. H. and KELLEY, S. O. (2014). *Three-dimensional, sharp-tipped electrodes concentrate applied fields to enable direct electrical release of intact biomarkers from cells*. Lab Chip, 14:1785.
- PRATER, B. D., ANUMULA, K. R. and HUTCHINS, J. T. (2007). *Automated sample preparation facilitated by PhyNexus MEA purification system for oligosaccharide mapping of glycoproteins*. Anal Biochem, 369:202.

REFERENCES

- PUSKEILER, R., KAUFMANN, K. and WEUSTER-BOTZ, D. (2005). *Development, parallelization, and automation of a gas-inducing milliliter-scale bioreactor for high-throughput bioprocess design (HTBD)*. Biotechnol Bioeng, 89:512.
- RAKEL, N., BAUM, M. and HUBBUCH, J. (2014). *Moving through three-dimensional phase diagrams of monoclonal antibodies*. Biotechnol Prog, 30:1103.
- RAMEEZ, S., MOSTAFA, S. S., MILLER, C. and SHUKLA, A. A. (2011). *High-Throughput Miniaturized Bioreactors for Cell Culture Process Development: Reproducibility, Scalability, and Control*. Biotechnol Prog, 30:497.
- RAMOS, C. R. R., ABREU, P. A. E., NASCIMENTO, L. A. T. O. and HO, P. L. (2004). *A high-copy T7 Escherichia coli expression vector for the production of recombinant proteins with a minimal N-terminal His-tagged fusion peptide*. Braz J Med Biol Res, 37:1103.
- RATCLIFFE, E., GLEN, K. E., WORKMAN, V. L., STACEY, A. J. and THOMAS, R. J. (2012). *A novel automated bioreactor for scalable process optimisation of haematopoietic stem cell culture*. J Biotech, 161:387.
- REGE, K., PEPSIN, M., FALCON, B., STEELE, L. and HENG, M. (2005). *High-Throughput Process Development for Recombinant Protein Purification*. Biotechnol Bioeng, 93:618.
- ROHE, P., VENKANNA, D., KLEINE, B., FREUDL, R. and OLDIGES, M. (2012). *An automated workflow for enhancing microbial bioprocess optimization on a novel micro-bioreactor platform*. Microb Cell Fact, 11:1.
- ROOS, P. H. (1999). *Ion Exchange Chromatography*, pages 3–88. Elsevier Science B. V.
- SCHEIDLE, M., KLINGER, J. and BUECHS, J. (2007). *Combination of On-line pH and Oxygen Transfer Rate Measurement in Shake Flasks by Fiber Optical Technique and Respiration Activity MONitoring System (RAMOS)*. Sensors, 7:3472.
- SCHEIN, C. H. and NOTEBORN, M. H. M. (1988). *Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperature*. Nat Biotechnol, 6:291.
- SHIRANO, Y. and SHIBATA, D. (1990). *Low temperature cultivation of Escherichia coli carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level*. FEBS Lett, 271:128.
- SHRIVASTAVA, S. (2013). *Genetics of Bacteria*, pages 103–105. Springer Verlag.
- SHUKLA, A. A., HUBBARD, B., TRESSEL, T., GUHAN, S. and LOW, D. (2007). *Downstream processing of monoclonal antibodies - application of platform approaches*. J Chrom B, 848:28.
- SHUKLA, A. A. and THOEMMES, J. (2010). *Recent advances in large-scale production of monoclonal antibodies and related proteins*. Trends Biotechnol, 28:253.

- SMOOT, L. M., SMOOT, J. C., GRAHAM, M. R., SOMERVILLE, G. A., STURDEVANT, D. E., MIGLIACCIO, C. A., SYLVA, G. L. and MUSSER, J. M. (2001). *Global differential gene expression in response to growth temperature alteration in group A Streptococcus*. Proc Natl Acad Sci USA, 98:10416.
- SOETAERT, K. E. R. and PETZOLDT, T. (2010). *Inverse modelling, sensitivity and Monte Carlo analysis in R using package FME*. J Stat Softw, 33:1.
- STANCIK, L. M., STANCIK, D. M., SCHMIDT, B., BARNHART, M., YONCHEVA, Y. N., BARNHART, D. M. and SLONCZEWSKI, J. L. (2002). *pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in Escherichia coli*. J Bacteriol, 184:4246.
- STARCK, J., KÄLLENIOUS, G., MARKLUND, B., ANDERSSON, D. I. and AKERLUND, T. (2004). *Comparative proteome analysis of Mycobacterium tuberculosis grown under aerobic and anaerobic conditions*. Microbiology, 150:3821.
- STIEBER, D., GABANT, P. and SZPIRER, C. Y. (2008). *The art of selective killing: plasmid toxin/antitoxin systems and their technological applications*. Biotechniques, 45:344.
- TABOR, S. and RICHARDSON, C. C. (1984). *A bacteriophage t7 rna polymerase/promoter system for controlled exclusive expression of specific genes*. Proc Natl Acad Sci USA, 82:1074.
- THIEMANN, J., JANKOWSKI, J., RYKL, J., KURZAWSKI, S., POHL, T., WITTMANN-LIEBOLD, B. and SCHLUETER, H. (2004). *Principle and applications of the protein-purification-parameter screening system*. J Chrom A, 1043:73.
- TREIER, K., HANSEN, S., RICHTER, C., DIEDERICH, P., HUBBUCH, J. and LESTER, P. (2012). *High-throughput methods for miniaturization and automation of monoclonal antibody purification processes*. Biotechnol Prog, 28:723.
- VALLEJO, L. F. and RINAS, U. (2004). *Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins*. Microb Cell Fact, 3:1.
- VAN LIEROP, M. J. C., WIJNANDS, F., LOKE, Y. W., EMMER, P. M., LUKASSEN, H. G. M., BRAAT, D. D. M., VAN DER MEER, A., MOSSELMAN, S. and JOOSTEN, I. (2002). *Detection of HLA-G by a specific sandwich ELISA using monoclonal antibodies G233 and 56B*. Mol Hum Reprod, 8:776.
- VINCENTELLE, R., CANAAN, S., CAMPANACCI, V., VALENCIA, C., MAURIN, D., FRASSINETTI, F., SCAPPUCINI-CALVO, L., BOURNE, Y., CABBILLAU, C. and BIGNON, C. (2004). *High-throughput automated refolding screening of inclusion bodies*. Protein Sci, 13:2782.
- VOEDISCH, B., MENZEL, C., JORDAN, E., EL-GHEZAL, A., SCHIRRMANN, T., HUST, M. and JOSTOCK, T. (2005). *Heterologe expression von rekombinanten proteinpharmazeutika*. Laborwelt, 6:26.

REFERENCES

- WALSH, G. (2002). *Proteins: Biochemistry and Biotechnology*, pages 123–130. Wiley VCH.
- WALSH, G. (2005). *Biopharmaceuticals: recent approvals and likely directions*. Trends Biotechnol, 23:553.
- WALSH, G. and MURPHY, B. (1999). *Biopharmaceuticals, and Industrial Perspective*, pages 2–5. Kluwer Academic Publishers.
- WANG, J., ZHAO, C., MENG, B., XIE, J., ZHOU, C., CHEN, X., ZHAO, K., SHAO, J., XUE, Y., XU, N., MA, Y. and LIU, S. (2007). *The proteomic alterations of Thermoaerobacter tengcongensis cultured at different temperatures*. Proteomics, 7:1409.
- WASSAF, D., KUANG, G., KOPACZ, K., WO, Q., NGUYEN, Q., TOEWS, M., COSIC, J., JACQUES, J., WILTSHIRE, S., LAMBERT, J., PAZMANY, C. C., HOGAN, S., LADNER, R. C., NIXON, A. E. and SEXTON, D. J. (2006). *High-throughput affinity ranking of antibodies using surface plasmon resonance microarrays*. Anal Biochem, 351:241.
- WEINER III, J., ZIMMERMANN, C. U., GOEHLMANN, H. W. and HERRMANN, R. (2003). *Transcription profiles of the bacterium Mycoplasma pneumoniae grown at different temperatures*. Nucleic Acids Res, 31:6306.
- WEN, E. P., ELLIS, R. and PUJAR, N. S. (2015). *Vaccine Development and Manufacturing*, pages 182–183. Wiley VCH.
- WENK, P., HEMMERICH, J., MUELLER, C. and KENSY, F. (2012). *Hochparallele Bioprozessentwicklung in geschuettelten Mikrobioreaktoren*. Chem Eng Technol, 84:704.
- WIENDAHL, M., SCHULZE WIERLING, P., NIELSEN, J., FOMSGAARD CHRISTENSEN, D., KRARUP, J., STABY, A. and HUBBUCH, J. (2008). *High throughput screening for the design and optimization of chromatographic processes and miniaturization, automation and parallelization of breakthrough and elution studies*. Chem Eng Technol, 21:893.
- WIENDAHL, M., VOELKER, C., HUSEMANN, I., KRARUP, J., STABY, A., SCHOLL, S. and HUBBUCH, J. (2009). *A novel method to evaluate protein solubility using a high throughput screening approach*. Chem Eng Sci, 64:3778.
- WINK, M. (2011). *Molekulare Biotechnologie*, pages 159–186. Wiley VCH.
- WOLLERTON, M. C., WALES, R., BULLOCK, J. A., HUDSON, I. R. and BEGGS, M. (2006). *Automation and Optimization of Protein Expression and Purification on a Novel Robotic Platform*. J Lab Autom, 11:291.
- WOO, P. C. Y., LAU, S. K. P., TSE, H., TENG, J. L. L., CURREEM, S. O. T., TSANG, A. K. L., FAN, R. Y. Y., WONG, G. K. M., HUANG, Y., LOMAN, N. J., SNYDER, L. A. S., CAI, J. J., HUANG, J., MAK, W., PALLAN, M. J., LOK, S. and YUEN, K. (2009). *The complete genome and proteome of Laribacter hongkongensis reveal potential mechanisms for adaptations to different temperatures and habitats*. PLoS genetics, 5:1.

- YE, R. W., TAO, W., BEDZYK, L., YOUNG, T., LI, L. and CHEN, M. (2000). *Global Gene Expression Profiles of Bacillus subtilis Grown under Anaerobic Conditions These include : Global Gene Expression Profiles of Bacillus subtilis Grown under Anaerobic Conditions*. J Bacteriol, 182:4458.
- YU, Z. and MORRISON, M. (2004). *Improved extraction of PCR-quality community DNA from digesta and fecal samples*. BioTechniques, 36:808.
- ZHAO, G., DONG, X. and SUN, Y. (2009). *Ligands for mixed-mode protein chromatography: Principles, characteristics and design*. J Biotech, 144:3.
- ZHOU, Z., SCHNAKE, P., XIAO, L. and LAL, A. A. (2004). *Enhanced expression of a recombinant malaria candidate vaccine in Escherichia coli by codon optimization*. Protein Express Purif, 34:87.

REFERENCES

Akademischer Lebenslauf

Pascal Baumann



Aktuelle Dienstadresse

Pascal Baumann
KIT Campus Süd
Institut für Bio- und Lebensmitteltechnik
Bereich IV: Molekulare Aufarbeitung von Bioprodukten
Engler-Bunte-Ring 3, 76131 Karlsruhe

Telefon: +49 721 608-47530
Fax: +49 721 608-46240
Email: pascal.baumann@kit.edu

Persönliche Daten

Geburtsdatum/-ort: 9. Nov. 1985, Bruchsal, Deutschland
Familienstand: ledig

Akademischer Werdegang

- 01/2012 - jetzt **Doktorand** am Institut für Bio- und Lebensmitteltechnik, Karlsruher Institut für Technologie
Thema der Doktorarbeit: *„Entwicklung einer integrativen Plattform zur Auslegung von Kultivierungs- und Aufarbeitungsprozessen“*
- 2011 **Diplom Bioingenieurwesen** (1,1 - mit Auszeichnung) am Karlsruher Institut für Technologie
Thema der Diplomarbeit: *„Optimierung der Partikel basierten Festphasensynthese hochdichter Peptidarrays durch Laser Fusing“*
- 05/2010 - 11/2010 **Studienarbeit** am Karlsruher Institut für Technologie
Thema der Studienarbeit: *„Fehlerursprungs- und Fehlerfortpflanzungsanalyse mit Monte-Carlo Simulationen am Beispiel robotergestützter Bindungsstudien für Proteinadsorption“*
- 10/2006 - 10/2011 **Studium des Bioingenieurwesens** am Karlsruher Institut für Technologie
Schwerpunkte: *„Biopharmazeutische Aufarbeitung & Industrielle Biotechnologie“*

Akademische Tätigkeiten

- 01/2012 - jetzt **Wissenschaftlicher Mitarbeiter** am Institut für Bio- und Lebensmitteltechnik, Karlsruher Institut für Technologie (u.a. Praktikumsleitung, S1 Projektleiter, Sicherheitsbeauftragter, Betreuung von Abschlussarbeiten, Betreuung von Doktoranden)
- 01/2011 – 05/2011 **Wissenschaftliche Hilfskraft** am Institut für Bio- und Lebensmitteltechnik, Karlsruher Institut für Technologie
- 10/2009 - 03/2010 **Industriepraktikum** bei CSIRO Food and Nutritional Sciences in Werribee (Melbourne), Australien
Thema: *„Pilot-scale Validierung und CFD Simulationen im Bereich Pulsed Electric Fields (PEF) Processing/ Extraction sowie Mikrowellen Sterilisation“*

Curriculum Vitae

Pascal Baumann



Address

Pascal Baumann
KIT Campus South
Institute of Process Engineering in Life Sciences
Section IV: Biomolecular Separation Engineering
Engler-Bunte-Ring 3, 76131 Karlsruhe

Phone: +49 721 608-47530
Fax: +49 721 608-46240
Email: pascal.baumann@kit.edu

Education

- 01/2012 - now **PhD student** at the Institute of Process Engineering in Life Sciences, Karlsruhe Institute of Technology
PhD thesis: *„Development of an integrated platform combining up- and downstream processing features“*
- 2011 **Diploma in Bioengineering** (1.1 - with honors) at the Karlsruhe Institute of Technology
Diploma thesis: *„Optimization of a particle based solid-phase synthesis of highly dense peptide arrays by Laser-Fusing“*
- 05/2010 - 11/2010 **Student research project** at the Karlsruhe Institute of Technology
Thesis: *„Error-Origin- and Error-Propagation-Analysis with Monte-Carlo-Simulations using the example of robot-based binding studies for protein adsorption“*
- 10/2006 - 10/2011 **Studies in Bioengineering** at the Karlsruhe Institute of Technology
Focus: *„Biopharmaceutical Downstream Processing & Industrial Biotechnology“*

Work experience

- 01/2012 - now **Research associate** at the Institute of Process Engineering in Life Sciences, Karlsruhe Institute of Technology (i.a. Supervision of practical courses, genetical safety level 1 (S1) project leader, safety officer, Supervisor for bachelor, master and PhD students)
- 01/2011 – 05/2011 **Scientific member** at the Institute of Process Engineering in Life Sciences, Karlsruhe Institute of Technology
- 10/2009 - 03/2010 **Practical industrial term** at CSIRO Food and Nutritional Sciences in Werribee (Melbourne), Australia
Topic: *„Pilot-scale validations and CFD simulations in the fields of pulsed electric fields (PEF) processing/ extraction and microwave disinfestations“*

Awards

“*Malcolm Lilly Award*” from the European Society of Biochemical Engineering Sciences (ESBES) at the ESBES-10 conference 2014 for an outstanding contribution of a promising young scientist or engineer working in the field of biochemical engineering

“*Dechema Student Award*” for excellent achievements and exceptional short and efficient studies in the field of biotechnology in 2013

“*Emil Kirschbaum Award*” from the Karlsruhe Institute of Technology (KIT) for excellent achievements in the academic year 2011/2012

Scientific Publications

Peer-reviewed articles

P. Baumann, T. Huuk, T. Hahn, A. Osberghaus, J. Hubbuch
Deconvolution of High-throughput Multi-component Isotherms Using Multi-variate Data Analysis of Protein Spectra
Eng. Life Sci., accepted manuscript

C. Ladd Effio, P. Baumann, P. Vormittag, C. Weigel, A. Middelberg, J. Hubbuch
Establishment of a High-throughput Platform for the Production of Virus-like Particles in E. coli
Biotechnol. Bioeng., submitted manuscript

P. Baumann, T. Hahn, J. Hubbuch
High-throughput Micro-scale Cultivations and Chromatography Modeling: Powerful Tools for Integrated Process Development
Biotechnol. Bioeng. (2015), <http://dx.doi.org/10.1002/bit.25630>

T. Hahn, P. Baumann, T. Huuk, V. Heuveline, J. Hubbuch
UV absorption-based inverse modeling of protein chromatography
Eng. Life Sci. (2015), <http://dx.doi.org/10.1002/elsc.201400247>

P. Baumann, K. Baumgartner, J. Hubbuch
Influence of Binding pH and Protein Solubility on the Dynamic Binding Capacity in Hydrophobic Interaction Chromatography
J. Chrom. A (2015), <http://dx.doi.org/10.1016/j.chroma.2015.04.001>

P. Baumann, N. Bluthardt, S. Renner, H. Burghardt, A. Osberghaus, J. Hubbuch
Integrated development of up- and downstream processes supported by the Cherry-Tag for real-time tracking of stability and solubility of proteins
J. Biotechnol. (2015), <http://dx.doi.org/10.1016/j.jbiotec.2015.02.024>

P. Baumann, A. Osberghaus, J. Hubbuch
Systematic purification of salt-intolerant proteins by ion-exchange chromatography: The example of Human α -Galactosidase A
Eng. Life Sci. (2015), <http://dx.doi.org/10.1002/elsc.201400210>

A. Osberghaus, P. Baumann, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch
Detection, quantification, and propagation of uncertainty in high-throughput experimentation by monte carlo methods
Chem. Eng. Technol. (2012), <http://dx.doi.org/10.1002/ceat.201100610>

K. Knoerzer, P. Baumann, R. Buckow
An iterative modelling approach for improving the performance of a PEF treatment chamber
Comput. Chem. Eng. (2012), <http://dx.doi.org/10.1016/j.compchemeng.2011.09.002>

R. Buckow, P. Baumann, S. Schroeder, K. Knoerzer
Effect of dimensions and geometry of co-field and co-linear pulsed electric field treatment chambers on electric field strength and energy utilisation
J. Food Eng. (2011), <http://dx.doi.org/10.1016/j.jfoodeng.2011.03.019>

R. Buckow, S. Schroeder, P. Berres, P. Baumann, K. Knoerzer
Simulation and Evaluation of Pilot-Scale Pulsed Electric Field (PEF) Processing
J. Food Eng. (2010), <http://dx.doi.org/10.1016/j.jfoodeng.2010.06.010>

Talks

P. Baumann, T. Hahn, J. Hubbuch
Integrated bioprocess development based on micro-scale cultivations and feed stock characterization by chromatography modeling
249th ACS National Meeting Exposition, Denver, USA, Mar. 2015

P. Baumann, K. Baumgartner, J. Hubbuch
Influence of pH and protein solubility on the dynamic binding capacity in hydrophobic interaction chromatography
11th Seminar on Chromatographic Separation Science, Sundern, Germany, Feb. 2015

P. Baumann, T. Hahn, N. Bluthardt, A. Osberghaus, J. Hubbuch
Combined optimization of up- and downstream processes based on high-throughput cultivations and chromatography modelling
34th ISPPP, Würzburg, Germany, Nov. 2014

P. Baumann, T. Hahn, N. Bluthardt, A. Osberghaus, J. Hubbuch
Glutathione-S-Transferase from Escherichia coli: high-throughput micro-scale cultivations and chromatography modeling as powerful tools for integrated process development
3rd Internat. Conference on High-Throughput Process Development (HTPD), Siena, Italy, Oct. 2014

P. Baumann, T. Hahn, N. Bluthardt, A. Osberghaus, J. Hubbuch
High-throughput micro-scale cultivations and chromatography modeling: powerful tools for integrated process development (Malcolm Lilly Award Presentation)
10th ESBES and 6th IFIBiop, Lille, France, Sep. 2014

P. Baumann, A. Osberghaus, J. Hubbuch
A systematic HTS strategy in ion exchange chromatography for protein purification by salt gradient or pH shift
10th Seminar on Chromatographic Separation Science, Egmond aan Zee, Netherlands, Feb. 2014

P. Baumann
Optimization of a particle based solid-phase synthesis of highly dense peptide arrays by laser-fusing
Tag der Fakultät - KIT Graduation Ceremony, Karlsruhe, Germany, Jun. 2013

P. Baumann, F. Kroener, A. Osberghaus, S. Hansen, J. Hubbuch
Elution pH screening in ion exchange chromatography: a global tool for downstream development by pH shift
9th Seminar on Chromatographic Separation Science, Weggis, Switzerland, Feb. 2013

Posters

P. Baumann, N. Bluthardt, A. Osberghaus, J. Hubbuch
Optimization of up- and downstream processes using Cherry-tagged proteins – product tracking by simple absorbance measurements
10th ESBES and 6th IFIBiop, Lille, France, Sep. 2014.

P. Baumann, N. Bluthardt, T. Hahn, A. Osberghaus, J. Hubbuch
Integral optimization of up- and downstream processes by high throughput cultivations and in silico chromatography predictions
10th ESBES and 6th IFIBiop, Lille, France, Sep. 2014.

P. Baumann, F. Kroener, A. Osberghaus, S. Hansen, J. Hubbuch
Development of a universal downstream strategy: pH shift in ion exchange chromatography
2nd European Congress of Applied Biotechnology (ECAB), The Hague, Netherlands, Apr. 2013.