High-Throughput Process Development in the Field of Protein Purification

–

Method Development, Application, and Characterization

zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften (Dr.-Ing.)

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

genehmigte

DISSERTATION

von
Dipl.-Ing. Patrick Diederich
aus Kelberg, Rheinland-Pfalz

Referent: Prof. Dr. Jürgen Hubbuch
Korreferent: Prof. Dr. Matthias Franzreb
Tag der mündlichen Prüfung: 15.06.2015

Danksagung

Folgenden Personen möchte ich ausdrücklich meinen Dank aussprechen:

- Prof. Dr. Jürgen Hubbuch für die gegebene Möglichkeit an seinem Lehrstuhl meine Doktorarbeit anfertigen zu können, für die ausgesprochen guten Arbeitsbedingungen, für die von ihm geförderte offene Institutskultur, für die fachliche Unterstützung sowie das mir entgegengebrachte Vetrauen.
- Prof. Dr. Matthias Franzreb für die freundliche Übernahme des Korreferats und dafür, dass er während meines Studiums mein Interesse an Proteinaufarbeitung geweckt hat.
- Marion Krenz, Margret Meixner, Susanne Haid und Erika Sonnenburg, die mir nicht nur in bürokratischen Dingen eine riesige Hilfe waren, sondern auch für ein funktionierendes Institut gesorgt haben.
- Robert Fehling, Katrin Lehmann-Barlen und Jens Gebken, die auf Seiten des Projektpartners interessiert und unterstützend zur Verfügung standen, und mir jedes Jahr die Möglichkeit zum fachlichen Austausch auf Doktorandenseminaren gegeben haben.
- Sven Amrhein, Katharina Drechsel, Frank Hämmerling, Simon Luxem, Anja Paulus und Marie-Luise Schwab ohne deren großen Einsatz und wertvolle Mitarbeit als Studien-/Diplomarbeiter oder Hiwi die vielfältigen Aufgaben nicht zu stemmen gewesen wären.
- Mark Hoffmann und Kristina Schleining für das Ausführen so mancher Laborversuche, die freundschaftliche Atmosphäre im gemeinsamen Büro und das Erwidern ironischer Bemerkungen.
- Anna Osberghaus und Stefan Oelmeier für die große fachliche Unterstützung, das offene Ohr bei Problemen im Labor sowie die schönen privaten Unternehmungen.
- Jörg Kittelmann für seine oftmals helfende Automatisierungsexpertise, seine humorvolle Art und für das vollste Verständnis dafür, dass Simulationsdaten für Konferenzbeiträge auch noch nachts im Hotelzimmer generiert werden können.
- Kollegen der ersten Stunde, Benjamin Maiser, Florian Dismer, Frieder Kröner, Katharina Lang und Natalie Rakel, mit denen ich viele schöne Erinnerungen ans Labor, Seminarfahrten, Kaffepausen und auch an Grillabende verbinde.
- Katrin Treier-Marxen, die gelegentliches (Ver-)Zweifeln an der Proteinanalytik in unserem gemeinsamen Büro mit Gelassenheit ertragen hat. Der Umzug des Analyselabors unter Koordination von Patrick Pionneau bleibt unvergessen.
- dem IT-Team und insbesondere Jan C. Müller, die zu jeder Tag- und Nachtzeit bei soft- und hardware Problemen erreichbar waren und außerordentlich hilfsbereit waren.

- allen weiteren Kollegen am “Mab”-Lehrstuhl, die namentlich nicht alle aufgefhrt sind, aber nichts desto minder durch die Schaffung einer offenen und gemeinschaftlichen Arbeitsatmosphäre einen großen Anteil daran haben, dass ich mich in der Gruppe so wohl gefühlt habe.
- meinen Freunden, die mir teilweise aus großer Entfernung stets treu geblieben sind, und mir immer wieder einmal verraten haben, dass es wichtigere Dinge im Leben als eine Doktorarbeit gibt.
- im besonderen Maße meiner Schwester und von ganzem Herzen meinen Eltern, die stets an mich geglaubt haben und mich während meines gesamten Lebens auf jegliche Art und Weise unterstützt haben.
- und schließlich von ganzem Herzen Sigrid Hansen, die mich unermüdlich und mit unendlicher Geduld begleitet hat. Du und unser Sohn seid mein größtes Glück!

Abstract

This thesis makes contributions to a modern approach of process development for the purification of biopharmaceutical ingredients termed 'High-Throughput Process Development' (HTPD). HTPD is based on the principles of miniaturization, parallelization and automation of methods and processes for protein expression and protein purification as well as protein analytics. The methods are mainly applied for the fast investigation of protein properties, the evaluation of process parameters, and for narrowing-down of a parameter space to an optimal region for final process optimization. In the field of protein purification, experimental methods have been developed in the last decade for different types of low-pressure liquid chromatography, protein precipitation, extraction, crystallization, enzymatic or chemical modifications and for the investigation of refolding conditions, biophysical properties (e.g. hydrophobicities, protein-protein-interactions), solubilities, and stabilities.

With regard to the development and implementation of HTPD in industries over the last decade, open questions and remaining areas for optimization were identified and discussed in the introduction of this work. This work deals with aspects of these questions and can be divided in three parts:

1. development/optimization of analytical methods providing high-throughput as well as fulfilling high demands concerning accuracy, precision and robustness

2. application of HTPD for the purification of the hen's egg-white protein avidin

3. characterization and quantification of method-specific effects and their influence on data quality for one of the most-common high-throughput column chromatography method.

The first part of this work presents the development of analytical methods providing both a high-throughput of samples and data quality comparable to traditional methods. High-Performance Liquid Chromatography (HPLC) is well-known for highest accuracy and precision and is the prevalent technique used for protein analysis. The main disadvantage, however, is the typically long analysis time (e.g. 20 – 60 min) making HPLC technology ineligible for the use in HTPD, in which a three-digit number of samples can be generated per day. In order to nevertheless be able to use HPLC methods in HTPD, a search for a more efficient utilization of current HPLC technology (resins, equipment, setup) was conducted.

In a first application, the analysis time in size-exclusion chromatography (SEC) was optimized for the separation of antibody aggregates and monomer. The lag phases before and after the elution of the species of interest are one feature of SEC-analysis which can sum up to a significant proportion of the total analysis time. First, different SEC-columns were evaluated regarding the best ratio of resolution to analysis time. Secondly, by two interlaced operating programs for a time-independent control of autosampler and detector, the lag phase between sample injection and the first eluting aggregates could be eliminated. Lastly, in order to save the lag phase between elution of antibody monomer and the latest eluting buffer species, two identical columns were operated in parallel but time-shifted by the

elution time of aggregates and monomer. In order to direct the injected samples to the columns and the protein elution phase independently from injection to the detector, two multi-port switching valves were built in at column in- and outlet. Despite of interlacing and time-shift, the correct assignment of injected sample to the corresponding chromatogram was enabled by the developed control programs. In comparison to a traditional mode of operation, the sample throughput was almost tripled by using two columns in parallel and interlaced injection. This is remarkable when considering that, except for two switching valves, no further major investment (e.g. detector) was necessary. Furthermore, identical data quality regarding accuracy, precision and robustness as obtained in traditional mode was achieved.

The developed mode of operation (interlaced sample injection in parallel) is only applicable for isocratic elutions, i.e. when using the same solvent composition throughout the assay. Therefore, it applies mainly to SEC-analysis. The majority of methods for protein quantification, however, is based on reversed phase (RP) HPLC which provides the highest resolutions. For the considered protein system consisting of avidin (AV), ovalbumin (OV), ovomucoid (OM), ovotransferrin (OT), and lysozyme (LYS) a high-throughput RP HPLC method was therefore developed featuring a three-segment gradient elution. Based on the development of the method for a single-column application, a tandem-application was established by using two identical columns, two multi-port switching valves, and by splitting the assay in two programs operating in alternation for injection/elution and regeneration/re-equilibration. Two versions of the method were set up allowing for operation either with inline split-loop or pulled-loop autosampler. In comparison to published application notes of the tandem-application using an inline split-loop autosampler, the analysis time could be reduced by additional $\sim$1 min by scheduling the valve switching differently. Furthermore, it was shown that for fast gradient applications of a few minutes and low flow rates, a high-pressure gradient pump is to be favored over a low-pressure gradient pump due to a lower gradient residence time. With the presented tandem-application for this challenging separation, very short analysis time of close to four minutes, high precision (relative standard deviation of $<$1% within the relevant concentration range), and high robustness with respect to changes in the sample matrix (e.g. buffers, concentration) were achieved.

The main disadvantage of a fast HPLC method is the time-costly method development. Chip-based capillary electrophoresis (LabChip®) as quantitative high-throughput method was therefore evaluated as an alternative to HPLC analysis. Advantages and disadvantages of this technology were discussed based on the AV-protein system and in comparison to the HPLC technology. A major advantage is that no molecule-specific method development is required. Furthermore, advantages are the high degree of automation and the small amount of sample volumes required. Major disadvantages are the limited use for proteins of similar size, potentially poor staining (in the considered system OM was not detected), high sensitivity against changes in the sample matrix, and low robustness of automated peak integration. The latter required manual correction of the integration which significantly reduced the overall sample throughput.

Besides methods for the determination of concentration and purity of AV-samples by RP HPLC, an AV activity assay (HABA-assay) was implemented as high-throughput method. The assay, which is based on photometric detection and is typically performed manually in cuvettes, was scaled-down to a 96-well micro-titer plate format and performed on a liquid handling station. The determined operation range and the obtained precision were in accordance with results from reference measurements in cuvettes. The sample volume and analysis time were reduced to minimal 20 μL and approx. 45 min for 96 samples, respectively.

The developed high-throughput analytical methods for the AV-system posed the pre-requisite for the development of an alternative purification process by applying HTPD which is presented in the second and main part of this work. Starting from pre-purified egg white, the goal was to remove the remaining impurities OV, OM, OT, and LYS by fewest possible, cost-effective and easily-combined process steps. First, extraction in aqueous two-phase systems (ATPS) was evaluated as an alternative to solid-liquid chromatography. One major goal was to separate AV from LYS effectively, which was, due to their similar isoelectric points, described as challenging in literature. The characterization of an ATPS by binodal curve and tie-lines and the screening of various system compositions for the separation of AV from its contaminants was performed in 96-well micro-titer plate format automated on a liquid handling station with subsequent analysis of protein concentrations in upper and bottom phase using the established RP HPLC method.

By automation and with sample consumption of only 65 μL per ATPS a large parameter space consisting of polyethylene glycol (PEG), salt concentrations, PEG molecular weight, pH, and sodium chloride concentration was tested and the effects on the distribution behavior of AV and contaminants investigated. For a PEG molecular weight of 600 kDa and the addition of 3% w/w sodium chloride, an effective separation was achieved in a PEG/phosphate ATPS resulting in a reduction of LYS by 99.7% and OM by 99.4%. For this ATPS composition, AV was enriched in the bottom phase with a total purification factor of approx. 6. The results obtained in high-throughput experiments were confirmed in an ATPS in laboratory scale.

Despite the efficient separation achieved by an optimized ATPS, further purification was required in order to increase the purity of AV to >99%. Furthermore, with the single-step extraction, an overall volume reduction was not achieved which is disadvantageous for a process in large-scale. Lastly, a competitive process for the recovery of AV from the bottom phase and transfer to an aqueous solution for protein formulation was sought, as re-extraction into another two-phase system or dia- and ultrafiltration were not considered as optimal solutions due to the low concentration of AV. An investigation of possible solutions to the mentioned challenges by applying HTPD was therefore included in the second part of this work.

Differences in solubility were detected in the ATPS-screening experiments, and thus, the use of selective protein precipitation for volume reduction and further purification of AV seemed promising. By performing high-throughput precipitation experiments, solubility curves for AV and all contaminants were generated. From these data, optimal conditions for the precipitation and further purification of AV at elevated PEG-concentrations could be determined. In order to provide a seamless integration of AV-precipitation in PEG-solutions with a subsequent extraction via ATPS, the precipitate was completely re-dissolved in the homogeneously mixed ATPS prior to the phase separation. By this integrated process, a purity of AV of 78% and a total yield of 91% was achieved. Finally, for the separation of the remaining impurity in the bottom phase (OV) from AV, chromatographic resins were tested in high-throughput column chromatography experiments. In order to enable direct processing of the salt-rich bottom phase via chromatography, resins for hydrophobic interaction and mixed-mode (MM) chromatography were evaluated. A separation could be achieved using a MM-resin and a linear reduction of the salt concentration during elution. The concentration of the impurities in the AV elution pool were below the lower limit of quantification (LLOQ). For the reached concentration of AV and the LLOQ of all impurities, the corresponding achieved purity of AV was 97.5%. During development, data evaluation methods were developed, in order to fully take advantage of the time savings gained by the experimental high-throughput methods. In conclusion, HTPD was successfully employed and an effective alternative process for the purification of AV was identified and optimized

in small-scale.

In the last part of this work, the method for performing high-throughput column chromatography (HTCC) was investigated in detail, and limitations of this method were presented. Since commercialization in 2007, the miniaturized columns investigated have been used in the field of protein purification for resin screenings and for optimizations of load and elution conditions. However, the application of HTCC for process development still goes along with questions, e.g. at which conditions and to what extent the gained data are comparable to data obtained in traditional lab-scale or production-scale. In order to eventually answer these questions, besides an investigation of scale-effects (e.g. via fluid dynamic simulations), a detailed characterization of method-specific experimental parameters is required. In the presented study, the single and combined influences of pipetting accuracy, absorption measurement in micro-titer plates, relative fraction size, flow interruptions, and salt step concentrations in gradient elutions on qualitative and quantitative results were investigated experimentally and by using simulations. It was shown that when applying suitable peak fitting, the influence of the salt steps, by which the gradient is approximated, can almost be neglected. The cause for apparent higher noise when performing salt gradient elution experiments was explained by using simulations for the chromatographic separation of AV and LYS. The influence of the chosen flow rate was demonstrated in conjunction with the flow interruptions, which are inherent in HTCC. The shown effects of various method parameters are important for researchers aiming the determination of model parameters for the calibration of chromatographic models via HTCC. The gained knowledge of method-specific effects complement future comparability studies in which scale-effects are investigated.

Zusammenfassung (deutsche Fassung)

Die vorliegende Arbeit ist in das Themengebiet Verfahrens- und Prozessentwicklung zur Aufreinigung von biopharmazeutischen Wirkstoffen einzuordnen. Im Speziellen geht es um einen neuartigen Ansatz zur Prozessentwicklung, der unter dem Sammelbegriff 'Hochdurchsatz-Prozessentwicklung' (HTPD) zusammengefasst wird. Die Grundsätze von HTPD basieren auf Miniaturisierung, Parallelisierung und Automation von Verfahren zur Proteinexpression und Proteinaufarbeitung sowie zur Proteinanalytik. Vorwiegend dienen die Methoden zur Untersuchung von Moleküleigenschaften, der Evaluierung von Prozessparametern sowie der Eingrenzung möglsicher Parameterkombinationen und Identifizierung eines optimalen Bereichs für die finale Prozessauslegung. Im Bereich der Proteinaufarbeitung sind Anwendungsbeispiele für Niederdruck-Flüssigchromatographie verschiedenster Trennprinzipien, Proteinpräzipitation, Extraktion, Kristallisation, enzymatische oder chemische Modifikation sowie zur Untersuchung von Löslichkeiten, Renaturierungsbedingungen, biophysikalischer Proteineigenschaften (z.B. Hydrophobizitäten, Protein-Protein-Wechselwirkungen) und Stabilitäten bekannt.
In Analyse der Entwicklung von Hochdurchsatzmethoden und Umsetzung des HTPD-Konzepts in der Industrie der letzten zehn Jahre wurden in der Einleitung zu dieser Arbeit derzeit noch offene Fragestellungen und wesentliche Optimierungsbereiche identifiziert und erläutert. In der vorliegenden Arbeit wurden Aspekte der dargelegten Fragestellungen in folgenden drei Teilbereichen bearbeitet:

1. Entwicklung/Optimierung von Analytikmethoden, die sowohl hochdurchsatzfähig sind als auch hohen Anforderungen an Richtigkeit, Präzision und Robustheit erfüllen,

2. Anwendung von HTPD zur Aufreinigung des Hühnereiweißes Avidin,

3. Charakterisierung und Quantifizierung methodenspezifischer Phänomene und deren Einfluss auf die Datenqualität anhand der aktuell gängigsten Methode zur Hochdurchsatz-Säulenchromatographie.

Im ersten Teil der Arbeit wurden Analytikmethoden entwickelt, mit denen sich der Probendurchsatz erhöhen lässt und gleichzeitig eine zu traditionellen Methoden vergleichbare Datenqualität erzielen lässt. Zur Proteinanalytik, die höchste Richtigkeit und Präzision leisten kann, wird vorwiegend Hochleistungs-Flüssigchromatographie (HPLC) eingesetzt. Meist erfordern HPLC Methoden jedoch lange Analysezeiten (z.B. 20 – 60 min). Sie sind somit als Analytikmethode für HTPD kaum geeignet, da hier eine dreistellige Anzahl an Proben pro Tag generiert werden kann. Um die Vorteile von HPLC Methoden dennoch für HTPD nutzen zu können, wurde versucht, derzeit verfügbare Technik (Trennmedien, Gerät und Konfiguration) effizienter auszunutzen.

Als erste Anwendung wurde die Analysezeit in der Größenausschluss-Chromatographie (SEC) am Beispiel der Trennung von Antikörper-Aggregaten und -monomer optimiert. Charakteristisch bei der SEC-Analyse sind die Verzögerungszeiten vor und nach Elution der zu bestimmenden Proteinspezies, welche einen erheblichen Anteil der gesamten Analysezeit ausmachen können. Zunächst

wurden verschiedene Trennsäulen hinsichtlich des besten Verhältnisses von Auflösung zu Analysezeit evaluiert. Das Einsparen der inhärenten Verzögerungszeit zwischen Probeninjektion und der zuerst eluierenden Aggregate konnte durch vorzeitige Probeninjektion in zwei ineinander verschachtelten Programmsequenzen zur zeitlich getrennten Steuerung von Autosampler und Detektor erzielt werden. Um darüber hinaus die Verzögerungszeit zwischen Elution von Antikörper-Monomer und den zuletzt eluierenden Pufferkomponenten einzusparen, wurden zwei identische Trennsäulen gleichzeitig, jedoch mit einer um die Elutionszeit von Aggregaten und Monomer verschobenen Probeninjektion betrieben. Zwei Mehrkanal-Schaltventile wurden hierzu an Säulenein- und Säulenausgänge geschaltet, um die injizierten Proben alternierend auf die Trennsäulen sowie die Elutionsphase alternierend zum Detektor zu leiten. Die Steuerungsprogramme konnten so gestaltet werden, dass trotz der Zeitverschiebungen jeder Probeninjektion das entsprechende Einzelchromatogramm für die Auswertung zugeordnet werden konnte. Der Probendurchsatz konnte durch diese Betriebsweise im Vergleich zum traditionellen Betriebsmodus nahezu verdreifacht werden, wobei bis auf zwei Säulenschaltventile keine weiteren Investitionen für die Parallelisierung (z.B. zweiter Detektor) erforderlich wurde. Gleichzeitig konnte die gleiche Datenqualität hinsichtlich Richtigkeit, Präzision und Robustheit wie im traditionellen Betriebsmodus erzielt werden.

Das Prinzip der verschachtelten Probeninjektionen kombiniert mit parallelem Betrieb von zwei Säulen ist nur bei isokratischer Elution, d.h. unter Verwendung desselben Laufmittels möglich und ist daher vorwiegend für SEC-Anwendungen geeignet. Die Vielzahl der Protein-Quantifizierungsmethoden basieren jedoch auf der reversed phase (RP) HPLC, mittels welcher meist die höchsten Auflösungen erzielt werden können. Für die in dieser Arbeit zu analysierenden Proteine Avidin (AV), Ovalbumin (OV), Ovomucoid (OM), Ovotransferrin (OT) und Lysozym (LYS) wurde daher eine Hochdurchsatz-Analytik mittels RP HPLC entwickelt. Eine Trennung konnte durch einen in drei Segmente unterteilten Gradienten erzielt werden. Basierend auf der entwickelten Methode für den Einsäulenbetrieb wurde eine Tandem-Anwendung unter der Verwendung von zwei Säulen, zwei Mehrkanal-Schaltventilen und der Aufteilung der Methode in zwei alternierend ablaufende Programme für die Probeninjektion/Elution und Elution/Regeneration entwickelt. Zwei Versionen der Methode wurden etabliert, um die Verwendung mit sowohl einem 'inline-split loop' als auch einem 'pulled-loop' Autosampler zu ermöglichen. Im Unterschied zu veröffentlichten Anwendungsbeispielen des Tandem-Betriebs für die 'inline split-loop' Konfiguration konnte durch Änderung der Ventilschaltzeiten die Analysezeit um ca. eine Minute verkürzt werden. Des Weiteren zeigte sich, dass für schnelle Gradientenmethoden im Bereich weniger Minuten und bei geringen Flussraten eine Hochdruckgradientenpumpe aufgrund der geringeren Gradientenverweilzeit gegenüber einer Niederdruckgradientenpumpe zu bevorzugen ist. Mit der Tandem-Methode konnte für das gegebene anspruchsvolle Trennproblem eine geringe Analysezeit von knapp über vier Minuten, eine niedrige Streuung im erwarteten Messbereich von <1 % relativer Standardabweichung und gleichzeitig hohe Robustheit gegenüber Änderungen in der Probenmatrix (z.B. Puffer, Konzentration) erzielt werden.

Als größter Nachteil einer schnellen HPLC Methode ist die aufwendige Methodenentwicklung zu nennen. Als Alternative wurde daher die Verwendung von chip-basierter Kapillargelelektrophorese (LabChip®) als quantitative Hochdurchsatzanalytik evaluiert. Vor- und Nachteile der Technologie wurden im Vergleich zur HPLC Technologie anhand der generierten Daten mit dem Avidin-Proteinsystem aufgezeigt. Ein Hauptvorteil ist, dass keine molekül-spezifische Methodenentwicklung nötig ist. Als weitere wesentlichen Vorteile sind der hohe Automatisierungsgrad sowie das geringe benötigte Probenmaterial zu nennen. Erhebliche Nachteile sind die limitierte Anwendbarkeit bei größenähnlichen oder schlecht anfärbbaren Proteinen (im untersuchten Modellsystem wurde OM nicht

detektiert), die Sensitivität gegenüber Änderungen in der Probenmatrix sowie die relativ geringe Robustheit bei der automatisierten Peakintegration. Durch nötige manuelle Korrekturen bei der Auswertung erniedrigte sich insgesamt betrachtet der Probendurchsatz signifikant.

Im Zuge der Methodenentwicklung für das Avidin-System wurde neben der Betrachtung der Reinheit und Konzentration mittels RP HPLC zudem eine Methode zur Aktivitätsbestimmung (HABA-Assay) automatisiert. Der auf photometrischer Detektion basierende und typischerweise in Küvetten ausgeführte Aktivitätsessay wurde im 96-well Mikrotiterplatten-Format für die Anwendung auf einer Pipettierstation skaliert. Die Bestimmung des Messbereichs und Präzision zeigte zu Referenzmessungen in manuellen Küvettenversuchen vergleichbare Ergebnisse. Mit dieser Methode reduzierte sich das benötigte Probenmaterial auf ein Minimum von 20 μL sowie die Analysezeit auf ca. 45 Minuten für 96 Proben.

Die Etablierung einer geeigneten Analytik für das Avidin-System war die Voraussetzung für die Entwicklung eines alternativen Aufreinigungsprozesses mittels HTPD, deren Ergebnisse im Haupt- und zweiten Teil dieser Arbeit präsentiert werden. Ausgehend von vorgereinigtem Eiweiß sollten die verbliebenen Verunreinigungen OV, OM, OT und LYS in möglichst wenigen, kostengünstigen und leicht kombinierbaren Prozessschritten von AV getrennt werden. Als Alternative zur Fest-Flüssig-Chromatographie wurde zunächst die Extraktion in wässrigen Zweiphasensystemen (ATPS) evaluiert. Ein Hauptziel war es, die aufgrund des ähnlichen isoelektrischen Punkts mittels Ionenaustausch-Chromatographie als schwierig beschriebene Trennung von AV und LYS mittels ATPS effektiver erzielen zu können. Die Charakterisierung des Zweiphasengebiets über Binodale und Konnoden sowie das Ansetzen und Testen verschiedenster ATPS zur Trennung von AV und dessen Verunreinigungen wurde in 96-well Mikrotiterplatten automatisiert auf einer Pipettierstation ausgeführt und die Proteinkonzentrationen in Ober- und Unterphase mittels der etablierten Tandem RP HPLC Analytik bestimmt.

Durch die Automation und den geringen Verbrauch an Probenmaterial von nur 50 μL Proteingemisch pro ATPS konnte ein großer Parameterraum bestehend aus Polyethylenglykol (PEG)- und Salzkonzentrationen, PEG-Molekulargewicht, pH-Wert und Natriumchlorid-Konzentrationen getestet und deren Einflüsse auf das Verteilungsverhalten von AV und Verunreinigungen ermittelt werden. Mit Reduzierung des molekularen Gewichts von PEG auf 600 kDa sowie der Addition von Natriumchlorid zu 3 % m/m konnte in einem PEG/Phosphat-System eine für einen einstufigen Extraktionsschritt äußerst effektive Reduktion von LYS um 99.7 % und von OM um 99.4 % erzielt werden. Für dieses ATPS wurde eine Aufreinigung von AV etwa um den Faktor 6 in der Unterphase erzielt. In ATPS-Versuchen im Labormaßstab konnten die Ergebnisse der Hochdurchsatzversuche bestätigt werden.

Obwohl mittels ATPS eine sehr gute Trenneffizienz erzielt werden konnte, bedurfte es einer weiteren Aufreinigung, um eine angestrebte Reinheit von Avidin mit > 99 % zu erzielen. Zudem konnte in dem einstufigen ATPS insgesamt betrachtet keine Volumenreduktion erzielt werden, was für eine großtechnische Realisierung des Prozesses nachteilig ist. Des Weiteren musste ein kostengünstiger Weg für die Gewinnung von AV aus der Unterphase und den Transfer in eine für die Formulierung geeignete wässrige Phase gefunden werden, wobei eine Rückextraktion oder Dia- und Ultrafiltration wegen der geringen AV-Konzentration nicht als optimale Lösungen betrachtet wurden. Daher wurde im zweiten Teil der Arbeit auch nach Lösungen für die genannten Fragestellungen unter Anwendung von HTPD gesucht.

Bereits in den ATPS-Versuchen wurden Unterschiede in den Protein-Löslichkeiten festgestellt, so dass der Ansatz einer Aufkonzentrierung von AV sowie eine weitere Abreicherung der Verunreinigun-

gen über selektive Präzipitation nahe lag. Mittels Präzipitationsversuchen im Hochdurchsatz konnten über die Aufnahme von Löslichkeitskurven optimale Bedingungen zur Präzipitation von AV bei erhöhten PEG-Konzentrationen ermittelt werden und eine weitere Aufreinigung erzielt werden. Um eine nahtlose Integration von AV-Präzipitation in einer PEG-Lösung mit anschließendem ATPS zu erzielen, wurde das Präzipitat direkt in der Zusammensetzung des sich anschließenden ATPSs rückgelöst und anschließend die Phasenseparation durchgeführt. Durch den integrierten Prozess konnte eine AV-Reinheit von 78 % bei einer Gesamtausbeute von 91 % erzielt werden. Abschließend wurden in Hochdurchsatz-Chromatographieversuchen auf einer Pipettierstation Chromatographiemedien für die Trennung der in der Unterphase verbliebenen Verunreinigung OV von AV getestet. Um ohne Konditionierung der salzreichen Unterphase eine chromatographische Trennung durchführen zu können, wurden Harze zur hydrophoben Interaktionschromatographie und Mixed-Mode (MM) Chromatographie untersucht. Letztlich konnte mit einem geeigneten MM-Harz durch lineare Erniedrigung des Salzgehalts bei der Elution eine Trennung erzielt werden. Verunreinigungen ließen sich im AV-Elutionspool mit der gewählten HPLC Analytik nicht mehr nachweisen, so dass unter Berücksichtigung der Konzentrationen und Nachweisgrenze der Analytik die erzielte AV-Reinheit mindestens 97.5 % betrug. Während der Arbeit wurden Methoden zur automatisierten Datenauswertung entwickelt, um den Zeitgewinn durch schnelle Labormethoden ohne Leerlaufzeiten voll ausnutzen zu können. Durch erfolgreiche Anwendung des HTPD-Ansatzes konnte somit eine effektive Prozessalternative für die Aufreinigung von AV identifiziert und im miniaturisierten Maßstab optimiert werden.

Im abschließenden Teil dieser Arbeit wurden die Methode der Hochdurchsatz-Säulenchromatographie genauer untersucht und Limitierungen der Methode aufgezeigt. Die verwendeten, miniaturisierten Trennsäulen werden seit der Kommerzialisierung in 2007 zum Testen verschiedener Trennmedien zur Proteinaufreinigung als auch zur Optimierung von Beladungs- und Elutionsbedingungen auf einer Pipettierstation eingesetzt. Einher mit dem Gebrauch der Methodik gehen die immer noch offenen Fragestellungen, inwiefern und zu welchen Bedingungen die erzielten Daten mit den im traditionellen Labormaßstab oder auch großtechnischen Maßstab erzielten Ergebnisse vergleichbar sind. Um diese Frage letztlich beantworten zu können, bedarf es, neben der eigentlichen Untersuchung von Maßstabseffekten (z.B. über Strömungssimulationen), der genauen Charakterisierung der methodenspezifischen Parameter wie z.B. des gewählten Fraktionierungsschemas. Sowohl experimentell als auch unter Verwendung von Simulationen wurden qualitativ und quantitativ die Einflüsse von Pipettiergenauigkeit, Absorptionsmessungen in Mikrotiterplatten, relative Fraktionsgröße, Flussunterbrechungen, Salzstufenkonzentrationen in Gradientenapplikationen als Einzelparameter und in Kombination erfasst. Die Ursachen für vermeintlich höheres Messrauschen bei Gradientenelutionen konnten in Chromatographiesimulationen am Beispiel der Trennung von AV und LYS erklärt werden. Des Weiteren wurde gezeigt, dass durch geeignetes Peakfitting der Einfluss der Salzstufen, mit denen der Gradient approximiert wird, nahezu vernachlässigbar wird. Der Einfluss der gewählten Flussrate wurde im Zusammenspiel mit den bei Salzgradientenelutionen nötigen Flussunterbrechungen dargelegt. Die gezeigten Einflüsse verschiedenster Parameter sind für Anwender, die mittels dieser Hochdurchsatz-Methodik Adsorptionsparameter bestimmen und ein chromatographisches Modell kalibrieren möchten, von Bedeutung. Die erzielten Erkenntnisse über die rein methodenspezifischen Effekte komplementieren zukünftige Arbeiten, in denen Effekte, die alleinig auf Maßstabsunterschiede basieren, untersucht werden.

Contents

Abstract **I**

Zusammenfassung (deutsche Fassung) **V**

1 Introduction **1**
1.1 Protein purification 1
1.2 Purification process development for therapeutic proteins 2
1.3 High-Throughput Process Development 5
1.3.1 Terminology 5
1.3.2 The evolution of HTPD - a literature review 6
1.3.3 Implementation and challenges of HTPD 8
1.4 The used avidin protein system 10
1.4.1 The interest in avidin 10
1.4.2 Properties of the avidin protein system 11
1.4.3 Purification of avidin 11
1.5 References in section Introduction 13

2 Research Proposal **21**

3 Publications & Manuscripts **23**
A Sub-Two Minutes Method for MAb-Aggregate Quantification using Parallel Interlaced Size Exclusion HPLC 25
High-Throughput Analytical Methods for the Egg White Protein Avidin 45
Evaluation of PEG/Phosphate Aqueous Two-Phase Systems for the Purification of the Chicken Egg White Protein Avidin by using High-Throughput Techniques 71
High-Throughput Process Development of Purification Alternatives for the Protein Avidin . 95
High-Throughput Column Chromatography performed on Liquid Handling Stations – Process Characterization and Error Analysis 129

4 Conclusion **167**

5 Glossary **171**

6 Bibliography **177**

1 Introduction

1.1 Protein purification

Proteins play an important role in our everyday lives. Industrially manufactured proteins are used in food, chemicals, cosmetics, and pharmaceutics either as active compound (e.g. enzymes, therapeutic proteins) or for aiding product formulation (e.g. gelatine, whey proteins). A large group of applied proteins originates from plants or animals. With the introduction of recombinant protein expression techniques, however, an increasing number of proteins are manufactured in large-scale using genetically modified organisms. Regardless whether the protein of interest (POI) is derived from natural sources or is expressed by cultivated organisms, purification is challenging as the number of impurities is vast. The removal of these impurities is achieved by a sequence of separation operations referred to as downstream process. Most process-related impurities originate from the host organism, such as cell fragments, DNA, RNA, host cell proteins (HCP), lipids, glycosides, endotoxins, etc. Other process-related impurities originate from the cultivation media or are introduced during the downstream process by the applied techniques. Product-related impurities are proteins differing only slightly in the amino acid sequence, by charge, or by post-translational modifications, such as glycosylation, phosphorylation, (de-)amidation, hydroxylation as listed in [17, 127].

From the variety of impurities it becomes clear that the efforts required in order to obtain pure POI can be very high and costly. Due to the high complexity of extracts from natural sources or cultivation fluids, a sequence of several purification steps is typically required to achieve the desired purity. For the purification of a POI (including the removal of water/solvent), common mechanical and thermal engineering process steps are applied, such as filtration, centrifugation, solid-liquid chromatography and liquid-liquid extraction. In contrast to separation of (bio-)chemical compounds of low molecular weight, protein bio-separations are usually performed at moderate conditions regarding pH, temperature, shear stress, and content of organic solvent, in order to maintain the biological functionality of the POI. An overview of the prevalent separation techniques along with the underlying biophysical parameter and the common purpose of the process step is given in Table I.

The selection of a purification technique and its process parameters is based on the biophysical properties of both the POI and the impurities. These properties are given by the sequence of amino acids the POI is composed of resulting in a three-dimensional structure of certain size, hydrophobicity, number and distribution of charged groups. As the charge of the amino acids changes with pH, proteins feature a titration curve with a net charge of zero at the isoelectric point (pI). Differently charged proteins can therefore be separated by using ion exchange adsorption when selecting suitable pH and salt conditions. The composition and sequence of amino acids of the POI results in hydrophobicities unique for each protein. The hydrophobicity depends on the distribution of nonpolar and polar patches on the protein surface. Due to structural flexibility and the buffering property of proteins, apparent hydrophobicity and solubility of a protein are functions of the solvent properties, such as content of salts, organic compounds, pH and temperature. The attempt to identify correlations between solubility

and hydrophobic character of a protein with the retention behavior on hydrophobic chromatographic resins [60, 77, 121, 88] and the distribution in aqueous two-phase systems [4, 60, 3, 91] is part of extensive research. Besides charge and hydrophobicity, differences in the molecular sizes of proteins can be utilized for separation. However, in comparison to other purification techniques separation processes based on filtration or size-exclusion feature either low resolution or low capacity.

Table I Overview of prevalent techniques used in protein purification processes, the underlying biophysical parameter, and the purpose of the step. V: liquid volume.

Unit operation	Parameter	Purpose
Centrifugation	Density	Removal of cells (+debris)
Filtration:		
Depth	Hydrodynamic radius	Removal of cells (+debris)
Micro	Hydrodynamic radius	Removal of cells (+debris)
Ultra	Hydrodynamic radius	Classification
Nano		Removal of small molecules, reduction of V
Precipitation	Solubility	Purification, reduction of V
Crystallization	Solubility	Purification, reduction of V, formulation
(Aqueous) two-phase extraction	Solubility, hydrophobicity	Purification, removal of cells (+debris)
Chromatography:		
Ion exchange	Electrostatic properties	Purification, reduction of V
Reversed phase	Hydrophobicity	Purification, reduction of V, buffer exchange
Hydrophobic interaction	Solubility, hydrophobicity	Purification, reduction of V
Size-exclusion	Hydrodynamic radius	Purification, buffer exchange
Affinity	Specific affinity	Purification, reduction of V
Mixed-mode	Electrostatic properties, hydrophobicity	Purification, reduction of V

Within the last decades, new integrated protein purification techniques have been evolved that are based on new principles at both molecular level regarding biophysical interactions and process engineering level regarding new equipment. For example, mixed-mode chromatography resins feature ligands with both polar and unpolar characteristics. With such resins, a separation of proteins which elute similarly from ion exchange or hydrophobic interaction chromatographic resins can be achieved. Improved engineering solutions for integrated processing and increased productivity are expanded bed adsorption [19, 59], magnetic separation [104, 39], and a variety of techniques for continuous processing (e.g. simulated moving bed [107] and counter-current liquid-liquid chromatography [118]). Which technique is applied depends first of all on the purity requirements and targeted ratio of production costs to sales pricing. In this work, focus was laid on the costly therapeutic proteins and their purification process development, in which bead-based chromatography is still the prevalent technique used.

1.2 Purification process development for therapeutic proteins

Therapeutic proteins are often referred to as biopharmaceuticals. This class of drugs cover a wide range of peptides, enzymes, hormones and larger proteins, such as antibodies and coagulation factors. They are classified due to their mode of effect as given by Leader *et al.* [74]. The major advantage of biopharmaceuticals certainly is their specific interaction and mode of effect within the human body. Most biopharmaceuticals feature such high complexity that chemical synthesis is not possible with current methods. In the second half of the 20th century the increasing understanding of biochemical processes in the human body paired with recombinant DNA technology and significant advances in large-scale cultivation technology led to a revolution within the pharmaceutical market. The industrial production of proteins which naturally occur in the human body became possible. Thus, not only extraction processes from human or animal blood or tissue could be replaced, but also the production

volume of these products significantly increased. The first recombinantly produced insulin licensed by Eli-Lilly in 1979 certainly poses the most prominent example of this development. Furthermore, technological advances and the increasing understanding of human diseases resulted in new types of therapies of human diseases, such as the use of humanized antibodies for cancer therapy and auto-immune disorders.

'Eroom's law'

The market of biopharmaceuticals has been growing strongly since the last quarter of the 20th century. The share of biopharmaceuticals within all licensed drugs was 18% in 2012 with US$ 169 billion and is forecast to increase to 19–20% in 2017, hence, achieving about US$ 220 billion in sales per year [62]. Remarkably, five of the top-ten selling biopharmaceuticals are antibodies making alone US$ 42.4 billion [112]. However, despite growth and strong sales numbers, the pharmaceutical industry is struggling with increasing costs per drug marketed. A market review of Scannell *et al.* [106] revealed that over the last 60 years the number of approved drugs per billion US$ spent in R&D has halved every nine years as shown in Figure 1. This trend is referred to as the Eroom's law (Moore's law backwards [106]) relating to the known phenomenon in IT industries, in which the number of transistors that can be placed on an integrated circuit for similar costs is doubled every two years. As reviewed by Morgan in 2011 [86], numbers stating the costs of one single drug entering the market vary much. In 2013, costs were estimated to possibly sum up to US$ 1.3 – 5.5 billion [55, 83] by taking into account the costs for drug candidates that fail in clinical trials. Among all industries, spendings for product development in R&D is by far the highest in pharmaceutical industries making 14.4% of net sales followed by software and computer services (9.9%), and the automobile industries with 4.2% (data from 2012, [34]).

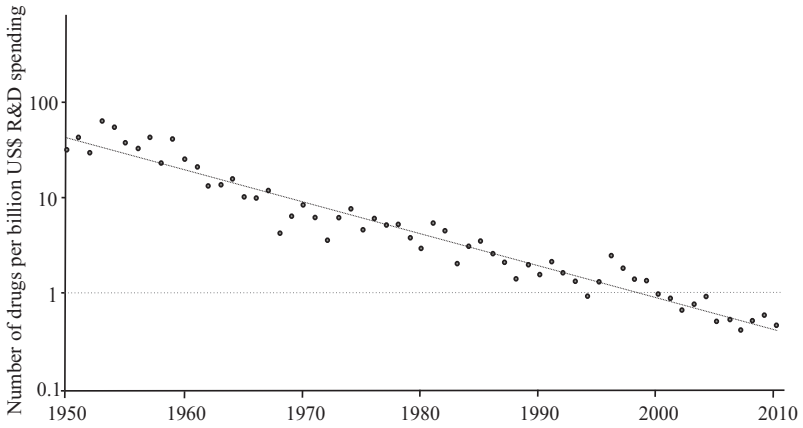

Figure 1: Development of expenses spent in Research and Development (inflation-adjusted) per marketed drug in billion US$. Reproduced from [106].

Among the reasons diagnosed for this development ([106, 85]) are the increasing development costs for processes and clinical trials and the increasing requirements from regulatory authorities regarding documentation of safety and efficacy. In particular, expenses for testing drugs in patients during clinical trials are high. Ten to fifteen years pass from the identification of a potential active pharmaceutical ingredient until entering the market. Companies aim to reduce 'time-to-market' for obvious reasons. This is attempted by reducing the time spent during drug discovery, process and formulation development, as time-lines during clinical phases can often not be further decreased. However, maybe equally important to the reduction of 'time-to-market' is the 'time-to-failure', because

the late recognition of insufficient efficacy or occurring side-effects in late clinical phases is extremely costly. As described in the first section, the development of a purification process is a complex task due to high number of impurities in the starting material (e.g. cultivation broth or tissue extract) and due to an impurity profile being unique for each expression strain and molecule. Costs for downstream processing development and production do not increase linearly with increasing purity requirements but rather increase exponentially making production costs of biopharmaceuticals with purities of 90% up to >99.9% particularly expensive. Another cost-driver are modifications during the downstream process, such as pegylation or ligation of functional side chains causing additional costs for side chains, reaction agents and require further purification due to new side products generated.

Pressed by the need of decreasing time-lines and costs, the way processes are developed has been questioned. In the past, processes were developed mainly driven by heuristic rules and experimental work on a trial-and error basis. Current alternatives and new approaches address an optimization of throughput and efficiency of downstream process development[49, 50]:

- *Fusion protein technology*
  By recombinant fusion technology the amino acid sequence of the POI is prolonged during expression by an additional tag. This tag sequence features specific affinity to another molecule thus enabling a highly-selective purification of the tagged POI via affinity chromatography (large-scale) or functionalized magnetic bead separation (lab-scale). A removal of the tag after capturing of the POI can be achieved by enzymatic cleavage. Pairs of tag-ligand exist for protein–protein, protein–small biological ligands, peptide–protein, or peptide–metal chelating ligands as recently reviewed by Pina *et al.* [95]. The technology is mostly applied in the screening phase of new proteins, since sufficient purity for subsequent efficiency tests can be achieved by one standardized step. This is of major significance in the early-phase of drug development where only a few milligrams of each protein are needed. The advantage of a more selective capture of the POI might, however, also be economical for the final production process. Furthermore, by fusion tags biophysical properties, e.g. stability, can be beneficially altered, and thus, the downstream process improved.

- *Platform-approach*
  The platform-approach describes the attempt to standardize the sequence of a purification process which is applied for a class of molecules. The most prominent example of this approach are the antibody purification platform processes [35, 109]. By standardization of the sequence of the process steps, further development is limited to fine-tuning of the process parameter ranges while the steps and techniques are set, and thus, experimental development work can be significantly reduced. Furthermore, pilot and production facilities can be employed for a whole class of antibodies without the need for rebuilding the facility. With an increasing number of molecules manufactured via a given platform process, another benefit from this approach is, that a data base on the process steps used is generated which can be used in risk evaluation of process parameters criticality for new and similar molecules. However, the development of a platform is challenging for other molecule classes than monoclonal antibodies because this process is mainly based on the outstanding selectivity of the first Protein A capture step achieving antibody purities of >95%.

- *Model-based process development*
  The increasing computational power seen within the last two decades enabled a more detailed modeling of molecular interactions as well as transport phenomena. Molecular dynamics simulations are used to mechanistically describe differences in the chemical potential of proteins,

solvent, and adsorbents at varying physico-chemical conditions. Attractive and repulsive forces on molecular level are correlated to aggregation, adsorption and desorption on a macro-level [73, 27, 125, 72]. Often, more simplified models, such as the Langmuir model or the steric-mass action model, are used to describe protein adsorption in dependence of the ionic strength. Extended models take several parameters, such as the modifier concentration and pH, into account [10, 22]. Chromatographic adsorption models are combined with models describing the transport phenomena known from chemical engineering science and in some cases by also accounting for diffusional hindrance of proteins in the resin pores [48, 42, 84, 108, 25]. The description of processes (model) and their predictive application (simulation) for process optimization and robustness analysis can result in a significant reduction of costs and time, however, is obviously depending on both the suitability of the model and the accuracy of the experimental data used for model calibration. The determination of model parameters of reasonable quality often poses a major challenge when applying modeling. Besides the determination of model parameters by experiments or via numerical fit routines, statistical methods such as quantitative structure-property relationships (QSPR) are used to describe a chemical property based on the molecule structure [64]. The identified descriptors are then correlated to macroscopic behavior, e.g. adsorption and partitioning.

- *High-throughput process development*
  The miniaturization of experiments and their automation by using liquid handling stations enable the screening of more parameter combinations while requiring less material and time in comparison to traditional experimental methods. This results in both a reduction of costs and increased process knowledge at early stages of the process development, and thus, in fast(er) identification of process conditions and optimization. This approach is in particular focus of this thesis, and therefore, terminology and further information will be elucidated in detail in the following.

While each of the above stated methodologies have the potential to improve process development in terms of reducing time-lines and costs, state-of-the-art bioprocess development is a combination of technologies, also referred to as 'hybrid' process development [90, 50].

1.3 High-Throughput Process Development

1.3.1 Terminology

Within the last decade, several terms have been used in literature describing the application of high-throughput methods in process development. However, the terminology used is not consistent, and therefore, the author gives an overview on the terms used in this work:

- *High-throughput*
  The term high-throughput screening originates in the 1990s and describes major advances of experimental methods in drug discovery [11, 15, 110]. Improvements in computational techniques and robotics in combination with (bio-)chemical assays enabled the rapid screening of potential drug candidates of low molecular weight (high-throughput screening, HTS). In the field of drug discovery, therefore, the term 'high-throughput' refers to a screening of several thousands of samples per day. A throughput of even more than 100.000 samples per day was outlined by Hook in 1996 [57], anticipating the development of micro-array technologies.

In this work and in the field of protein purification, the term 'high-throughput', however, is already used when performing dozens or hundreds of experiments per day rather than thousands. The reason for these differences in sample throughput lies in the complexity of processes applied for protein purification and constraints to the minimal reasonable scale. As in comparison to traditional laboratory method such experimental methods lead to a significant gain in throughput, using the term 'high-throughput' might nevertheless be justified.

- *High-throughput experimentation (HTE)*
  High-throughput experimentation (HTE) refers to experimental methods that are based on miniaturization, parallelization, and automation leading to a significant increase in sample throughput. In this work, HTE is considered only in the field of protein purification and refers to small-scale models for typical process unit operations. Liquid handling stations and micro-titer plate (MTP) formats (e.g. 96-well, 384-well plates) are considered as essential elements in HTE.

- *High-throughput analytics (HTA)*
  In the analytical field, the counterpart of HTE is referred to as high-throughput analytics (HTA). In this work, HTA refers to all kinds of methods for protein analysis featuring a high sample throughput, a high degree of automation, and the possibility of handling small sample volumes.

- *High-throughput process development (HTPD)*
  In this work, HTPD is considered as the methodology of developing manufacturing processes for peptides and proteins by combining the elements of HTE and HTA, and thus, reducing expenses for material, manual work, and time. While the term HTPD comprises the full chain of process development consisting of up-, downstream, and formulation processes, in this work, the focus lies on downstream processing. The experimental part is supplemented by the application of computational tools, e.g. mathematical modeling. Although rarely mentioned in literature, for a functional HTPD-approach, suitable techniques for data evaluation, data managing and time-management tools for synchronization of working routines are essential elements complementing HTE and HTA.

1.3.2 The evolution of HTPD - a literature review

In the beginning of a development resulting in the current form of HTPD, principles of HTS as known from drug discovery were adapted for the use in bioprocess development. In first publications in the beginning of the 21st century, it was the use of multi-well micro-titer plates rather than the use of liquid handling stations which was adapted first and stated as HTS. Methods were developed for strain selection, protein production for biochemical assays, and enzymatic conversions [113, 126, 79, 2]. In the field of downstream processing, the first publications on HTS were presented for sample conditioning in protein crystallization screenings [114] or for chromatographic separation by batch adsorption [119, 100]. The development of new hardware for HTE in the field of liquid chromatography certainly poses an important milestone. Chromatographic processes are, on one hand, most-often used for the purification of biopharmaceuticals, but, on the other hand, are also very time-consuming to develop. With the introduction of commercially available hardware and its application described in several publications [111, 9, 23, 14, 129, 20, 99, 123], a market was served, that desired a revolution of the traditional experimental approach. Interestingly, a significant fraction of initial studies were published from industries, a fact which demonstrates the industrial need for speeding up the time-consuming purification process development.

An increasing focus on HTE led to the development of methods covering almost the whole spectrum of protein separation science by transferring traditional laboratory methods into a micro-plate format, which enabled parallelization and automation [21, 69] on liquid handling stations. Besides chromatography, the range of published methods covers solubility and precipitation [130, 66, 93, 8], protein refolding [124, 7], crystallization [114, 61], two-phase separation [6, 92, 91, 93], and filtration [18, 63, 98, 65].

From the advances in HTE it very soon became clear that increasing the number of generated samples does not necessarily lead to an overall acceleration of the process development work. The access to analytical methods matching the speed of HTE was identified as a major bottleneck. In recent years, new types of bioanalytical devices have entered the market which are capable of processing a high number of samples in short time and can be combined with automation platforms (e.g. chip-based capillary electrophoresis, immunoassays in MTP or on centrifugal disc, MTP confocal imaging, ultra high performance liquid chromatography coupled with mass spectrometry, MTP dynamic light scattering, surface plasmon resonance [13, 120, 40]). Besides new hardware, various new analytical methods have been developed addressing the need for high-throughput and also taking the low sample material available into account [36, 37, 52, 51, 102]. A generic approach to select suitable analytical methods to complement HTE has recently been presented by [67]. With HTE and HTA going hand in hand and a synchronization of working routines, the term HTPD became eventually meaningful.

Another aspect of HTPD is the increasing use of computational methods for modeling of structural protein properties and protein separation, as modeling provides the possibility of both increasing throughput and saving experimental work. In an optimal scenario, experimental data obtained by HTE and HTA are used to calibrate empirical or mechanistic models that can then be used for process optimization and prediction of scale-up scenarios. Several studies have been published from researchers in academia demonstrating such 'hybrid' approaches on protein model systems for single unit operations [101, 116, 89, 88, 94]. Furthermore, a high-throughput of samples enables the use of strategies for the design of experiments that would not be feasible using traditional methods. Besides the use of experimental designs of higher resolution, the application of genetic algorithm or simplex algorithm combined with HTE has been reported [117, 122, 67].

The evolvement of methods applied in both academia and industries is accompanied by a dedicated conference series which started in 2010 ('HTPD', www.htpdmeetings.com). Here, in discussion across institutions and industries, other potential bottlenecks in HTPD have been identified, namely generic methods for data evaluation, data handling and data management. Processing of large data sets can not be performed adequately without automated evaluation procedures and, furthermore, instrument software packages need to be compatible to the software that is used for data processing and storage [75].

In summary, HTPD has evolved from a set of single, customized screening methods to a toolbox containing items for both experimentation and analytics. The success of HTPD in biopharmaceutical industries is based on the major advantages of providing a high degree of automation, a high number of experiments, and, most importantly, a reduction of time for the identification of suitable process parameters. Furthermore, HTPD fits into the 'Quality-by-Design' initiative [38] by providing the opportunity to exploit the parameter space to greater extent, and thus, to gain better process understanding [12, 69].

1.3.3 Implementation and challenges of HTPD

The methodology of HTPD is considered state-of-the-art in early-phase bioprocess development [50]. However, the implementation on HTPD can pose a complex task. A schematic of an HTPD implementation process sketched from the author's point of view in relation to the net benefit achievable is shown in Figure 2. A 'benefit' is considered the improvement in at least one of the following: a reduction of development time and hence costs, or an increase in process knowledge relevant for identifying robust and close to optimal process conditions.

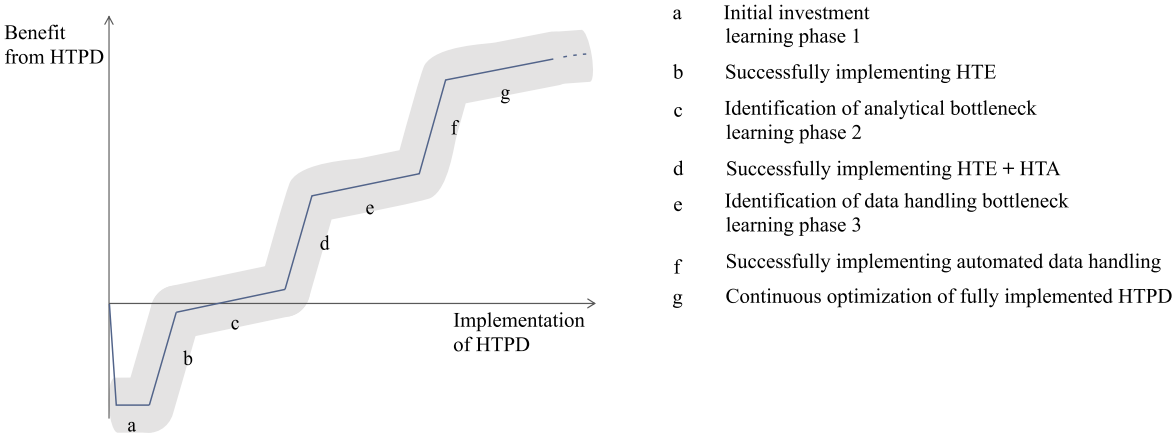

Figure 2: Illustration of qualitative increase in the benefits from HTPD with different implementation phases. The gray shaded line outline the variability in this qualitative trend curve given by differences from case to case.

The start-up phase of the implementation of HTPD (Figure 2-a) is characterized by a series of investments, typically the acquisition of a liquid handling station and other supplementing hardware. HTE requires consumables that are typically for single-use and more expensive than traditional labware. In a first learning phase, published experimental methods are used as starting points, and training and time for gaining experience are required. The development of generic methods applicable to a broad spectrum of protein systems while avoiding over-optimized and too complex methods might require a large initial investment until eventually HTE is successfully applied (Figure 2-b). If sample preparation (e.g. ultra-, diafiltration) cannot be automated in small-scale, or treatments specific to the protein system are required, degree of automation and speed of the adjusted methods are compromised. Whether the application of HTE at this stage already poses a benefit to the organization might depend on the chosen application and the willingness to invest in method development and dedicated personnel. A potential gap between the achieved level of HTE and the expectations at this phase might result in disappointment and potentially in even stopping further investment in the HTPD-program.

In the case where HTE has been implemented successfully, the bottleneck caused by low analytical throughput becomes instantly obvious. Further investments in both analytical method development and analytical equipment compatible with HTE are required (Figure 2-c). Resources need to be available for adapting published methods or developing adequate methods from scratch. Again, there is a risk of developing methods that are too specific to be applied to other molecules in the pipeline. The selection of generic analytical techniques and a reduction of method development times significantly contribute to successful implementation of HTA (Figure 2-d). If experimentation and analytics are

performed in different parts of an organization, the experimental workflow needs to be coordinated, a fact which needs to be considered in the early planning phase.

The increasing amount of data that comes along with HTE and HTA requires evaluation procedures unknown from typical laboratory experiments. While, for example, three samples are obtained from a common chromatography experiment (application, wash, and elution pool), a comparable experiment in HTE-mode using 96-well filter plates would generate about 6–12 samples per experiment (well), because wash and elution are performed in repetitive cycles [23]. Considering a total of approximately 600 – 1200 data points from one 96-well plate, potentially several experiments per day, and furthermore, the need for cherry-picking and sorting samples in preparation for applying different types of analysis, manual analysis is not feasible any longer. Methods for the automated data evaluation and data storage need to be developed requiring further investment in method development (Figure 2-e). Accessibility of data from various control software packages might pose a first hurdle to overcome. Suitable evaluation methods would feature automated sorting of data and calculation of relevant information with subsequent storage of data in common formats. Implementing sufficient flexibility in order to handle data also when experimental methods change might be one of the essential learnings (Figure 2-f). With all relevant elements of HTPD implemented, continuous improvement can be expected (Figure 2-g) by increasing experience in HTPD and a growing method toolbox, but also due to growing data bases. These contain knowledge applicable to molecules of the next generation.

Considering current topics in the field of HTPD (www.HTPDmeetings.com) and by taking available literature into account, the following issues in the field of HTPD still require further improvement:

1. Demonstrating the application of HTPD using various molecule classes
   A large fraction of published HTE- and, in particular, HTA-methods were developed and used for the development of antibodies or antibody fragments purification processes (e.g. [36, 66, 92, 123]). Other published methods were developed using artificial model systems consisting of two or three model proteins [117, 94, 51]. Demonstration of HTE- and HTA-methods to industrial-relevant systems apart from the well-known antibody platform is therefore highly desired.

2. Extending the HTA-toolbox
   Due to the demonstration of HTE-methods using simple protein systems, the requirements on the analytical techniques were often low and easily fulfilled by using photometric measurements in plate readers. In order to completely remove the analytical bottleneck, the portfolio of HTA-methods needs to be further expanded for more complex protein systems and by methods with highest precision.

3. Understanding scale-down effects
   Since the introduction of HTPD, questions and concerns are raised regarding the degree of comparability of HTE-data to data obtained by traditional laboratory methods. In general, data comparability is stated sufficient for fulfilling the needs in early-phase process development, that is, a reduction of the parameter space by screenings of process parameters and chromatographic phases. With the intention of expanding the use of HTPD to later phases of process development the requirements on data quality are, however, significantly higher. For detailed process optimization or process characterizations using HTE-methods, it is essential to know whether differences seen in data from different scales/methods are caused by the experimental parameters (method-specific) or by physical effects (scale-specific). Further studies of scale-effects are therefore required in order to better understand chances and limitations of the miniaturized methods [68, 41].

4. Understanding experimental errors of scaled-down methods
   In order to assist the mentioned scale comparability studies and hence make HTPD an acknowledged tool in a Quality-by-Design strategy, the variances introduced by the HT-method need to be characterized. With this knowledge, scaling effects would become distinguishable from method-specific effects. A better understanding of the experimental errors paired with further knowledge on scale-effects eventually lead to further acceptance of HTPD in industries and by regulatory authorities.

5. Establishing standardized methods for generic data evaluation and data management
   While commercial software applications for the control of HTE- and HTA-instruments are typically easy to use, user-friendly applications of data transfer, data evaluation and data storage are commonly developed in-house. Increasing standardization of data formats as well as development of generic evaluation routines flexible enough for a variety of experimental approaches would facilitate simpler access to methods and data [75] and hence, reduce reservations towards HTPD.

This work presents contributions to challenges 1, 2, and 4 described above. The identification and development of a suitable HTA-method applicable to the chosen, industrially-relevant protein system of avidin represents an essential part of this work. Furthermore, achieving both high accuracy and high precision was prioritized during the development of HTA-methods. In order to shed light on the method-specific variances obtained in HTE, a detailed characterization of a HTE-method for column chromatography was performed.

1.4 The used avidin protein system

As model system, the purification of avidin from pre-treated hen's egg white was chosen.

1.4.1 The interest in avidin

The protein avidin has drawn attention of researchers in the field of biology, nutrition science, and pharmacology due to two outstanding properties. First, avidin is very stable at rather extreme pH-values and temperatures [45]. Second, avidin features very high affinity to biotin. In the 1940s, observations of toxic reactions in animals when administrating a diet of raw or dried egg white was explained by the presence of avidin in egg white binding to the essential nutrient biotin (vitamin H) [31, 32, 115, 71]. In the 1960s and 1970s, research groups of Green and DeLange crucially contributed to the knowledge on the amino acid sequence and structure of avidin, and the biotin-binding site. They also published methods on the fractionation and purification of avidin from egg white [43, 82, 44, 45, 26, 58]. In the following decades, assays based on the avidin-biotin interaction were developed for the use in bioanalytical methods for diagnostics, immunology or affinity-based separation [131, 132, 133, 87, 81]. Pioneering work was done by the working group of Bayer and Wilchek providing a comprehensive list of publications, reviews, and innovative applications of the avidin-biotin-technology [33]. In a pharmaceutical context, avidin-biotin technology has gained interest for intelligent drug delivery [105, 134] or targeting biotin-labeled drugs for controlled inactivation [53]. Recombinantly manufactured streptavidin and modified forms were found an alternative to avidin featuring very similar properties regarding the binding to biotin [78, 70]. Recombinant manufacturing of avidin appears desirable and was reported feasible however with low titers (<10 mg/L [1]). Extraction of avidin from egg white might still be the most economical method, as several other egg white proteins that are obtained during the fractionation process can also be commercially used. The growing field of

applications of the avidin-biotin technology led to an increasing demand of purified avidin. In order to fulfill the demands, existing manufacturing processes are reviewed and optimized aiming a reduction of purification costs of avidin.

1.4.2 Properties of the avidin protein system

The avidin dry mass fraction of egg white protein is only 0.05%. Pre-treated egg white (after removal of insoluble mucin, lipids, water, and a large fraction of albumins) was used for the intermediate purification and polishing of the avidin (AV). The major impurities in this starting material were ovalbumin (OV), ovotransferrin (OT), ovomucoid (OM), and lysozyme (LYS). In this work, it was therefore focused on the separation of avidin from these four impurities.

Avidin consists of four identical subunits each composed of 128 amino acids [26]. The four subunits have a total weight of 57.4 kDa. The difference to the experimental weight of the naturally occurring tetrameric avidin (66.0 – 68.3 kDa [47, 46, 76]) is explained by the high degree of oligosaccharides attached to amino acid asparagin 17 of each subunit. Each subunit binds non-covalently one biotin molecule with extraordinarily high affinity (dissociation constant of $k_d \approx 10^{-15}$ [43]). The biotin-binding activity was found the same for native glycosylated avidin compared to non-glycosylated avidin [56]. The transition temperature indicating thermal denaturation of avidin and the avidin-biotin complex is 85°C and 132°C, respectively [28].
In Table II, biophysical properties of avidin and the major impurities in ovomucin-free egg white are summarized.

Table II Selected properties of egg white proteins used in this study. Stated data are considered average values as isoforms are not taken into account. Data from [16]

Protein	Mass fraction in total protein (%) dry mass	Molecular weight (kDa)	Isoelectric point (-)
Avidin	0.05	66.0–68.3	10.0
Ovalbumin	54	45–46.0	4.5
Ovotransferrin (Conalbumin)	12.0–13.6	76.0–80.0	6.1–6.6
Ovomucoid (Ovotrypsin-Inhibitor)	11.0	28.0	3.9–3.5
Lysozyme	3.5	14.3–14.6	10.7

1.4.3 Purification of avidin

Extraction and purification of avidin in large-scale is cost-intensive, in particular when considering the high purity (>99%) demanded for the use as active pharmaceutical ingredient. Commonly, fractionation of egg white proteins is achieved by precipitation using inorganic salts, such as ammonium sulfate, or organic solvents, such as ethanol, followed by ion exchange chromatography and gel filtration [128, 82, 5]. With the given differences in the isoelectric points of avidin and its impurities (Table II), a charged-based separation appears promising and is, in fact, the pre-dominantly used technique for the purification of AV reported in literature. A review on avidin purification processes based on ion exchange chromatography was given by Durance [29]. When using ion exchange chromatography, the separation of avidin and lysozyme poses the main challenge to be overcome by the selection of suitable ligands and solvent systems. Piskarev *et al.* [96] obtained an AV purity of approx. 99%, however, employed a concave gradient and high-pressure cation exchange chromatography, which comes with limitations when scaling-up. Limitations regarding scaling-up is also the disadvantage of the approach Rothemund and co-workers presented [103]. In their work, avidin was effectively separated

from lysozyme in a two-step filtration process using an electrophoresis instrument with a recovery of 60%-65%.

The high affinity of avidin to biotin provides the possibility to capture effectively avidin by using biotin-tagged ligands. The application of affinity-based techniques was reported in several studies already in the 1960s [80, 24, 54]. However, a preparative application of biotinylated resins is not favorable due to both harsh elution conditions required and high resin manufacturing costs. An alternative affinity-based method was presented by Rao *et al.* in 2002 [97] by selectively eluting avidin from a weak cation exchanger with 4'-hydroxy-azobenzene-2-carboxylic acid (HABA), that binds to avidin similarly selectively as biotin. In contrast to the biotin-avidin complex, the HABA-avidin complex can be disrupted under less harsh conditions (e.g. 0.2 M acetic acid), however, a subsequent dialysis step is required.

As a separation of lysozyme and avidin via charged-based separation is challenging, other techniques employing differences in protein hydrophobicities could pose suitable alternatives. However, only one study has been published demonstrating the use of metal-chelate interaction and hydrophobic interaction for the separation of avidin from lysozyme in comparison to the use of cation exchange chromatography [30]. While the latter was shown more effective, it was found worthwhile to re-evaluate alternative techniques and, in this work, explore different methods aiming at a high purity of the final avidin fraction and an easy-scalable, cost-effective purification process.

1.5 References in section Introduction

[1] Airenne, K. J., Marjoma, V. S., Kulomaa, M. S., 1999. Recombinant avidin and avidin - fusion proteins. *Biomol. Eng.* 16, 87–92.

[2] Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G., Kell, D. B., 2003. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* 21 (6), 692–6.

[3] Andrews, B. A., Schmidt, A. S., Asenjo, J. A., 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: effect of surface hydrophobicity and charge. *Biotechnol. Bioeng.* 90 (3), 380–90.

[4] Asenjo, J., Schmidt, A., F., H., B.A., A., 1994. Model for predicting the partition behaviour of proteins in aqueous two phase systems. *J. Chrom. A* 668, 47–54.

[5] Awade, A. C., 1996. On hen egg fractionation: applications of liquid chromatography to the isolation and the purification of hen egg white and egg yolk proteins. *Z. Lebensm. Untersuchung Forsch.* 202 (1), 1–14.

[6] Bensch, M., Selbach, B., Hubbuch, J., 2007. High throughput screening techniques in downstream processing: Preparation, characterization and optimization of aqueous two-phase systems. *Chem. Eng. Sci.* 62 (7), 2011–2021.

[7] Berg, A., 2008. Development of High-Throughput Screening Methods for the automated optimization of inclusion body protein refolding processes. Ph.d., Karlsruhe Institute of Technology (KIT).

[8] Berg, A., Schuetz, M., Dismer, F., Hubbuch, J., 2013. Automated measurement of apparent protein solubility to rapidly assess complex parameter interactions. *Food Bioprod. Process.* 92 (2), 133–142.

[9] Bergander, T., Nilsson-Välimaa, K., Öberg, K., Łacki, K. M., 2008. High-Throughput Process Development: Determination of Dynamic Binding Capacity Using Microtiter Filter Plates Filled with Chromatography Resin. *Biotechnol. Progr.* 24 (3), 632–639.

[10] Bernardi, S., Gétaz, D., Forrer, N., Morbidelli, M., 2013. Modeling of mixed-mode chromatography of peptides. *J. Chrom. A* 1283, 46–52.

[11] Bevan, P., Ryder, H., Shaw, I., 1995. Identifying small-molecule lead compounds: The screening approach to drug discovery. *Trends Biotechnol.* 13 (3), 115–121.

[12] Bhambure, R., Kumar, K., Rathore, A. S., 2011. High-throughput process development for biopharmaceutical drug substances. *Trends Biotechnol.* 29 (3), 127–35.

[13] Bielefeld-Sevigny, M., 2009. AlphaLISA immunoassay platform – The "nowash" high throughput alternative to ELISA. *Assay Drug Dev. Techn.* 7 (1), 90–92.

[14] Britsch, L., Schroeder, T., Friedle, J., 2008. Automated, High-Throughput Chromatographic Separation of Biological Compounds. *American Biotechnology Laboratory* 26 (6), 20–23.

[15] Broach, J., Thorner, J., 1996. High-throughput screening for drug discovery. *Nature* 384 (6604 Suppl), 14–16.

[16] Campbell, L., Raikos, V., Euston, S. R., 2003. Review Modification of functional properties of egg-white proteins. *Food/Nahrung* 47 (6), 369–376.

[17] Carta, G., Jungbauer, A. (Eds.), 2010. *Preparative Chromatography: Process Development and Scale-Up*, 1st Edition. WILEY-VCH Verlag, Weinheim.

[18] Chandler, M., Zydney, A., 2004. High throughput screening for membrane process development. *J. Membr. Sci.* 237 (1–2), 181–188.

[19] Chase, H. A., 1994. Purification of proteins by adsorption chromatography in expanded beds. *Trends Biotechnol.* 12, 296–303.

[20] Chhatre, S., Titchener-Hooker, N. J., 2009. Microscale methods for high-throughput chromatography development in the pharmaceutical industry. *J. Chem. Technol. Biotechnol.* 84 (7), 927–940.

[21] Chhatre, S., Titchener-hooker, N. J., 2011. Comprehensive Biotechnology, second ed. Edition. Vol. 1. Elsevier.

[22] Chilamkurthi, S., Sevillano, D. M., Albers, L. H. G., Sahoo, M. R., Verheijen, P. J. T., van der Wielen, L. A. M., den Hollander, J. L., Ottens, M., 2014. Thermodynamic description of peptide adsorption on mixed-mode resins. *J. Chrom. A* 1341, 41–9.

[23] Coffman, J. L., Kramarczyk, J. F., Kelley, B. D., 2008. High-throughput screening of chromatographic separations: I. Method development and column modeling. *Biotechnol. Bioeng.* 100 (4), 605–18.

[24] Cuatrecasas, P., Wilchek, M., 1968. Single-step purification of avidin from egg white by affinity chromatography on biocytin-sepharose columns. *Biochem. Biophys. Res. Commun.* 33 (2), 1967–1968.

[25] de Neuville, C. B., Tarafder, A., Morbidelli, M., 2013. Distributed pore model for bio-molecule chromatography. *J. Chrom. A* 1298, 26–34.

[26] Delange, R. J., 1970. Egg White Avidin: I. Amino Acid Composition; Sequence of the Amino- and Carboxyl-Terminal Cyanogen Bromide Peptides. *J. Biol. Chem.* 245 (5), 907–916.

[27] Dismer, F., 2009. Characterization of Transport and Adsorption Mechanism in Chromatographic Media. Ph.d., Karlsruhe Institute of Technology (KIT).

[28] Donovan, J. W., Ross, K. D., 1973. The Stability of Avidin Produced by Binding of Biotin. A Differential Scanning Calorimetric Study of Denaturation by Heat. *Biochemistry* 12 (3), 512–517.

[29] Durance, D. T., 1994. Separation, purification and thermal stability of lysozyme and avidin from chicken egg white. In: Sim, J. S., Nakai, S. (Eds.), *Egg Uses and Processing Technologies. New Developments.* International CAB, Wallingford, UK, pp. 77–93.

[30] Durance, T. D., Nakai, S., 1988. Purification of Avidin by Cation Exchange, Gel Filtration, Metal Chelate Interaction and Hydrophobic Interaction Chromatography. *Can. Instit. Food Sci. Technol. J.* 21 (3), 279–286.

[31] Eakin, R. E., McKinley, W. A., Williams, J., 1940. Egg-white injury in chicken and its relationship to a deficiency of vitamin H (biotin). *Science* 92, 224–225.

[32] Eakin, R. E., Snell, E. E., Williams, J., 1941. The concentration and assay of avidin, the injury-producing protein in raw egg white. *J. Biol. Chem.* 140, 535–543.

[33] Ed Bayer Group, . *Avidin-biotin publications*. November 2014.
URL `http://wws.weizmann.ac.il/Biological_Chemistry/scientist/Bayer/pubAvidin`

[34] EFPIA, 2014. The Pharmaceutical Industry in Figures. Tech. rep., European Federation of Pharmaceutical Industries and Association.

[35] Fahrner, R. L., Knudsen, H. L., Basey, C. D., Galan, W., Feuerhelm, D., Vanderlaan, M., Blank, G. S., 2001. Industrial Purification of Pharmaceutical Antibodies: Development, Operation, and Validation of Chromatography Processes. *Biotechnol. Gen. Eng.* 18 (1), 301–327.

[36] Farnan, D., Moreno, G. T., Stults, J., Becker, A., Tremintin, G., van Gils, M., 2009. Interlaced size exclusion liquid chromatography of monoclonal antibodies. *J. Chrom. A* 1216 (51), 8904–9.

[37] Farnan, D., Rea, J. C., Lou, Y., Parikh, R., Moreno, G. T., Nov 1, 2010. High-Throughput Multi-Product Liquid Chromatography for Characterization of Monoclonal Antibodies. *Biopharm International*.

[38] Food and Drug Administration, 2009. Guidance for Industry - Q8(R2) Pharmaceutical Development. URL `http://www.fda.gov/downloads/Drugs/Guidances/ucm073507.pdf`

[39] Franzreb, M., Siemann-Herzberg, M., Hobley, T. J., Thomas, O. R. T., 2006. Protein purification using magnetic adsorbent particles. *Appl. Microbiol. Biotechnol.* 70 (5), 505–16.

[40] Fraser, S., Cameron, M., O'Connor, E., Schwickart, M., Tanen, M., Ware, M., 2014. Next generation ligand binding assays-review of emerging real-time measurement technologies. *AAPS journal* 16 (5), 914–24.

[41] Gerontas, S., Shapiro, M. S., Bracewell, D. G., 2013. Chromatography modelling to describe protein adsorption at bead level. *J. Chrom. A* 1284, 44–52.

[42] Golshan-Shirazi, S., Guiochon, G., 1994. Modeling of preparative liquid chromatography. *J. Chrom. A* 658, 149–171.

[43] Green, N. M., 1963. Avidin 1. The use of ^{14}C Biotin for kinetic studies and for assay. *Biochem. J.* 89 (1959), 585–591.

[44] Green, N. M., 1963. Avidin 3. The nature of the biotin-binding site. *Biochem. J.* 89, 599–609.

[45] Green, N. M., 1963. Avidin 4. Stability at extremes of pH and dissociation into subunits by guanidin hydrochloride. *Biochem. J.* 89, 609–620.

[46] Green, N. M., 1975. Avidin. In: C.B. Anfinsen, J. T. E., Richards, F. M. (Eds.), *Advances in Protein Chemistry*. Vol. 29. Academic Press, pp. 85–133.

[47] Green, N. M., Toms, E. J., 1970. Purification and Crystallization of Avidin. *Biochem. J.* 118, 67–70.

[48] Gu, T., Tsai, G.-J., Tsao, G. T., 1993. Modeling of Nonlinear Multicomponent Chromatography. *Adv. Biochem. Eng. Biot.* 49, 46–69.

[49] Guiochon, G., Beaver, L. A., 2011. Separation science is the key to successful biopharmaceuticals. *J. Chrom. A* 1218 (49), 8836–8858.

[50] Hanke, A. T., Ottens, M., 2014. Purifying biopharmaceuticals: knowledge-based chromatographic process development. *Trends Biotechnol.* 32 (4), 210–20.

[51] Hansen, K. S., 2013. High Throughput Analytics for Protein Purification Process Development and Process Control. Ph.d., Karlsruhe Institute of Technology (KIT).

[52] Hansen, S. K., Skibsted, E., Staby, A., Hubbuch, J., 2011. A label-free methodology for selective protein quantification by means of absorption measurements. *Biotechnol. Bioeng.* 108 (11), 2661–2669.

[53] Harenberg, J., 2009. Development of idraparinux and idrabiotaparinux for anticoagulant therapy. *Thrombosis and Haemostasis* 102 (5), 811–5.

[54] Heney, G., Orr, G. A., 1981. The Purification of Avidin and Its Derivatives on 2-Iminobiotin-6-aminohexyl-Sepharose 4B. *Anal. Biochem.* 114, 92–96.

[55] Herper, M., 2013. The Cost Of Creating A New Drug Now $5 Billion, Pushing Big Pharma To Change. forbes magazine. \http://www.forbes.com/sites/matthewherper/2013/08/11/how-the-staggering-cost-of-inventing-new-drugs-is-shaping-the-future-of-medicine/, accessed Nov 10, 2014.

[56] Hiller, Y., Gershoni, J. M., Bayer, E. A., Wilchek, M., 1987. Biotin binding to avidin Oligosaccharide side chain not required for ligand association. *Biochem. J.* 248, 167–171.

[57] Hook, D., 1996. Ultra high-throughput screening – a journey into Nanoland with Gulliver and Alice. *Drug Discovery Today* 1 (7), 267–268.

[58] Huang, T. S., DeLange, R. J., 1971. Egg White Avidin II. Isolation, composition, and amino acid sequences of the tryptic peptides. *J. Biol. Chem.* 246 (3), 686–697.

[59] Hubbuch, J., Thömmes, J., Kula, M.-R., 2005. Biochemical Engineering Aspects of Expanded Bed Adsorption. *Adv. Biochem. Eng. Biot.* 92, 101–123.

[60] Huddleston, J., Abelaira, J. C., Wang, R., Lyddiatt, A., 1996. Protein partition between the different phases comprising poly(ethylene glycol)-salt aqueous two-phase systems, hydrophobic interaction chromatography and precipitation: a generic description in terms of salting-out effects. *J. Chrom.. B* 680 (1-2), 31–41.

[61] Hui, R., Edwards, A., 2003. High-throughput protein crystallization. *J. Struct. Biol.* 142 (1), 154–161.

[62] IMS Institute for Healthcare Informatics, accessed Nov 21, 2013. The Global Use of Medicines: Outlook through 2017.
URL http://www.imshealth.com

[63] Jackson, N., Liddell, J., Lye, G., 2006. An automated microscale technique for the quantitative and parallel analysis of microfiltration operations. *J. Membr. Sci.* 276 (1-2), 31–41.

[64] Kaliszan, R., Baczek, T., 2010. QSAR in Chromatography: Quantitative Structure-Retention Relationships (QSRRs). In: Puzyn, T., Leszczynski, J., Cronin, M. T. (Eds.), *Recent Advances in QSAR Studies*, 1st Edition. Springer Science, Heidelberg London New York, pp. 223–251.

[65] Kazemi, A. S., Latulippe, D. R., 2014. Stirred well filtration (swf) – a high-throughput technique for downstream bio-processing. *J. Membr. Sci.* 470 (0), 30 –39.

[66] Knevelman, C., Davies, J., Allen, L., Titchener-Hooker, N. J., 2009. High-throughput screening techniques for rapid PEG-based precipitation of IgG4 mAb from clarified cell culture supernatant. *Biotechnol. Progr.* 26 (3), 697–705.

[67] Konstantinidis, S., Chhatre, S., Velayudhan, A., Heldin, E., Titchener-Hooker, N., 2012. The hybrid experimental simplex algorithm – an alternative method for 'sweet spot' identification in early bioprocess development: case studies in ion exchange chromatography. *Analytica Chimica Acta* 743, 19–32.

[68] Łacki, K. M., 2012. High-throughput process development of chromatography steps: Advantages and limitations of different formats used. *Biotechnol. J.* 7 (10), 1192–1202.

[69] Łacki, K. M., 2014. High throughput process development in biomanufacturing. *Curr. Opin. Chem. Eng.* 6, 25–32.

[70] Laitinen, O. H., Nordlund, H. R., Hytönen, V. P., Kulomaa, M. S., 2007. Brave new (strept)avidins in biotechnology. *Trends Biotechnol.* 25 (6), 269–77.

[71] Landy, M., Dicken, D. M., Bicking, M. M., Mitchell, W. R., 1942. Use of Avidin in Studies on Biotin Requirement of Microorganisms. *Exp. Biol. Med.* 49 (3), 441–444.

[72] Lang, K. M., Kittelmann, J., Dürr, C., Osberghaus, A., Hubbuch, J., 2015. A comprehensive molecular dynamics approach to protein retention modeling in ion exchange chromatography. *J. Chrom. A*.

[73] Latour, R. A., 2008. Molecular simulation of protein-surface interactions: Benefits, problems, solutions, and future directions. *Biointerphases* 3 (3), 1–12.

[74] Leader, B., Baca, Q. J., Golan, D. E., 2008. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discover* 7, 21–39.

[75] Li, M., 2015. Bioanalytical laboratory automation development: why should we and how could we collaborate? *Bioanalysis* 7 (2), 153–155.

[76] Li-Chan, E., Nakai, S., 1989. Biochemical basis for the properties of egg white. *CRC Critical Rev. Poultry Biol.* 2, 21–58.

[77] Lienqueo, M. E., Mahn, A., Asenjo, J. A., 2002. Mathematical correlations for predicting protein retention times in hydrophobic interaction chromatography. *J. Chrom. A* 978, 71–79.

[78] Livnah, O., Bayert, E. A., Wilchek, M., Sussman, J. L., 1993. Three-dimensional structures of avidin and the avidin-biotin complex. *Proc Natl Acad Sci USA* 90, 5076–5080.

[79] Lye, G. J., Ayazi-Shamlou, P., Baganz, F., Dalby, P. A., Woodley, J. M., 2003. Accelerated design of bioconversion processes using automated microscale processing techniques. *Trends Biotechnol.* 21 (1), 29–37.

[80] McCormick, D. B., 1965. Specific Purification of avidin by column chromatography on biotin-cellulose. *Anal. Biochem.* 13, 194–198.

[81] McMahon, R. J. (Ed.), 2008. *Avidin-biotin interactions: methods and applications*, 1st Edition. Humana, Totowa, NJ.

[82] Melamed, M. D., Green, N. M., 1963. Avidin 2. Purification and composition. *Biochem. J.* 89 (1959), 591–599.

[83] Mestre-Ferrandiz, J., Sussex, J., Towse, A., 2012. The R&D cost of a new medicine. Tech. rep., Office of Health Economics, London.
URL `http://www.ohe.org`

[84] Mollerup, J. M., 2007. The thermodynamic principles of ligand binding in chromatography and biology. *J. Biotechnol.* 132 (2), 187–95.

[85] Moors, E. H. M., Cohen, A. F., Schellekens, H., 2014. Towards a sustainable system of drug development. *Drug Discovery Today* 19 (11), 1–10.

[86] Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., Greyson, D., 2011. The cost of drug development: a systematic review. *Health Policy* 100 (1), 4–17.

[87] Nau, F., Guérin-Dubiard, C., Croguennec, M., 2007. Avidin-biotin technology. In: Huopalahti, R., López-Fandiño, R., Anton, M., Schade, R. (Eds.), *Bioactive Egg Compounds*. Springer, Berlin Heidelberg, pp. 287–292.

[88] Nfor, B. K., Hylkema, N. N., Wiedhaup, K. R., Verhaert, P. D. E. M., van der Wielen, L. A. M., Ottens, M., 2011. High-throughput protein precipitation and hydrophobic interaction chromatography: salt effects and thermodynamic interrelation. *J. Chrom. A* 1218 (49), 8958–73.

[89] Nfor, B. K., Noverraz, M., Chilamkurthi, S., Verhaert, P. D. E. M., van der Wielen, L. A. M., Ottens, M., 2010. High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *J. Chrom. A* 1217 (44), 6829–50.

[90] Nfor, B. K., Verhaert, P. D. E. M., van der Wielen, L. A. M., Hubbuch, J., Ottens, M., 2009. Rational and systematic protein purification process development: the next generation. *Trends Biotechnol.* 27 (12), 673–679.

[91] Oelmeier, S. A., 2012. Aqueous two phase extraction of proteins: From molecular understanding to process development. Ph.d., Karlsruhe Institute of Technology (KIT).

[92] Oelmeier, S. A., Dismer, F., Hubbuch, J., 2011. Application of an aqueous two-phase systems high-throughput screening method tBo evaluate mAb HCP separation. *Biotechnol. Bioeng.* 108 (1), 69–81.

[93] Oelmeier, S. A., Ladd-Effio, C., Hubbuch, J., 2012. High throughput screening based selection of phases for aqueous two-phase system-centrifugal partitioning chromatography of monoclonal antibodies. *J. Chrom. A* 1252 (2010).

[94] Osberghaus, A., Drechsel, K., Hansen, S., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012. Model-integrated process development demonstrated on the optimization of a robotic cation exchange step. *Chem. Eng. Sci.* 76, 129–139.

[95] Pina, A. S., Lowe, C. R., Roque, A. C., 2014. Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnol. Adv.* 32 (2), 366–81.

[96] Piskarev, V. E., Shuster, A. M., Gabibov, A. G., Rabinkov, A. G., 1990. A novel preparative method for the isolation of avidin and riboflavin-binding glycoprotein from chicken egg-white by the use of high-performance liquid chromatography. *Biochem. J.* 265 (1), 301–304.

[97] Rao, M., Gupta, M., Roy, I., . Process for the isolation and purification of the glycoprotein avidin. *European Patent Office* EP 1 505 882 B1, Filed Oct 14 2003.

[98] Rayat, A. C., Lye, G. J., Micheletti, M., 2014. A novel microscale crossflow device for the rapid evaluation of microfiltration processes. *J. Membr. Sci.* 452, 284–293.

[99] Rege, K., Heng, M., 2010. Miniaturized parallel screens to identify chromatographic steps required for recombinant protein purification. *Nat. Protoc.* 5 (3), 408–17.

[100] Rege, K., Ladiwala, A., Cramer, S. M., 2005. Multidimensional high-throughput screening of displacers. *Anal. Chem.* 77 (21), 6818–6827.

[101] Rege, K., Tugcu, N., Cramer, S. M., 2003. Predicting Column Performance in Displacement Chromatography from High Throughput Screening Batch Experiments. *Sep. Sci. Technol.* 38 (7), 1499–1517.

[102] Rey, G., Wendeler, M. W., 2012. Full automation and validation of a flexible ELISA platform for host cell protein and protein A impurity detection in biopharmaceuticals. *J. Pharm. Biomed. Anal.* 70, 580–586.

[103] Rothemund, D. L., Thomas, T. M., Rylatt, D. B., 2002. Purification of the basic protein avidin using Gradiflow technology. *Protein Expression Purif.* 26 (1), 149–52.

[104] Safarik, I., Safarikova, M., 2004. Magnetic techniques for the isolation and purification of proteins and peptides. *Biomagnetic Research Technol.* 2 (1), 7.

[105] Sakahara, H., Saga, T., 1999. Avidin-biotin system for delivery of diagnostic agents. *Adv. Drug Delivery Rev.* 37, 89–101.

[106] Scannell, J. W., Blanckley, A., Boldon, H., Warrington, B., 2012. Diagnosing the decline in pharmaceutical R D efficiency. *Nat. Rev. Drug Discovery* 11 (3), 191–200.

[107] Seidel-Morgenstern, A., Kessler, L. C., Kaspereit, M., 2008. New developments in simulated moving bed chromatography. *Chem. Eng. Technol.* 31 (6), 826–837.

[108] Seidel-Morgenstern, A., Schmidt-Traub, H., Michel, M., Epping, A., Jupke, A., 2012. *Modeling and Model Parameters*. In: Schmidt-Traub, H., Seidel-Morgenstern, A., M, S. (Eds.), *Preparative Chromatography*, 2nd Edition. WILEY-VCH, Weinheim, Ch. 6, pp. 321–424.

[109] Shukla, A. A., Hubbard, B., Tressel, T., Guhan, S., Low, D., 2007. Downstream processing of monoclonal antibodies–application of platform approaches. *J. Chrom. B* 848 (1), 28–39.

[110] Sittampalam, G. S., Kahl, S. D., Janzen, W. P., 1997. High-throughput screening: advances in assay technologies. *Curr. Opin. Chem. Biol.* 1, 384–391.

[111] Smith, C., 2005. Striving for purity: advances in protein purification. *Nat. Methods* 2 (1), 71–77.

[112] Statista, 2014. Top-selling biotech drugs based on revenue in 2013. Accessed Nov 21, 2014. URL `http://www.statista.com/statistics/299138/top-selling-biotech-drugs-based-on-revenue/`

[113] Stevens, R. C., 2000. Design of high-throughput methods of protein production for structural biology. *Structure* 8 (9), 177–185.

[114] Stevens, R. C., 2000. High-throughput protein crystallization. *Curr. Opin. Struct. Biol.* 10 (5), 558–563.

[115] Sullivan, M., Nicholls, J., 1942. Nutritional dermatoses in the rat: V. signs and symptoms resulting from a diet containing unheated dried egg white as the source of protein. *Arch. Dermatol. Syph.* 45 (2), 295–314.

[116] Susanto, A., Knieps-Grünhagen, E., von Lieres, E., Hubbuch, J., 2008. High throughput screening for the design and optimization of chromatographic processes: Assessment of model parameter determination from high throughput compatible data. *Chem. Eng. Technol.* 31 (12), 1846–1855.

[117] Susanto, A., Treier, K., Knieps-Grünhagen, E., von Lieres, E., Hubbuch, J., 2009. High Throughput Screening for the Design and Optimization of Chromatographic Processes: Automated Optimization of Chromatographic Phase Systems. *Chem. Eng. Technol.* 32 (1), 140–154.

[118] Sutherland, I. A., 2007. Recent progress on the industrial scale-up of counter-current chromatography. *J. Chrom. A* 1151 (1–2), 6 –3.

[119] Thiemann, J., Jankowski, J., Rykl, J., Kurzawski, S., Pohl, T., Wittmann-Liebold, B., Schlüter, H., 2004. Principle and applications of the protein-purification-parameter screening system. *J. Chrom. A* 1043 (1), 73–80.

[120] Tia, S., Herr, A. E., 2009. On-chip technologies for multidimensional separations. *Lab on a Chip* 9 (17), 2524–36.

[121] To, B. C., Lenhoff, A. M., 2007. Hydrophobic interaction chromatography of proteins: ii. solution thermodynamic properties as a determinant of retention. *J. Chrom. A* 1141 (2), 235–243.

[122] Treier, K., Berg, A., Diederich, P., Lang, K., Osberghaus, A., Dismer, F., Hubbuch, J., 2012. Examination of a genetic algorithm for the application in high-throughput downstream process development. *Biotechnol. J.* 7 (10), 1203–1215.

[123] Treier, K., Hansen, S., Richter, C., Diederich, P., Hubbuch, J., Lester, P., 2012. High-throughput methods for miniaturization and automation of monoclonal antibody purification processes. *Biotechnol. Progr.* 28 (3), 723–32.

[124] Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., Bignon, C., 2004. High-throughput automated refolding screening of inclusion bodies. *Prot. Sci.* 13 (10), 2782–2792.

[125] Vlachakis, D., Bencurova, E., Papangelopoulos, N., Kossida, S., 2014. Current State-of-the-Art Molecular Dynamics Methods and Applications. *Adv. Prot. Chem. Struct. Biol.* 94, 269–313.

[126] Wahler, D., Reymond, J.-L., 2001. High-throughput screening for biocatalysts. *Curr. Opin. Biotechnol.* 12 (6), 535–544.

[127] Walsh, G., 2010. Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today* 15 (17–18), 773–80.

[128] Warner, R. C., 1954. . In: Neurath, H., Bailey, K. (Eds.), *The proteins*, vol 2 part Edition. Academic Press, New York, Ch. 17, p. 440.

[129] Wiendahl, M., Schulze Wierling, P., Nielsen, J., Fomsgaard Christensen, D., Krarup, J., Staby, A., Hubbuch, J., 2008. High Throughput Screening for the Design and Optimization of Chromatographic Processes – Miniaturization, Automation and Parallelization of Breakthrough and Elution Studies. *Chem. Eng. Technol.* 31 (6), 893–903.

[130] Wiendahl, M., Völker, C., Husemann, I., Krarup, J., Staby, A., Scholl, S., Hubbuch, J., 2009. A novel method to evaluate protein solubility using a high throughput screening approach. *Chem. Eng. Sci.* 64 (17), 3778–3788.

[131] Wilchek, M., Bayer, E. A., 1988. The Avidin-Biotin Complex in Bioanalytical Applications. *Anal. Biochem.* 171, 1–32.

[132] Wilchek, M., Bayer, E. A., 1989. Avidin-biotin technology ten years on: Has it lived up to its expectations? *Trends Biochem. Sci.* 14 (10), 408–412.

[133] Wilchek, M., Bayer, E. A., Livnah, O., 2006. Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunol. Letters* 103 (1), 27–32.

[134] Zavaleta, C. L., Phillips, W. T., Soundararajan, A., Goins, B. A., 2007. Use of avidin-biotin-liposome system for enhanced peritoneal drug delivery in an ovarian cancer model. *Int. J. Pharm.* 337 (1–2), 316–28.

2 Research Proposal

The purpose of this work is to contribute to the recent progress in the field of HTPD by demonstrating the feasibility of HTPD for the development of an alternative purification process for the high value protein avidin used for biopharmaceutical purposes. In contrast to previous works, in which single methods have been presented, the linkage of various experimental and analytical high-throughput methods is to be demonstrated on this industrial-relevant case study. Furthermore, detailed insight in the methods will generate a better understanding of the achievable data quality and limitations.

Part 1: As essential building block for beneficial HTPD, analytical methods need to be developed which enable a sample throughput that matches the speed of the experimentation without compromising analytical precision and accuracy. This conflict in particular represents the major reason for HTPD having been mostly considered as pure screening methodology for which requirements on precision are significantly lower compared to later phases in process development. Furthermore, the higher the demands on purity are in the later process steps, the more challenging the requirements on the analytical techniques become in terms of sensitivity and precision. An important aspect of this work will therefore be the development of suitable analytical techniques required for the development of an avidin purification process.

Method development of suitable assays should focus on the analysis of avidin and its four major contaminants in egg white, however, should also consider the general applicability of the methods to other protein systems, and thus, contribute to the portfolio of HTA-methods available.

Part 2: With suitable high-throughput analytical methods at hand, a purification process for avidin is to be identified by developing and applying high-throughput methods. As the highest potential for process optimization was identified downstream the initial recovery steps, development work will start from ovomucin-free, pre-purified avidin rather than from raw egg white. As an alternative to the currently used ion exchange chromatography, aqueous two-phase extraction and selective precipitation will be investigated for the intermediate purification of avidin. In particular, focus will be on the separation of lysozyme and avidin which is currently challenging when using ion exchange chromatography at preparative scale.

The development of aqueous two-phase extraction processes is commonly driven by empirical rules which makes the use of HTE particular advantageous. Data on protein distribution and solubility generated in high-throughput screenings will build the basis for further development that aims for the combination of single process steps in the most beneficial way. Optionally, the use of chromatography based on hydrophobic interaction will be considered , thus supplementing the identified alternative process in order to reach a purity of pharmaceutical grade. Methods for automated data evaluation need to be developed and included in the overall workflow. It is the intention of this case study to demonstrate the benefits from using the HTPD-toolbox in a synchronized interplay of experimenta-

tion, analytics, and data evaluation.

Part 3: In another aspect of this work, a detailed insight into one of the most complex high-throughput methods, i.e. high-throughput column chromatography (HTCC), will be generated. HTTC provides the possibility to obtain data on selectivity and binding capacity in a dynamic system, and therefore, data are expected to be very similar to the data obtained on commonly used preparative liquid chromatography systems. HTCC is used for different purposes: While some researchers apply HTCC only as screening method for resin comparison, others consider HTCC for detailed process optimization, model parameter determination and process robustness studies. However, uncertainty exists to what extent differences between data from HTCC and traditional FPLC experiments are due to scale-effects or the experimental method. In order to conclude on effects only caused by the scale differences, fundamental understanding of the influence of the robotic experimental parameters needs to be generated first. This part of the work aims furthermore to explore the limitations of HTCC in terms of data quality. Thus, it serves the judgment whether HTCC is suitable for tasks in which the requirements on precision are high, e.g. process robustness studies, model parameter calibration, or development of polishing steps.

3 Publications & Manuscripts

1. **A Sub-Two Minutes Method for Monoclonal Antibody-Aggregate Quantifcation using Parallel Interlaced Size Exclusion High Performance Liquid Chromatography**

   Patrick Diederich, Sigrid K. Hansen, Stefan A. Oelmeier, Bianca Stolzenberger, Jürgen Hubbuch

   In this paper, a method is presented enabling the quantification of monoclonal antibody aggregate content by analytical High-Performance Liquid Chromatography within only two minutes representing the maximal throughput feasible for this type of equipment and chromatographic column. The developed high-throughput method was shown to provide both high accuracy and high precision, and hence poses a superior tool for high-throughput process development for antibody purification processes.

   published in: Journal of Chromatography A. 1218 (2011): 9010–9018

2. **High-Throughput Analytical Methods for the Egg White Protein Avidin**

   Patrick Diederich, Simon Luxem, Sven Amrhein, Jürgen Hubbuch

   Three different high-throughput techniques for purity, content and activity determination of avidin were developed and characterized. Purity and content analysis comprised the analysis of the impurities ovalbumin, ovomucoid, lysozyme and ovotransferrin. Methods for analysis via LabChip®(GXII, Caliper LS), Reversed Phase HPLC method as well as photometric analysis in 96-well plate format were developed. Using the avidin protein system as model, the benefits and limitations of these techniques were demonstrated resulting in general conclusions for their use in HTPD.

   Parts of this manuscript were presented at the 1st International Conference on High-Throughput Process Development, Kraków, Poland, Oct 4-7, 2010

3. **Evaluation of PEG/Phosphate Aqueous Two-Phase Systems for the Purification of the Chicken Egg White Protein avidin by using High-Throughput Techniques**

   Patrick Diederich, Sven Amrhein, Frank Hämmerling, Jürgen Hubbuch

   This paper presents distribution characteristics for avidin and its four major contaminants in PEG/phosphate aqueous two-phase systems. Both extraction experiments and analytics were performed in high-throughput mode. Conditions were identified for an effective distribution of lysozyme, ovomucoid, and ovotransferrin in the upper phase and the distribution of avidin and

the remaining impurity ovalbumin into the bottom phase.

published in: Chemical Engineering Science. 104 (2013): 945–956

4. **High-Throughput Process Development of Purification Alternatives for the valuable Protein Avidin**

Patrick Diederich, Marc Hoffmann, Jürgen Hubbuch

Purification process alternatives to the commonly used ion-exchange chromatography techniques were identified and evaluated by using the HTPD-methodology. From resulting high-throughput data, a process layout is proposed consisting of selective precipitation with polyethylene glycol, re-solubilisation and extraction in an PEG/salt/salt aqueous two-phase system, and seamless processing of the bottom phase via mixed-mode chromatography.

published in: Biotechnology Progress. 31 (2015): 957-973

5. **High-Throughput Column Chromatography performed on Liquid Handling Stations – Process Characterization and Error Analysis**

Patrick Diederich, Jürgen Hubbuch

A detailed investigation of method-specific parameters and error sources in high-throughput column chromatography (HTCC) is presented. Influence of the parameters are evaluated both qualtitatively and quantitatively by using experimental data and chromatography simulations providing a better understanding of the observed experimental variances.

submitted for publication in: Preparative Chromatography for Separation of Proteins, edited by A. Staby, S. Ahuja, and A. Rathore, Wiley & Sons, 2016

MANUSCRIPT 1

A Sub-Two Minutes Method for MAb-Aggregate Quantification using Parallel Interlaced Size Exclusion HPLC

Patrick Diederich[+], Sigrid K. Hansen[+], Stefan A. Oelmeier[+], Bianca Stolzenberger, Jürgen Hubbuch

[+] *These authors contributed equally to this publication*
Institute of Engineering in Life Sciences, Section IV:
Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

published in: *Journal of Chromatography A, 1218 (2011), 9010–9018*

Abstract

In process development and during commercial production of monoclonal antibodies (mAb) the monitoring of aggregate levels is obligatory. The standard assay for mAb aggregate quantification is based on size exclusion chromatography (SEC) performed on a HPLC system. Advantages hereof are high precision and simplicity, however, standard SEC methodology is very time consuming. With an average throughput of usually two samples per hour, it neither fits to high throughput process development (HTPD), nor is it applicable for purification process monitoring. We present a comparison of three different SEC columns for mAb-aggregate quantifi cation addressing throughput, resolution, and reproducibility. A short column (150 mm) with sub-two micron particles was shown to generate high resolution ($\sim$ 1.5) and precision (relative standard deviation (RSD) $<$ 1) with an assay time below six minutes. This column type was then used to combine interlaced sample injections with parallelization of two columns aiming for an absolute minimal assay time. By doing so, both lag times before and after the peaks of interest were successfully eliminated resulting in an assay time below two minutes. It was demonstrated that determined aggregate levels and precision of the throughput optimized SEC assay were equal to those of a single injection based assay. Hence, the presented methodology of paralell interlaced SEC (PI-SEC) represents a valuable tool addressing HTPD and process monitoring.

Keywords: Monoclonal antibody, aggregates, Size Exclusion Chromatography, high throughput analytics, interlaced injection, PI-SEC

1 Introduction

Aggregate levels in monoclonal antibody drugs are a critical quality attribute due to their potential immunogenicity [1, 2]. Aggregates of monoclonal antibodies are often the most abundant product related impurity. The purification process needs to ensure that aggregate levels are reduced to an acceptable level in the final drug product. While the first two steps in a standard mAb downstream process are readily capable of depleting three highly abundant process related impurities, host cell protein, DNA, and water, the reduction of aggregate levels to acceptable levels is often challenging. Thus, monitoring aggregate levels is critical in process development.

One way to reduce process development costs is to increase development throughput. Various process steps have been scaled down to fit into a high throughput process development (HTPD) scheme [3, 4, 5, 6]. Additionally, platform processes have been implemented for monoclonal antibody based products, further reducing the efforts needed from process development down to process verification [7]. These improvements have created an analytical bottleneck in process development. To match throughput of the experimentation, reasonably short analysis times need to be achieved.

Size exclusion chromatography (SEC) is the standard method for mAb-aggregate analysis. The standard SEC assay with a throughput of two samples per hour [8, 9] does however not suit a HTPD approach. Several measures are thus in the spotlight to increase throughput in HPLC without changing the analytical technique as such: parallelization and interlacing sample injection. While parallelization using multiple HPLC stations is currently the most often used approach, it is for obvious reasons also the most expensive. Parallelization of multiple columns on a single detector via column switching valves is a way to reduce parallelization cost and has been successfully demonstrated [10]. Most often in this approach, the elution and the regeneration of a chromatographic analysis are separated such that one column regenerates while the other column performs an analysis [11]. In contrast to gradient elution, column regeneration is however not necessary in SEC. Another approach to improve throughput is to run a single column in an interlaced mode. In interlaced chromatography a sample is injected onto the column before the preceding analysis has been completed. This approach requires isocratic conditions. Farnan *et. al* [12] successfully demonstrated its use for aggregate analysis of mAbs and were able to reduce assay time per sample by more than a factor of two from 30 minutes to 14 minutes.

Finally, HPLC equipment capable of higher back pressures has been implemented (most often termed UltraHPLC) [13]. Shorter columns with smaller column volume and smaller particle sizes can be used with this equipment, thus reducing assay time without sacrificing resolution. While one of the most often used columns for mAb-aggregate analysis has a pressure limit of 7.2 MPa (Tosoh TSKgel ®3000 SWxl), two new SEC columns suitable for higher back pressures of 24.1 MPa (Zenix ™SEC-250 (Sepax Technologies)) to 41.4 MPa (ACQUITY UPLC ®BEH200 SEC (Waters Corporation)) recently became commercially available.

In this paper, we compare mAb-aggregate analysis performed on these three SEC columns. The columns are compared in terms of assay throughput, resolution, and precision. We demonstrate the application of ACQUITY UPLC ®BEH200 SEC columns (Waters Corporation) in an interlaced mode as well as by interlaced injections on two columns run in parallel. We demonstrate how throughput can be increased by a factor of 10 – 15 compared to a standard analysis using a TSKgel ®3000 SWxl column. Advantages and disadvantages of the methodology are discussed.

1.1 Theory - Increasing Throughput by Interlacing and Parallelization

While the presented methodology can be applied universally to any type of SEC-column, differences arise in the use of (Ultra)HPLC equipment and the actual pressure rating of the respective SEC-columns and adsorbents. To implement the method developed in this study to its full potential, a prerequisite lies in the use of an (Ultra)HPLC system which is equipped with two independent flow switching valves. An *inlet valve* directs the flow to the columns and autosampler and an *outlet valve* directs the flow from the column outlets to the detector and waste. For maximum throughput two SEC columns can thus be run in parallel applying interlaced injections on each of the two identical columns. The idea of parallel interlaced (PI-) SEC methodology is to eliminate every region of a chromatogram which is not providing any relevant data (e.g. antibody aggregate and monomer). In a first step, data of a single chromatographic SEC analysis therefore serve as a benchmark for the estimation of analysis time and method development as described in the following:

Single Injection

In Figure 1 A and 2 A typical chromatograms of common mAb SEC analysis are displayed. The chromatograms can be divided into four main phases. The first phase after sample injection is the initial lag phase (t_{lag}). The time span in which aggregate species and monomer elute is referred to as information phase (t_{inf}). In this work, protein fragments are not considered as species of interest and are not included in t_{inf}. The third phase between monomer peak and the eluting salt fraction is referred to as hold phase (t_{hold}). It is assumed that no protein elute later than the salt fraction of the injected sample. The elution region of salt species is refered to as t_{salt}.

An single chromatogram of the sample material provides the user with the retention times of every elution phase for the column used at the specific flow rate. The total time required for the analysis of n samples can be stated as:

$$t_{total} = n \cdot (t_{lag} + t_{inf} + t_{hold} + t_{salt}) \tag{1}$$

Given these retention times, the first step to increase analysis throughput is to eliminate t_{lag} from the resulting chromatograms as explained below.

Interlaced Injection

Farnan *et al.* [12] has described the methodology of interlaced SEC in detail. In a brief, the methodology is based on injecting a subsequent sample before the ongoing analysis of a sample has completed. The subsequent information phase begins immediately after the salt fraction of the preceding sample has eluted. Figure 3 A and B show the transition from a mode of single injection to interlaced injection. By the use of a second timebase (see section 2.2), a separate control program for data aquisition ("program DAD") facilitates distinct chromatograms for each injection and corresponding sample. In Figure 3 B it is demonstrated that the lag phase can thus be eliminated from analysis. The total time required for the analysis of n samples can be stated as:

$$t_{total} = t_{lag} + n \cdot (t_{inf} + t_{hold} + t_{salt}) \tag{2}$$

Parallel Interlaced Injection

A further increase in throughput can be achieved when applying interlaced injections on two columns which are operated in parallel. Starting from interlaced chromatogaphy, in parallel interlaced SEC the assay time is further reduced by t_{hold}, as is demonstrated in Figure 3 B and C. Two switching valves are used to direct the flow alternately between autosampler, two columns and the detector, thus enabling the elimination of t_{lag}, t_{hold} and t_{salt}. In Figure 3 D a scheme of the valve switching is displayed. The use of two columns and switching valves require two distinct programs assigned to *timebase 1*, on which pumps, autosampler and column compartment including the switching valves are controlled. The programs contain the same commands, but differ in the direction of both valves switching. As for interlaced chromatography, data aquisition is performed separately by using a second timebase (*timebase 2*) for the detector, now only recording phase t_{inf} of each injected sample.

For programming PI-SEC, three possible cases need to be considered, since elution profiles of a single injection analysis differ in t_{lag}, t_{hold} and t_{salt} depending on column type and sample material. For reason of simplicity, it is assumed that $t_{hold} > t_{salt}$, which is the common case in SEC analysis of antibody samples.

Case-1. $t_{inf} > t_{hold}$:

The first sample is injected on column 1 at:

$$t_1 = 0 \tag{3}$$

The second sample is injected on column 2 at:

$$t_2 = t_1 + t_{inf} \tag{4}$$

The subsequent samples are alternately injected on column 1 and column 2 at times:

$$t_{n,inj} = t_{n,inj-1} + t_{inf} \tag{5}$$

The total assay time for the analysis of n samples can hence be calculated by equation 6. This equation gives the theoretically possible increase in throughput which can be gained via PI-SEC using one single detector.

$$t_{total} = t_{lag} + n \cdot (t_{inf}) + t_{hold} + t_{salt} \tag{6}$$

The *outlet valve* is switched as soon as the information phase of a sample from one column has passed the detector. At that time, the salt peak has completely eluted from the other column. Samples are alternately injected on the two columns and analyzed without any interference of eluting salt fractions. As an example, Figure 1 shows a schematic drawing of PI-SEC methodology for the case of $t_{hold} < t_{inf}$.

Case-2. $k \cdot t_{inf} < t_{hold}$:

If $k \geq 1$, one or more informational phases fit into t_{hold} and k additional injections (rounding down of k to whole numbers) on one column become feasible before switching to the second

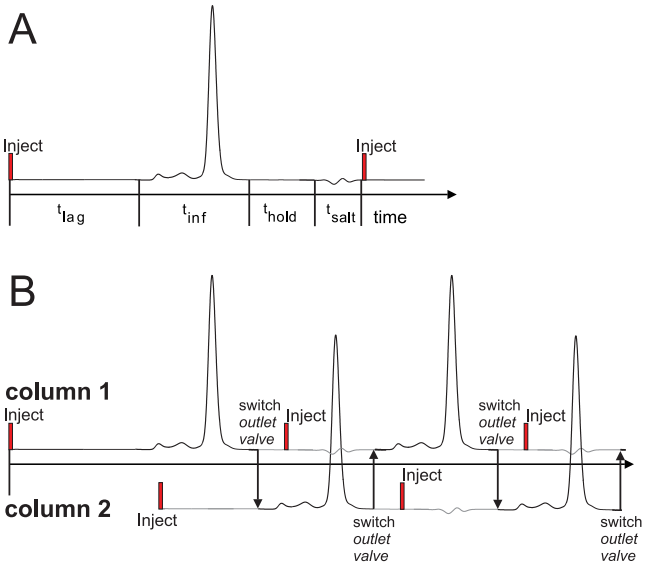

Figure 1: Schematic of PI-SEC methodology applicable in the case of $t_{inf} > t_{hold}$. A: chromatogram of a mAb sample analyzed in single injection mode. Using the elution phases $t_{lag}, t_{inf}, t_{hold}$, and t_{salt}, a PI-SEC program can be set up (B). In this case, samples are injected alternately on two columns, while the *outlet valve* directs the flow from the column outlet to the detector.

column. The injection times and the time needed for the analysis of n samples can be estimated using the same equations 3 - 6 as given in case one. Figure 2 shows a schematic drawing of the PI-SEC methodology applied for a case 2 elution profile where $1 < k < 2$. Now, two salt peaks elute from one column within the time two information phases elute from the other column.

Although time benefit is the same as in case one, it should be noted that in this mode proteins of multiple, subsequently injected samples pass the salt fraction of the preceding injected samples, whereas for case one the salt fraction of each sample always elute earlier from the column than does the information phase. Multiple injections on one column is further only applicable, if no species of lower molecular weight than the monomer species is present in the sample material. Otherwise the species of lower molecular weight will elute within the information phase of the subsequent sample injected on the same column.

In the case that $k < 1$ and the *outlet valve* is switched instantly after the information phase of a sample from one column has passed the detector, the salt fraction of the preceding sample has not eluted yet from the second column. Therefore, some additional time (t_{add}) must be added before switching the *outlet valve*. The sum of $t_{add} + t_{inf}$ needs to be greater than $t_{hold} + t_{salt}$. The time needed for the analysis of n samples can be estimated using equation 6, while including t_{add} (9). This delay needs also to be factored in the injection times of the interlaced mode of each column. When the first injection at t_1 is performed, the second injection takes place at:

$$t_2 = t_1 + t_{inf} + t_{add} \tag{7}$$

The injection time of sample n can be hence given by:

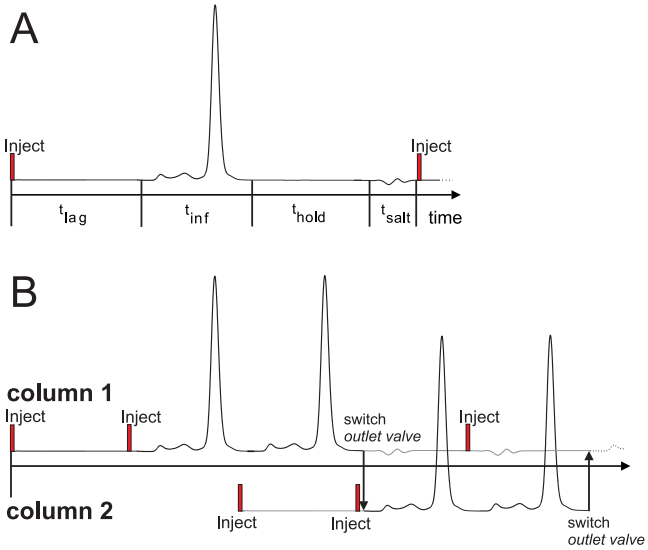

Figure 2: Schematic of PI-SEC methodology applicable in the case of $t_{inf} < t_{hold}$. A: chromatogram of a mAb sample analyzed in single injection mode. Using the elution phases $t_{lag}, t_{inf}, t_{hold}$, and t_{salt}, a PI-SEC program can be set up (B). In this case, two samples are subsequently injected per column before switching to the second column.

$$t_{n,inj} = t_{n,inj-1} + t_{inf} + t_{add} \tag{8}$$

The total assay time for n samples can be calculated using:

$$t_{total} = t_{lag} + n \cdot (t_{inf} + t_{add}) + t_{hold} \tag{9}$$

From a practical aspect it should be mentioned that, if t_{inf} is slightly smaller or exactly equals the sum of $t_{hold} + t_{salt}$, the *outlet valve* is switched just when salt is detected or just arrives at the detector. The baseline determination and an autozero processing of the absorbance signal is hence affected and might lead to unprecise peak integration.

Regarding all described scenarios case one marks the optimal condition for PI-SEC since information phases of samples injected alternately on two columns neither interfere with eluting salt fractions nor are additional times required. With an increasing ratio of t_{inf}/t_{hold}, the benefit of using two columns in parallel over interlaced injection decreases. For the purpose of method robustness, in any of the above described cases additional time for switching the inlet and outlet valves should be implemented: Switching the *inlet valve* should occur a few seconds before the injection takes place and switching of the *outlet valve* should occur a few seconds before the high molecular weight species elute. Thus, baseline determination and peak integration become more precise. To set up the control program, sampling and washing times need to be taken into account. The duration of sampling and washing depends strongly on the used (Ultra)HPLC equipment and might significantly slow down the assay if it exceeds the duration of the information phase. Furthermore, differences in column packing and hence retention times need to be considered.

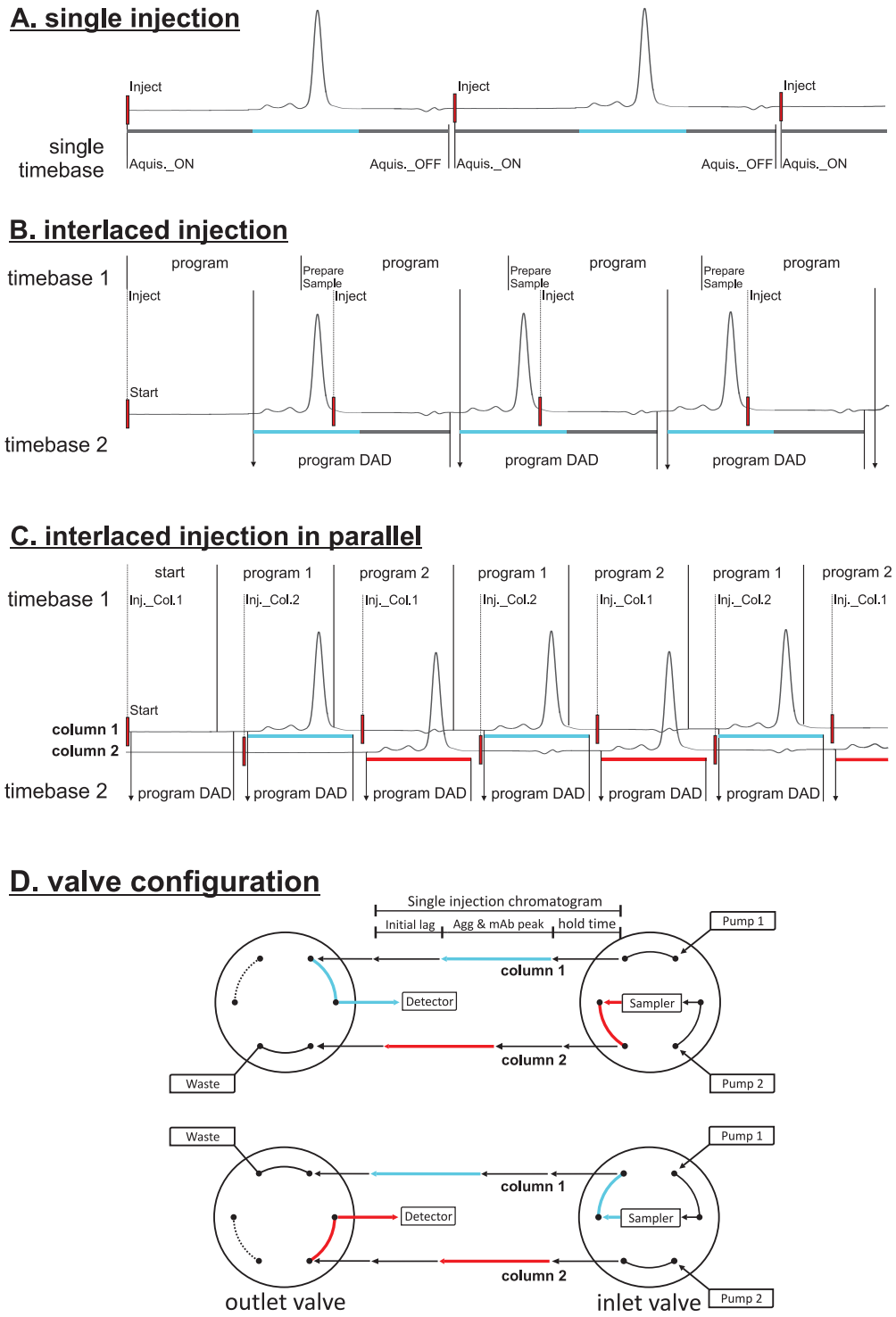

Figure 3: Three modes of operating SEC analysis are displayed. Based on single run chromatograms **(A)** throughput can be improved by interlacing sample injections **(B)** on one SEC column. By using a second timebase (*timebase 2*) for data acquisition, a dedicated chromatogram is generated for every sample injection. Each timebase is controlled by separate programs. Using a second column run in parallel and two timebases **(C)**, throughput can pushed to its theoretical maximum by performing interlaced injections on both columns. Hereby, two programs on *timebase 1* are implemented differing only in the switching direction of the switching valves. A schematic of the configuration of two six-port valves **(D)** demonstrates the switching procedure which has to be implemented in the control programs 1 and 2.

2 Materials & Methods

2.1 SEC columns

SEC columns from three vendors were used in this work: 1. TSKgel 3000 SWxl (Tosoh Corporation, Tokio, Japan) 2. ACQUITY UPLC®BEH200 SEC (Waters Corporation, Milford, MA, USA) 3. Zenix SEC-250 (Sepax Technologies, Newark, DE, USA). Columns were fitted with 0.2 μm inlet filter (Opti-Solv®EXP™, Optimize Technologis, Oregon City, OR, USA). In table I the column properties are listed. The columns differ in macroscopic as well as microscopic dimensions.

Table I: Specifications of the HPLC SEC columns used in this study.

vendor description	column dimension	pore size	particle size	maximum pressure	volume column	volume void
ACQUITY UPLC BEH200 SEC	4.6x150 mm	200 Å	1.7 μm	41.5 mPa	2.5 mL	1.97 mL
Zenix SEC-250	4.6x250 mm	300 Å	3.0 μm	24.1 mPa	4.2 mL	3.45 mL
TSKgel 3000 SWxl	7.8x30 0mm	250 Å	5.0 μm	7.8 mPa	14.3 mL	12.23 mL

2.2 UHPLC setup

An UltiMate3000 RSLC x2 Dual system from Dionex (Sunnyvale, CA, USA) was used for UHPLC analysis. The system was composed of two HPG-3400RS pumps, a WPS-3000TFC-analytical autosampler and a DAD3000RS detector. The autosampler was equipped with a sample loop of 5 μl or 20 μl, respectively. The volume of the injection needle was 15 μl, the syringe size was 250 μl. In all experiments, full loop injections were performed. The system included a TCC-3000RS column thermostat to enclose two columns, which were connected to two six-port column switching valves. The *inlet valve* directs the flow between autosampler outlet and column inlets, hence controlling to which column a sample is injected. The *outlet valve* directs the flow between column outlets and UV-detector, hence controlling from which column outlet the UV signal is measured. All column experiments were conducted at 25 °C. For SEC analysis performed in interlaced and parallel-interlaced mode, the system was split in two virtual parts by using two separate timebases. *Timebase 1* controlled pumps, autosampler, valves and column compartment and *Timebase 2* controlled the UV detector. The two timebases were physically linked by connecting a relay assigned to *timebase 1* with an input assigned to *timebase 2*. Switching of the relay in *timebase 1* triggered an input signal in *timebase 2*. This input signal was then used to trigger the UV signal acquisition. By this setup, it was possible to record the information phase of each sample separately.

2.3 Software

Matlab2010a (The Mathworks Natick, ME, USA) was used for data analysis. Chromeleon®(6.80 SR10) was used to control the UHPLC equipment and to integrate the elution peaks in the chromatograms. The Chromeleon software was extended to include two timebases.

2.4 Buffer and Sample

SEC analysis were performed using a 0.2 M potassium phosphate buffer at pH 6.2 containing 0.25 M potassium chloride. Buffers were filtered through 0.2 μm filters (Sartorius, Germany) prior to use. When two pumps were used simultaneously (parallel-interlaced protocol), the same buffer preparation was apportioned in two bottles. A proteinA pool of a CHO expressed IgG was used as mAb sample. The concentration was set to a concentration of 1 g/L by dilution with dH2O.

2.5 Aggregate Level and Chromatographic Resolution

For each single injection run, the aggregate level and the resolution was determined. For all interlaced and parallel-interlaced runs only the aggregate level was determined. The aggregate level was defined as the percentage of the species in the mAb sample eluting prior to the monomer. The achieved chromatographic resolution of the mAb monomer and the smallest aggregate (dimer) was calculated based on the EP norm:

$$R = 1.18 \cdot \frac{t_{monomer} - t_{dimer}}{W_{50\%,monomer} + W_{50\%,dimer}} \tag{10}$$

2.6 Single Injection SEC Protocols

The TSKgel column was loaded with 20 μL of sample and the analysis was run at flow rates between 0.235 mL/min and 1.5 mL/min (30 – 188 cm/h). The ACQUITY column was loaded with 5 μL of sample and run at flow rates between 0.05 mL/min and 0.5 mL/min (18 – 181 cm/h). The Zenix column was loaded with 5 μL of sample and run at flow rates between 0.05 – 0.96 mL/min respectively (18 – 347 cm/h). The exact flow rates are listed in table II.

2.7 Interlaced SEC protocol

For interlaced SEC experiments the chromatography system was split in two virtual parts as described in section 2.2. It should be noted, that this is not a necessary prerequisite in interlaced chromatography, but rather a convenience for the experimenter. By splitting the instrument and running dedicated programs for UV signal acquisition, the relation of chromatogram and injected sample is facilitated. The methodology described in section 1.1 was applied to the use of ACQUITY columns. A single chromatographic run at a flow rate of 0.4 ml/min was used to determine the initial lag phase (t_{lag}) (see Figure 1 A).

In the adapted method, the data acquisition program on *timebase 2* was triggered by switching a relay on *timebase 1* at $t = t_{lag}$ after injection. The withdrawal of the subsequent sample (pulled-loop mechanism) was triggered 51 seconds prior to injection by using the "PrepareNextSample"- command. This avoided additional hold phases between subsequent control programs.

2.8 Parallel-Interlaced SEC Protocol

To improve throughput further, a second column was run in parallel to the first column using two switching valves directing the flow to the columns and to the detector, respectively. The eluate of one column was directed to the waste right after the monomer peak has passed the detector. The eluate of the second column was then directed to the detector, while the salt peak eluted from the first column into the waste. By running both columns simultaneously in an interlaced mode, the maximum

possible throughput of the system was realized (section 1.1). In this work, two ACQUITY columns were used at a flow rate of 0.4 mL/min. The time for sample withdrawal was adjusted to 27 seconds (pulled-loop mechanism). Thoroughly washing of the sample loop and the injection needle was set to be performed within 90 seconds.

2.9 Aggregate Spiking Studies

Aggregate spiking studies were conducted in order to evaluate the linearity of aggregate determination of the presented parallel-interlaced methodology. Two solutions containing different levels of aggregate were mixed to control the level of aggregate in the samples. In order to obtain a solution with a high aggregate content, aggregate was isolated from the proteinA pool. This was done by loading the mAb sample onto a Poros 50 HS (GE Healthcare, Germany) column. Before loading the column, the mAb sample had been adjusted to a conductivity of 15 mS/cm and a pH of 5.5. These conditions had been found to provide high selectivity for mAb aggregates compared to mAb monomer. The elution was performed with a sodium chloride gradient from 10 – 150 mM in 20 mM MES buffer at pH 5.5. The eluate was collected in fractions, analyzed by SEC and merged to create an aggregate pool with approximately 50% aggregate. Seventeen aggregate levels were tested ranging from 2.1 to 48.7%. The samples were first analyzed on two different ACQUITY columns in single injection mode, where each sample was measured sixfold. Subsequently, the presented parallel-interlaced assay was applied, using the same two columns and the same samples which were measured sixfold each. The results were compared in terms of coincidence of the linear regression between expected aggregate level and aggregate level determined via the different approaches.

3 Results & Discussion

SEC columns from three different vendors with different particle size, pore size, and length were applied for mAb aggregate quantification. In contrast to the TSKgel column, the ACQUITY and the Zenix columns have entered the market recently. The TSKgel column has been on the market for almost 25 years and a literature survey revealed a marked preference for this particular column in relation with mAb analysis (data not shown). The chosen columns were compared in terms of generated chromatographic resolution, throughput and precision of aggregate quantification. Based on the results, the best suited column and flow rate was chosen and used to establish a in throughput optimized assay by combining interlaced injections with parallel operation of two SEC columns.

3.1 Single Injections

Three different columns were used to analyze identical mAb samples. Figure 4a shows all three resulting chromatograms. The applied flow rates were 108 cm/h for the ACQUITY, 116 cm/h for the Zenix column, and 126 cm/h for the TSKgel column. For comparability, the chromatograms were normalized with respect to void volume of the respective column (Figure 4b). The void volume of each column was defined as the elution volume of the sample buffer. These are listed in table I.

The normalized chromatograms revealed similar elution patters for all columns in which the mAb species eluted over a range from approximately 0.45 to 0.85 void volumes. The elution order, based on normalized elution volume of the monomer species from the three different columns ($V_{ACQUITY} <$

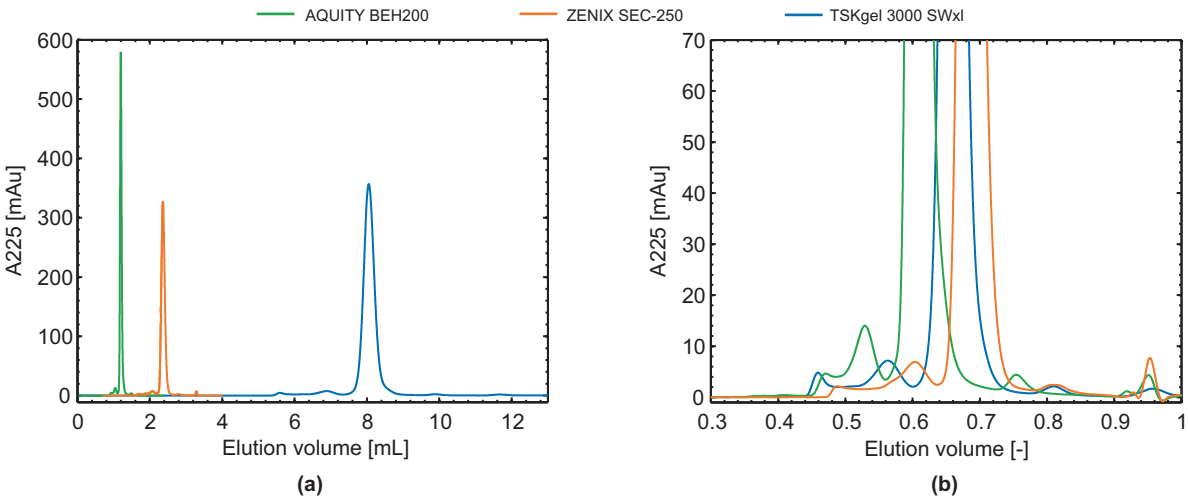

Figure 4: a: Overlay of single injection chromatograms of the mAb sample (1.0 g/L) analyzed on three different SEC columns. b: For comparability, elution volumes were normalized to column void volumes.

$V_{TSKgel} < V_{Zenix}$) correlated with the decreasing pore size of the column matarial (ACQUITY: 200 μm, TSKgel: 250 μm, Zenix: 300 μm). The elution profiles generated by the Zenix and the TSKgel column exhibited a more widely stretched elution of the aggregate species. At very low flow rates, these two columns also revealed a third aggregate species in the mAb sample which eluted in between the two main aggregate species (data not shown). However, if an analytical assay aims for the total aggregate level, a resolution of single aggregate species is not necessary. In such a case, the most important parameter is the resolution of the smallest mAb aggregate species (dimer) and the mAb monomer. Hence, in the following the term resolution will refer only to the resolution of mAb monomer and dimer species.

3.1.1 Aggregate Levels and Precision

The determined resolution, aggregate level, and RSD for each applied flow rate and column are listed in table II. The columns were shown to generate different results regarding aggregate level, even though the same mAb sample was analyzed. Using the TSK column, the highest and most stable aggregate level (4.80% ± 0.08) over the tested range of flow rates was determined. Using the ACQUITY column, a lower mean aggregate level was determined (4.17% ± 0.44) and further the determined aggregate levels exhibited an increase with increasing flow rate (3.79% – 5.02 %). However, the precision resulting from each tested flow rate was comparable to the accuracy obtained with the TSKgel column ($RSD_{mean,TSKgel}$ = 0.91, $RSD_{mean,ACQUITY}$ = 0.87). The overall aggregate level determined using the Zenix column (4.20% ± 0.35) was similar to the one obtained with the ACQUITY column, however the accuracy of the results was lower compared to both other columns ($RSD_{mean,Zenix}$ = 1.38). As for the ACQUITY column, the aggregate level determined with the Zenix column exhibited an increase with increasing flow rate (3.69% – 4.54%). For all columns, a tendency of higher precision at medium flow rates was observed.

Table II: Aggregate levels determined for a mAb sample using three different columns. Each column was operated at several different flow rates. All displayed results are based on six replicates.

TSKgel® 3000 SWxl				
Flow rate (cm/h)	Flow rate (mL/min)	Aggregate (%)	RSD (%)	resolution (-)
30	0.235	4.74	1.91	1.85
44	0.352	4.87	0.85	1.77
63	0.50	4.87	0.60	1.71
94	0.75	4.84	0.27	1.59
126	1.00	4.83	0.52	1.50
157	1.25	4.79	0.48	1.41
188	1.50	4.64	1.75	1.34
ACQUITY UPLC® BEH200 SEC				
Flow rate (cm/h)	Flow rate (mL/min)	Aggregate %	RSD %	resolution (-)
18	0.05	3.79	1.94	1.66
27	0.075	3.90	1.00	1.60
36	0.10	3.90	0.99	1.61
72	0.20	4.00	0.48	1.56
108	0.30	4.16	0.27	1.52
144	0.40	4.36	0.94	1.47
181	0.50	5.07	1.52	1.45
Zenix™ SEC-250				
Flow rate (cm/h)	Flow rate (mL/min)	Aggregate %	RSD %	resolution (-)
18	0.05	3.69	2.33	1.35
27	0.075	3.96	0.97	1.33
36	0.10	4.11	0.58	1.30
116	0.32	4.28	0.97	1.14
231	0.64	4.62	1.53	1.01
347	0.96	4.54	1.91	0.92

3.1.2 Resolution vs. Analysis Time

The main objective of the presented work, was to establish an ultra-rapid SEC assay for mAb aggregate quantification. Due to the different column dimensions, the correlation between resolution and flow rate does not transmit directly to analysis time. To give an overview of the direct relation between analysis time and chromatographic resolution, the resolution generated for each flow rate and column was plotted as a function of the required time per analysis (Figure 5). The evaluation was performed in sequential mode, thus time per analysis equals time needed for processing a single column volume (CV). In generel, the decrease in resolution correlated with the particle size of the column material. We found that at assay times above 20 min, the TSKgel column achieved the highest resolution of the columns tested. The resolution achieved under these conditions ranged from 1.59 to 1.85. However, in most cases, a resolution of 1.5 is sufficient for precise quantification. Hence, the high resolution achieved by the TSKgel column at the lower end of the tested flow rates will in some cases be disadvantageous as an unnecessary low throughput is the consequence of the achieved yet dispensable resolution. At lower assay times (increased flow rates) the resolution achieved with the TSK column was shown to decrease faster compared to the ACQUITY column. Of all columns, the ACQUITY column was shown to generate the highest resolution at assay times below 20 min. This finding correlates with the smaller particle size of the ACQUITY column. The tested Zenix column was outperformed by the TSKgel and ACQUITY columns with respect to resolution at all tested assay times. One advantage of the Zenix column was the potentially lower assay time, but the low resolution under these conditions were shown to generate imprecise results (see table II) . However, assay times

down to 13 min generated adequate precision (RSD < 1) despite the low resolution. Hence, taking the relative low cost for the Zenix column compared to the TKSgel and the ACQUITY column into consideration (which exhibits a factor of 1:1.5:2), this column could pose a favorable alternative to the otherwise comprehensive use of the TSKgel column.

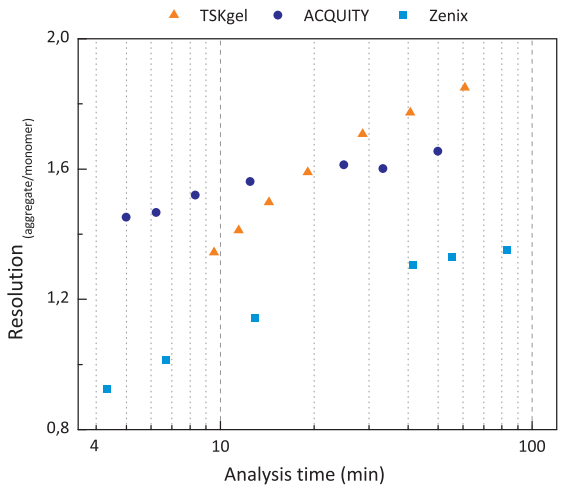

Figure 5: Achieved chromatographic resolution for each tested column displayed as a function of the analysis time. Each data point represents the mean value of six measurements.

Sufficient resolution ($\sim$1.5) and precision (RSD$<$1) was shown feasible with the ACQUITY column even at very low analysis times. This clearly favours the ACQUITY column for development of a high throughput parallel-interlaced SEC assay. A flow rate of 0.4 ml/min was chosen, both to guarantee sufficient accuracy and also not to operate the column close to maximal flow rate.

The findings presented above are based on measurements performed with only one column per column type. Hence, the conclusions do not take batch and packing variability into consideration. This influence is shown in the studies below. Further, a buffer optimization was not in the scope of this work and changes in performance under other buffer conditions can not be ruled out.

3.2 Interlaced SEC

Twenty five injections of the same load material were performed on three ACQUITY columns in interlaced mode. Average analysis time per sample was 3:27 minutes at a flow rate of 0.4 ml/min. Figure 6A shows the resulting A225 trace from the detector. It can be seen that the initial lag time was successfully cut from the analysis time. In this mode of operation 1.43 samples were analyzed per column volume. While aggregate levels resulting from all three column were in the same range and normally distributed around their mean, pairwise t-tests ($\alpha = 0.01$) showed that all results differed statistically significantly from one another. The first column resulted in a mean aggregate level of 5.08% with a standard deviation of 0.04. The second column yielded mean 5.02% with a standard deviation of 0.05. The third column yielded mean 4.91% with a standard deviation of 0.04.

By interlacing injections and switching to a column of smaller volume and particle size, the assay time was reduced from 14 minutes reported by Farnan [12] to 3:27 minutes. The obvious advantage of using interlaced injections lies in the improved throughput. However, special care has to be taken in

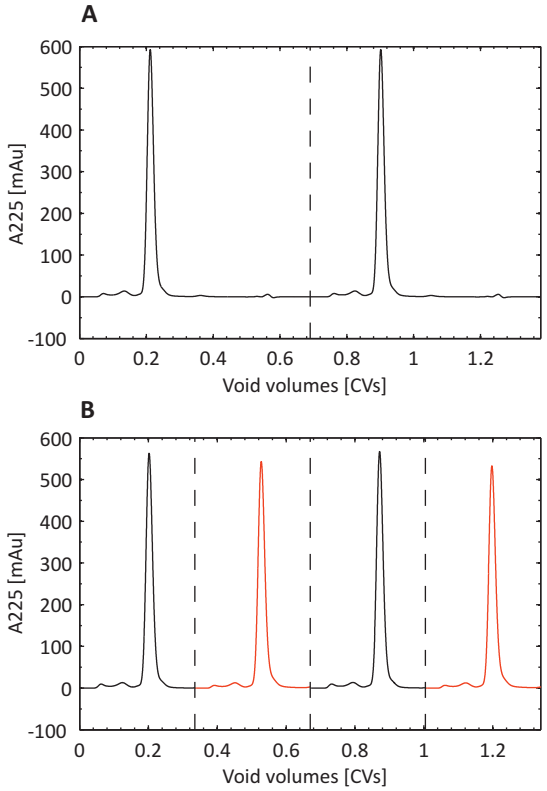

Figure 6: A: A225 absorption data of two injections run in interlaced mode on the ACQUITY column. The dashed line represents the limit between the two samples. 1.43 samples could be analyzed per CV in this mode of operation. B: A225 absorption data of four injections run in parallel-interlaced mode on two ACQUITY columns. The dashed lines represent the moments of switching the column outlet valve to the detector for the subsequent sample. Separate result files are generated for each sample as delimited by the dashed lines. Equally colored lines represent samples analyzed over the same column. Three samples were analyzed per CV in this mode of operation.

order to correctly relate sample and chromatogram. By splitting the instrument into two virtual parts (timebases) a comfortable solution to this problem can be achieved. While throughput was increased, there was still room for optimization. First, column utilization is not optimal as only the initial lag phase is eliminated by interlaced injections. Second, the next sample was not injected until 15 seconds after the salt fraction of the preceding sample had eluted.

3.3 Parallel-Interlaced SEC

Program Parameters

By parallelization of two ACQUITY columns operated with interlaced sample injections, chromatograms containing only the aggregate and monomer areas could be generated. As described in chapter 1.1, the control program was set up based on a single run at a flow rate of 0.4 ml/min. The operation commands of the Chromeleon®software and the corresponding times in the control programs of *timebase 1* and *timebase 2* are summarized in table III. t_{lag} was set to 2:00 minutes. t_{inf}^{min}, the minimal possible analysis time was 1:12 minutes. Twenty-four seconds were added to t_{inf}^{min} to make the method more robust against changes in sample composition. t_{inf} used for programming the method was thus

1:36 minutes. The determined t_{hold} was 1:18 minutes. A sequence of samples was first started with a dummy run in which the first sample is injected but no protein elutes. DAD data acquisition thus generated a blank sample. Immediately after DAD data acquisition has ended, the *outlet valve* was switched. Fifteen seconds were added to the method to ensure a stable baseline after switching the *outlet valve* (t^1_{add} = 0:15). Next, the *inlet valve* was switched. Three seconds were added to the method to flush the autosampler prior to injection (t^2_{add} = 0:03). Triggering the data acquisition was performed three seconds after the sample injection by using the following commands: after the `Inject` command triggered sample injection in *timebase 1*, a `Relay.State = ON` command switched a relay which was connected to an input via cable. A `wait Input.State = ON` as first command in the control method for *timebase 2* triggered the start of this control method and thus of DAD data acquisition as soon as relay 3 was switched. 1:27 minutes later the next sample withdrawal was started using the `PrepareThisSample` command. 0:09 minutes afterwards, DAD data acquisition was stopped thus closing one cycle of sample injection and detection. The process of sample withdrawal took 27 seconds and was performed during the last nine seconds of t_{inf} of the preceding sample and the t^1_{add} and t^2_{add} after switching the *outlet valve* and *inlet valve*.

In general, the operating speed of the autosampler was found to be an important factor when programming the control method. Slower autosampling equipment might hinder the implementation of the method. Compared to the data presented, faster autosampling procedures, for example by using a inline split-loop autosampler instead of the used pulled-loop would take the method closer to its theoretical minimum of 1:12 minutes.

To analyze a batch of samples, two batch files were created, one for each timebase. The batch file for *timebase 1* contained two different control programs with each used for every other sample. The two control programs were equal but for the valve switching commands. The batch file for *timebase 2* consisted of a sequence of the DAD control program. The two batch files were started simultaneously.

Method Performance

Fifty injections (25 on each column) of the same mAb load material were performed in parallel-interlaced mode. The analysis time for this batch was 1:57 minutes per sample. Figure 6B shows the resulting detector signal at a wavelength of 225 nm of four consecutive samples. Compared to the standard analytic (single injections, TSKgel column), throughput was improved by 10x – 15x. Compared to single injections on the same column type, throughput was increased approximately 3x. In accordance to equation 6 the analysis time per sample for n samples can be calculated as follows:

$$t_{analysis} = \frac{t_{lag}}{n} + (t_{inf} + \sum t_{add}) + \frac{t_{hold}}{n} \tag{11}$$

which in our case amounts to

$$1{:}57 = \frac{2{:}00}{50} + (1{:}36 + (0{:}15 + 0{:}03)) + \frac{1{:}18}{50} \tag{12}$$

It is obvious that t_{lag} and t_{hold} do not contribute substantially to the overall analysis time when running the columns in parallel interlaced mode.

A statistical analysis of the results was performed and two data points differing more than 3 standard deviations from the mean value were excluded from further analysis. Average aggregate content detected was 5.03% with a standard deviation of 0.26 This rather large standard deviation was due to differing results from the two separate columns used. Mean aggregate level determined on

the first column was 5.27% with a standard deviation of 0.06. Mean aggregate level determined on the second column was 4.78% with a standard deviation of 0.05. While both columns yielded aggregate levels normally distributed around their mean value, results from both column differed statistically significantly as determined by a t-test ($p < 0.1\%$).

Table III: Control parameters used to control *timebase 1* (TB1; autosampler, pumps, column compartment including switching valves) and *timebase 2* (TB2; DAD). The commands for injecting five samples are shown. The initial flow path was: sampler → column 1 → DAD. Column "Time" shows the actual time during the analysis. Columns "TB 1" and "TB 2" show the time points programmed into the control programs for timebase 1 and timebase 2. The "action" columns adjacent to the "TB 1" and "TB 2" columns contain the commands used at the corresponding time point. Column "Sample" shows the time during which a sample is on a specific column. The first data acquisition on *timebase 2* generates a chromatogram ('dummy #') that only contains the t_{lag} of the first sample. (The two control programs of *timebase 1* differ only in switching valve commands. The data acquisition program on *timebase 2* is started by switching a relay ON.)

Time	Sample	TB 1	Action	Flow path	Action	TB 2
00:00		-0:27	Prepare sample #1			
00:27		0:00	Inject + Start Wash		Wait Input.state = ON	00:00
00:30		0:03	Relay.State = ON		Data Acquisition On	
01:54	**sample #1**	1:27	Pump Acquisition OFF			
01:57	**/ column 1**	1:30	method end		**dummy #**	
01:57		-0:27	Prepare sample #2			
02:06		-0:18	switch outlet valve	column 1 → DAD	Data Acquisition Off	1:36
02:12					method end	1:42
02:21		-0:03	switch inlet valve	sampler → column 2		
02:24		0:00	Inject + Start Wash		Wait Input.state = ON	00:00
02:27		0:03	Relay.State = ON		Data Acquisition On	
03:51	**sample #2**	1:27	Pump Acquisition OFF			
03:54	**/ column 2**	1:30	method end		**sample #1**	
03:54		-0:27	Prepare sample #3			
04:03		-0:18	switch outlet valve	column 2 → DAD	Data Acquisition Off	01:36
04:09					method end	01:42
04:18		-0:03	switch inlet valve	sampler → column 1		
04:21		0:00	Inject + Start Wash		Wait Input.state = ON	00:00
04:24		0:03	Relay ON		Data Acquisition On	
05:48	**sample #3**	1:27	Pump Acquisition OFF			
05:51	**/ column 1**	1:30	method end		**sample #2**	
05:51		-0:27	Prepare sample #4			
06:00		-0:18	switch outlet valve	column 1 → DAD	Data Acquisition Off	01:36
06:06					method end	1:42
06:15		-0:03	switch inlet valve	sampler → column2		
06:18		0:00	Inject + Start Wash		Wait Input.state = ON	00:00
06:21		0:03	Relay.State = ON		Data Acquisition On	
07:45	**sample #4**	1:27	Pump Acquisition OFF			
07:48	**/ column 2**	1:30	method end		**sample #3**	
07:48		-0:27	Prepare sample #5			
07:57		-0:18	switch outlet valve	column 2 → DAD	Data Acquisition Off	01:36
08:03					method end	01:42
08:12		-0:03	switch inlet valve	sampler → column 1		
08:15		0:00	Inject + Start Wash		Wait Input.state = ON	00:00
08:18	**sample #5**	0:03	Relay.State = ON		Data Acquisition On	
09:42	**/ column 1**	1:27	Pump Acquisition OFF			
09:45		1:30	method end		**sample #4**	

The presented method was shown to achieve large improvements of throughput for the particular analysis investigated. Certain prerequisites for achieving these improvements for any given chromatographic assay should be noted. First, the method works for isocratic elutions only, which is the case for SEC and some IEC/HIC analytics. Second, the improvement in assay throughput is related to the ratio of the information to the non-information phases of the chromatogram as only those parts containing no valuable information can be eliminated from the chromatogram. In the case described here, the information phase was approximately 24% of the entire chromatogram. Samples and analysis tasks making use of a larger portion of the chromatogram are amenable to the methodology as described in section 1.1 but might not yield throughput improvements as high as those reported here.

Reliability, robustness, and quantitativeness are the hallmarks of analytical SEC chromatography for mAb-aggregate quantification. Thus, it is preferred over other, even faster analytical methods such as capillary gel electrophoresis. The presented methodology increased sample throughput to an extend that it matches the speed of high throughput experimentation without changing the robust, underlying analytical principle. More detailed studies of aggregation and aggregate depletion during process development and production of mAb based pharmaceuticals can thus be performed.

3.4 Aggregate spiking studies

Aggregate spiking studies resulted in a linear response of the detected aggregate level to the expected aggregate level in the sample throughout the entire range tested (2.1% to 48.7%). The linear regression of measured aggregate level versus expected aggregate level were compared for the two separate columns used and two modes of operation (single and parallel-interlaced injection mode). The linear regression results were found to coincide, slope and intercepts were found to be statistically not different. The overall regression of expected versus measured value was resulted in a R^2 value of 0.9993 with an intercept fixed at 0 and a resulting slope of 1.01. This underlines our conclusion that the method presented herein can replace the standard method of running SEC columns for mAb-aggregate analysis and that the column used is well suited for the analysis task investigated. In theory, increasing aggregate levels could have increased the aggregate peak area to an extend where either monomer-aggregate peak resolution would decrease or where column valve switching times might have had to be adjusted. However, neither was found leading to the conclusion that the presented method is robust regarding aggregate levels of up to 48.7%. Aggregate levels below 2.1% were not investigated owing to the sample material at hand. However, the authors find no reason to believe that lower aggregate levels would pose a problem to the method.

4 Conclusion

In case of total mAb aggregate quantification, we find the ACQUITY column to be the best suited choice of the tested columns, as it enables more than a two fold improvement in throughput when compared to the TSKgel column (assay time comparison at a resolution of 1.5, see Figure 5 and table II). Further, due to the relatively low influence of flow rate on the separation which was found for the ACQUITY column, assay throughput can be increased further without compromising resolution significantly. The ACQUITY column also offers the benefits of lower buffer consumption and lower sample volume, latter being of great importance when performing HTPD.

A new methodology to improve throughput for SEC mAb analysis applied in biopharmaceutical science was demonstrated in this paper. By combining interlaced injections with parallel operation of two columns, near optimum utilization of SEC columns for the quantification of monomer and aggregate of a monoclonal antibody solution was achieved. Assay time was reduced to 1:57 minutes per sample as compared to 20 – 30 minutes using standard analytical protocols. Resulting aggregate levels were found to be comparable between different columns and different modes of operation. As an added benefit, heterogeneity between separate columns is factored into the results by using this method. With analysis times in the range of 2 minutes per sample the method presented in this paper is well suited for current high throughput pharmaceutical process development and process monitoring.

References

[1] M. Vázquez-Rey, D. a. Lang, Aggregates in monoclonal antibody manufacturing processes, *Biotechnol. Bioeng.* 108 (7) (2011) 1494–508.

[2] A. S. Rosenberg, Effects of protein aggregates: an immunologic perspective, *AAPS J.* 8 (3) (2006) E501–7.

[3] J. L. Coffman, J. F. Kramarczyk, B. D. Kelley, High-Throughput Screening of Chromatographic Separations: I. Method Development and Column Modeling, *Biotechnol. Bioeng.* 100 (4) (2008) 605–618.

[4] P. S. Wierling, R. Bogumil, E. Knieps-Grünhagen, J. Hubbuch, High-Throughput Screening of Packed-Bed Chromatography Coupled With SELDI-TOF MS Analysis : Monoclonal Antibodies Versus Host Cell Protein, *Biotechnol. Bioeng.* 98 (2) (2007) 440–450.

[5] M. Wiendahl, P. S. Wierling, J. Nielsen, D. F. Nielsen, J. Krarup, A. Staby, J. Hubbuch, High Throughput Screening for the Design and Optimization of Chromatographic Processes - Miniturization, Automation, and Parallelization of Breakthrough and Elution Studies, *Chem. Eng. Technol.* 6 (31) (2008) 893–903.

[6] S. Oelmeier, F. Dismer, J. Hubbuch, Application of an Aqueous Two-Phase Systems High-Throughput Screening Method to Evaluate mAb HCP Separation, *Biotechnol. Bioeng.* 108 (1) (2011) 69–81.

[7] B. Kelley, Industrialization of mab production technology: the bioprocessing industry at a crossroads, *mAbs* 1 (5) (2009) 443–52.

[8] J. F. Kramarczyk, B. D. Kelley, J. L. Coffman, High-throughput screening of chromatographic separations: Ii. hydrophobic interaction, *Biotechnol. Bioeng.* 100 (4) (2008) 707–20.

[9] Tosoh Bioscience LLC, SEC analysis of mAbs (2010), *Application note A13L07A*.

[10] R. King, C. Miller-Stein, D. Magiera, J. Brann, Description and validation of a staggered parallel high performance liquid chromatography system for good laboratory practice level quantitative analysis by liquid chromatography/tandem mass spectrometry, *Rapid communications in mass spectrometry: RCM* 16 (1) (2002) 43–52.

[11] M. Nelson, J. Dolan, Parallel chromatography - double your money, *LC GC North America* 22 (4) (2004) 338–343.

[12] D. Farnan, G. Moreno, J. Stults, A. Becker, G. Tremintin, M. van Gils, Interlaced size exclusion liquid chromatography of monoclonal antibodies, *J. Chrom. A* 1216 (51) (2009) 8904–9.

[13] M. Swartz, Uplc: An introduction and review, *J. Liquid Chrom. Related Technol.* 28 (7) (2005) 1253–1263.

MANUSCRIPT 2

High-Throughput Analytical Methods for the Egg White Protein Avidin

Patrick Diederich, Simon Luxem, Sven Amrhein, Jürgen Hubbuch

Institute of Engineering in Life Sciences, Section IV:
Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

parts of this work were presented at the *HTPD, Krakow, Poland (2010)*

Abstract

Three different high-throughput methods for quantitative purity, content and activity analysis were established and characterized for the analysis of the egg white protein avidin and its major contaminants in egg white ovalbumin, ovomucoid, lysozyme, and ovotransferrin. The robustness of the automated capillary electrophoresis instrument LabChip®(GXII, Caliper LS) was investigated regarding interference from changes in pH, salt and buffer species. A fast reversed phase HPLC method for purity and concentration determination using a UltraHPLC-System (Ultimate 3000 RSLC, Dionex) was developed and shown to provide highly precise data with a run-time of only 4.25 minutes. Lastly, a quantitative assay for the avidin-biotin interaction was scaled-down to a 96-well plate format enabling automated performance on a liquid handling station. Using the avidin protein system as model, the benefits and limitations of these three analytical methods were elucidated and general conclusions for their use in high-throughput process development are presented.

Keywords: high-throughput screening, LabChip® gel electrophoresis, tandem-HPLC, avidin, egg white proteins

1 Introduction

Robotic-based high-throughput experimentation (HTE) in protein purification process development features major advantages, e.g. a reduction of sample material needed and an automation of workflows. However, in order to fully utilize high-throughput techniques advantageously, it is essential to have implemented analytical methods for automated and rapid protein analysis. Improved technologies for purity measurements and protein quantification were developed in the last years which in combination with high-throughput experimentation on liquid handling stations led to an significantly increase of the experimental throughput [1, 2, 3, 4]. However, due to the novelty of these techniques, experience and knowledge regarding a broad range of applications and method robustness is scarce. Furthermore, the gain in throughput is often stated by comparing the analysis times rather than evaluating the overall time spent including preparation steps and data evaluation.

We present in this study the learnings gained during a search for fast analytical methods for quantitative and qualitative analysis of the valuable protein avidin. The goal was to develop methods that can match the speed of high-throughput experimentation when applied for process development for the purification of avidin (AV) from egg white. The major impurities in this process are lysozyme (LYS), ovomucoid (OM), ovalbumin (OV), and ovotransferrin (OT). While providing short analysis times, the aimed analytical techniques should not compromise precision and accuracy when compared to traditional methods with lower throughput. Furthermore, the challenge was to develop methods also applicable for samples of high purities of 90% – 99.5%. As a consequence, the quantification of the impurities down to 1 mg/L was aimed in presence of high concentrations of AV.

For protein quantification and purity determination, the high-throughput capillary gel electrophoresis instrument GXII from Caliper Life Sciences was evaluated. Based on the principle of SDS PAGE [5] proteins are electrophoretically separated through a gel matrix due to differences in size. In the GXII, the separation takes place in a micro channel on a chip which is loaded automatedly with samples from a 96-well or 384-well micro-titer plate. The large benefit of this technique is that, in addition to a virtual gel view of the protein bands indicating the protein size, fluorescence measurement is used for a correlation to protein mass, thus, providing also a quantitative measurement. The fluorescence is caused by a dye which is injected into the separation channel and there binds to the SDS-protein complex. The combination of qualitative and quantitative analysis, the high degree of automation (integration on liquid handling station possible) and the very low analysis time (≈50 s per sample) are ideal features for high-throughput concentration and purity analysis.

As alternative approach for the quantification of AV and the determination of its purity, we investigated whether traditional HPLC technology could fulfill the demands on a high-throughput method. While high precision and accuracy should be rather easily obtained when using HPLC, the challenge in this case was to achieve a reasonable short analysis time. A run time <5 five minutes was set as target, which would match the average number of generated samples per day.

In addition to methods for the quantification and determination of purity a high-throughput compatible activity assay was established. Avidin features an exceptionally strong affinity to the vitamin biotin ($kd = 10^{-15)}$. The avidin-biotin interaction is a valuable tool for biochemical assays and diagnostics [6] as well as for therapeutic use [7]. High and well-defined activity is therefore a quality attribute which needs to be determined. A method for the automated and parallelized performance of the biotin-avidin binding assay using 4'-hydroxyazobenzene-2-carboxylic acid (HABA) in miniaturized micro-titer plate (MTP) format was developed. The HABA-assay is based on the competing binding between HABA and biotin to the same binding site in avidin [8, 9]. A change in absorption

intensity at wavelength (λ) 500 nm can be determined by photometric measurement after the addition of HABA to an avidin solution. Biotin is then added and displaces HABA resulting in a change of absorption which can be linearly correlated to the amount of biotinylated avidin. A photometric assay is in particular suitable for high-throughput applications because photometric analysis can easily be performed in plate readers fully integrated on LHSs. Furthermore, this kind of analysis is typically performed within seconds and does not require expensive instruments and materials.

With this study we intend to demonstrate the potential of different analytical techniques as well as the limitations we came across. By purpose, rather high requirements on the quantification methods such as as the high avidin purity and hence low quantification levels for the impurities, were set. This was done, because an evaluation on the applicability of high-throughput methods at challenging conditions was considered necessary, in order to further increase the acceptance and understanding of high-throughput process development.

2 Material and methods

2.1 Materials

2.1.1 Proteins and Chemicals

The egg white proteins, ovalbumin (prod.-no. A5503), ovotransferrin (conalbumin, prod.-no. C0755), ovomucoid (trypsin-inhibitor, prod.-no. T9253), lysozyme (prod.-no. 6290), phosphorylase b (prod.-no. P6635), and β-galactosidase (prod.-no. G5160) were purchased from Sigma-Aldrich Co. (Taufkirchen, GER). Avidin with a purity $> 90\%$ was provided by an industrial partner.
The salts H_2NaPO_4, HNa_2PO_4, MES, MES sodium salt, sodium acetate, and NaCl were of analytical grade and purchased from Merck KGaA (Darmstadt, DE). For pH adjustments, hydrochloric acid (Merck KGaA (Darmstadt, DE)) and acetic acid and sodium hydroxide (Sigma-Aldrich Co. (Taufkirchen, DE)) were used. Water was purified using an Arium®pro UV system (Sartorius Stedim Biotech, Göttingen, D). Acetonitrile (ACN) of gradient-grade was purchased from Merck KGaA (Darmstadt, DE). Trifluoroacetic acid (TFA) of sequencing grade supplied by Thermo-Fisher Scientific (Bonn, DE) was used as modifier. D-Biotin and dithiothreitol (DTT) were purchased from Sigma-Aldrich Co. (Taufkirchen, DE).

Protein stock solutions of AV, LYS, OV, and OM were prepared in water or 100 mM phosphate buffer, pH 7, while OT was solubilized in 100 mM phosphate buffer, pH 4.5, due to its low solubility at neutral pH. D-Biotin and 4-Hydroxyazobenzene-2-carboxylic acid (HABA) were purchased from Sigma-Aldrich Co. (Taufkirchen, DE).

2.1.2 Consumables

HT PROTEIN EXPRESS (version 2) LabChip®chips from Caliper LS (Hopkington, US) were applied for the analysis on the GXII. Buffers and solutions (HT PROTEIN EXPRESS Reagent Kit and HT Protein 200 Sample Buffer) were purchased readily prepared from CaliperLS (Hopkington, US). UV Star 96 polysterol micro-titer plates from Greiner Bio-one and skirted 96-well polypropanol PCR plates from Eppendorf (Hamburg, DE) were used. For sample desalting, Zeba™Spin Desalting 96-well plates with a MWCO of 7 KDa were used. For HPLC analysis, customized Vydac®ProZap™Ex-pedite C18 columns (ID x L = 2.0 x 30 mm and ID x L = 2.0 x 20 mm) from Grace (Deerfield, IL, USA), an UPLC®BEH300 C4 column (ID x L = 2.1 x 50 mm) from Waters (Milford, MA, USA), and a

Chromolith Performance RP C18 (ID x L = 2.0 x 100 mm) purchased from Merck Millipore KGaA (Darmstadt, DE) were used.

2.2 Methods

2.2.1 LabChip®electrophoresis

Protein samples were analyzed using the HT Protein Express 100 High Sensitivity assay and, for comparison, the HT Protein Express 100 (standard) assay. The LabChip®preparation, sample preparation and analysis was performed according to the manufacturer's protocol [10]. Regarding the sample preparation, the procedure for the Protein Express or Protein Express High Sensitivity assay were identical except for the volume ratio of sample to sample buffer. In brief, 10 μL or 4 μL sample solution (for the Protein Express High Sensitivity assay or Protein Express standard assay, respectively) were mixed with 14 μL sample buffer in a MTP. The MTP was sealed with adhesive aluminum sealing foil and incubated on a MTP thermo-mixer with lid heating (Ditabis AG, Pforzheim, DE) at 95°C for 5 min while shaking at 300 rpm. The MTP was then cooled down to room temperature and centrifuged shortly prior to removing the plate sealing foil. To every well, 64 μL purified water was added (70 μL when using the Protein Express standard assay). The plate was shaken for 5 min at 800 rpm followed by short centrifugation prior to loading the plate into the GXII instrument. The protein reference standard ('ladder') was prepared in one well of the sample plate incubating 12 μL ladder and adding 120 μL water after incubation. The ladder was transferred into a single vial, placed into the GXII instrument and used for maximal 8 hours.

For testing method robustness, a full-factorial experimental design was used and samples were analyzed using the HS-assay. AV samples with a concentration of 1.5 g/L were prepared in two stock solutions of 0 mM and 600 mM of each 100 mM sodium phosphate, 100 mM sodium MES and 100 mM sodium acetate buffers at pH 4.5, 5.5, 6.5, 7.5, and 8.5. Further salt concentrations were obtained by mixing the two salt stock solutions of the corresponding buffer and pH.

The first 12 injections on a plate were used as 'dummy' samples and were not considered for the evaluation. This was done due to the observation that after instrument downtimes the first injections showed a significantly higher signal compared to the rest of the samples. For all samples the peak integration was manually corrected if needed.

In order to test desalting of the samples prior to analysis on the GXII device, protein sample volumes between 25 μL and 50 μL followed by the same volume of water ('stacker') were pipetted into Zeba™ Desalting plates. Subsequently, the plate was centrifuged for 2 min at 1000 g and the desalted eluate was collected in a 96-well MTP.

Factors influencing the determination of the concentration

The GXII software calculates protein concentrations based on peak areas. The area values, however, are corrected for differences in injection volume and shifts in retention time, because later eluting proteins move slower through the gel matrix and therefore generate a broader and larger peak. The correction for the injection volume is performed by the injection of an internal standard of low molecular weight (lower marker) which is mixed with the sample on the chip during injection. If the area of the lower marker is different from the default area, the protein peak areas are corrected by the reciprocal of this ratio. The relation of area to concentration is calibrated by the average area of the ladder proteins for which concentrations are known. The correction due to shifts in retention time is

performed using the retention times of the ladder proteins which are aligned by the retention time of the lower marker. Thus, the determination of the concentration depends on the performance of the ladder injection and the correct elution time of the lower marker while assuming that if the lower marker area would be affected (e.g. lower injection volume), the protein peak areas would be affected with the same factor. A full description of the algorithm performed by the software was, however, not available in the software manual [10].

2.2.2 RP HPLC analysis

RP HPLC analysis was performed on an UltiMate®3000 RSLC x2 Dual system from Thermo Fisher Scientific (Waltham, MA, USA). The system was composed of two HPG-3400RS pumps, a WPS-3000TRS analytical autosampler or a WPS-3000 TFC autosampler, a column thermostat TCC-3000RS, and a DAD3000RS detector. The column compartment was equipped with two six-port column switching valves. The column temperature was 25 °C. The flow rate was set to 0.4 mL/min. An injection volume of 2 – 5 μL was used. Purified water containing 0.1% (v/v) TFA (solvent A) and ACN containing 0.1% (v/v) TFA (solvent B) were used as solvents. The method featured a multi-segment increase of solvent B in solvent A. Details on the established program are given in the results section, Figure 7. The method was used as single and tandem application, the latter by operation of two columns simultaneously. Protein absorption was measured at λ=225 nm. A solution of 2 mM D-Biotin dissolved in 100 mM sodium phosphate buffer, pH 7, was added in excess (volume ratio: 1 part d-Biotin to 10 parts of sample) to each sample prior analysis. Addition of biotin prevented the elution of avidin in several peaks.

2.2.3 HABA-assay

Reference analysis in cuvettes

A protein sample volume of 250 μL was diluted two-fold with 100 mM phosphate buffer, pH 7.0. Subsequently, 12.5 μL of 10 mM HABA dissolved in 100 mM sodium phosphate buffer, pH 7.0, was added and mixed by aspiration and dispension several times using the pipette. The absorption (A_{+HABA}) was measured using a Tecan®Magellan reader at a wavelength of 500 nm and a path length of 1 cm. Subsequently, 12.5 μL of 2 mM d-biotin dissolved in 100 mM sodium phosphate buffer, pH 7.0 was added, mixed thoroughly, and the absorption was measured at λ=500 nm using the Magellan reader.

Automated assay in MTP format

Automated experiments were performed on a Freedom Evo®200 station (Tecan, Crailsheim, DE) equipped with an 8-channel liquid handling arm (fixed steel tips), a centric gripper, a shaker (Teleshake 95, inheco, Munich, DE), an integrated centrifuge (Rotanta 46RSC, Hettich, Tuttlingen, DE), and an Infinite200 plate reader (Tecan, Crailsheim, DE). Sample volumes between 20 – 100 μL were pipetted with the 8-channel pipetting arm into a 96-well MTP and filled up to 250 μL with 100 mM sodium phosphate buffer, pH 7.0. A volume of 25 μL HABA solution was added to each sample. Prior to the absorption measurement at λ=500 nm in the plate reader, the plate was shaken for 1 minute with ≈1500 rpm. Subsequently, 25 μL d-biotin solution was pipetted into each sample well, the plate was shaken again for 1 min at 1500 rpm, and the absorption at λ=500 nm was measured again.

Determination of concentration of active avidin

The concentration of biotin-binding AV $c_{AV,active}$ was calculated by

$$c_{AV,active} = \left(\frac{M_{AV}}{\epsilon_{A,HABA-AV}}\right) \cdot \left(\frac{A_{+HABA} \cdot \frac{V_{+HABA}}{V_{sample}}}{L_{+HABA}} - \frac{A_{+biotin} \cdot \frac{V_{+biotin}}{V_{sample}}}{L_{+biotin}}\right) \quad (1)$$

where M_{AV} is the molecular weight of AV monomer (16.39 g/mmol), $\epsilon_{A,HABA-AV}$ is the molar extinction coefficient of the HABA-AV-complex at pH 7.0 (34 L/mmol·cm), A_{+HABA} and $A_{+biotin}$ are the absorption values at 500 nm after addition of HABA and biotin solution, respectively, to the AV sample (V_{sample}). V_{+HABA} and $V_{+biotin}$ are the nominal volumes of the sample after addition of HABA and biotin solution, respectively. In the case of measurements in cuvettes, the path lengths L_{+HABA} and $L_{+biotin}$ were constant and were to 1 cm whereas in the automated method L_{+HABA} and $L_{+biotin}$ in the MTPs changed. In order to determine the liquid levels and hence the path lengths in each well, the absorption at wavelengths 990 nm and 900 nm were measured in a plate reader. The absorption difference was correlated to the liquid level as presented by McGown and Hafemann **(author?)** [11].

2.3 Software

The GXII was operated by the LabChip®GX software version 2.2.417.0 SP1. Chromeleon® (6.80 SR10, Thermo Fisher Scientific (Waltham, MA, USA)) was used to operate the UltiMate®3000 RSLC x2 Dual system and for peak integration. MODDE 8.0 from UMETRICS (Uppsala, SE) was used for the evaluation of full-factorial experiments. The liquid handling station was operated with Evoware 2.4 SP3 (Tecan, Crailsheim, DE). UV-VIS absorption measurement were performed using the Magellan 6.5 or the i-control 1.3.3 software. Import of pipetting values and export of photometric data was realized via Excel®(Microsoft, Redmond, WA, USA).

3 Results & discussion

3.1 High-throughput capillary gel electrophoresis

The potential of the LabChip®GXII for a high-throughput analysis of AV and its major impurities in egg white LYS, OV, OM, and OT was investigated using different mixing ratios of these proteins. Besides the possibility to use this device as an alternative to common SDS-PAGE analysis for qualitative analysis, we were interested in the performance regarding a determination of the concentration, quantification limits and robustness towards changes in the sample matrix. For an initial evaluation, we used both the standard Protein Express 100 protocol (in the following referred to as ST-assay) and the Protein Express 100 High Sensitivity assay (in the following referred to as HS-assay).

3.1.1 Qualitative analysis

A mixture of egg-white proteins with 1.0 g/L of AV and concentrations varying between 10 – 200 mg/L of LYS, OM, OV, and OT in 10 mM phosphate buffer, pH 7 were analyzed using the GXII device. The virtual gels and an example of an electropherogram are shown in Figure 1. Only four clearly separated peaks were obtained. A comparison to single-protein samples revealed that OM was not detected by this analysis. When using reducing conditions by adding DTT to the sample OM was detected as broad peak ranging from 69 kDa to 120 kDa. Thus, OM would elute together with OT and using reducing conditions was therefore not feasible. Regular SDS-PAGE using a 16% Bis-Tris gel was performed in order to evaluate whether the lack of detection of OM was instrument specific or in

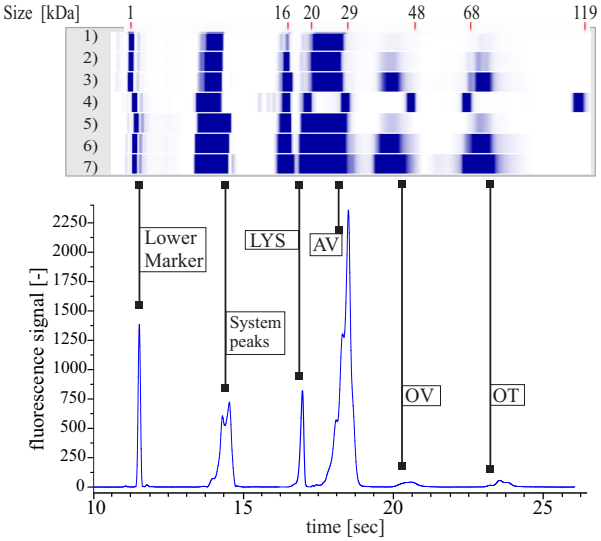

Figure 1: Virtual gel view (rotated) and example of corresponding electropherogram (sample lane 2) of egg protein samples analyzed using the LabChip®ST-assay (lane 1-3) and HS-assay (lane 5-7). The AV concentration in each sample was 1.0 g/L. The impurities LYS, OT and OV were spiked with concentrations of 10 mg/L (lane 1 and 5), 50 mg/L (lane 2 and 6), and 200 mg/L (lane 3 and 7). Lane 4: ladder reference sample.

general challenging by methods based on the staining of protein-SDS complexes with a dye. Although no reducing agent was added, OM was detected with regular SDS PAGE, however, the OM band was of low intensity and, similar to the result from the GXII, very broad with a range between 60 – 80 kDa. Difficulties in the determination and quantification of OM by the principle of SDS gel electrophoresis is described in literature [12]. Due to the high degree of glycosylation of OM (25%, [13]), sufficient binding of SDS to the molecule is hindered and thus staining and electrophoretic mobility affected. The results obtained with the ST-assay and the HS-assay differed significantly in the signal intensity. This is in agreement with the larger sample volume used for the HS-assay.

3.1.2 Quantitative analysis

Determination of concentration

For both the ST-assay and the HS-assay, AV concentrations of 100 – 1500 mg/L and LYS, OV and OT concentrations of 10 – 100 mg/L were analyzed. Each sample was prepared as replicate on a second plate measured directly after the first plate. Each plate was measured twice. The mean of the four concentration values determined by the LabChip®GXII software are plotted in Figure 2 as a function of the prepared sample concentrations. The latter are referred to as nominal concentration in the following.

From these data it can be seen that the absolute concentrations determined by the LabChip®GXII software did not match the nominal concentrations. Thus, each protein would require a separate calibration if a correct determination of the concentration was desired. In the lower concentration range, the determined and nominal concentrations correlated linearly as expected. A linear calibration curve was applicable for the ST-assay for concentrations of AV and the impurities up to 1000 mg/L and 100 mg/L, respectively. When using the HS-assay, a fit of the response using a quadratic function was required for concentrations >100 mg/L (Figure 2a). According to the different protein sample

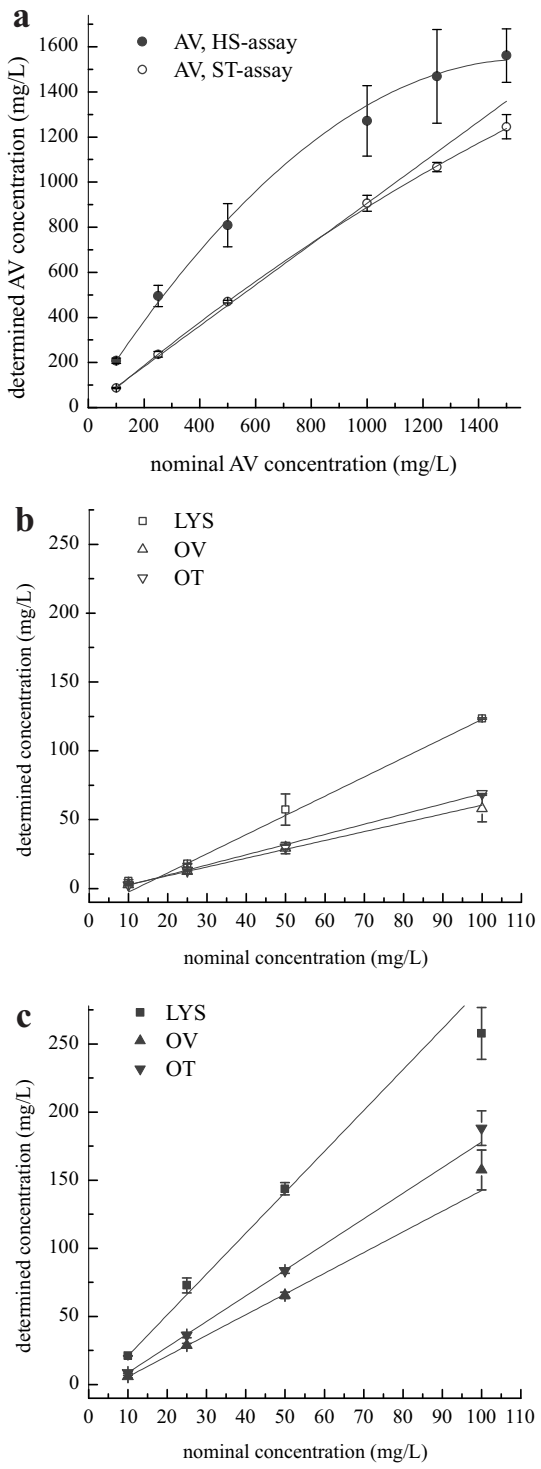

Figure 2: Comparison of nominal and determined concentration for AV (a) and the impurities LYS, OV, and OT (b,c) when using the LabChip®ST-assay (a,b) and HS-assay (a,c). Concentration ranges were selected based on ranges expected for the application of the method. Correlations were fitted using a quadratic polynomial function (a) or linear function (a (for comparison, ST-assay), b,c). Error bars represent +/- one standard deviation (no. of samples: 4).

volumes used for ST- and HS-assay, the protein signals are about 2-3 times higher when using the HS-assay.

As indication of the lower limit of quantification (LLOQ) of the impurities, the signal to noise ratio was estimated from the noise of the baseline. The concentration at which a signal to noise ratio of 10 is achieved is commonly considered as LLOQ [14]. When using the ST-assay, signal to noise ratios for sample concentrations of 10 mg/L of approx. 2.5 were obtained for OV and OT, whereas a ratio of approximately 100 was obtained for LYS. When using the HS-assay for sample concentrations of 10 mg/L signal to noise ratios were 500, 21, and 17 for LYS, OV, and OT, respectively. For the HS-assay, the corresponding LLOQs were hence estimated with 0.2 mg/L, 4.8 mg/L, and 5.8 mg/L for LYS, OV and OT, respectively. As the purpose of this analysis was to determine purities of >90% of AV at concentrations ranging from 0.2 – 2 g/L, the application of the HS-assay would therefore be more suitable than using the ST-assay.

Determination of molecular weight

The molecular weights (MWs) determined by LabChip®GXII software for the proteins LYS, AV, OV and OT were compared with values retrieved from literature (Table 1). While the determined MW for OT was in agreement with the literature values for the theoretical size, the MWs for AV, LYS, and OV deviated from literature values by +30%, +12.5%, and 5.6%, respectively. Values obtained from ST- and HS-assay were similar. An influence of the protein concentration on the MW determination was not observed.

Table 1: Molecular weight determination using the LabChip ST- and HS-assay in comparison to literature values. Theoretical MWs were given in [15]. Values obtained in SDS-PAGE with silver staining were stated in [13]. All values have the unit kDa.

protein	**literature**		**ST-assay**		**HS-assay**	
	theoretical	by SDS PAGE	mean	relSTD (%)	mean	relSTD (%)
LYS	14.5	14	16.3	2.5	16.4	1.8
AV monomer	15.5	17-20	24.6	2.8	24.0	3.2
OV	45.0	50	42.6	1.0	42.2	1.2
OT	76.0	74	74.8	1.4	73.9	1.3

AV consists of four identical subunits. In literature, a total mass of 55.0–68.3 kDa is stated [15]. The determined size of 24 kDa indicates a penetration of the matrix most-likely as monomer which has been proposed from SDS PAGE results ([13, 16]). In comparison to the deviation between SDS-PAGE analysis and literature values, the deviations obtained when using the LabChip®were larger.

3.1.3 Aspects of method precision

The quantitative precision of the LabChip®was of great interest and was intended to be compared to the quality of HPLC data. With the benefit of no method development and much shorter analysis times, one would possibly accept the precision to be lower compared to HPLC. However, the question was rather to what extent the experimenter would need to compromise precision when using the LabChip®instrument. This was investigated in two studies. In a first study using the ST-assay, 20 samples of 200 mg/L avidin in 100 mM sodium phosphate buffer, pH 7, were prepared according to the original procedure described in the user manual [10] and by using a single-channel pipette. These samples were analyzed randomly distributed with other egg protein samples and blank samples in between. Three hours after the first analysis (hereafter measurement-1), the same sample plate was analyzed again (hereafter measurement-2). The obtained area values and concentrations determined

by the software are shown in Figure 3.

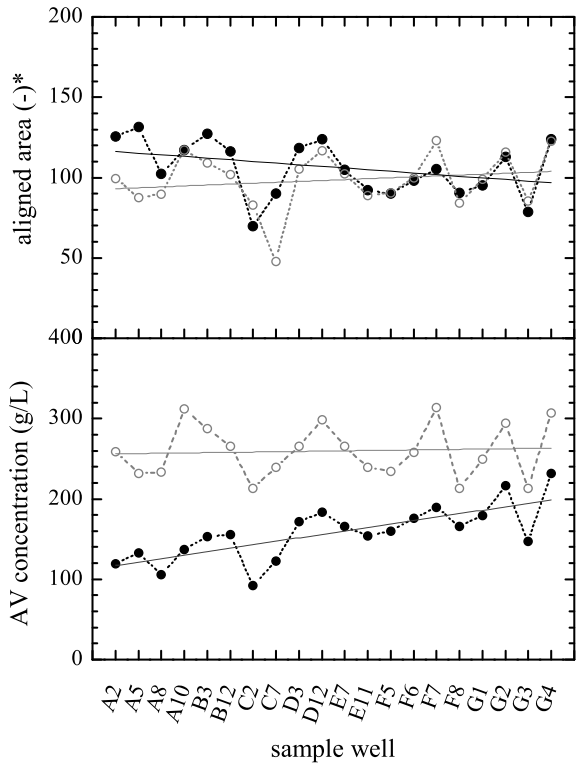

Figure 3: Area values and corresponding concentrations of 20 samples of 200 mg/L AV. The first measurement-1 (measurement-1, black) was repeated after three hours (measurement-2, gray). Samples were randomly distributed on a plate with various calibration samples in between (data not shown). Straight lines indicate the average trend with increasing sample number and hence analysis time. * The aligned area gives the peak area corrected after alignment of the internal standards (lower markers) with the markers in the ladder.

For measurement-1 and measurement-2, the determined mean avidin concentration was 168 mg/L and 260 mg/L with relative standard deviations (RSDs) of 21.8% and 12.9%, respectively. The concentrations increased in average throughout measurement-1 from ∼100 mg/L to ∼220 mg/L. The area values did not show this trend but were rather constant throughout the measurement. In order to explain the increasing concentrations for measurement-1, the standard protein ladder areas and sizes were plotted (Figure 4). The area signal of all standard proteins in the ladder decreased in measurement-1 with increasing analysis time, whereas the ladder areas were rather stable in measurement-2. The determined ladder protein MWs were exactly the same in both measurements. Besides the decreasing ladder area signals, the elution times of the lower marker in the ladder samples increased with increasing sample number (data not shown). Both factors contribute to the trend of increasing AV concentrations in measurement-1, however, an exact description of the algorithm underlying the calculation of the concentration was not available in the software documentation. Furthermore, an experimental reason for the decrease in ladder area signals could not be identified. This observation was made regularly in the first measurements of a day and were therefore considered as an systematic

error. This was regardless whether the applied chip was new or reused.

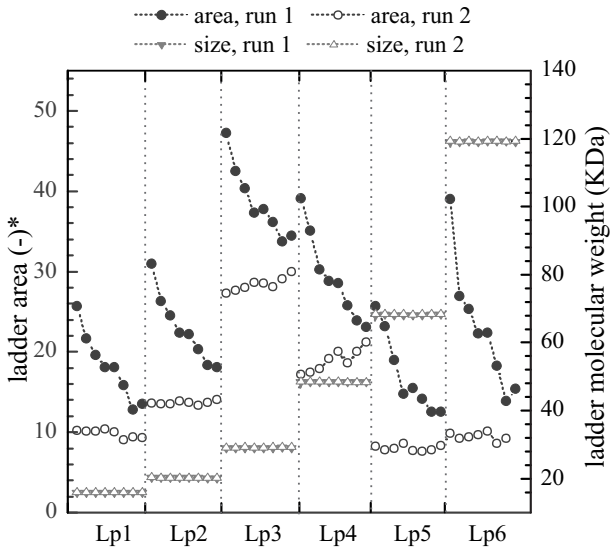

Figure 4: Area and molecular weight (size) of all ladder proteins (Lp1-Lp6) determined during measurement-1 and measurement-2 (from Figure 3) displayed in chronological order. For every 12 protein samples, the ladder sample was analyzed resulting in 8 data points for measurement-1 and measurement-2, respectively.

When considering only the aligned area values, it can be concluded from Figure 3 that the deviations between each of the 20 pairs of samples (measurement-1 and measurement-2) are lower than the deviations within one measurement of 20 prepared samples (caused by errors in sample preparation).

In a follow-up study, the sample preparation was optimized by doubling the sample and buffer volumes as described in section 2.2.1, by using a multi-channel pipette, and priming the instrument twice prior to starting the analysis. For each of the five concentration levels between 100 mg/L–1500 mg/L, eleven AV samples were prepared on one samples plate (in total 55 samples). All 55 samples on this plate were analyzed with the HS-assay in three subsequent measurement series (hereafter series-1, series-2 and, series-3). The ladder areas were inspected after measurement of series-1 and showed very consistent area values (data not shown). When comparing the three measurement series, however, a deviation of up to 24% in area values was determined for each ladder protein, and again, the largest deviation was identified in series-1. For a comparison of the instrument precision with the precision of the measurement including sample preparation, the data set was evaluated in two ways. For each sample measured in triplicate (3 series), the RSD was calculated. The mean of these eleven RSDs were used as indication of the instrument precision. For an estimation for the variance including the sample preparation, the mean of the RSDs calculated for each series of 11 samples were calculated. These two means of the RSDs are displayed for each concentration level in Figure 5.

The triplicate measurements (series-1, series-2, and series-3) of 55 AV samples (11 samples of 5 concentration levels) showed that for concentrations up to 500 mg/L the sample preparation error is significantly higher compared to the instrument error. With increasing sample concentration, the mean RSD for each concentration level (11 samples), however, decreases to values below 2%, while the instrument error appears to be independent of the sample concentration. The RSDs of the mean RSD values for the three replicate measurements are significantly higher compared to the RSDs regarding a series of 11 samples. Although this might have been due to the different number of samples (3 vs. 11)

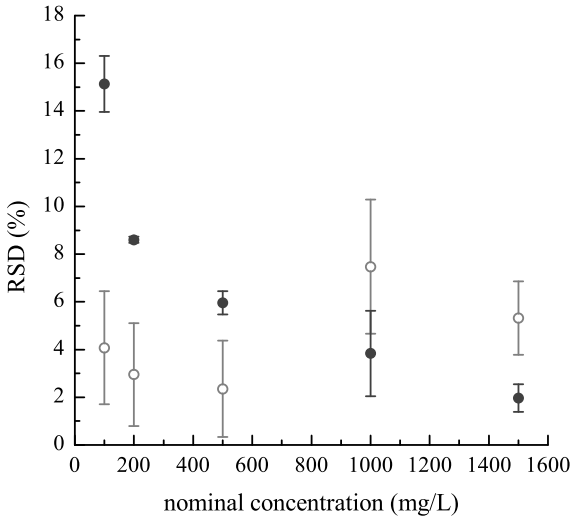

Figure 5: Means of eleven RSDs for each identical sample measured in triplicate (○,gray) indicating the instrument precision, and means of three RSDs calculated for each series of eleven samples (●, black) indicating the overall precision including errors in sample preparation. The error bars display ± one RSD of the means. The AV samples were prepared in 100 mM sodium phosphate buffer at pH 7 and were analyzed using the HS-assay.

for which the RSDs were calculated, it was found that the concentrations determined in series-1 were higher compared to series-2 and series-3. The deviations between the last two series were between 1.7% – 0.9%. This confirmed the observation made previously that the first measurement deviated stronger from the following measurements, although chip preparation and instrument handling was performed as given in the user manual. When comparing only the last two measurements, it was shown that the instrument can provide a reproducible analysis and the main error source is the sample preparation (e.g. pipetting error, sample loss due to the heating/cooling during preparation etc.). However, difficulties with the first measurements after instrument down time without any indications from the standard protein ladder signal, poses a question mark on the overall repeatability of the measurements.

3.1.4 Method robustness

The previously described measurements had all been performed with proteins dissolved in the same buffer system. However, the purpose of this method was to analyze samples obtained from a variety of purification process steps, and hence, to analyze samples of significantly varying sample matrices. In order to determine the method robustness regarding the influence of the sample matrix, the effect of NaCl concentration, type of buffer ion (inorganic: phosphate, organic: acetate, synthetic: MES), and pH was therefore investigated in a full-factorial study. Samples were prepared in duplicates and analyzed using the HS-assay. The area, concentration and MW values determined by the GXII software (after manual control of the peak integration) were evaluated regarding the main effects and interaction effects, and the results are shown in Figure 6. In electropherograms of several samples with phosphate and 0 or 200 mM NaCl, the elution profile was strongly shifted to earlier elution times. The suspected lower marker areas eluted with the system peaks resulting in undefined lower marker peak areas , and hence, in apparently false concentration values. Obviously incorrect concentration values (e.g. >2.5 g/L) were excluded from further data evaluation.

Effect of salt concentration

The NaCl concentration had a large effect on the determined area, concentration, and MW. The area values decreased with an increasing NaCl concentration (Figure 6a,d) regardless of the pH or buffer species. The determined concentrations increased for all pH values and buffers in the range of 30% – 55% (Figure 6b,e). The opposite trend was expected for the concentration values, because larger peak areas should result in larger concentrations. This trend was, however, caused by a decrease in the area of the lower marker with an increasing NaCl concentration. The software therefore assumed lower injection volumes indicated by the lower marker areas and corrected the concentration values for the apparently lower injection volumes. As the lower marker areas were more affected by salt than the protein peaks, the software overestimated the correction factor for calculating the concentrations resulting in increasing concentrations with an increasing salt concentration.

The determined MW increased with an increasing salt concentration, however, only in the case of using acetate and phosphate, whereas an influence of the NaCl concentration in MES samples was not observed. For samples in phosphate and acetate buffer, the average increase of the AV MW across all pH values was between 18% and 29%.

Effect of buffer species

The choice of buffer species clearly showed an effect on both the concentration and the MW (Figure 6d-i). The determined AV concentrations were lowest in presence of phosphate and highest for samples in MES buffer. The average deviation from the nominal concentration value across all pHs and salt concentrations was ∼25%. The difference in the determined MWs for samples with different buffer species was significant with approximately ∼19 kDa, ∼23 kDa, and ∼28.5 kDa for samples in phosphate, acetate, and MES buffer, respectively.

Effect of sample pH

The change in pH showed little effect on the determined concentration, whereas area values decreased with an increasing pH (Figure 6g,h). In this case, the software was able to correct for the decreasing area and calculated rather constant concentration values for all pHs. The determination of the MW was not affected significantly by changes in pH for acetate and phosphate samples (within 1 kDa), however, for MES samples the MW of AV increased with pH by approximately 11%.

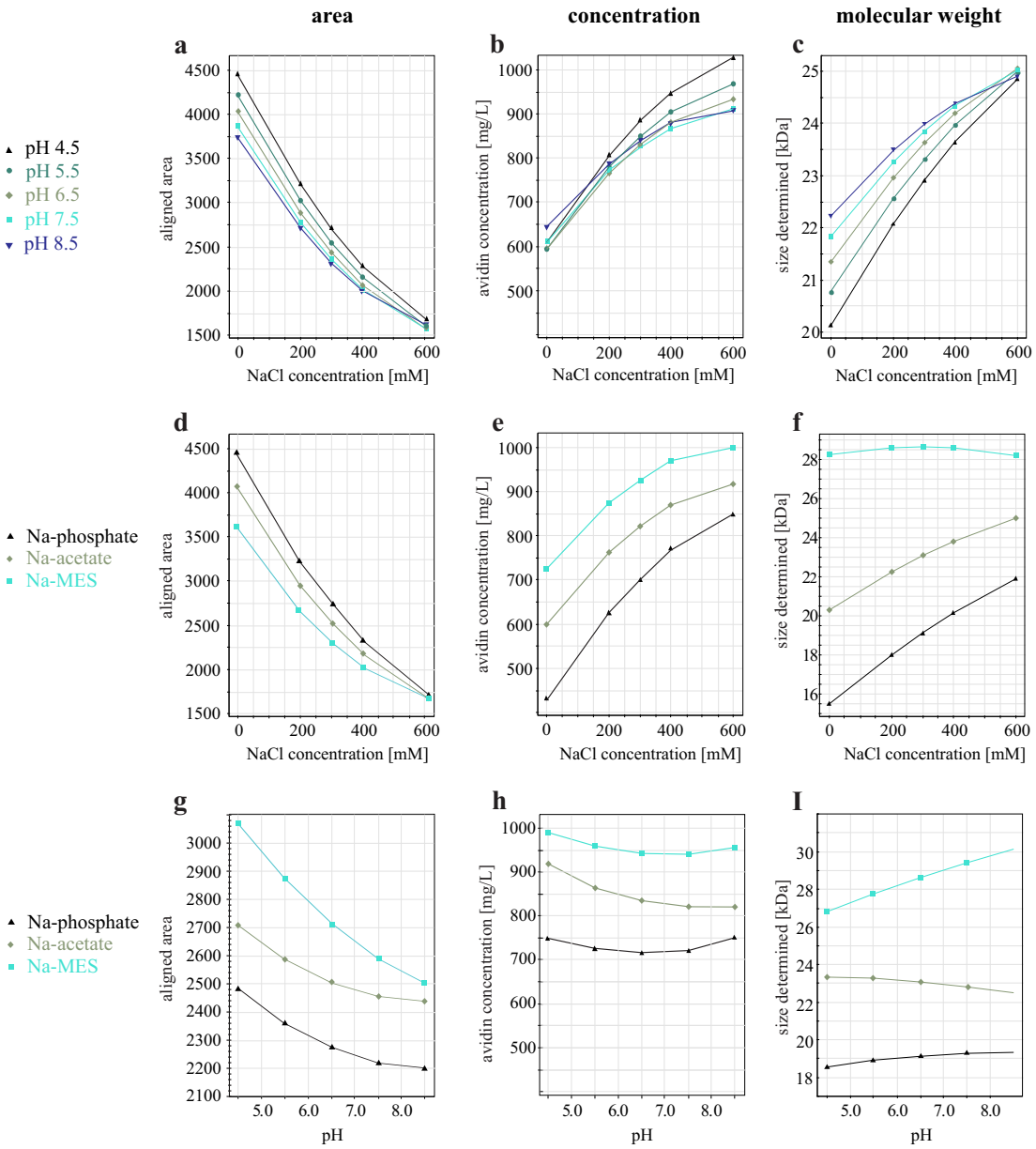

Figure 6: Interaction plots for the effect of pH, buffer ion, and NaCl concentration on area, concentration and molecular weight. The values display the results of duplicates of 1.5 g/L AV dissolved in the corresponding buffer. The effects of two parameters are displayed at average level of the third parameter (evaluation in MODDE). All samples were analyzed using the HS-assay.

Parameter interactions

An interaction of parameters was indicated when the trends in each subplot of Figure 6 had different slopes. Regarding the obtained area values, the slopes were very similar for all combinations. However, the changes of pH and buffer species had a larger effect when no NaCl was present compared to when using 600 mM NaCl. With regards to the determination of the concentration, it was clearly shown that the influence of NaCl was almost independent from the buffer species or the pH. The sensitivity against NaCl was slightly higher at low pH as can be seen in Figure 6b. For the determination of the MW, it was shown that acetate and phosphate were influenced by a change of the NaCl concentration, but rather unaffected when changing the pH. The opposite was seen when using MES.

In conclusion, as seen from the relation of area to concentration values, the necessity of a correction algorithm and internal standard of known concentration in each sample was demonstrated. However, the changes in retention time and area caused by a varying NaCl concentration or the buffer species could not be corrected by the software resulting in varying concentrations for samples of identical nominal AV concentration.

3.1.5 Potential method improvements

Desalting of samples

The robustness study revealed a high sensitivity towards changes in the buffer matrix. In order to circumvent this influence, we tested the use of 96-well desalting plates in order to generate identical buffer conditions prior to sample preparation for the LabChip®analysis. Chen and Flynn [17] had reported a benefit from desalting samples from cell culture supernatants for the determination of antibody aggregates on the LabChip®and obtained a recovery of $93\% \pm 2.9\%$. Desalting was tested for solutions of AV and the impurity with lowest MW (LYS). It was found that $40\,\mu$L of protein solution was required to provide total recovery of the AV in the sample. Total recovery was, however, only obtained using at least $40\,\mu$L of water/sample buffer ('stacker' volume) added to the sample prior to centrifugation. Without stacker volume the recovery decreased below 50%. The recovery of LYS could not be increased to over 50%, a fact, which makes desalting by this simple and high-throughput compatible method not applicable for this analysis task.

Different internal standard

The use of a second internal marker would be a feasible option supported by the GXII software as potential remedy for the incorrect correction due to the lower marker. The usage of an upper marker is, however, only possible in cases in which there is a sufficiently large range of MW in which no proteins of the analyzed sample elute (in our case approx. >85 kDa). We tested both phosphorylase b (97 kDa) and β-galactosidase (119 kDa), which were mixed with the lower marker solution and injected with each sample. While phosphorylase b partly co-eluted with OT, β-galactosidase was found suitable for the use as upper marker. For eight identically prepared AV samples (750 mg/L) in 100 mM phosphate, pH 4.5, with 0 mM and 300 mM NaCl, the achieved RSDs were 28% and 34%, respectively, when using the upper marker and 67% and 187% when using the lower marker (data not shown). In the latter case, lower marker detection and integration appeared incorrect due to the low area and shift in retention time of the lower marker peak.

In summary, the LabChip®technology was shown to provide a very easy and fast way to qualitatively determine the protein MWs. In comparison to the analysis using traditional SDS-PAGE, the huge benefit of the high throughput is complemented by additional information given by electrophero-

grams in which peak overlaps can be detected. In terms of protein quantification, an intermediate precision with a RSD of 2–15% was determined which might be sufficient for many high-throughput applications. However, deviations between replicate measurements of identical sample plates differed for unknown reasons giving the experimenter some uncertainty regarding the results.
When using the HS-assay, the sensitivity to buffer species, salt concentration, and pH was very large with changes in protein concentration up to 55%. The shown influence of buffer and NaCl however is likely lower when using the ST-assay, because the ratio of sample volume (user buffer matrix) to sample buffer (pre-defined buffer from manufacturer) used in the HS-assay is 2.5 fold compared to the ratio in the ST-assay. Therefore, if the LLOQ for a given application is about 20–30 mg/L or higher, the use of the ST-assay would be recommended. Alternatively, desalting samples prior to analysis as a very fast way to equalize buffer conditions can be used to remove effects caused by the buffer matrix. However, this was not applicable in this case as the recovery of LYS from the desalting procedure was poor.

As additional aspect, the efforts for evaluating the large amount of data should be mentioned (up to four 96-well MTPs per day seems realistic). Despite of several software options it was not possible to generate a robust integration procedure for all peaks in our application. The integrated peaks were either split into several peaks or the baseline was not set appropriately. This required manual correction of the integration for the majority of the peaks, and thus, the time-benefit from the fast analysis was significantly reduced.

3.2 RP HPLC analysis

3.2.1 Initial development

The LabChip®analysis did not fulfill all the challenging requirements we had set for the analysis of the egg white protein system consisting of AV and the impurities OV, OM, OT, and LYS. In particular, the missing quantification of OM and the sensitivity towards NaCl led us continue the search for a suitable high-throughput method for the given analytical task. HPLC analysis is by far the most commonly used technique when it comes to high requirements on precision and accuracy. The major disadvantage of HPLC analysis is the required time for sample analysis, hence the challenge was to develop a HPLC method with both sufficient precision and low analysis time.

It was observed that AV eluted from RP columns in three major peaks with one broad peak eluting at retention times similar to OT and OV. By adding d-biotin in excess to the samples prior to analysis, AV eluted as one peak having the shortest retention time which made the analysis applicable for the quantification of AV and all impurities. However, AV which was not biotinylated eluted shortly after LYS. Therefore, special focus of the gradient optimization was the separation of LYS from the unbiotinylated peak which was in most cases <2% of the total AV peak area.

In a first step, three different reversed phase columns were tested with a generic gradient from 5% ACN (solvent B) to 80% ACN in purified water (solvent A) over 30 CV at approximately 70% of the maximal allowed pressure for each column (data not shown). Peak separation for all impurities was obtained with all tested columns (Waters BEH300 C4, Grace ProZap™ C18 (L=20 mm), and Merck Millipore Chromolith C18). However, the chromatograms obtained with the Chromolith column showed an abrupt step in the absorbance signal during the elution of AV and was therefore not taken into further consideration. It was decided to continue development work and gradient optimization

using the Grace column as the generated back pressure and the expenses for column and required solvent volumes were lower than for the Waters column.

Prior to gradient optimization, different TFA mass fractions between 0.05–0.2% were tested regarding the effect of the peak shape and resolution (data not shown). It was found that a TFA mass fraction of 0.1% in both solvents significantly decreased peak tailing. A further increase to 0.2% TFA did not result in further improvement of the elution profile.

For sufficient resolution between all peaks, a sequence of three linear gradients of solvent B were optimized. A total method time of 3.2 min was achieved using the 20 mm Grace ProZap™ column at a flow rate of 0.5 mL/min. However, when column load was >5 μg, the AV peak split and eluted partly with the injection peak. In order to increase the binding capacity, a Grace ProZap™ column with 30 mm column length was chosen instead. For this column, a load of up to 8 μg protein was possible without a breakthrough of AV. At a flow rate of 0.4 mL/min, a total assay time of 6.1 min was achieved with the 30 mm Grace column by applying a multi-segment gradient of solvent B of 8%–32.5% in 1.05 min, 32.5%–37.% in 1 min, 37.5%–58% in 0.8 min, and subsequent regeneration phase (100% solvent B for 1.25 min), and re-equilibration phase at 8% B (2.0 min).

3.2.2 Tandem application

The initial development had shown that a fast HPLC method for the quantification of AV and its impurities was possible. With a throughput of about ten samples per hour, the analysis time was already acceptable. However, almost 35% of the analysis time was spent for column regeneration and re-equilibration in which no relevant sample information was generated. In order to save this time, a second column was operated in tandem mode.

In Figure 7, the principle of this tandem application is displayed with the resulting chromatograms. The main principle of a tandem application is that the separation via gradient elution is performed in alternation to the regeneration and re-equilibration. While the gradient program is performed on column-1, the regeneration and re-equilibration phase takes place on column-2. As soon as all relevant peaks have eluted from column-1, the subsequent sample is injected on column-2 et cetera.

The tandem application was realized in two method programs which were running in alternation and only differed in the valve switching commands and the pump running the gradient program. With the use of a 6-port valve at the column inlets and a 6-port valve at the column outlets, the flow was directed from the injection needle in the autosampler to either column-1 or column-2 (Figure 7a). The 6-port valve at the column outlets was switched accordingly, gating the column effluent either to waste or the detector. Hence, the resulting chromatograms consisted only of relevant analytical information (Figure 7b).

The established method differs from the tandem application suggested by the manufacturer [18]. There, the aspiration of the subsequent sample, and thus, also flushing of the sample loop (100 μL) with equilibration buffer starts after elution of the previous sample has been completed. This flushing process for the used 'inline split-loop' autosampler took up to one minute in order to fully equilibrate the flow paths and, in general, becomes a significant factor when using low flow rates. In our alternative procedure, flushing the sample loop with equilibration buffer and aspiration of the subsequent sample ('prepare next sample' command) was enabled while the gradient program was still running. This was done by switching the A-valve shortly after the injection had been completed and by operating the gradient without passing the autosampler (Figure 7a,b) but directing the regeneration and re-equilibration phase through the autosampler. Thus, the sample loop was readily flushed with

equilibration buffer after about 2.5 min allowing the preparation/aspiration of the next sample (Figure 7b).

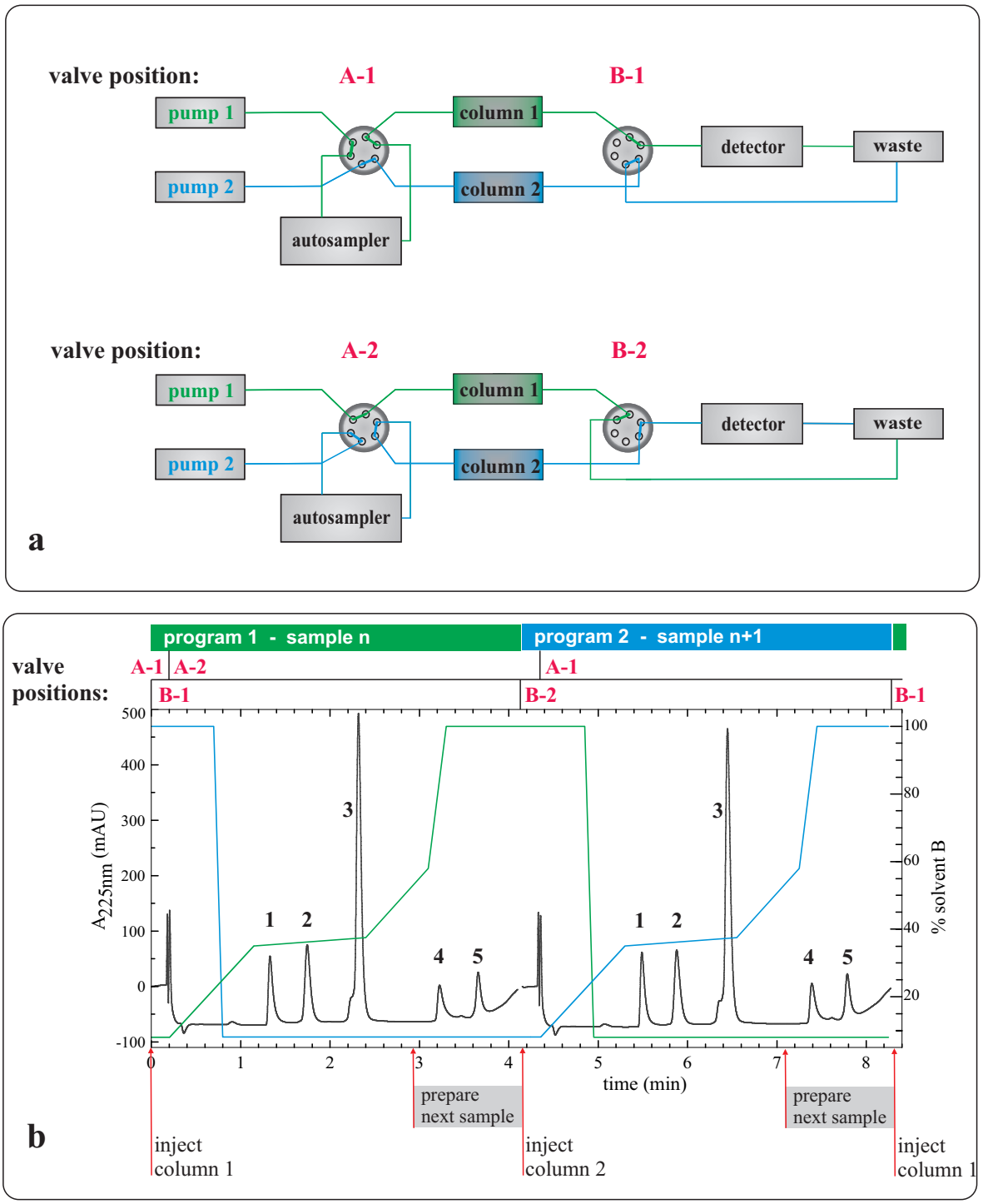

Figure 7: a: Flow scheme of the HPLC set up for a tandem application for the example of an inline split-loop autosampler and two independently switching six-port valves. b. Chromatogram of two samples analyzed in tandem mode including valve positions and gradients: mixture with concentrations of each protein of 200 mg/mL. 1: biotinylated AV, 2: OM, 3: LYS, 4: OT, 5: OV. See text for further explanation.

Method performance

The assay time was reduced from 6.10 min to 4.15 min per sample while achieving identical precision and accuracy compared to the single column method. The average relative standard deviation for protein concentrations >50 mg/L was less than 1% and increased to 2.1% for a concentration of 18 mg/L. A LLOQ was determined as ten times the detector noise at λ=225 nm and corresponded to 0.3 mAU. For AV, OV, OM, OT, and LYS the LLOQ was 0.48 mg/L, 0.94 mg/L, 0.94 mg/L, 0.6 mg/L, and 0.29 mg/L, respectively. For a linear calibration for AV in the range of 0.05–1.5 g/L (9 levels, n=5) a coefficient of determination R^2 of 0.9997 was obtained.

The sensitivity of the method regarding changes in the sample matrix was tested for 0.5 g/L AV dissolved in water, 100 mM sodium phosphate buffer with 0 mM or 500 mM NaCl. Further, diluted top and bottom phases from aqueous two-phase extraction experiments, in which salt or polyethylene glycol (PEG)1000 and PEG4000 were present with mass fractions of up to 10%, were included in this investigation. Changes in NaCl concentration and the presence of PEG4000 and salt (phosphate and ammonium sulfate) did not show a statistical effect on the mean values for AV (significance level 95%, n=6). The RSDs ranged within 1% for 500 mM NaCl and up to 2% at mass fractions of 10% ammonium sulfate or 10% PEG4000. When using PEG1000, a breakthrough of AV was observed at mass fractions of $\geq$5% PEG1000. As it was not within the scope of this work to further investigate this, it is only concluded that samples containing PEG1000 should be diluted to mass fractions of $\leq$5% prior analysis.

When using an autosampler based on a pull-loop principle instead of the inline split-loop principle, deviations in the elution signal occurred depending on the aspiration speed. Maximum precision for samples with high salt or PEG content was only achieved for full loop injections and aspiration speeds below 2 μL/s.

3.3 Method comparison

In this section, the advantages and disadvantages of both the LabChip®and the HPLC assay are discussed based on the experiences gained during the development of a quantification method for AV and its impurities and with respect to the potential use as high-throughput application. In Table 2, attributes of both methods are summarized and compared.

One major advantage of LabChip®analysis compared to HPLC analysis is the property of being a 'plug and play' tool. Method development of HPLC methods can be cumbersome and challenging if the assay times shall be reduced while still providing the required resolution and capacity. In this respect, the LabChip®clearly shows its benefits as a quick evaluation tool that certainly finds its application in early development and rapid screening experiments. The throughput requirements, however should be carefully evaluated. While run times on the LabChip®were significantly lower, the effort for integration and data evaluation were significantly higher compared to the evaluation with the HPLC software. This might partly be due to the decades of development work that has been invested in HPLC software compared to a rather new software for the LabChip®. A fast and robust integration and data evaluation is obligatory in order to fully gain the benefit of fast analysis times. For a comparison of the overall throughput one must also keep in mind that changing a chip needs to be performed manually after in average every four plates (384 samples). Thus, on a common workday 4 runs (for manual preparation of chip and samples) or up to 8 plates (when using a liquid handling station for sample preparation and using two chips) appear as a realistic maximal feasible throughput. For comparison, the HPLC could be operated continuously 24 h per day (the

Table 2: Comparison of LabChip® and RP HPLC analysis for the quantification of proteins based on the experiences gained during this study.

	LabChip®	RP HPLC/UltraHPLC
	method development	
effort	low[1]	high
considerations	reducing vs. non-reducing conditions incubation temperature	column (solid phase?) solvent system and modifier (mobile phase?) gradient program
minimal assay time	50 s	$\geq$ 2 min[2]
	data quality	
calibration	required[2]	required
mean deviation	1 – 15 % 2 – 50 % rel STD (when varying buffer, salt, pH)	0.1 – 1 % 0.1 – 2 % (when varying buffer, salt, polymer)
working range[3]	0.05 – 1.5 g/L, ST-assay 0.01 – 1.0 g/L, HS-assay	0.001 – 1.5 g/L
LLOQ	1 – 40 mg/L, ST-assay 0.5 – 20 mg/L, HS-assay	0.2 – 5 mg/L
method robustness	medium, ST-assay low, HS-assay	high
	costs	
instrument[4]	∼100.000 €	80.000 – 120.000 €
spare parts, consumption	low, low	low, medium
consumables	chips [4], 500 € reagents[4], 220 – 250 €	columns[4], 300 – 1000 € solvents[4,5], 30 – 50 €
cost per sample[6]	1.8 – 2.0 €	0.1 – 0.4 €

[1] further development for specific applications can be required (e.g. antibody aggregate determination [17])
[2] estimation of the minimal analysis time possible for gradient applications and UltraHPLC equipment
[3] concentration values provided by the software give an estimation for a concentration based on average values from ladder proteins, calibration recommended based on the results from this study, using a quadratic function
[4] status 2010
[5] for the presented method at flow rate of 0.4 mL/min
[6] LabChip®: 400 samples per chip and reagent kit/ HPLC column price 500 €, for 2000 – 10.000 injections

autosampler fitted three plates). In this study, it required about 6 h to perform the integration and evaluation of 400 samples from LabChip®analysis compared to 2 h for 400 samples with occasional re-integration using Chromeleon and further processing using a Matlab script. The actual throughput with the LabChip®therefore is only higher when also the sample preparation and incubation is fully automated.

In terms of accuracy and precision, HPLC/UltraHPLC technology poses without doubt the 'Gold-standard' for protein analysis. High precision was also achieved in this study regardless the decrease in column length and achieved reduction in assay time. Thus, a high-throughput compatible analytical method was developed which would be applicable also for later stages during process development where the requirements on an analytical method are high. In comparison to the HPLC method, the precision of the LabChip®was lower. Furthermore, the results were influenced on randomly occurring variations (e.g. ladder signals). The reason for a ladder signal changing from measurement to measurement even when using the same sample plate was not found. Relative deviations below 1% were obtained between two sample plates, whereas a third measurement showed a continuous increase of the concentration values with increasing injection number resulting in deviations up to 15%.

The major advantage of the HPLC analysis compared to the LabChip®analysis is the robustness against changes in the sample matrix such as buffer salt species, pH and high salt concentrations. It is recommended to perform a buffer exchange prior to LabChip®analysis. This might in certain cases,

however, not be possible or introduce unacceptable deviations due to insufficient recovery or sample dilution, and thus, limiting the possible applications of the LabChip®.

The sensitivity of both methods was high with LLOQ down to 0.5–10 mg/L. These values obviously depend on the protein type and other method parameters (e.g. injection volume). However, with the HPLC method, a sensitivity was achieved even higher than achieved with the HS-assay.

Concluding on the data quality, we found that LabChip®analysis could not compete with HPLC technology. The price for higher precision and robustness of the HPLC assay, however was a five-fold longer assay time. Throughput in HPLC assays might be even further increased by means of higher flow rates and tandem applications, however, assay times below 2 minutes appear unrealistic, in particular for gradient applications. It is important to note, that for fast gradient assays high pressure gradient pumps should be favored over low pressure gradient pumps in order to minimize the gradient residence time.

In terms of costs, the investment on the instrument is rather equal whereas higher costs per sample were estimated for the LabChip®application compared to costs per HPLC analysis. However, HPLC sample costs might be higher when columns are used that are more expensive or have a shorter life-time.

In summary, the LabChip®might be favorable for broad screenings and early process development, since time for method development can be saved and requirements on precision might be lower in comparison to later stages of process development. The risk of insufficient protein separation or the lack of detection of certain proteins must be taken into consideration. High accuracy and reasonable throughput can be obtained using (Ultra)HPLC with suitable columns and by taking advantage of the well-developed software packages. Robustness, precision and resolution is gained while paying time for method development.

3.4 Automation of avidin activity assay

With the HABA-assay the binding activity of avidin to biotin is quantified. The experimental steps of the HABA assay were adjusted for the 96-well MTP format and performed on a liquid handling station as described in section 2.2.3. In contrast to measurements in cuvettes, in which only the dilution by adding HABA and biotin solutions needs to taken into account, in the MTP based method, the path length changed as well after both the addition of HABA and biotin. This change was included in the calculation of the concentration of active AV (equation 1) with the path lengths determined by the method adapted from McGown [11].

3.4.1 Comparison to measurements in cuvettes

The accuracy and precision of the automated method were compared with measurements in cuvettes. Six solutions of 0.24 g/L of partly purified AV in sodium phosphate buffer at pH 7.0 with either 0 mM or 200 mM NaCl were analyzed. The identical solutions were used for both measurements in order to exclude errors introduced by preparation of sample solutions.

The results of the measurements are shown in Figure 8 activity for AV in samples without NaCl were 76% with a RSD of 3.2% determined in cuvettes and 78% with a RSD of 1.5% determined for the automated method. In presence of 200 mM NaCl, the determined activity was 72 % with a RSD of 3.2% and 75% with a RSD of 1.1% for measurements in cuvettes and on the LHS, respectively.

The 99% confidence intervals of the means overlapped indicating no significant difference between the results obtained from both methods.

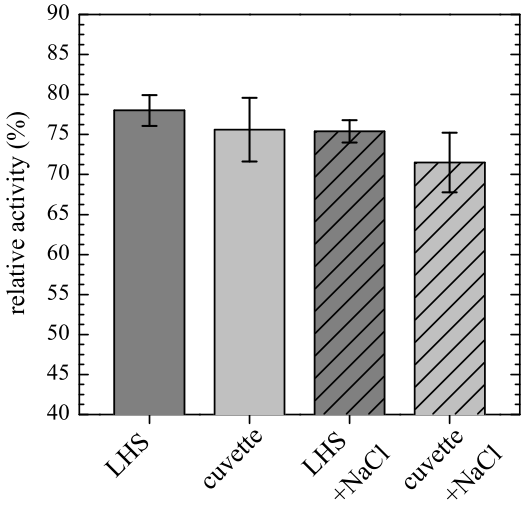

Figure 8: Comparison of AV activity determined by automated (LHS) and manual assay (cuvettes) both with and without 200 mM NaCl. Error bars display 99% confidence intervals of the mean (number of samples = 6).

Concentration working range

The working range of the HABA-assay were evaluated for AV samples of 0.048 – 0.74 g/L by performing sixfold replicates on the LHS. The determined absolute and relative activities of AV are shown as a function of the prepared (nominal) AV concentration in Figure 9. For nominal concentrations >0.15 g/L, the activity response was constant with a RSD of 3.0%. The RSD for the sixfold measurements of each concentration level was in average 1.6%. A clear decrease of both determined activity and the precision for the determination was obtained for AV samples of below a concentration of 0.15 g/L. The activity at 0.48 g/L was 54.2 % with a RSD of 34%, whereas the activity for a sample of this concentration determined in a cuvette was 66% with a RSD of 3.4%. An increase in standard deviations for sample concentrations below 0.2 g/L when measuring in MTPs has also been reported in Hansen *et al.* [19], however, does not account for the total variance. From the data shown in Figure 9, a range of 0.25 g/L – 0.75 g/L AV was considered as suitable working range for the automated assay. With the automation of all handling steps on the LHS and the parallel pipetting, 96 samples were analyzed within 45 minutes, while manual performance of this assay took about 2-3 minutes per sample. Due to the automation, not only the analysis time was reduced, but the overall manual handling was minimized to approx. 5 minutes spent for preparation of one MTP with 96 samples. The assay can be easily combined with high-throughput experiments performed on the LHS. In the current layout, the required sample volumes range within 20–100 μL resulting in a dilution of factor 12.5 – 2.5. A further reduction of the sample volumes was not performed in order to secure minimal pipetting errors. However, a further scale-down requiring only 5 – 20 μL of sample volume using half area 96-well plates or 384-well plates is possible. For further optimization of pipetting precision, coated and low-volume pipetting tips should be used.

The benefits of miniaturization and automation shown by this example can be obtained for all assays which have steps such as pipetting, mixing, incubation, and photometric analysis in common (e.g. immuno-assays, such as ELISA, or protein assays, such as Bradford assay or Biuret assay, etc.).

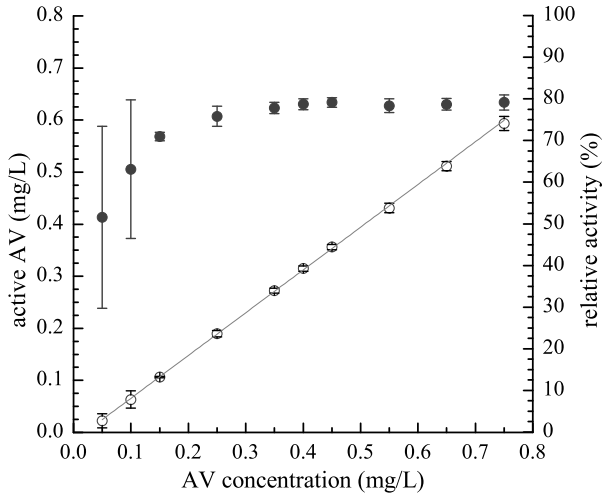

Figure 9: Absolute (open circles) and relative (black circles) activities of AV regarding the binding to biotin determined by using the HABA-assay automated on a LHS. Error bars display the 99% confidence intervals of the mean (number of samples = 8). AV samples were prepared in 100 mM sodium phosphate buffer, pH 7.0.

These steps can all be performed in a MTP format and by the use of liquid handling stations on which a plate reader for UV-VIS and fluorescence measurements are integrated.

4 Summary and conclusion

This study comprised the development and evaluation of three different analytical methods for avidin. First, chip-based capillary electrophoresis using the LabChip®GXII device was evaluated for the determination of the concentration and purity of AV. Both available assays were tested in calibration experiments showing that a quadratic correlation of nominal to measured concentration can be used in order to extend the working range of the method. The desired quantification of low-concentrated impurities LYS, OM, OV, and OT required the use of the HS-assay which was further evaluated regarding method precision and sensitivity for changes in the buffer matrix. Method precision was high considering measurements of only one plate, however, comparison of several measurements per day demonstrated the first measurement being influenced by changing ladder signals during the measurement, thus, decreasing the precision significantly. Further work should therefore focus on improving the start up procedure of the instrument and chip preparation. The method was demonstrated being sensitive towards changes in NaCl concentration, buffer species, and pH, and therefore, desalting procedures, such as using gel filtration, acetone precipitation or reversed phase trap columns, is highly recommended prior to analysis. The use of an upper marker could be beneficial when working in phosphate buffer ions, otherwise did not result in a higher precision. As furthermore OM could not be detected properly with this method, the high requirements on a quantification method set in this case study were not fulfilled by the LabChip®.

Second, a HPLC-assay was developed using a short column with small particle sizes operated on an UltraHPLC instrument. A method was developed in which two columns are operated in a tandem application. The method provided a precise and robust quantification of AV and the impurities LYS, OM, OV, and OT in approx. four minutes which makes this assay applicable to high-throughput

process development. In comparison to the LabChip®application, the costs per sample by HPLC were estimated approx. 30% of the costs for LabChip% analysis. However, the time spent for method development was significantly longer and need to be considered when deciding for a high-throughput analysis method.

Lastly, a photometric activity assay for the binding of biotin to AV was automated and performed in MTPs using a LHS. The assay provided similar results regarding accuracy and precision when compared to the original method performed in cuvettes. This assay showed the large benefit of automation when assays are simple and require only pipetting steps and photometric measurement. Future work should consider further miniaturization of the assay by working in 384-well MTPs in order to further increase sample throughput and to further decrease the required sample volume.

References

[1] C. Haber, Microfluidics in commercial applications; an industry perspective, *LabChip* 6 (9) (2006) 1118–1121.

[2] M. Bielefeld-Sevigny, AlphaLISA immunoassay platform – The "nowash" high throughput alternative to ELISA, *Assay Drug Dev. Techn.* 7 (1) (2009) 90–92.

[3] S. Chhatre, R. Francis, D. G. Bracewell, N. J. Titchener-Hooker, An automated packed protein G micro-pipette tip assay for rapid quantification of polyclonal antibodies in ovine serum, *J. Chrom. B* 878 (30) (2010) 3067–3075.

[4] S. K. Hansen, High Throughput Analytics for Protein Purification Process Development and Process Control, Ph.D. thesis, Karlsruhe Institute of Technology (2013).

[5] U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage t4, *Nature* 227 (5259) (1970) 680–685.

[6] R. J. McMahon (Ed.), *Avidin-biotin interactions: methods and applications*, 1st Edition, Humana, Totowa, NJ, 2008.

[7] C. L. Zavaleta, W. T. Phillips, A. Soundararajan, B. A. Goins, Use of avidin-biotin-liposome system for enhanced peritoneal drug delivery in an ovarian cancer model, *Int. J. Pharm.* 337 (1–2) (2007) 316–28.

[8] N. M. Green, A spectrophotometric assay for avidin and biotin based on binding of dyes by avidin, *Biochem. J.* 94 (23c-24c).

[9] N. M. Green, *Spectrophotometric determination of avidin and biotin. Methods in Enzymology*, Academic Press, New York, NY, 1970.

[10] Caliper LS, *HT Protein Express LabChip Kit, Version 2 LabChip GXII*, user Guide Rev 4 (2009).

[11] E. L. McGown, D. G. Hafeman, Multichannel pipettor performance verified by measuring pathlength of reagent dispensed into a microplate, *Anal. Biochem.* 258 (1) (1998) 155–7.

[12] M. D. Melamed, Electrophoretic properties of ovomucoid, *Biochem. J.* 103 (3) (1967) 805–810.

[13] C. Desert, C. Guérin-Dubiard, F. Nau, G. Jan, F. Val, J. Mallard, Comparison of Different Electrophoretic Separations of Hen Egg White Proteins, *J. Agric. Food Chem.* 49 (10) (2001) 4553–4561.

[14] F. T. Peters, H. H. Maurer, Bioanalytical method validation and its implications for forensic and clinical toxicology - a review, in: *Validation in Chemical Measurement*, Springer, 2005, pp. 1–9.

[15] L. Campbell, V. Raikos, S. R. Euston, Review Modification of functional properties of egg-white proteins, *Food/Nahrung* 47 (6) (2003) 369–376.

[16] E. Bayer, S. Ehrlich-Rogozinski, M. Wilchek, Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin, *Electrophoresis* 17 (1996) 1319–1324.

[17] X. Chen, G. C. Flynn, A high throughput dimer screening assay for monoclonal antibodies using chemical cross-linking and microchip electrophoresis, *J. Chrom. B* 877 (2009) 3012–3018.

[18] F. McLeod, F. Arnold, Computer Assisted Design of Column Switching Instrument Methods for Reduced Chromatography Run Times, `http://www.dionex.com/en-us/webdocs/35679-LPN%201795-02_Computer%20assisted.pdf` (2006).

[19] 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 (11) (2011) 2661–9.

MANUSCRIPT 3

Evaluation of PEG/Phosphate Aqueous Two-Phase Systems for the Purification of the Chicken Egg White Protein Avidin by using High-Throughput Techniques

Patrick Diederich, Sven Amrhein, Frank Hämmerling, Jürgen Hubbuch

Institute of Engineering in Life Sciences, Section IV:
Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

published in: *Chemical Engineering Science, 104 (2013), 945–956*

Abstract

The egg white protein avidin is known for its strong binding to biotin. Because of various applications in bioanalytics and biopharmaceutical industries, there is a great interest in the manufacturing of highly purified avidin. For the first time, the distribution behavior of avidin and the major contaminants occurring in egg white, ovalbumin, ovotransferrin (conalbumin), ovomucoid (ovotrypsin-inhibitor) and lysozyme in aqueous two-phase systems (ATPS) are presented. The results revealed a high capability of PEG/phosphate systems for the intermediate purification of the highly valuable protein avidin from chicken egg white solutions. Several PEG/phosphate systems varying in PEG molecular weight, pH and NaCl concentration were investigated in an automated small scale screening performed on a liquid handling station. The high throughput screening was used to determine binodal curve and tie-lines as well as protein distribution and recovery. Tandem Reversed Phase HPLC was applied for very fast quantification of both impurities and active (biotin binding) avidin within 4.25 minutes per sample. The main finding was that avidin partitioned in almost all ATPSs to the salt-rich bottom phase while distribution of impurities depended strongly on chosen experimental conditions. Starting from a pre-purified avidin solution containing 11% (c/c) avidin, in the most effective small-scale ATPS, a purification factor of 5.7 was obtained with an avidin yield of 92% in the bottom phase. In this particular system, a removal of ovalbumin of 65%, ovomucoid of 99.4%, lysozyme of 99.7% and ovotransferrin of 95.4% was achieved. The screening results were confirmed in common lab scale experiments representing a scale up by factor 15.

Keywords: aqueous two-phase systems, high throughput screening, tandem-HPLC, avidin, egg white proteins

1 Introduction

Avidin is a glycosylated, homo-tetrameric protein. Each subunit can bind one molecules of the vitamin biotin, which features one of the strongest non-covalent binding known in nature (dissociation rate constant ka=10^{-15} [1]). The avidin-biotin interaction has been widely used in bioanalytical assays such as ELISAs, but also makes avidin an interesting molecule for the use as receptor specific targeting or receptor specific transport of biopharmaceuticals [2]. Avidin is found in oviducts of birds, amphibians and reptiles and can be extracted from their egg whites. Besides of the extraction from egg white, the production of avidin using genetically modified organism, such as *e. coli* [3], *pichia pastoris* [4] or in transgenic plants [5] has been reported. Proteins constitute approximately 10% of the mass of chicken egg white. In this protein fraction, the predominantly occurring proteins are ovalbumin (54%), ovotransferrin (12-13%), ovomucoid (11%) and lysozyme (3.5%). Only 0.5% of the total protein is avidin [6, 7]. Extensive studies on the purification of these proteins can be found in literature as reviewed by Awade [7] or reported by Guerin-Dubiard [8] and Omana [9]. However, there is only a limited number of studies published addressing the purification of avidin.

For the purification of avidin, affinity chromatography e.g. using iminobiotin has been applied. This is, however, not favored regarding a large-scale production due to high costs and potentially short resin lifetimes. In most publications, an initial precipitation is followed by ion exchange chromatography [10]. The isoelectric point (pI) of avidin (10.0) and lysozyme (10.7) are similar [11] and therefore, an effective separation via ion exchange chromatography is challenging. Durance and Nakai [12] and Li-Chan *et al.* [13] reported of co-elution of avidin and lysozyme. However, besides precipitation, ion exchange chromatography and biotin-based affinity chromatography there has hardly been any work published in which separation techniques underlying different biochemical or biophysical mechanistics, e.g. hydrophobic interaction chromatography, were applied.

A separation using aqueous two-phase systems (ATPSs) which are composed of an aqueous polymer phase and an aqueous salt phase could pose an alternative to the chromatographic separation of avidin from egg white proteins. Protein separation in an ATPS is based on different affinity and hence distribution of proteins to either the salt-rich bottom phase or the more hydrophobic, polymer-rich top phase. ATPSs have been used effectively for the separation of a variety of proteins and are gaining increasing interest of industries for the purification of high valuable products as reviewed in literature [14, 15, 16, 17]. The technique provides several advantages such as good scalability, biocompatibility, tolerance of solid particles, solid particle removal, and selectivity. High selectivity, however, might only be identified by testing different ATPS compositions varying the polymer type and molecular weight, type of salt, tie-line length, phase ratio, concentration of neutral salts, pH or temperature. Furthermore, the protein concentration and solubility must be taken into account concerning high yields. Available models on phase formation and protein distribution do not comprehend all complexity inherent in the mechanistics of ATPS separations of proteins and therefore, the identification of suitable ATPSs is commonly still driven by heuristics. Even when applying Design of Experiments or optimization tools such as Simplex or Genetic Algorithm, ATPS process development still requires high experimental effort. The application of high throughput experimentation (HTE) enables automated screenings for phase compositions and protein distribution with a reduced amount of sample material. If combined with fast analytics, HTE offers the researcher a fast identification of suitable systems, the possibility to test more ATPSs within the same timeline and thus, a rapid, initial determination of process robustness.

This study comprises three aims. The main purpose of this study was to present an evaluation of

the distribution of the egg white protein avidin (AV) in polymer/phosphate ATPSs and focus on the differences in partitioning compared to the other egg white proteins ovomucoid (OM), lysozyme (LYS), ovotransferrin (OT) and ovalbumin (OV). While OM, LYS, OT and OV have been used as model proteins in aqueous two phase partitioning experiments in several studies [18, 19, 20, 21, 22], AV partitioning in ATPSs is reported for the first time. The second aim was to demonstrate the beneficial application of a high-throughput approach in order to investigate the effects of changes in PEG molecular weight, sodium chloride concentration and pH on the distribution and recovery. A robotic screening method modified from Oelmeier *et al.* [23] was therefore applied to perform the partitioning experiments on a liquid handling station. The third aim was to evaluate if extraction in ATPSs would pose a promising alternative to existing avidin purification process steps. Therefore, the screening was performed with a multi-component solution derived from partly purified egg white rather than using a model system of highly purified avidin. This required the application of an analytical technique different from photometric analysis, which otherwise would probably be the first choice for the analysis of the protein concentration. A tandem Reversed Phase (RP) HPLC method was developed for the fast quantitative analysis of AV and the impurities ovomucoid (OM), lysozyme (LYS), ovotransferrin (OT) and ovalbumin (OV).
With using a real egg white solution as starting material we highlight the relevance of such an HTE approach as suitable for industrial early process development. The screening identified extraction being very promising for the separation of avidin from its impurities. An effective separation of the basic proteins avidin and lysozyme is presented, which is in particular beneficial compared to the challenging separation via ion exchange chromatography. This work therefore should also be considered as a basis for a further process development for the intermediate purification of AV via two-phase separation.

2 Material and methods

Unless stated differently percentage data given in the following correspond to the mass fraction.

2.1 Materials

2.1.1 Protein Solutions

Egg white protein suspension with an avidin concentration of approximately 11% (c/c) was provided by an industrial partner. The suspension was filtered using cellulose acetate 0.45 μm filter (Sartorius, Göttingen, DE) prior to buffer exchange into 10 mM sodium phosphate (Sigma-Aldrich Co., Taufkirchen, DE) buffer (pH 7) and concentration using a PureTec CP-120 Tangential Flow Filtration system (SciLog, Middleton, WI, USA) equipped with a 3 kDa MWCO Omega™ (Minimate™ TFF Capsules, Pall, Port Washington, NY, USA). The retentate was filtered (sterile filter, 0.2 μm from Millipore, Schwalbach, DE) and stored in aliquots at -30 °C. This pre-purified avidin solution (PPAS) was equilibrated to room temperature prior use. The protein concentrations in the PPAS used in all experiments were: 0.70 g/L AV, 4.02 g/L OM, 1.54 g/L LYS, 0.26 g/L OT and 0.82 g/L OV. Standard proteins, ovalbumin (prod.-no. A5503), ovotransferrin (conalbumin, prod.-no. C0755), ovomucoid (trypsin-inhibitor, prod.-no. T9253) and lysozyme (prod.-no. 6290) were purchased from Sigma-Aldrich Co. (Taufkirchen, DE). Highly purified avidin was provided by an industrial partner.

For reference analysis standard protein stock solutions were prepared in 100 mM sodium phosphate buffer at pH 7.

2.1.2 Chemicals and stock solutions

Water was purified using an Arium®pro UV system (Sartorius Stedim Biotech, Göttingen, DE). Polyethylene glycol (PEG) was of synthesis grade and was purchased from Merck KGaA (Darmstadt, DE). PEG was dissolved in purified water to prepare following stock solutions: 60% PEG600, 60% PEG1000, 40% PEG1450 and 40% PEG4000. All salts used were purchased from Merck KGaA (Darmstadt, DE) and were of analysis grade. Salt stock solutions were prepared by dissolving the salt in water: 40% sodium phosphate (pH 3.5, pH 4 and pH 4.5), 40% potassium phosphate (pH 7), and 42 % ammonium sulfate (pH 7). The pH was adjusted by adding NaOH or HCl purchased from Merck KGaA (Darmstadt, DE). The added mass of sodium (< 1 g/L) was included into the calculation of the total mass ratio of phosphate solutions.

2.2 Protein analytics

Concentrations of AV, OV, OM, OT and LYS were determined via Reversed Phase (RP) HPLC performed on an UltiMate®3000 RSLC x2 Dual system from Thermo Fisher Scientific (Waltham, MA, USA). The system was composed of two HPG-3400RS pumps, a WPS-3000TRS analytical autosampler and a DAD3000RS detector. The column thermostat TCC-3000RS was equipped with two six-port column switching valves. Two customized Grace Vydac ProZAP!™ C18 columns (diameter x length = 2.0 mm x 30 mm) were run in tandem mode, thus providing total analysis time of 4.25 minutes per sample at a flow rate of 0.4 mL/min. Acetonitrile (ACN) containing 0.1% (v/v) trifluoracetic acid (TFA) was chosen as solvent B, while solvent A was purified water containing 0.1% (v/v) TFA. A multi-segment increase of solvent B in solvent A was applied as follows: 8%–32.5% in 1.05 min, 32.5%–37.5% in 1 min, 37.5%–58% in 0.8 min and 58%–100% in 0.15 min in to 100% was applied. While one column was regenerated and equilibrated, the next sample was analyzed on the other column (tandem application). The temperature of the column compartment was set to 25°C. A wavelength of 225 nm was used. Prior to analysis, 2 mM d-Biotin (Sigma-Aldrich Co., Taufkirchen, D) dissolved in 100 mM sodium phosphate buffer, pH 7, was added in excess (volume ratio: 1 part d-Biotin to 10 parts of sample) to every sample and standard. Thus, the purity and the concentration of biotin-binding (active) avidin were quantified in only one assay. In order to increase method accuracy, a one-phase mixture of corresponding PEG and salt (10% PEG, 5% salt) in purified water was analyzed to serve as ATPS-specific blank sample and was subtracted from sample chromatograms prior to integration. The detection limit and the lower limit of quantification (LLOQ) were estimated multiplying the noise signal from blank injections by three and ten, respectively. The LLOQ for the analyzed proteins were: AV: 0.48 mg/L, OV: 0.94 mg/L, OM: 0.94 mg/L, OT: 0.60 g/L and LYS: 0.29 mg/L. The average relative standard deviation at a concentration of 18 mg/L was 2.1% and for protein concentrations >50 mg/L less than 1%. A chromatogram of a PPAS sample is provided in the appendix of this article.

2.3 Liquid handling station

Automated ATPS experiments were performed on a Freedom Evo®200 station (Tecan, Crailsheim, DE), equipped with an 8-channel liquid handling arm (teflon coated fixed tips), a centric gripper,

a rotational shaker (Te-shake, Tecan Crailsheim, DE), an integrated centrifuge (Rotanta 46RSC, Hettich, Tuttlingen, DE), and an Infinite200 spectrophotometer (Tecan, Crailsheim, DE).

2.4 Software

The liquid handling station was controlled using Evoware 2.4 SP3. Import of pipetting values and export of photometric data was automatically performed in Evoware programs employing Excel® (Microsoft, Redmond, WA, USA) files. Chromeleon® (6.80 SR10) was used to control the UltiMate®3000 RSLC x2 Dual system and for peak integration. Data of both photometric measurements and RP HPLC analysis were further processed and evaluated using routines programmed in Matlab®R2011b (The Mathworks, Natick, ME, USA).

2.5 Two-phase characterization

2.5.1 Determination of binodal curves in microtiter plate format

All binodal curves were determined using the liquid handling station to prepare the two-phase systems. Purified water was pipetted into a standard 96-well polystyrene microtiter plate (Greiner bio-one, Frickenhausen, D), followed by salt and PEG stock solutions to obtain different combinations of PEG/salt concentrations in system volumes of 300 μL per well. The 96 solutions were mixed on a Te-shake at 1100 rpm for five minutes. Subsequently, phase formation was determined by visual inspection: Two-phase systems became opaque when being mixed while one-phase systems stayed clear. Starting from literature values or initial guesses system points were chosen at 12 different salt concentrations, each combined with eight PEG concentrations varying by maximal 1% in PEG. For each salt concentration level, a binodal point was defined as the mean value of two neighboring systems of which one exhibited the formation of two phases. Thus, sufficient precision was achieved to fit the binodal curve as a function f using following equation as given in [24] with x_{PEG} and x_{salt} respresenting the mass fractions of PEG and salt, respectively, and a, b, and c as fit parameters:

$$x_{PEG} = f_(x_{salt}) = a \cdot exp(b \cdot (x_{salt})^{0.5} + c \cdot (x_{salt})^3) \tag{1}$$

2.5.2 Determination of tie-lines in microtiter plate format

Due to the small volumes present in the microplates, phase volumes were determined similarly to the colorimetric method described in [25] and [23]. In the latter study, it was proven that tie-lines determined with the high-throughput compatible, colorimetric method were in accordance with published data that have been generated using different techniques such as gravimetric analysis, conductivity (salt) and refractive index (PEG) analysis [24, 26]. One mM methyl violet dissolved in purified water was added to every ATPS. This dye exclusively partitioned into the top phase and, thus, the exctinction measured at 586 nm in a sample of the top phase correlated to the phase volume V_{tp}. By applying the lever arm rule, which derives from mass balances (neglecting density differences of phase solutions), the tie-line lengths (L_{tp} and L_{bp}) were calculated by the ratio pr of top phase volume (V_{tp}) and bottom phase volume (V_{bp}) :

$$pr = \frac{V_{tp}}{V_{bp}} \approx \frac{L_{bp}}{L_{tp}}. \tag{2}$$

It must be noted that neglecting the different phase densities introduces an error on the determination of the exact phase compositions and slopes of the tie-lines. This error increases with increasing

mass concentrations of the phase forming components and differences in the densities of salt and polymer. The error on the concentrations in a PEG/salt system was in the range of approximately 1% – 7% (unpublished data) and therefore considered to be acceptable for the purpose of this study. A non-graphical evaluation method was established to determine tie-line slope and hence the compositions of top (x_{tp}) and bottom phase (x_{bp}) (see Figure 1). The system of equations (3-6) was iteratively solved using the *fsolve* operation in Matlab®. Equation 3 and equation 4 are true at the intersections of tie-line and binodal curve (represented as $f(x_{salt,tp})$ and $f(x_{salt,bp})$):

$$x_{PEG,tp} - f(x_{salt,tp}) \stackrel{!}{=} 0, \tag{3}$$

and

$$x_{PEG,bp} - f(x_{salt,bp}) \stackrel{!}{=} 0, \tag{4}$$

A third relation between the searched tie-line of the mixing point M and the binodal can be derived using pythagorian theorem applied to two triangles of which line segments of length L_{tp} and L_{bp} represent the hypotenuses (Figure 1):

$$\frac{(x_{salt,bp} - x_{salt,M})^2 + (x_{PEG,M} - x_{PEG,tp})^2}{(x_{salt,M} - x_{salt,tp})^2 + (x_{PEG,tp} - x_{PEG,M})^2} - pr^2 \stackrel{!}{=} 0 \tag{5}$$

The tie-line is determined when the slopes of the hypotenuses are equal:

$$\left(\frac{x_{PEG,bp} - x_{PEG,M}}{x_{salt,bp} - x_{salt,M}}\right) - \left(\frac{x_{PEG,M} - x_{PEG,tp}}{x_{salt,M} - x_{salt,tp}}\right) \stackrel{!}{=} 0 \tag{6}$$

Additional boundary conditions for the iteration procedure were set; $x_{salt,tp} < x_{salt,M} <_{salt,bp}$, $x_{PEG,tp} > x_{PEG,M} >_{PEG,bp}$.

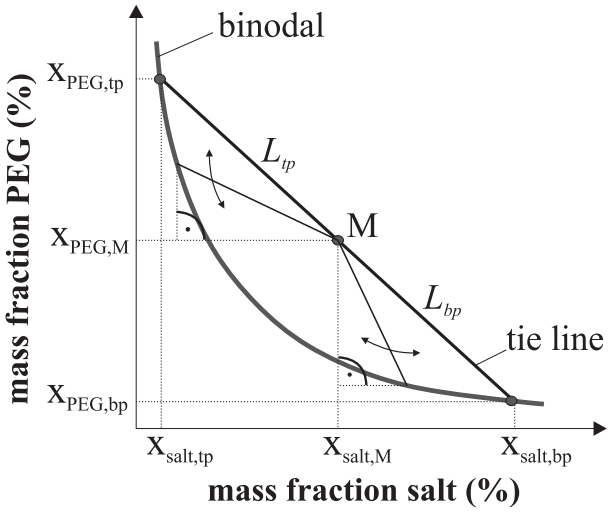

Figure 1: Schematic of tie-line determination. Two right-angled triangles can be stated using potential top and bottom phase points (intersections of binodal and tie-line) and the mixing point M. The tie-line is determined correctly, if the lengths of the hypothenuses equal the phase ratio and the slopes of the hypothenuses are hence equal. In order to determine the searched tie-lines, the phase compositions of top and bottom phase and hence the lengths of the hypothenuses are iteratively changed (as indicated by the arrows) until the solution of the system of equations stated in equations 3–6 reaches a tolerance value $< 10^{-8}$.

2.6 Two-phase separation experiments

All two-phase experiments were performed at room temperature (23°C).

2.6.1 Automated routine

Prior to experiments, the pipetting accuracy of all stock solutions was tested using an analytical balance (WXTS205DU, Mettler-Toledo, Greifensee, CH). The actual pipetted volume was calculated out of the measured mass and the density of the corresponding liquid. The deviations between pipetted volume and nominal volume were correlated by a linear function. In order to obtain the correct pipetting volumes, these relations were given to the liquid handling software. The automated procedure used to perform ATPS experiments was adapted from a method developed in our group [23]. ATPSs were prepared using the liquid handling arm to pipette stock solutions into 1.3 mL 96-well Deep Well plates (Nalgene Nunc, (Rochester, NY, USA) in the following order: water, salt solution(s) and PEG solution. The ATPSs were mixed at 1100 rpm for one minute using the Te-shake. Subsequently, 65 μL of PPAS were added to each well resulting in an ATPS volume of 650 μL. Protein dispensing was performed while mixing the plate in order to prevent phase separation prior to protein addition. The deep well plate was shaken for another 30 min at 1100 rpm. Phase separation was achieved by centrifugation for five minutes at 1800 g. Subsequently, 50 μL samples were taken from top and bottom phase using the pipetting arm according to an optimized sampling method [23] and dispensed into 200 μL purified water in sample plates. Finally, 25 μL of 2 mM d-Biotin solution were added to each sample and mixed by shaking the plates for 1 minute on the Te-shake at 1100 rpm. RP HPLC-analysis was started within two hours after the experiments had been finished. A one-phase reference system composed of 10% PEG and 5% phosphate was prepared eight-fold whenever PEG MW, NaCl concentration or pH was changed. The mean protein concentration determined in these systems served as reference for the calculation of protein recovery.
One PEG/salt-combination was analyzed within 24 hours. This included the determination of the binodal curve, tie-line determination, separation experiments in quadruplicates of 12 different PEG/salt concentrations and the RP HPLC-analysis of 96 samples in total overnight (2×4 samples of the one-phase reference system and 2×44 samples from top and bottom phase). The tie-line determination was performed in duplicate (two rows for blanks, two rows with dye) in parallel with the protein distribution experiments (four rows, i.e. 12 systems in quadruplicate) on the same plate. Furthermore, the large portion of automation in data processing by using specifically developed routines in Matlab®and Excel®templates enabled data evaluation within a few minutes. However, it must pointed out that automated integration of HPLC chromatograms within the HPLC Chromeleon®software often had to be corrected manually in order to ensure consistent integration of very small peak areas.

2.6.2 Manual experiments in lab-scale

To compare results gained from the screening study with data obtained from experiments in larger scale, some ATPSs were prepared manually in 15 mL polypropylene tubes (VWR International GmbH, Darmstadt, De). ATPSs with a total volume of 9.75 mL were pipetted manually in the same order as performed on the liquid handling station. The systems were mixed by vortexing. PPAS was then added while gently vortexing and the solutions were subsequently mixed for 30 minutes using an overhead shaker (Bibby Scientific, Stuart Rotato SB2, Stafforshire, UK). Phase separation was achieved by centrifugation at 1800 g for 5 minutes. Samples of bottom phase were taken using a

syringe cannula. Contamination of bottom phase with top phase throughout sampling was reduced to a minimum by discarding the first and last third of the sample volume taken. Sample dilution and analysis were equivalent to the procedures described in section 2.2 and 2.6.1. Phase volumes were determined in a second system prepared in a measuring cylinder, in which phase separation was obtained by gravity. As well as in robotic experiments, a reference system composed of 10% PEG and 5% phosphate was prepared and used for the calculation of recovery.

2.7 Determination of biological activity, distribution ratio, recovery and purity

The addition of d-biotin to samples prior to RP HPLC analysis allowed direct quantification of biotinylated and thus, biologically active avidin. Only the elution peak of biotinylated avidin was considered for the calculation of distribution ratio, yield and purity and is referred to as avidin in this work. The peak area of non-biotinylated relative to biotinylated avidin in the starting material PPAS was determined to be below 2%. The distribution ratio D of proteins between PEG-rich top phase and salt-rich lower phase was calculated by the ratio of protein peak area of top phase (a_{tp}) divided by peak area determined in bottom phase (a_{bp}):

$$D = \frac{a_{tp}}{a_{bp}}. \tag{7}$$

Protein mass balance, in the following referred to as recovery, was calculated by using equation 8:

$$Rec = \frac{a_{tp} \cdot V_{tp} + a_{bp} \cdot V_{bp}}{a_0 \cdot V_{system}}, \tag{8}$$

where a_0 represents the protein peak area determined by RP HPLC in a one-phase reference system composed of 10% PEG and 5% phosphate, V_{system} states the total volume of the ATPS, V_{tp} and V_{bp} represent the volumes of top and bottom phase, respectively. Recovery values give the total amount of soluble protein found in both phases. The protein yield Y in either top or bottom phase was calculated by the ratio of soluble protein mass determined in the respecting phase divided by the protein mass in the system (equation 9).

$$Y_{phase} = \frac{a_{phase,protein} \cdot V_{phase}}{a_0 \cdot V_{system}} \tag{9}$$

Purity Pu of avidin was calculated by dividing the avidin concentration ($c_{phase,AV}$) by the sum concentration of AV, OV, OM, OT, and LYS in the respecting phase:

$$Pu_{phase} = \frac{c_{phase,AV}}{c_{phase,AV} + c_{phase,OV} + c_{phase,OM} + c_{phase,OT} + c_{phase,LYS}}. \tag{10}$$

In the case of concentrations being below the level of quantification (LOQ) of the used analytical method, the corresponding LOQ was used in the calculation of Pu.

3 Results & discussion

In the following the effects of the PEG molecular weight, the NaCl concentration and the pH-value on the distribution of avidin and its impurities is presented. For each system a grid-like screening design was used. The capacity of the ATPSs, which would be an important factor

3.1 Binodal curves

The PEG/phosphate ATPSs investigated in this study varied in PEG molecular weight from PEG600 to PEG4000, NaCl concentrations from 0% to 3% and pH values between 3.5 and 7. For each system, the binodal curve function was determined and the determined coefficients of the fit equation are summarized in Table I. At least 12 data points obtained from the screening procedure described in section 2.5.1 were used to fit the binodal curve as described in [24]. The binodal curves were used to evaluate potential combinations of PEG and salt concentrations for two-phase bioseparation experiments. Furthermore, fitted binodal curves were used to determine the tie-lines as described in section (2.5.2). The effect of PEG molecular weight, the addition of NaCl and pH on binodal curves will be discussed along with the effect on protein distribution in the following sections.

Table I: Coefficients of the binodal curve for PEG/phosphate systems fitted with equation 1.

pH	PEGMW (kDa)	NaCl (% (w/w))	a	b	c	R^2
7	600	–	67.00	-0.320	$-1.1186 \cdot 10^{-4}$	0.998
7	600	3	62.42	-0.353	$-1.4255 \cdot 10^{-4}$	0.998
7	1000	–	86.11	-0.431	$-1.5615 \cdot 10^{-4}$	0.998
7	1000	3	63.21	-0.415	$-2.3168 \cdot 10^{-4}$	0.999
7	1450	–	68.67	-0.382	$-2.7594 \cdot 10^{-4}$	0.998
7	1450	1.5	56.71	-0.386	$-3.3762 \cdot 10^{-4}$	0.997
7	1450	3	52.37	-0.412	$-3.2627 \cdot 10^{-4}$	0.999
7	1450	5	45.41	-0.382	$-4.0033 \cdot 10^{-4}$	0.999
7	4000	–	83.83	-0.522	$-5.4803 \cdot 10^{-4}$	0.999
3.5	1000	–	60.00	-0.242	$-1.2873 \cdot 10^{-4}$	0.995
3.5	1000	3	71.17	-0.320	$-1.2762 \cdot 10^{-4}$	0.992
4.5	1000	–	86.50	-0.430	$-1.2168 \cdot 10^{-4}$	0.993
4.5	1000	3	59.53	-0.347	$-1.7750 \cdot 10^{-4}$	0.990

3.2 Effect of PEG molecular weight on protein separation

PEG/phosphate two-phase systems with varying PEG molecular weight (PEG600, PEG1000, PEG1450, PEG4000) were compared at pH 7 regarding the partitioning behavior of the egg white proteins present in PPAS. In Figure 2, the determined binodal curves of the tested systems along with selected tie-lines are displayed including the shortest and longest tie-line. With increasing molecular weight of PEG the amount of phosphate and PEG needed to obtain two phases decreases. This effect can be explained by a model of volume exclusion as described for polymer/polymer systems by [27]. The determined binodal curves agree very well with the binodal curves for these systems given in [27]. As expected, the determined slope of the tie-lines increased with increasing PEG MW [28]. The risk of protein precipitation is increased with increasing tie-line length, because the concentrations of phase forming components in the resulting top and bottom phases increase with increasing tie-line lengths. If PEG

MWs were changed, the experimental systems were therefore shifted along with the binodal curve to lower PEG/phosphate concentrations. The systems were chosen to feature only moderate volume phase ratios.

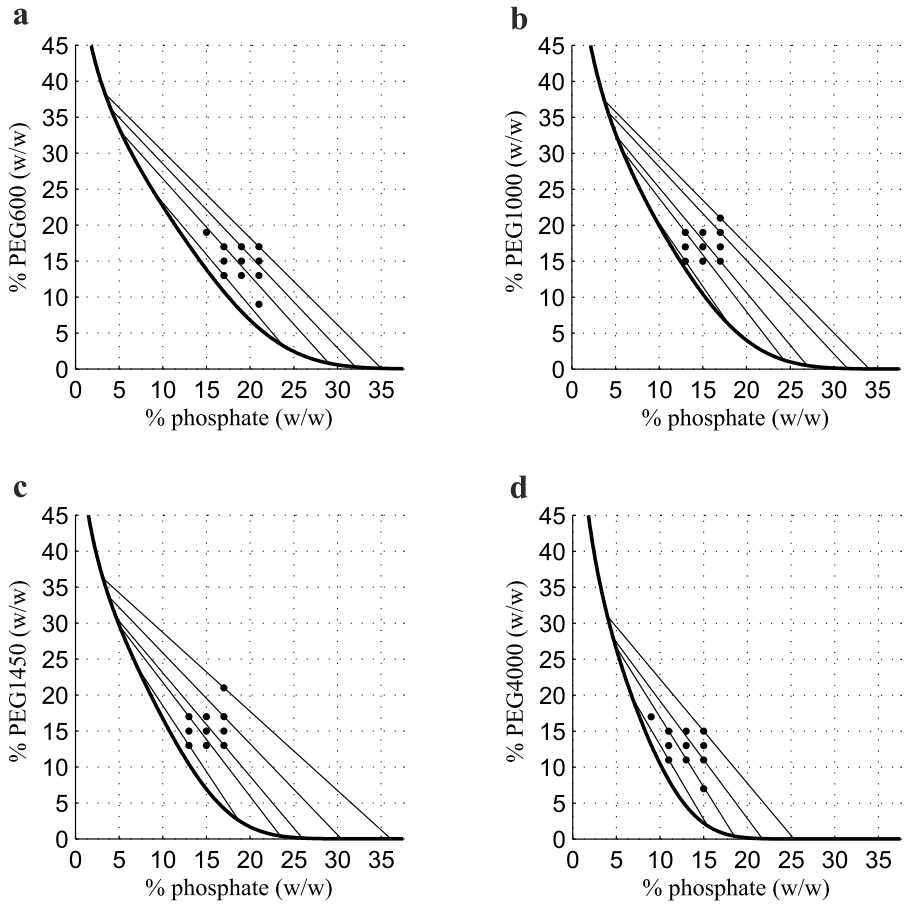

Figure 2: Binodal curves and tie-lines for ATPSs composed of phosphate and PEG600 (a), PEG1000 (b), PEG1450 (c) and PEG4000 (d) at room temperature and pH 7. Dots represent the tested ATPSs in protein separation experiments.

The difference in phase distributions and total recoveries of AV, OM, LYS, OT and OV are displayed in Figure 3. With increasing PEG MW the distribution of OM, LYS, OT decreased by factors up to 100 and partitioning was reversed from top into bottom phase. AV and OV strongly distributed into the bottom phase in all systems. However, while distribution values of OV decreased ten-fold when using PEG4000 instead of PEG600, partitioning of AV did not change significantly through all systems (mean $D < 0.04$ in all ATPS). For PEG600/phosphate, the largest difference in distribution ratios between AV and the other proteins was obtained. The observation of decreasing distribution ratio with increasing PEG MW is well known [29, 30, 31] and explained sufficiently in literature [32, 31, 33].

The distribution ratios of AV, OT and OV increased with increasing tie-line length (data not shown). This was, however, not only due to a change in partition behavior as described, but is also due to the decreased solubility of the proteins in the bottom phase with increasing salt concentrations. For example, D_{OV} increased up to a value of 1.63 while recovery was only 5%. Furthermore, the phase

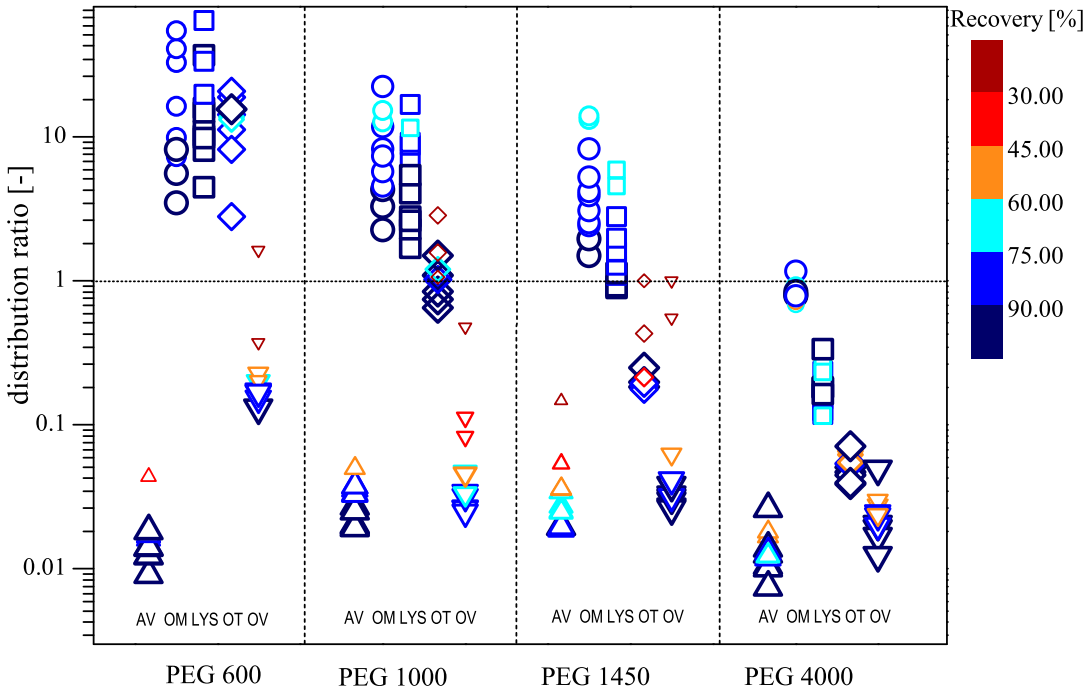

Figure 3: Distribution coefficients of avidin (AV, △), ovomucoid (OM, ○) lysozyme (LYS, □), ovotransferrin (OT, ◇) and ovalbumin (OV, ▽) in PEG/phosphate ATPSs. PEG MW was varied from 600 kDa to 4000 kDa. Differences in protein recovery values are indicated both in color code and a marker size according to the obtained recovery.

ratio showed an effect on the distribution ratio. With reducing bottom phase volume, precipitation increasingly occurred in the salt-rich phase (data not shown). For example, in an ATPS of 7% PEG4000 and 15% phosphate the recovery of all proteins was $> 89\%$, whereas at a higher phase ratio, provided in an ATPS of 15% PEG4000 and 13% phosphate, the recovery of all proteins was $< 62\%$, although both systems share the same tie-line as seen in Figure 2d.

Whether the determined differences in mean distribution values for avidin ($D_{AV} \approx 0.015$ for PEG600 and PEG4000 and $D_{AV} \approx 0.035$ for PEG1450 and PEG4000) are significant is questionable (see Figure 3). One explanation for this deviation are the different dilution ratios used to prepare samples of top and bottom phase for the subsequent HPLC analysis. Dilution of the samples was necessary, because the very high concentrations of either PEG or salt in the undiluted top or bottom phase, respectively, showed an influence on the analysis results. Samples of PEG600 and PEG4000 were diluted three-fold before analysis, while samples of experiments using PEG1450 and PEG1000 were diluted five-fold before analysis. Due to the strong affinity of avidin to the bottom phase, the avidin peak area in the top phase was often below the lower level of quantification (LLOQ) of the analytical method. In such cases, the concentrations were set to LLOQ, and thus, higher distribution ratios were calculated for the five-fold diluted samples. In the case of precipitation, concentration values were further reduced, which hence increased discrepancy between distribution ratios of five-fold diluted to three-fold diluted samples. It therefore must be kept in mind, that in all cases of low

recovery or extreme D-values, the determined distribution ratios are a function of the corresponding LLOQs, the phase ratio and the solubility.

3.3 Effect of sodium chloride concentration on protein distribution

Partitioning of hydrophobic proteins was reported to be directed into the top phase with increasing content of sodium chloride [34, 22]. Oelmeier *e al.* [23] have reported an increase in the distribution ratio of lysozyme with the addition of NaCl. The results described in section 3.2 showed avidin to strongly distribute into the bottom phase. Therefore, an increase of sodium chloride seemed promising in order to increase differences in protein distribution and to obtain higher purities of avidin in the bottom phase. The sodium chloride mass fraction was varied between 0 – 5% in ATPSs using phosphate and three different PEG molecular weights (PEG1000, PEG1450 and PEG4000). The experimental variables phosphate concentration, PEG concentration and NaCl concentration were varied according to a three level full-factorial design. A change in NaCl addition affected phase behavior and distribution coefficients in the same way independently of the PEG molecular weight (data not shown) and therefore, the effects are presented exemplary for the PEG1450/phosphate systems. The screened ATPSs, the determined binodal curves and two tie-lines of the investigated systems are displayed in Figure 4.

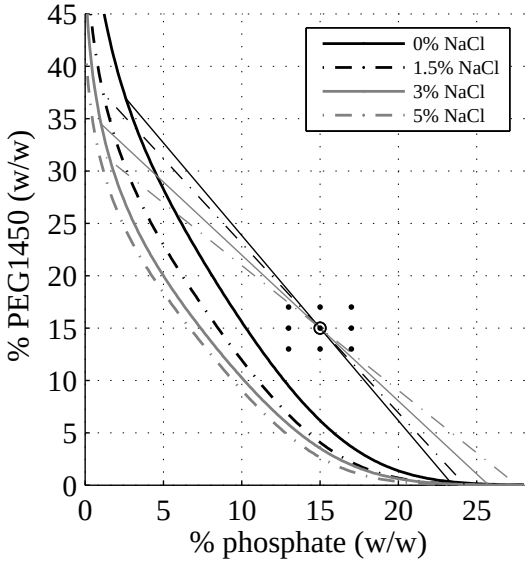

Figure 4: Influence of NaCl-concentrations on binodal curve and tie-line in PEG1450-phosphate systems at pH7, 23°C. The ATPS compositions used in protein separation experiments are displayed as dots.

The binodal curve shifted to lower PEG and phosphate concentrations with increasing content of NaCl in the solutions. Accordingly, average tie-line slopes are decreasing with increasing NaCl concentration. The shift of the binodal curve to the left within the phase diagram is in accordance with the observation reported by [24] who used PEG4000/phosphate systems at NaCl mass fractions of 0%–10%.

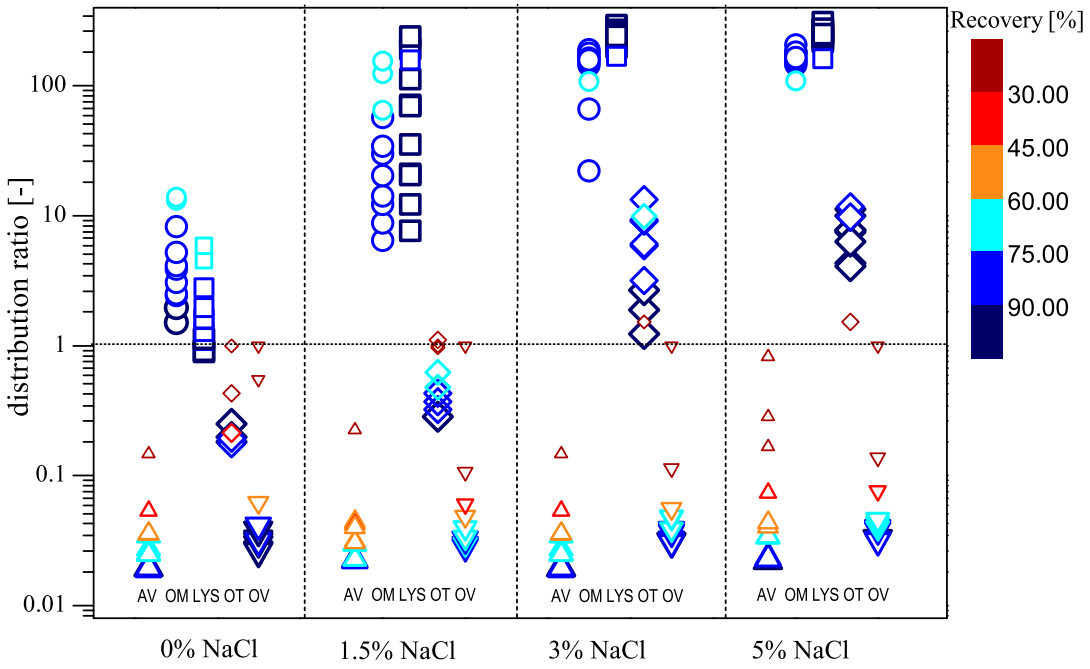

Figure 5: Influence of NaCl-concentrations on distribution ratios of avidin (AV, △), ovomucoid (OM, ○) lysozyme (LYS, □), ovotransferrin (OT, ◇) and ovalbumin (OV, ▽) in PEG1450phosphate ATPSs. Differences in protein recovery values are indicated both in color code and a marker size according to the obtained recovery.

In Figure 5, the influence of the NaCl-concentration on recovery and distribution values of AV, OM, LYS, OT and OV is displayed. In all cases, AV and OV distributed into the bottom phase providing average D-values lower than 0.1. The distribution of OM, LYS, and OT was significantly influenced by the NaCl-content. Distribution ratios for these proteins increased with increasing NaCl-concentration in average by a factor of 20. The distribution ratio of LYS increased the most from a median $D_{LYS} = 1.44$ at 0% NaCl up to a median $D_{LYS} = 240$ at 3% NaCl. A further significant increase of D_{LYS} at 5% NaCl ($D_{LYS} = 244$ (median)) was not determined. This was due to the fact that the concentration in the bottom phase was below the LLOQ at both 3% and 5% NaCl concentration (compare 3.2). The distribution values for OM reached values of about 200, but at the same time a lower AV recovery was obtained (median 80%) in comparison to LYS (median 95%). The D_{OM}-values reported in this work were hence more than 33 times higher compared with the highest D_{OM}-values de Oliveira and co-workers obtained in their work by using a PEG1500/sodium carbonate system [18]. OT distributed into the bottom phase at 0% NaCl. With addition of NaCl, the affinity of OT to the top phase increased and a median value of $D_{OT} = 7.6$ at 5% NaCl was reached. Similar to the results shown in Figure 3, the interplay of reduced protein solubility and the distribution ratios are reflected in the recovery values for the proteins AV, OT and OVA which all distributed into the salt-rich bottom phase. As can be seen from Figure 4, tie-line slope decreases with increasing NaCl concentrations resulting in increased phosphate concentration in the bottom phase.

3.4 Effect of pH on protein distribution

In the previous section it was demonstrated that by choosing low molecular weight PEG and the addition of at least 3% NaCl AV was effectively separated from OM and LYS and distributed into the bottom phase. OV was the only protein which showed the same distribution characteristics as AV. Varying the pH will in general lead to a change in proteins' surface and might induce changes in the distribution of proteins as reported for polymer/polymer two-phase systems [27] and polymer/salt ATPSs [33, 35]. Besides experiments at pH 7, partitioning of the egg white proteins was therefore also investigated in PEG1000/phosphate ATPSs at pH 3.5 and pH 4.5. Low pH values were selected based on the isoelectric point of OM (4.1) and OV (4.5) [7]. All systems were also investigated adding 3% NaCl. The determined binodal curves and a comparison of tie-lines at 17% PEG and 17% phosphate concentration are displayed in Figure 6. A decreasing pH led to a shift of the binodal curve to increasing PEG and phosphate concentrations. While the difference of the binodal curve for pH 4.5 and pH 7.0 was small, the binodal curve at pH 3.5 significantly differed regarding slope and shape. For example, at 15% phosphate, 10 % PEG was required to form a second phase at pH 4.5 and 7.0, while 15% PEG was required at pH 3.5. The observed shifting of the binodal curve to lower concentrations with increasing pH is in accordance with the findings of [29] who used a PEG1450/phosphate system at pH 7.0 and 9.0. However, [36] reported the opposite trend in ATPSs composed of PEG8000-phosphate over the range of pH 6.0–9.0. Tie-line slopes as shown for 17%/17% PEG/phosphate system did not seem to be affected by a change in pH. However, the average tie-line slope changed with increasing pH from -1.27 to -1.54 (data not shown).

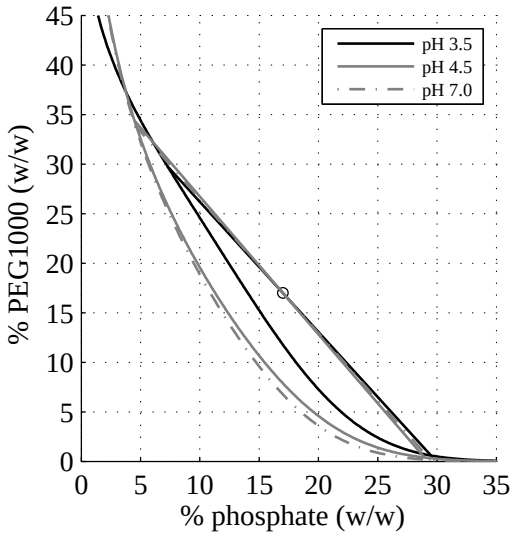

Figure 6: Influence of pH on binodal curve and tie-line in PEG1000/phosphate systems at 23°C and without NaCl-addition. Tie-line slopes and lengths are compared for an ATPS composed of 17% PEG and 17% phosphate (circle).

For the investigation of the influence of the pH on the protein distribution the system points were chosen to shift along the observed shift of the binodal curve with varying pH and 0% and 3% NaCl. In Table II, the screening results at three different pH values are summarized. In experiments without NaCl, AV and OV strongly distributed into the bottom phase and D-values increased with decreasing

pH, whereas D-values of OM, OT and LYS did not change. Although pH 4.5 was chosen near the isoelectric points of OV and OM D_{OV} and D_{OM} were not increased compared to values obtained at pH 7. Remarkably, with addition of 3% NaCl AV and OVA distributed into the top phase at pH 3.5. Here, D-values > 50 were obtained with recoveries $> 69\%$. Hence, the increase of D_{AV} and D_{OV} could not be correlated to an increased precipitation as described in previous sections. However, as seen in Table II on D-values for pH 7.0, 3% NaCl, 17% phosphate and increasing mass fractions of PEG (17%, 19 % and 21%), precipitation increased with increasing PEG concentration and significantly affected the determined D-values.

Table II: Distribution ratios and recoveries of egg white proteins in PEG1000/phosphate ATPS varying in pH, NaCl-concentration and phase ratio (pr)

system point				**avidin**		**ovomucoid**		**lysozyme**		**ovotransferrin**		**ovalbumin**	
NaCl % (w/w)	PO_4 % (w/w)	PEG1000 % (f)	pr –	D –	rec %	D –	rec %	D –	rec %	D –	rec %	D –	rec %
pH 3.5													
0	17	17	1.28	0.02	54	4.58	100	3.99	88	1.86	76	0.05	53
0	19	11	0.61	0.06	83	3.15	100	2.42	91	1.16	81	0.11	90
0	21	13	0.69	0.04	41	11.0	100	8.04	90	6.02	67	0.12	36
0	23	7	0.52	0.72	100	2.66	69	1.94	69	1.39	67	0.72	98
0	23	15	0.78	0.07	20	22.2	96	16.4	86	15.6	65	0.75	13
3	15	19	1.40	0.71	69	45.4	96	40.9	95	15.7	77	1.67	75
3	17	13	0.83	0.46	69	18.3	100	17.7	100	20.4	79	1.30	76
3	17	17	1.12	1.89	69	48.8	98	46.7	96	16.5	73	3.33	71
3	19	15	0.84	3.47	78	58.7	96	90.7	96	19.2	75	6.61	74
3	21	9	0.41	0.81	89	97.1	95	310	94	34.1	86	1.90	90
3	21	17	0.96	85.2	86	247	99	321	97	18.9	78	51.4	84
pH 4.5													
0	13	19	1.71	0.01	94	2.35	91	3.88	91	0.35	83	0.01	100
0	15	13	0.92	0.01	86	2.54	96	3.25	96	0.47	79	0.02	92
0	17	17	0.96	0.02	45	6.93	88	8.76	91	1.31	50	0.04	43
0	19	9	0.40	0.02	89	4.48	93	5.47	93	0.54	89	0.02	94
0	19	17	0.87	0.04	32	11.2	83	13.7	88	2.47	25	0.12	16
3	11	19	2.44	0.01	88	9.20	95	38.6	96	0.78	83	0.08	91
3	13	13	1.35	0.07	86	4.08	100	7.21	100	0.96	91	0.26	92
3	15	15	0.92	0.02	61	33.5	88	80.2	90	3.49	68	0.16	61
3	17	9	0.48	0.02	90	26.4	100	80.2	100	3.62	82	0.12	91
3	17	17	1.00	0.24	39	119	91	264	93	11.2	60	0.44	35
pH 7.0													
0	13	15	1.38	0.02	96	2.26	93	1.69	94	0.65	99	0.03	88
0	15	17	1.14	0.03	80	5.71	86	4.04	89	1.09	92	0.03	79
0	17	17	1.01	0.03	69	11.8	80	8.72	88	1.20	66	0.05	60
0	17	19	1.14	0.03	66	12.7	72	9.14	81	1.58	26	0.08	32
0	17	21	1.29	0.05	42	15.2	67	11.5	74	1.06	14	0.48	8
3	13	15	1.08	0.02	100	146	91	220	100	5.68	94	0.03	84
3	15	17	1.03	0.03	89	147	87	216	97	4.76	93	0.04	78
3	17	17	0.93	0.05	59	156	87	231	98	11.1	83	0.06	56
3	17	19	1.12	0.05	50	130	78	195	89	8.49	69	0.11	29
3	17	21	1.24	0.95	5	119	75	182	88	7.81	66	0.52	8

In order to compare only the effect of pH and NaCl the D-values and recoveries in ATPSs composed of 17% PEG/17% phosphate are displayed in Figure 7. D-values of AV, OM and LYS increased with increasing pH at 0% NaCl, whereas D-values of OT were highest at the lowest pH. The distribution values of OV did not change and were slightly higher than D_{AV} (0.047 compared to 0.027). An approximately 10-fold increase of D-values was observed for OM, LYS and OT with the addition of 3% NaCl, whereas the trends of D_{OM}, D_{LYS} and D_{OT} obtained with the variation of the pH were similar and independant from NaCl concentration. An interaction of the parameters pH and NaCl concentration was identified only for AV and OV. With additional 3% NaCl present in solution D_{AV} and D_{OV} increased approximately by a factor of 100 with decreasing pH. At pH 4.5 and pH 7 recoveries of AV and OV decreased when NaCl was present, whereas the opposite was true for proteins OM, LYS and

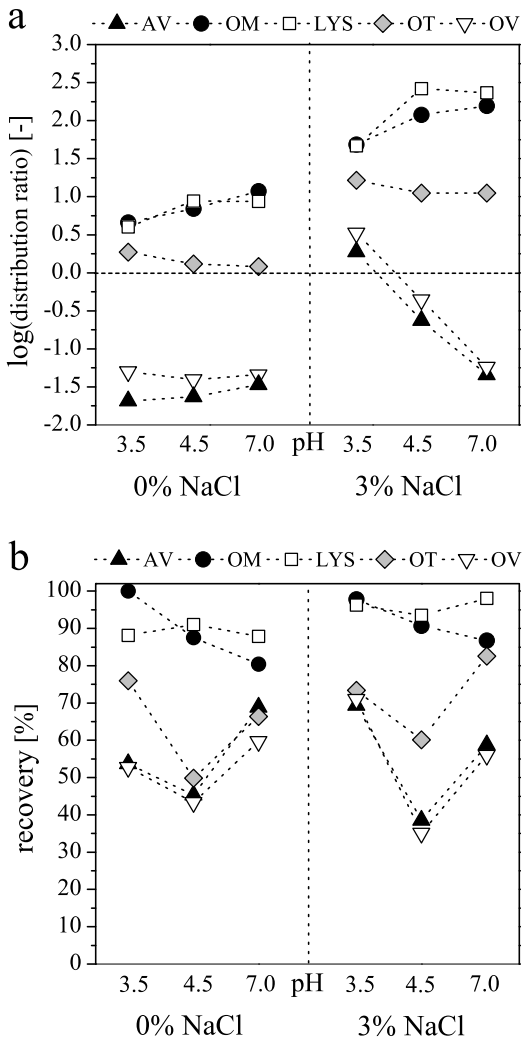

Figure 7: Influence of pH and NaCl-concentration on distribution ratio (a) and recovery (b) of egg white proteins in a 17% PEG1000/17% phosphate ATPS at 23°C. Systems were investigated at three pHs (3.5, 4.5 and 7.0), each at NaCl mass fractions of 0% and 3%. Data points are connected for better visualization.

OT. Possible explanations for the change in partitioning of AV and OV mainly into the top phase at pH 3.5 are an increased salting-out in the bottom phase and an increased shielding of proteins' surface charges by Na^+ and Cl^- ions. At this low pH conformational changes of AV and OV might lead to an increased exposure of hydrophobic surface areas and, hence, a favored distribution of the proteins into the top phase. An increase in structural flexibility for OV at low pH was reported in literature [37].

The very similar behavior regarding the distribution and recovery of AV and OV lead to the assumption of a potential interaction between these proteins. Interactions between basic protein LYS and OV has been reported [38] and might occur between the basic protein AV and OV as well. In order to test this hypothesis, experimentally derived distribution values obtained in the following two ATPSs composed of 15% PEG1000 and 13% phosphate at pH 7 using a) 0.7 g/L OV and b) 0.7 g/L AV + 0.7 g/L OV were compared. The obtained D_{OV}-values were a) 0.29 and b) 0.32 and thus, in both systems about ten times higher than the values obtained using PPAS protein solution.

The D_{AV} obtained in experiment b) was equal to the values obtained in experiments using PPAS. In an ATPS composed of 19% PEG1000 and 17% phosphate at pH 7 D_{OV} of a) 3.93 and b) 3.23 were determined, while in experiments using PPAS D_{OV} was 0.08. An interaction between AV and OV therefore seems unlikely. However, the OV-distribution observed differs between single protein experiments and PPAS experiments. In future studies, analytical tools such as static light scattering could be applied to test for egg white protein-protein interactions. Furthermore, molecular modeling might enable investigations on distribution characteristics and potential interactions between the egg white proteins on a molecular basis. The findings demonstrate that the effect of a change in pH should be evaluated in multiparameter studies rather than being investigated in a single parameter study. The high number of experiments required for multiparameter studies clearly justifies the application of high-throughput screening technology as presented in this study.

3.5 Purification of avidin

The screening results revealed large differences in the distribution of AV, OM, LYS, OT and OV in ATPSs. In almost all investigated systems AV distributed strongly into the bottom phase, whereas with decreasing PEG molecular weight and with the addition of NaCl > 3%, OM, LYS, and OT increasingly distributed into the top phase. In a final study, we investigated if a combination of low PEG molecular weight and NaCl-addition would cause a further increase of the distribution ratios of the impurities OM, LYS, OT, and OV. ATPSs composed of PEG600/phosphate with 3% NaCl at pH 7 were chosen in order to evaluate the effectiveness of a purification of AV via two-phase separation.

3.5.1 Results of robotic screening

In total, eleven different two-phase compositions regarding PEG600 and phosphate were prepared in quadruplicates on the liquid handling station. In Figure 8, a the tested ATPSs are displayed in the corresponding phase diagram. In the contour plots in Figure 8b-f, the resulting distribution ratios coded by color and the corresponding recovery values are displayed.
AV distributed strongly into the bottom phase (D_{AV} <0.15) except for the three systems with the longest tie-line lengths. The recovery values were about 100% and decreased with increasing tie-line length. This explains the higher D-values calculated for these systems. The impurities LYS and OM strongly distributed into the top phase showing distribution ratios >150 and recovery values >84% in most of the ATPSs. As observed in all previously investigated systems, the distribution pattern of OV followed that of AV resulting in the largest D-values for the systems with longest tie-line lengths. In a system composed of PEG600/phosphate/3% NaCl, D_{OV}-values >1 were obtained. In an ATPS composed of 13% PEG600 and 19% phosphate, the avidin yield in the bottom phase was 89% ± 1.2% and the purity of avidin could be increased from 11% in the starting material up to 63% ± 0.2%. In almost all systems, the AV concentration in the top phase did not succeed the LLOQ and was consequently set to the LLOQ. Therefore, the calculated D_{AV} varied only with the concentration determined in the bottom phase. Thus, D_{AV} becomes a function of the solubility in the bottom phase and the phase ratio. In fact. D_{AV} could even be smaller than the determined values. The concentrations of the impurities OM, OT, and LYS in the bottom phase were often below the LLOQ, and the D-values hence depend only on the concentration in the top phase. This concentration is decreasing with decreasing phase ratio and solubility in the top phase. Therefore, the presented D-values for LYS, OM and OT only give the minimal distribution ratios and could possibly be higher.

Nevertheless, with the ATPS conditions found, in only one batch extraction a removal of OV by 65.2%, OM by 99.4%, LYS 99.7% and OT of 95.4% was achieved.

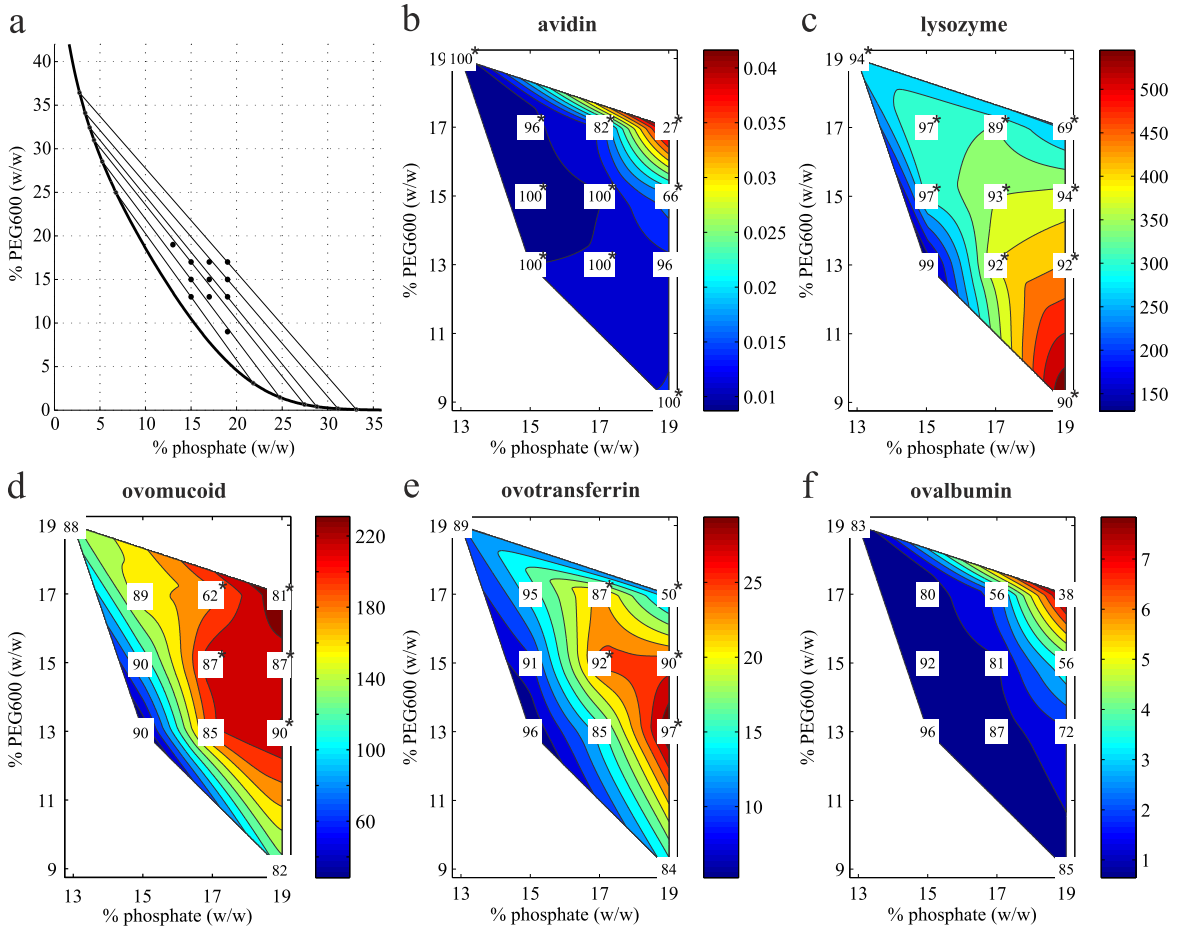

Figure 8: Screening results of PEG600/phosphate ATPSs with additional NaCl with a mass fraction of 3% NaCl at pH 7 and 23°C. a: phase diagram, tie-lines and system points (dots) chosen for screening. b-f: distribution ratios of the egg white proteins displayed in color coded contour plots. Numbers within contour plots represent the protein recovery values obtained in the corresponding ATPS. If concentration in one phase did not exceed the analytical limit of quantification, systems were marked with *. This was the case in all top phases for AV and in some bottom phases for LYS, OM and OT.

3.5.2 Comparison to lab-scale experiments

In order to verify the results obtained in miniaturized scale, the ATPS composed of 13% PEG600, 19% phosphate (pH 7) and 3% NaCl was investigated in lab-scale using a total system volume of 9.75 mL (scale-up by factor 15). In contrast to only 30 μL of protein solution required for one robotic experiment, 975 μL of protein solution was used for one experiment in falcon tubes. The ATPS experiment in lab scale was conducted in duplicate, whereas the robotic experiment was performed in quadruplicate. In Table III, the obtained distribution ratios, recoveries, and absolute errors are compared.

The overall trends of partitioning characteristics of all proteins seen in the miniaturized experiments were confirmed in larger scale. Both D-values and recoveries of LYS and OT were comparable

Table III: Comparison of AV purity and yield in the bottom phase and the distribution ratios, recoveries and absolute deviations. Results were obtained in robotic screenings using deep 96-well plates (0.65 mL) and lab scale experiments using falcon tubes (9.75 mL).

objective	**robot scale**	**lab scale**
AV purity (%)	63 ± 0.2	70 ± 0.7
AV yield (%)	92 ± 1.5	87 ± 1.2
distribution ratio (-)		
avidin	0.014 ± 0.0004	0.028 ± 0.0051
ovomucoid	234 ± 50.3	113 ± 12.4
lysozyme	442 ± 27.1	437 ± 9.6
ovotransferrin	31.6 ± 1.10	31.8 ± 0.56
ovalbumin	1.63± 0.125	4.18 ± 0.372
recovery (%)		
avidin	96 ± 4.5	89 ± 3.1
ovomucoid	94 ± 4.9	95 ± 1.4
lysozyme	92 ± 3.6	93 ± 0.9
ovotransferrin	97 ± 3.1	100 ± 1.4
ovalbumin	71 ± 14.2	83 ± 1.4

between both scales with sufficient precision (e.g. relative standard deviation of D-values in robotic experiments: 1.2% – 6.6%). AV recovery was 89% compared to 96% obtained in screening experiments. However, the determined D_{AV} in lab scale experiments was 0.028 and, thus, double as high as the value obtained in robotic experiments (0.014). The D_{OM}-value obtained in lab scale experiments was significantly smaller compared to the value determined in the robotic experiment. The phase ratio determined in robotic experiments was 0.65, whereas $pr = 0.58$ was obtained in lab scale experiments. This deviation might be due to the different methods used to determine phase volumes in both scales. The purity of AV in the bottom phase was 70% with a yield of 87%±1.3%. These values are in good agreement with the values obtained in the robotic experiments ($Pu = 63\%$ and $Y = 89\%$). It must be pointed out that in this study a natural feedstock was used, in which a large concentration range was present. The very strong partitioning into one phase results in very low protein concentrations in the other phase. This low protein concentrations mainly caused the deviations and imprecision. Small changes in the integration of the obtained HPLC peak areas (< 0.2 mAU) had a large effect on the determined value, because the peak areas for these proteins were often very small in one of the two phases. Nevertheless, changes in interfaces and mixing performance due to different container materials and geometry might also cause deviations in partitioning. Kinetic effects of precipitation competing with mass transfer into the top phase could be another reason for the determined differences for Rec_{OV} and D_{OV}. Assuming that for OV the solubility limit in the salt-rich bottom phase is reached, the protein either precipitates or favors the polymer-rich phase. The overhead shaking used for mixing in falcon tube experiments generated larger interfaces. Thus, mass transfer might be increased compared to conditions in a deep well plate, resulting in higher D-values and increased recovery. A transfer of the method to process scale should be performed in following studies, in order to evaluate the quality of the data obtained in the miniaturized experiments.

3.6 Aspects on the screening procedure

The robotic screening included the initial determination of the binodal curve, the tie-line determination and the extraction experiments. It must be noted that the binodal determination was not necessary

for the evaluation of the protein distribution. However, this proceeding supported the choice of the ATPS compositions to be tested and prevented a preparation of systems which were located in the one-phase region of the phase diagram. The determination of the tie-lines was performed in parallel to the protein partitioning experiments (see section 2.6.1). The tie-lines were not required for the calculation of the distribution coefficients and recovery values. However, the information of the top and bottom phase compositions and concentrations are relevant concerning the development of following process steps.
For the extraction experiments, a grid-like screening design was applied. Thus, first data on a possible effect of the phase ratio and the robustness of the system were obtained. This information is valuable for a comparison of the systems regarding an industrial application of the process. However, if the experimental effort needs to be further reduced, it is suggested to vary only the tie-line lengths and keep the phase ratio constant [17].

4 Conclusion and outlook

In this paper, the distribution of AV in ATPS was investigated for the first time. The study revealed large differences in partitioning of avidin and its impurities in egg white, ovomucoid, lysozyme, ovotransferrin, and ovalbumin in PEG/phosphate ATPSs. AV and OV strongly distributed into the salt-rich bottom phase, while OM, LYS, and OT could be distributed into the top phase when using low PEG MW. By the addition of 3% NaCl, a ten-fold increase in distribution values was obtained. Separation of OV from AV with a pH-value near the isoelectric point of OV was not achieved, but rather increased distribution ratios of both proteins. It was clearly demonstrated that recovery values as well as phase ratios need to be considered when comparing distribution ratios, because this value only comprises soluble protein. An alteration of the protein concentrations in the starting material can affect the degree of protein precipitation and hence protein distribution [34]. Furthermore, the capacity of the ATPS poses an important factor concerning the economic evaluation of a potential industrial process. Therefore, it will be necessary to investigate different loadings in future development studies.
The evaluation of ATPSs was performed using high-throughput experimentation combined with fast analytics. The presented screening procedure enabled a complete characterization of one ATPS within 24 h. This included the determination of the binodal curve and tie-lines, automated ATPS experiments of 12 system points in quadruplicate, RP HPLC analysis and data evaluation automated in Matlab®. Streamlining the timelines of experimentation (during the day), analytics (overnight) and rapid, largely automated data evaluation poses a concept of efficient work. With the shown example of using a complex mixture and selective analytics different from photometric analysis commonly used in HTE, we demonstrate the applicability for such screenings also for industrial process development. The automated screening identified ATPSs composed of PEG600, phosphate at pH 7 with additional 3% NaCl resulting in purities >60% and high yields (>90%) of avidin. One single batch extraction step is already remarkably effective regarding the removal of the impurities LYS, OM and OT. By applying a sequence of multiple extraction steps or continuous processing by using centrifugal partitioning chromatography, AV purity can probably be further increased. The results from batch ATPS were confirmed by the findings in commonly used lab scale experiments. The effective separation of lysozyme from avidin underlines ATPSs being a valuable alternative to ion-exchange chromatography, which according to literature is widely used for the purification of avidin. Following work will focus

on the separation of the high valuable protein AV from the mainly remaining impurity OV from the salt-rich bottom phase.

Appendix

The following Figure serves as supplementary data for section 2.2.
In Figure 1 a/b two sample chromatograms obtained with one column by the developed tandem Reversed Phase HPLC assay are displayed. In Figure 1a, the chromatogram represents the analysis of a ten-fold diluted pre-purified avidin solution (PPAS). In Figure 1b, the chromatogram represents the analysis of a one-phase reference system prepared for each corresponding ATPS. The concentration levels of the one-phase reference system correspond to the total dilutions prepared during an automated experiments. As an example for the concentrations quantified in the ATPS experiments, the area signals obtained in the reference system (Figure 1b) for avidin and ovotransferrin correspond to concentrations of 21.1 mg/L and 5.3 mg/L, respectively.

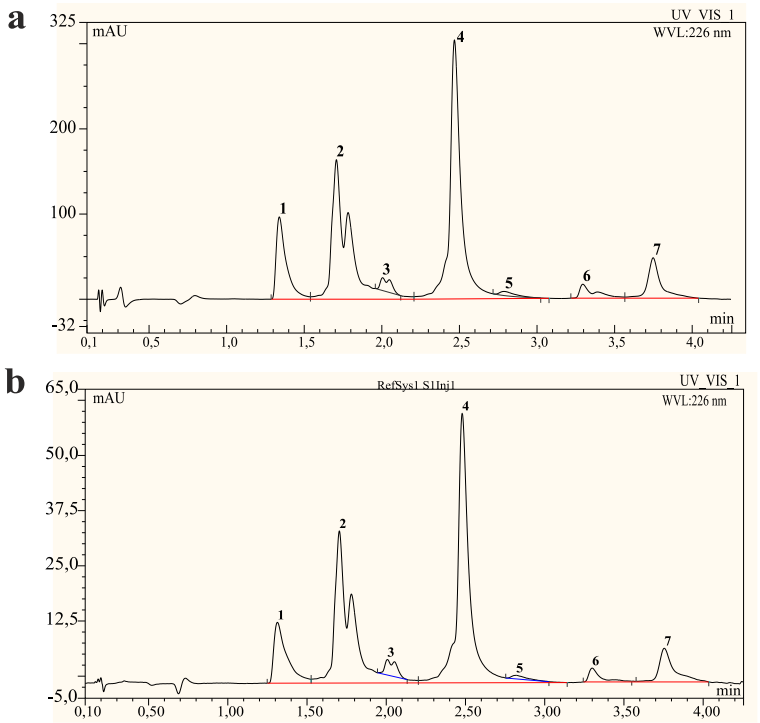

Figure 9: a: analysis of PPAS, dilution: 1:10. b: analysis of PPAS obtained in a reference system for the ATPS separation experiments (5% phosphate, 10% PEG600), dilution: 1:65. Peaks: 1: biotinylated avidin (refers to as active biotin), 2: ovomucoid, 3: non-identified component, 4: lysozyme, 5: non-biotinylated avidin, 6: ovotransferrin, 7: ovalbumin. The UV-signals shown in both figures were previously corrected by the subtraction of a blank sample.

References

[1] M. Melamed, N. Green, Avidin, *Biochemistry* 89 (1963) 591–599.

[2] R. J. Boado, Y. Zhang, Y. Zhang, C.-F. Xia, Y. Wang, W. M. Pardridge, Genetic engineering, expression, and activity of a chimeric monoclonal antibody-avidin fusion protein for receptor-mediated delivery of biotinylated drugs in humans, *Bioconjugate Chem.* 19 (2008) 731–739.

[3] V. P. Hytönen, O. H. Laitinen, T. T. Airenne, H. Kidron, N. J. Meltola, E. J. Porkka, J. Hörhä, T. Paldanius, J. a. E. Määttä, H. R. Nordlund, M. S. Johnson, T. a. Salminen, K. J. Airenne, S. Ylä-Herttuala, M. S. Kulomaa, Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli, *Biochemical J* 384 (2004) 385–90.

[4] J. Schenk, K. Balazs, C. Jungo, J. Urfer, C. Wegmann, A. Zocchi, I. W. Marison, U. V. Stockar, B. Bioscience, Influence of Specific Growth Rate on Specific Productivity and Glycosylation of a Recombinant Avidin Produced by a Pichia pastoris Mut R Strain, *Biotechnol. Bioeng.* 99 (2008) 368–377.

[5] E. E. Hood, D. R. Witcher, S. Maddock, T. Meyer, C. Baszczynski, M. Bailey, P. Flynn, J. Register, L. Marshall, D. Bond, E. Kulisek, A. Kusnadi, R. Evangelista, Z. Nikolov, C. Wooge, R. J. Mehigh, R. Hernan, W. K. Kappel, D. Ritland, C. P. Li, J. A. Howard, Commercial production of avidin from transgenic maize : characterization of transformant, production, processing, extraction and purification, *Mol. Breeding* 3 (1997) 291–306.

[6] Y. Mine, Recent advances in the understanding of egg white protein functionality, *Trends Food Sci. Technol.* 6 (1995) 225–235.

[7] A. C. Awade, On hen egg fractionation: applications of liquid chromatography to the isolation and the purification of hen egg white and egg yolk proteins, *Z. Lebensm. Untersuchung Forsch.* 202 (1996) 1–14.

[8] C. Guérin-Dubiard, M. Pasco, A. Hietanen, A. Quiros del Bosque, F. Nau, T. Croguennec, Hen egg white fractionation by ion-exchange chromatography, *J. Chrom. A* 1090 (2005) 58–67.

[9] D. A. Omana, J. Wang, J. Wu, Co-extraction of egg white proteins using ion-exchange chromatography from ovomucin-removed egg whites, *J. Chrom. B* 878 (2010) 1771–1776.

[10] D. T. Durance, Separation, purification and thermal stability of lysozyme and avidin from chicken egg white. In: Sim, J. S., and Nakai, S. (Eds.) *Egg Uses and Processing Technologies. New Developments*, International CAB, Wallingford, UK, 1994.

[11] L. Campbell, V. Raikos, S. R. Euston, Review Modification of functional properties of egg-white proteins, *Food/Nahrung* 47 (6) (2003) 369–376.

[12] T. D. Durance, S. Nakai, Simultaneous Isolation of Avidin and Lysozyme, *J. Food Sci.* 53 (1988) 1096–1102.

[13] E. Li-Chan, S. Nakai, J. Sim, D. B. Bragg, K. V. Lo, Lysozyme Separation from Egg White by Cation Exchange Column Chromatography, *J. Food Sci.* 51 (1986) 1032–1036.

[14] J. A. Asenjo, B. A. Andrews, Aqueous two-phase systems for protein separation: phase separation and applications, *J. Chrom. A* 1238 (2012) 1–10.

[15] P. A. J. Rosa, A. M. Azevedo, S. Sommerfeld, W. Bäcker, M. R. Aires-Barros, Aqueous two-phase extraction as a platform in the biomanufacturing industry: economical and environmental sustainability, *Biotechnol. Adv.* 29 (2011) 559–567.

[16] A. M. Azevedo, P. A. J. Rosa, I. F. Ferreira, M. R. Aires-Barros, Chromatography-free recovery of biopharmaceuticals through aqueous two-phase processing, *Trends Biotechnol.* 27 (2009) 240–247.

[17] J. Benavides, M. Rito-palomares, Practical experiences from the development of aqueous two-phase processes for the recovery of high value biological products, *J. Chem. Technol. Biot.* 142 (2008) 133–142.

[18] F. C. de Oliveira, J. S. dos Reis Coimbra, L. H. M. da Silva, E. E. G. Rojas, M. do Carmo Hespanhol da Silva, Ovomucoid partitioning in aqueous two-phase systems, *Biochem. Eng. J.* 47 (2009) 55–60.

[19] R. D. C. S. de Sousa, J. S. D. R. Coimbra, L. H. M. da Silva, M. D. C. H. da Silva, E. E. G. Rojas, A. A. A. Vicente, Thermodynamic studies of partitioning behavior of lysozyme and conalbumin in aqueous two-phase systems, *J. Chrom. B* 877 (2009) 2579–2584.

[20] C.-K. Su, B. H. Chiang, Partitioning and purification of lysozyme from chicken egg white using aqueous two-phase system, *Process Biochem.* 41 (2006) 257–263.

[21] J. Asenjo, A. Schmidt, H. F., A. B.A., Model for predicting the partition behaviour of proteins in aqueous two phase systems, *J. Chrom. A* 668 (1994) 47–54.

[22] B. A. Andrews, A. S. Schmidt, J. A. Asenjo, Correlation for the partition behavior of proteins in aqueous two-phase systems: effect of surface hydrophobicity and charge, *Biotechnol. Bioeng.* 90 (2005) 380–390.

[23] S. A. Oelmeier, F. Dismer, J. Hubbuch, Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation, *Biotechnol. Bioeng.* 108 (2011) 69–81.

[24] J. C. Merchuk, B. A. Andrews, J. A. Asenjo, Aqueous two-phase systems for protein separation: Studies on phase inversion, *J. Chrom. B* 711 (1998) 285–293.

[25] M. Bensch, B. Selbach, J. Hubbuch, High throughput screening techniques in downstream processing: Preparation, characterization and optimization of aqueous two-phase systems, *Chem. Eng. Sci.* 62 (2007) 2011–2021.

[26] R. Hatti-Kaul, *Aqueous Two-Phase Systems - Methods and Protocols*, Humana Press, Totowa, New Jersey, 2000.

[27] P.-A. Albertsson, *Partition of Cell Particles and Macromolecules*, 3rd Edition, Wiley-Interscience, New York, 1986.

[28] B. Y. Zaslavsky, *Aqueous Two-Phase Systems - Physical Chemistry and Bioanalytical Applications*, Marcel Dekker, Inc., New York, NY, 1994.

[29] J. G. Huddleston, K. W. Ottomar, D. M. Ngonyani, A. Lyddiatt, Influence of system and molecular parameters upon fractionation of intracellular proteins from Saccharomyces by aqueous two-phase partition, *Enzyme Microb. Technol.* 13 (1991) 24–32.

[30] H. Walter, D. E. Brooks, D. Fisher, *Partitioning in aqueous two-phase systems : theory, methods, uses, and applications to biotechnology*, Academic Press, Orlando, 1985.

[31] J. Huddleston, A. Veide, K. Köhler, J. Flanagan, S. O. Enfors, A. Lyddiatt, The molecular basis of partitioning in aqueous two-phase systems, *Trends Biotechnol.* 9 (1991) 381–388.

[32] J. N. Baskir, T. A. Hatton, U. W. Suter, Protein partitioning in two-phase aqueous polymer systems, *Biotechnol. Bioeng.* 34 (1989) 541–548.

[33] J. Huddleston, J. C. Abelaira, R. Wang, A. Lyddiatt, Protein partition between the different phases comprising poly(ethylene glycol)-salt aqueous two-phase systems, hydrophobic interaction chromatography and precipitation: a generic description in terms of salting-out effects, *J. Chrom. B* 680 (1996) 31–41.

[34] A. S. Schmidt, B. A. Andrews, J. A. Asenjo, Correlations for the partition behavior of proteins in aqueous two-phase systems: effect of overall protein concentration, *Biotechnol. Bioeng.* 50 (1996) 617–626.

[35] J.-P. Chen, Partitioning and Separation of a-Lactalbumin and fl-Lactoglobulin in PEG / Potassium Phosphate Aqueous Two-Phase Systems, *J. Ferment. Bioeng.* 73 (1992) 140–147.

[36] L. H. M. da Silva, J. S. R. Coimbra, A. Jose, D. A. Meirelles, Equilibrium Phase Behavior of Poly (ethylene glycol) + Potassium Phosphate + Water Two-Phase Systems at Various pH and Temperatures, *J. Chem. Eng. Data* 9568 (1997) 398–401.

[37] T. Koseki, N. Kitabatake, E. Doi, Conformational changes in ovalbumin at acid pH, *J Biochem.* 103 (1988) 425–430.

[38] N. Matsudomi, Y. Yamamuraand, K. Kobayashi, Heat-induced Aggregation between Ovalbumin and Lysozyme, *Agricultural Biological Chem.* 50 (1986) 1389–1395.

MANUSCRIPT 4

High-Throughput Process Development of Purification Alternatives for the Protein Avidin

Patrick Diederich, Marc Hoffmann, Jürgen Hubbuch*

Institute of Engineering in Life Sciences, Section IV:
Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

published in *Biotechnology Progress, 31 (4) (2015), 957-973*

Abstract

With an increased number of applications in the field of the avidin-biotin technology, the resulting demand for highly-purified protein avidin has drawn our attention to the purification process of avidin that naturally occurs in chicken egg white. The high-throughput process development (HTPD) methodology was exploited, in order to evaluate purification process alternatives to commonly used ion-exchange chromatography. In a high-throughput format, process parameters for aqueous two-phase extraction, selective precipitation with salts and polyethylene glycol, and hydrophobic interaction and mixed-mode column chromatography experiments were performed. The HTPD strategy was complemented by a high-throughput tandem High-Performance Liquid Chromatography assay for protein quantification. Suitable conditions for the separation of avidin from the major impurities ovalbumin, ovomucoid, ovotransferrin, and lysozyme were identified in the screening experiments. By combination of polyethylene glycol precipitation with subsequent resolubilization and separation in a polyethylene glycol/sulfate/sodium chloride two-phase system an avidin purity of 77% was obtained with a yield >90% while at the same time achieving a significant reduction of the process volume. The two-phase extraction and precipitation results were largely confirmed in larger scale with scale-up factors of 230 and 133, respectively. Seamless processing of the avidin enriched bottom phase was found feasible by using mixed-mode chromatography. By gradient elution a final avidin purity of at least 97% and yield >90% was obtained in the elution pool. The presented identification of a new and beneficial alternative for the purification of the high value protein thus presents a successful implementation of HTPD for an industrial relevant purification task.

Keywords: high-throughput process development, downstream processing, avidin, egg white protein

1 Introduction

Biopharmaceutical process development is particularly driven by tightly set timelines in order to reach toxicological studies and clinical phases as fast as possible. Therefore, within the last decade, process development scientists have adapted the methodology of 'high-throughput screening' known from drug discovery departments for their purposes. By miniaturization, parallelization and automation of experiments various process unit operations were scaled down in order to fit the micro-plate format. This enabled the operation on robotic liquid handling stations. In comparison to traditional methods, only fractional amounts of material was now required for the experiments. In the field of downstream processing, screening methods for chromatographic batch binding conditions [1–3], parallelized column chromatography in μL-scale [4, 5], solubility and precipitation screenings [6–8], and aqueous two-phase extraction [8, 9] have been developed. The evolving demand for analytical techniques matching the speed of the experimentation was responded within the last years by both academia and industry. The commercialization of hardware solutions led to wide acceptance of the technologies in industries and by regulatory authorities. The popular buzzword 'high-throughput process development' (HTPD) that Rege *et al.* stated in 2006 [10], nowadays represents the combination of high-throughput screening techniques and fast analytics and is steadily augmented with methods of mechanistic modeling, process simulation and data processing tools. Besides the application of HTPD to newly develop processes, an identification of process alternatives is clearly a promising application of the existing toolbox.

The globular glycoprotein avidin (AV) is well-known for its exceptionally strong and non-covalent binding with biotin (vitamin H) with a dissociation rate constant reported of kd=10^{-15} [11]. The pronounced interaction of biotin and AV has been applied for the detection and quantification of certain targets both in biochemical assays and for biomedical applications. Commonly, avidin and biotin are chemically conjugated to specific binders (e.g. antibodies, oligonucleotides, enzymes, etc.) and reagents such as fluorophores or enzymes to enable the detection of the target. In McMahon [12] and the reviews by Diamandis and Wilchek [13, 14] many improved methods and applications of the affinity-based tool either using avidin or the microbially-derived avidin homolog streptavidin are summarized. Additionally, an increasing interest in using the biotin-avidin technology for biopharmaceutical applications is reported [15–18]. The resulting demand on highly purified avidin asks for a re-evaluation of the manufacturing process and particularly the purification processes. Commonly, the tetrameric protein avidin is extracted from chicken egg whites, which is therefore the process we focused on in this work. Avidin constitutes only 0.05% (w/w) of the total protein in egg white [19, 20] which is one reason for the high manufacturing costs. The largest fractions of the protein mass is represented by ovalbumin (OM, 54%), ovotransferrin (OT, 12-13%), ovomucoid (OM, 11%), and lysozyme (LYS, 3.5%).

For the separation of avidin from egg white protein solutions, biotinylated ligands and iminobiotin have been used for affinity chromatography [21, 22]. The achieved high selectivity, however, comes with common disadvantages of affinity chromatography such as ligand leakage and high resin costs. The most commonly used techniques for egg white protein purification, precipitation and ion exchange chromatography [20, 23], have also been applied for the purification of avidin as reviewed by Durance et al. [24]. The published methods include a sequence of ion-exchange chromatography steps (Mealmed and Green [11]: purity 93%, yield 27%), the extension of this protocol by using gel filtration and crystallization [25], and a process using ion exchange chromatography reported by Durance and Nakai [26] (purity: 41%, yields: 74 – 80%). When using ion exchange chromatography, co-elution of avidin and lysozyme was reported [26, 27]. This is due to the similar isoelectric points of avidin

(10.0) and lysozyme (10.7) [20]. Piskarev *et al.* [28] obtained a separation of lysozyme employing a concave gradient in a final cation exchange High-Performance Liquid Chromatography (HPLC) step yielding 75%–85% avidin and purity of approximately 99% determined by SDS-PAGE. Rothemund and co-workers [29] achieved a separation of avidin from lysozyme in a two-step process using an electrophoresis instrument and obtained a recovery of 60% – 65%. However, the application of both HPLC and electrophoresis in large-scale is limited by currently available equipment.

Extraction using aqueous two-phase systems (ATPS) have been successfully applied for the purification of proteins [30–33] and promise good scalability, tolerance to solid particles and high selectivities. Recently, we have investigated the partitioning behavior of egg white proteins in polyethylene glycol (PEG)/phosphate systems regarding the effects of varying the PEG molecular weights, pH and NaCl concentrations [34]. The screening results showed favored partitioning of AV and OV into the bottom phase and distribution of LYS, OM and OT mainly into the upper phase when a low PEG molecular weight was chosen and NaCl was added to the systems.

This paper pursues two objectives. First, we intend to demonstrate the application of HTPD methodology to rapidly develop a process consisting of several steps. This approach features only little protein sample consumption per experiment and a large portion of automated experimental routines linked with speed-matching and selective tandem-HPLC analytics and automated data evaluation. Second, we present a new, alternative process layout for the purification of the high valuable protein avidin (AV) from a chicken egg white solution aiming high purities. By this combination, we hence address the lack in literature of demonstrating HTPD to other applications than monoclonal antibody platform optimizations or studies on artificial mixtures on two or three model proteins.

Based on the recently gained knowledge of the distribution of AV in ATPSs [34], in this work, we focused on the development of an integrated purification process for AV and hence on a possible separation of AV and OV out of an ATPS bottom phase. The efficiency of two-phase extraction in PEG/phosphate, PEG/ammonium sulfate ATPSs, and selective AV precipitation with PEG is demonstrated both in high-throughput experiments (300–600 μL) and in scale-up experiments (40–150 mL). The importance of tracking protein solubility in ATPSs for the correct judgment of distribution ratios is elucidated in detail. We further link the development of a purification step by aqueous two-phase separation and selective precipitation with a screening for a suitable final chromatography step (mixed-mode (MM) and hydrophobic interaction chromatography (HIC)).

2 Materials and Methods

All experiments were performed at room temperature (23°C). In the following, fractions given in percentage correspond to weight per weight (w/w) unless stated differently.

2.1 Chemicals and stock solutions

Water was purified using an Arium®pro UV system (Sartorius Stedim Biotech, Göttingen, DE). PEG 600 and PEG 1000, synthesis grade, purchased from Merck KGaA (Darmstadt, DE) was dissolved in purified water to prepare following stock solutions: 60% PEG600 and 60% PEG1000. K_2HPO_4, NaCl and $(NH_3)_2SO_4$ of analysis grade were supplied by Merck KGaA (Darmstadt, DE). The salts were dissolved in purified water to obtain following stock solutions: 40% potassium phosphate (pH 7) and 40 % ammonium sulfate (pH 7). NaCl was dissolved in purified water obtaining a 25% mass fraction. NaOH solution or H_3PO_4 purchased from Merck KGaA (Darmstadt, DE) were used

for titrating solutions to the final pH. Acetonitrile used in RP HPLC was purchased from Merck (LiChrosolv®gradient grade, Merck KGaA, Darmstadt, DE).

2.2 Protein Solutions

An egg white protein suspension containing approximately 11% (c/c) avidin was supplied by an industrial partner. This material was filtered using 0.45 μm cellulose acetate filter from Sartorius (Göttingen, DE), the buffer was exchanged to 10 mM sodium phosphate, pH 7, and the protein solution was concentrated using a PureTec CP-120 Tangential Flow Filtration system (SciLog, Middleton, WI, USA) employing a 3 kDa MWCO Omega™ membrane (Minimate™ TFF Capsules, (Pall, Dreieich, DE). The filtered retentate (sterile filter, 0.2 μm from Millipore, Schwalbach, DE) was aliquotted and stored at -30 °C until use. The final protein concentrations in this protein stock solution which was used in all experiments were: 1.27 g/L AV, 6.32 g/L OM, 2.69 g/L LYS, 0.354 g/L OT and 1.30 g/L OV. In the following, the solution is referred to as pre-purified avidin solution (PPAS). Standard proteins ovalbumin (prod.-no. A5503), ovotransferrin (conalbumin, prod.-no. C0755), ovomucoid (trypsin-inhibitor, prod.-no. T9253) and lysozyme (prod.-no. 6290) were purchased from Sigma-Aldrich Co. (Taufkirchen, DE). Purified avidin (>90%) was provided by an industrial partner.

2.3 Liquid handling station

Automated ATPS experiments were performed on a Freedom Evo®200 station (Tecan, Crailsheim, DE), equipped with following components from Tecan: an 8-channel liquid handling arm (teflon coated fixed tips), a centric gripper, a plate storage and delivery device Te-Stack™, a rotational shaker (Te-Shake™), a Te-Chrom™ chromatography module, a Te-VacS™ vacuum module for filtration, and integrated Infinite200 spectrophotometer. An integrated centrifuge was used for liquid-liquid phase separation (Rotanta 46RSC, Hettich, Tuttlingen, DE).

2.4 Software

The liquid handling station was controlled using Evoware 2.4 SP3. Import of pipetting values and export of photometric data was automatically performed in Evoware programs employing Excel®(Microsoft, Redmond, WA, USA) files. Chromeleon® (6.80 SR10) was used to control the UltiMate®3000 RSLC x2 Dual system and for peak integration. Data of both photometric measurements and RP HPLC analysis were further processed and evaluated using developed routines in Matlab®R2011b (The Mathworks, Natick, ME, USA).

2.5 Protein analytics

A high-throughput Reversed Phase (RP) HPLC method developed in an earlier study (see manuscript 3, [34]) was used for protein quantification and purity analysis. In case of protein concentrations found below the lower limit of quantification (LLOQ), the corresponding LLOQ concentration was used for calculations of distribution ratio, yield, and purity.

2.6 Aqueous two-phase experimentation

2.6.1 Automated characterization of two-phase system

Binodal curves were determined by turbidity measurements using the liquid handling station for the preparation of the two-phase systems. Purified water was pipetted into a standard 96 well polystyrene micro-titer plate (Greiner bio-one, Frickenhausen, DE), followed by salt and PEG stock solutions to obtain different systems each of 300 μL. The 96 solutions were mixed on a Te-Shake™ at 1100 rpm for five minutes. Subsequently, phase formation was determined by visual inspection. While one-phase systems stayed clear, two-phase systems became opaque when being mixed. Starting with initial guesses for similar systems derived from literature, system points were chosen for 12 different salt concentrations, each combined with eight PEG concentrations varying by maximal 1% in PEG concentration. For each salt concentration level, a binodal point was determined when in one of two neighboring systems phase formation was indicated by turbidity. Thus, sufficient precision was achieved to fit the binodal curve as a function f using following equation as given in Merchuk et al. [35]:

$$x_{PEG} = f_(x_{salt}) = a \cdot exp(b \cdot (x_{salt})^{0.5} + c \cdot (x_{salt})^3) \quad (1)$$

Tie-lines and compositions of top and bottom phases were determined using a colorimetric method adapted from [9] and described in detail in [34]. In brief: One mM methyl violet dissolved in purified water was added to the ATPSs. Since the dye exclusively partitions into the top phase, the extinction measured at a wavelength of 586 nm was linearly correlated to the top phase volume (V_{tp}). The lever arm rule deriving from mass balances was used while neglecting the density differences of the phase solutions, in order to calculate tie-line lengths (L_{tp} and L_{bp}), bottom phase volume (V_{bp}), and hence the phase ratio pr:

$$pr = \frac{V_{tp}}{V_{bp}} \approx \frac{L_{bp}}{L_{tp}}. \quad (2)$$

Having a function of the binodal curve and volumes of both top and bottom phase at hand, the unknown salt and PEG concentrations were determined from an equation system based on mass balances as described elsewhere [34, 35]. Each tie-line was determined in duplicate. Thus, four wells were occupied for the determination of one tie-line; two without dye (blank), two with dye.

2.6.2 Automated and miniaturized two-phase separation

The automated screening procedure used to perform ATPS experiments was identical to the method used in [34]. ATPSs were prepared in 1.3 mL 96-well Deep Well plates (Nalgene Nunc, Rochester, NY, USA) using the liquid handling arm. Purified water, salt stock solutions and PEG were pipetted to obtain the desired mass ratios of total volume of 650 μL per well and subsequently mixed at 1100 rpm for one minute using the Te-Shake™. While mixing, 65 μL of PPAS were added to each well. The deep well plate was shaken for another 30 min at 1100 rpm before phase separation was achieved by centrifugation for five minutes at 1800 g. Subsequently, 50 μL were pipetted from top and bottom phase according to an optimized sampling method [9] and diluted five-fold with purified water. The samples were prepared for RP HPLC by adding 25 μL of 2 mM d-Biotin solution to each well and mixed by shaking for one minute at 1100 rpm. A one-phase system composed of 10% PEG and 5% phosphate was prepared eight-fold whenever PEG MW, NaCl concentration, or pH was

changed. The mean protein concentration determined in these systems served as reference for the calculation of protein recovery. Every ATPS containing protein was prepared in quadruplicate. Thus, for each combination of PEG/salt/salt, eight wells were occupied (four for the determination of the corresponding tie-line (see 2.6.1) and four wells for protein distribution experiments).

2.6.3 Scaled-up two-phase separation

ATPSs were prepared in 250 mL Schott Duran®laboratory glass bottles supplied from VWR International GmbH (Darmstadt, DE). In the same order as performed on the liquid handling station, purified water, stock solution of phosphate or ammonium sulfate, NaCl solution and PEG stock solution were manually pipetted into the bottle. The system was mixed using magnetic stirring bar and controller (VWR International GmbH (Darmstadt, DE). After 5 minutes 15 mL PPAS was added, thus obtaining the desired mass ratios of salts and PEG in a total system volume of 150 mL. The solution was kept stirring for another 30 minutes. Phase separation was achieved by gravitation within 30 – 60 minutes. Samples of bottom phase were taken using a 5 mL pipette by mimicking the same sampling procedure as performed in the automated routine on the liquid handling station. The samples were diluted threefold and fivefold using purified water prior the addition of biotin and HPLC-analysis. Phase volumes were determined in a second system prepared in a 250 mL measuring cylinder, in which phase separation was obtained by gravity. In this system, PPAS solution was replaced by 10 mM phosphate buffer, pH 7.

2.7 Precipitation studies

2.7.1 Automated and miniaturized precipitation experiments

A high-throughput screening procedure adapted from [8] was used for protein precipitation studies. Purified water, salt stock solutions, and/or PEG solutions were pipetted into polypropylene 96-well filter plates (Bio-Inert®, Pall GmbH, Dreieich, DE) and mixed on the Te-Shake™for 5 minutes at 1100 rpm. While shaking at 800 rpm 30 μL of PPAS was added resulting in a total volume of 300 μL per well. The micro-titer plate was then sealed and shaken for 3 hours. For solid-liquid separation, vacuum was applied for 15 minutes. In the case of resolubilization of precipitate, first, a 'wash-step' was performed by adding 200 μL of fresh, protein-free precipitation solution with identical concentrations of the precipitant, subsequent shaking for 15 minutes at 1100 rpm and removal of liquid by vacuum filtration for 10 min. Subsequently, 100 μL of resolubilization solution was added to the precipitate followed by shaking for 15 minutes at 1100 rpm and vacuum filtration for 10 minutes. Prior the addition of biotin and HPLC analysis, the filtrate and the resolubilized samples were diluted threefold and the filtrate samples from the washing step were diluted twofold in purified water. For each series of experiments using one precipitant, 12 concentrations were prepared in quadruplicate. Four different precipitants were hence tested in two micro-plates and analyzed within 24 h by employing tandem HPLC analysis.

2.7.2 Scaled-up precipitation experiments

Polypropylene centrifuge tubes (50 mL) purchased from VWR International GmbH (Darmstadt, DE) were used to investigate precipitation behavior in systems of 40 mL total volume. Water and stock solutions of precipitant and additives were manually pipetted and vortexed thoroughly, before 4 mL of PPAS were added. The solution was gently vortexed and placed on a rocking shaker for mixing.

Shaking was stopped after three hours and solid particles were sedimented by centrifugation for 10 minutes at 3000 g. The supernatant was carefully decanted and diluted threefold in purified water. The procedure was repeated using 10 mL of fresh precipitation solution, and an incubation time of 15 minutes. The supernatant was decanted, diluted twofold and prepared for later analysis. The precipitate was resolubilized with 5 mL of resolubilization buffer. The tube was gently vortexed until no precipitate was visually detected any more and incubated for another 15 minutes. Samples were diluted threefold prior analysis.

2.8 Automated and miniaturized column chromatography experiments

Four resins each for hydrophobic interaction and mixed-mode chromatography were tested regarding the binding and elution behavior of the proteins. The eight MediaScout®RoboColumns supplied by Atoll GmbH (Weingarten, DE) with a column volume (CV) of 200 μL were operated in parallel using the liquid handling station. In total, 60 elution buffers of either decreasing potassium phosphate or decreasing ammonium sulfate concentration were prepared in advance by mixing purified water and a highly concentrated salt solution with the corresponding ratios on the liquid handling station. The highly concentrated salt solution was previously generated in a separate two-phase experiment using the same phase compositions than used for the corresponding two-phase experiment. The same salt solution was used for the equilibration and wash steps for 8 CV and 3 CV, respectively. Afterwards, each column was loaded with 8 CV of the protein containing bottom phase that was obtained in the two-phase separation experiment. The flow rate was 3 μL/s (54 cm/h). The elution buffers were then consecutively aspirated and 200 μL of each was dispensed into the columns, thus achieving a gradient-like salt step elution. Column regeneration was performed by dispensing 6 CV of purified water followed by 6 CV of 0.2 M sodium hydroxide. Eluate was collected in 200 μL fractions in 96-well plates. The resulting sample plates were analyzed at wavelengths of 280 nm, 900 nm and 990 nm using the plate reader integrated on the liquid handling station. Fractions with absorption values $A_{280nm} - A_{320nm} > 0.06$ were further analyzed using the HPLC assay.

2.9 Quantitative parameters

The addition of d-biotin to samples prior to RP HPLC analysis allowed direct quantification of biotinylated and thus, biologically active avidin. Only the elution peak of biotinylated avidin was considered for the calculation of distribution ratios, yields and purities and is referred to as avidin in this work.

The distribution ratio D of each protein was calculated from HPLC data by the ratio of protein peak area of the PEG-rich top phase (a_{tp}) divided by peak area determined in salt-rich bottom phase (a_{bp}):

$$D = \frac{a_{tp}}{a_{bp}}. \tag{3}$$

Protein recovery was calculated by

$$Rec = \frac{a_{tp} \cdot V_{tp} + a_{bp} \cdot V_{bp}}{a_0 \cdot V_{system}}, \tag{4}$$

where a_0 represents the HPLC protein peak area determined in a 10% PEG and 5% phosphate one-phase reference system. V_{system} states the total volume of the ATPS. The protein recovery hence represents the total amount of soluble avidin found in both phases. Density differences of the resulting phases were neglected, since for such small volumes its measurement was not feasible.

The avidin yield Y in ATPS experiments was calculated by the ratio of soluble avidin determined in the corresponding phase divided by the avidin mass in the system:

$$Y = \frac{c_{phase} \cdot V_{phase}}{c_0 \cdot V_{system}}. \tag{5}$$

The purity Pu of avidin was calculated by dividing the avidin concentration (c_{AV}) by the sum concentration of AV, OV, OM, OT and LYS in the respecting phase (equation 6. In the case the composition of a mixture was stated using the purities for each protein, c_{AV} in the numerator was replaced by the concentration of the corresponding protein.

$$Pu = \frac{c_{AV}}{c_{AV} + c_{OV} + c_{OM} + c_{OT} + c_{LYS}}. \tag{6}$$

In precipitation and resolubilization studies the protein yield and AV purity were calculated using the concentrations in and the volume of the supernatant and resolubilization buffer, respectively. Avidin purity and yield achieved in column chromatography experiments refer to the purity determined in a virtual pool of collected elution fractions and the soluble avidin found in these selected fractions.

3 Results

The purpose of this study was to use high-throughput methods in order to develop an alternative process for the purification of avidin from egg white solution. The goal was to minimize linkage steps, such as volume reduction via ultra-filtration and time-intensive buffer exchanges between process steps. Therefore, a combination of precipitation with subsequent resolubilization and extraction in ATPSs and chromatography based on multi-modal interactions was chosen. In the following sections, ATPS representing the core of the process is evaluated in terms of achieved purities and yield. In order to evaluate the potential of further purification or concentration of the target molecule, precipitation screenings were performed and will be presented in the second part of the results and discussion section. Finally, the results of a resin screening for the direct processing of an ATPS bottom phase for the separation of the mainly remaining contaminant OV will be shown.

3.1 Aqueous two-phase experiments

PEG/ammonium sulfate ATPSs and PEG/phosphate ATPSs at pH 7 using PEG600 and PEG1000 molecular weights with varying phase ratio were investigated. The governing process parameters guiding the selection process for further process development were yield, purity and concentration factor. In the context of 'green processing' relating to using bio-degradable salts and additives, citrate was also investigated. The solubilities of AV obtained in PEG/citrate ATPSs were, however, significantly lower than in phosphate or sulfate systems and therefore was not followed during further development (data not shown).

3.1.1 High-throughput ATPS screenings

Binodal curves

Prior to any ATPS screening, binodal curves were determined to map the relevant screening and process space. Additionally, the intersection of tie-line and binodal line determines the salt/PEG composition of the obtained upper and lower phase. This information on phase composition serves

as link for a seamless interplay with prior and successive steps when considering a high degree of integration. The concentrations at which the formation of a second phase was observed were fitted according to the regression function for the binodal curve Merchuk and Huddleston used in their work [35, 36]. The fit coefficients are given in Table I. In accordance with literature, the addition of 3% NaCl led to a shift of the binodal curve to lower concentrations [35, 37, 38] whereas using lower PEG molecular weight caused the opposite effect [39, 40]. A significant difference between binodal curves of PEG600/phosphate and PEG600/$(NH_4)_2SO_4$ at pH 7 including the addition of NaCl to a mass fraction of 3% was not determined (see Fig. 1a, Fig. 1g, and Fig. 2c).

Table I Coefficients of the binodal curves for PEG/ammonium sulfate systems at pH 7 fitted with equation $f(x_{PEG})=a\cdot\exp(b\cdot x_{salt}^{0.5}-c\cdot x_{salt}^{3})$. Data for PEG600/$K_3PO_4$ were taken from [34].

PEGMW kDa	salt phase % (w/w)	NaCl % (w/w)	a	b	c	R^2
1000	$(NH_4)_2SO_4$	–	83.23	-0.365	$-1.8373\cdot10^{-4}$	0.998
1000	$(NH_4)_2SO_4$	3	72.54	-0.443	$-2.1047\cdot10^{-4}$	0.998
600	$(NH_4)_2SO_4$	3	57.33	-0.307	$-1.7524\cdot10^{-4}$	0.997
600	K_3PO_4	3	62.42	-0.353	$-1.4255\cdot10^{-4}$	0.998

General screening - system evaluation

The screening for a suitable ATPS for the purification of AV was split into two screening rounds. The first screening round aimed at an evaluation of protein distribution values, obtained recovery and the determination of tie-lines. At this point prior knowledge gained from a broad screening of ATPSs with varying the PEG molecular weight, NaCl concentration and pH-value [34] was taken into account. The outcome of the first screening round incorporating five system points with increasing PEG/salt concentrations for each of the selected ATPS is shown in Table II providing mean distribution values and protein recoveries of the selected PEG/salt systems.

The obtained D_{AV}-values ranged mostly between 0.10 and 0.017 representing a favored distribution into the bottom phase. OV distributed similarly into the lower phase using PEG1000, however, in PEG600 systems D_{OV}-values above 1 were obtained. D-values of the impurities LYS, OM and OT were significantly above 1 arising from a distinct partitioning into the top phase and increased with increasing tie-line length and lower PEG MW. The addition of NaCl led to a pronounced increase of the partitioning of LYS, OM and OT into the top phase. When comparing the values for 15% PEG1000 and 17% ammonium sulfate for example, the D_{OV} increased up to 108-fold. In contrast, the influence of a NaCl-addition on partitioning of AV and OV was low with an increase by factor 1.07 and 1.5, respectively. The positive effect of NaCl addition and lower PEG MW was also observed for the partitioning of these proteins in PEG/phosphate systems [34] and are in accordance with observations for different proteins and PEG/salt systems in literature [9, 41, 42]. With respect to a potential purification process using ATPS, both salt/PEG600/3% NaCl systems appeared promising as large differences in distribution between AV and its impurities were determined.

The median of the avidin recovery values was >90% and decreased with increasing PEG and salt concentrations which corresponds to an increasing tie-line length. With increasing PEG and salt concentration, the recovery values for the impurities decreased for salt and PEG concentrations above 17% and 15%, respectively. The recovery slightly decreased with addition of NaCl. When using phos-

Table II Distribution ratios and recoveries of egg white proteins avidin (AV), ovalbumin (OV), lysozyem (LYS), ovotransferrin (OT), and ovomucoid (OM) in ATPS with different phase ratios (*pr*) at pH 7. The systems are listed in the order of increasing tie-line length. Data for PEG600/potassium phosphate were taken from [34].

system point			AV		OV		LYS		OT		OM	
salt	PEG	pr	D	Rec	D	Rec	D	Rec	D	Rec	D	Rec
% (w/w)	% (w/w)	–	–	%	–	%	–	%	–	%	–	%
PEG1000/ammonium sulfate												
15	13	0.63	0.015	100	0.022	100	1.86	97	0.610	100	2.97	98
15	15	0.7	0.013	100	0.020	100	2.25	93	0.085	91	3.71	93
17	15	0.61	0.015	100	0.022	100	4.39	95	1.97	74	7.87	95
19	15	0.56	0.014	100	0.026	94	10.3	92	11.0	55	27.1	88
19	17	0.68	0.014	100	0.040	56	19.43	92	10.7	43	50.0	88
PEG1000/ammonium sulfate/3% NaCl (w/w)												
17	11	0.42	0.014	100	0.029	83	346	96	12.9	85	74.0	93
17	13	0.49	0.017	96	0.030	78	519	97	14.8	87	195	94
17	15	0.60	0.016	95	0.033	70	473	96	22.5	83	277	93
19	15	0.57	0.016	92	0.067	38	492	97	22.1	80	320	92
21	15	0.57	0.048	33	0.530	6	521	100	20.4	74	382	98
PEG600/ammonium sulfate/3% NaCl (w/w)												
17	11	0.45	0.016	100	1.46	92	91.0	100	32.1	100	68.7	95
17	13	0.57	0.016	95	2.73	88	427	96	26.5	100	193	92
17	15	0.69	0.015	93	6.45	87	433	98	22.8	100	324	95
19	15	0.62	0.013	88	1.28	75	453	96	23.8	100	329	89
21	15	0.57	0.018	58	4.01	74	466	94	23.2	100	323	91
PEG600/potassium phosphate/3% NaCl (w/w)												
15	13	0.83	0.11	100	0.66	96	130	99	5.23	96	28.1	90
15	15	1.00	0.010	100	0.81	92	313	97	8.99	91	107	90
17	15	1.17	0.010	100	0.87	81	324	93	15.1	92	172	87
19	15	0.78	0.018	66	3.90	56	386	94	26.2	90	285	87
19	17	0.94	0.044	27	8.50	38	262	69	12.3	50	249	81

phate the recovery values were slightly reduced in comparison to the sulfate systems.

Detailed screening - influence of phase ratio

The largest differences between the distribution values of AV and the impurities were obtained in systems of PEG600/sulfate/NaCl and were in the same range when compared to D-values reported for PEG600/phosphate/NaCl (table II). Therefore, both salt systems were used in combination with PEG600 and 3% NaCl in a second, more detailed screening. The purpose of this screening was to systematically investigate the effect of phase ratio variations on the distribution and recovery of AV compared to the impurities. Based on the initial screening results shown, three tie-lines for each of the two salt/PEG600/3% NaCl systems were selected and system points with phase ratios (*pr*) 4, 1, and 0.25 were analyzed for protein distribution and yield. This screening scheme was complemented by two system points with a *pr* ≈1 located on tie-lines shorter and longer than the other three.

The phase diagrams, distribution values, and recoveries of AV and the four impurities OV, LYS, OM, and OT are displayed in Figure 1. The affinity of the proteins to the top phase decreased in the following order: LYS > OM > OT > OV > AV. An increase in the tie-line length led in general to an increase in the D-values. With an increasing phase ratio the distribution values for all proteins were reduced. The largest D-values for all impurities were obtained at 7.9% PEG600 and 23.9% sulfate at $pr = 0.25$. The largest D_{AV} was determined in a PEG600/sulfate ATPS with a short tie-line length (composed of 7.6% PEG600, 20.5% sulfate) at $pr = 0.25$. The lowest D_{AV}-values (≤0.01) were obtained at $pr = 4$. The effects of varying the phase ratio and changing the tie-line length appeared to be additive, however, differently pronounced for each protein. Results from sulfate systems were in very similar range when compared to the results from phosphate systems. However, as can be seen in the contour plots for D_{LYS} in Figures 1d and 1j, the tie-line length had a more pronounced effect on

D_{LYS} in sulfate systems than in phosphate systems. Furthermore, the distribution values of AV at low phase ratios increased with increasing tie-line length when using phosphate, whereas the opposite trend is true for low phase ratios when using sulfate.

Matching the result from initial screenings (Table II), the recovery values decreased generally with increasing tie-line length. However, this effect was more pronounced for the proteins AV and OV when compared to the recoveries of LYS, OM, and OT which were >85% in all systems. In comparison to AV, even lower recovery values were obtained for OV when analyzing the results at $pr = 1$. In the phosphate ATPS with the longest tie-line (composed of 17.9% PEG600, 17.8% phosphate), only 25% AV was recovered. In comparison, the AV recovery values in the sulfate ATPS with the longest tie-line (composed of 20.0% PEG600, 16.6% phosphate) was 92%. The average relative standard deviation for the recovery values was 6.6%.

The data given in Figure 1 presenting distribution characteristics of AV and impurities were translated to the achieved AV purity and yield in order to guide further process development (see Figure 2a and b). The obtained contour plots depict a similar purity pattern for both PEG600/salt systems. The Pu was increased with increasing tie-line length and with increasing phase ratio. The Pu in sulfate systems was however about 10% higher when compared to phosphate systems. The highest Pu was achieved in the systems located on the longest tie-line. In these systems the lowest Y was obtained.

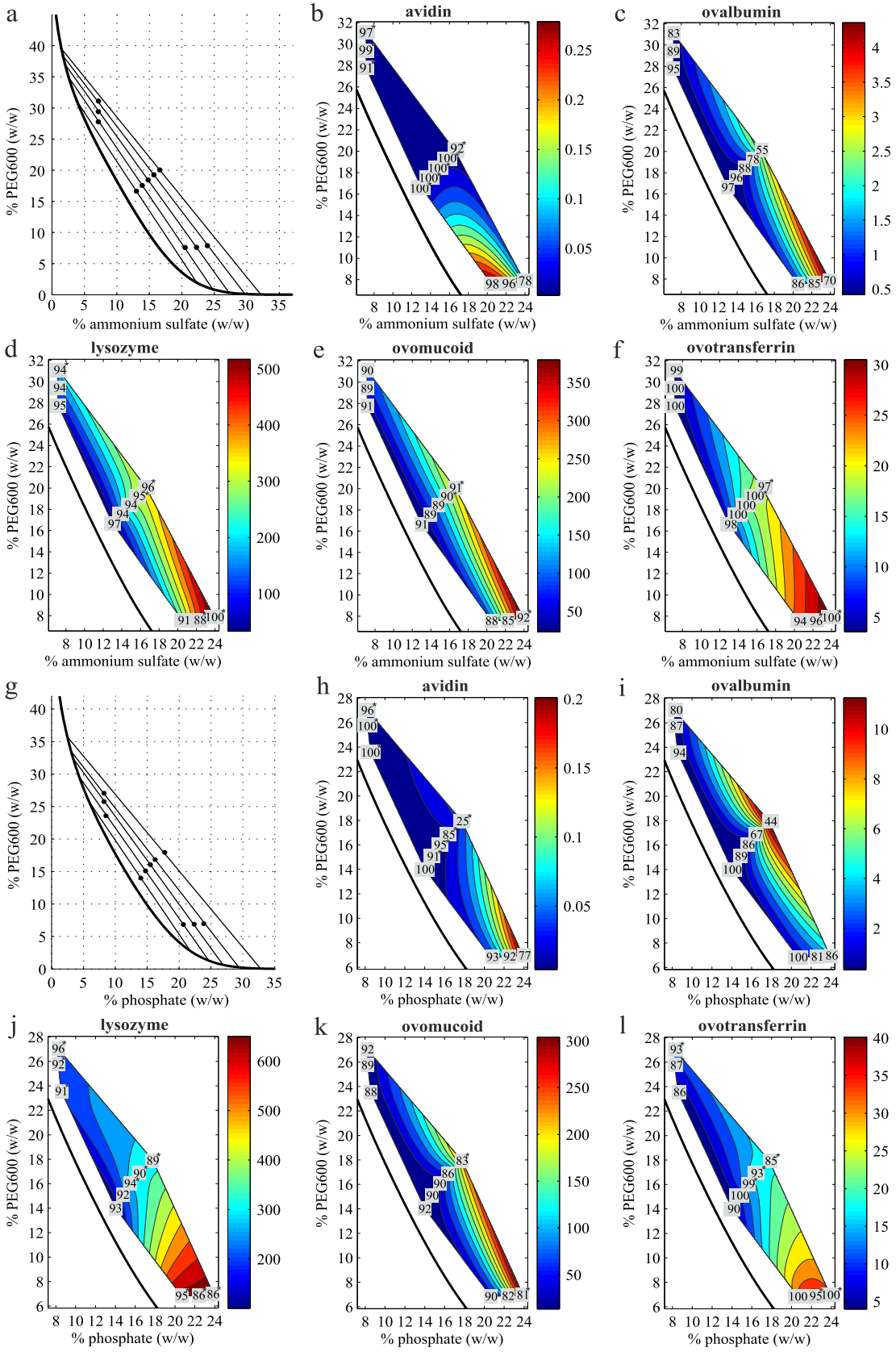

Figure 1: a and g: Phase diagrams for PEG600/sulfate and PEG600/phosphate ATPSs each with 3% NaCl, pH 7. Dots represent the investigated systems. b–f and h–l: Contour plots of distribution values (color coded) and the corresponding recovery values in percent obtained in the corresponding system. Systems marked with * indicate that the concentration in one phase did not exceed the analytical LLOQ.

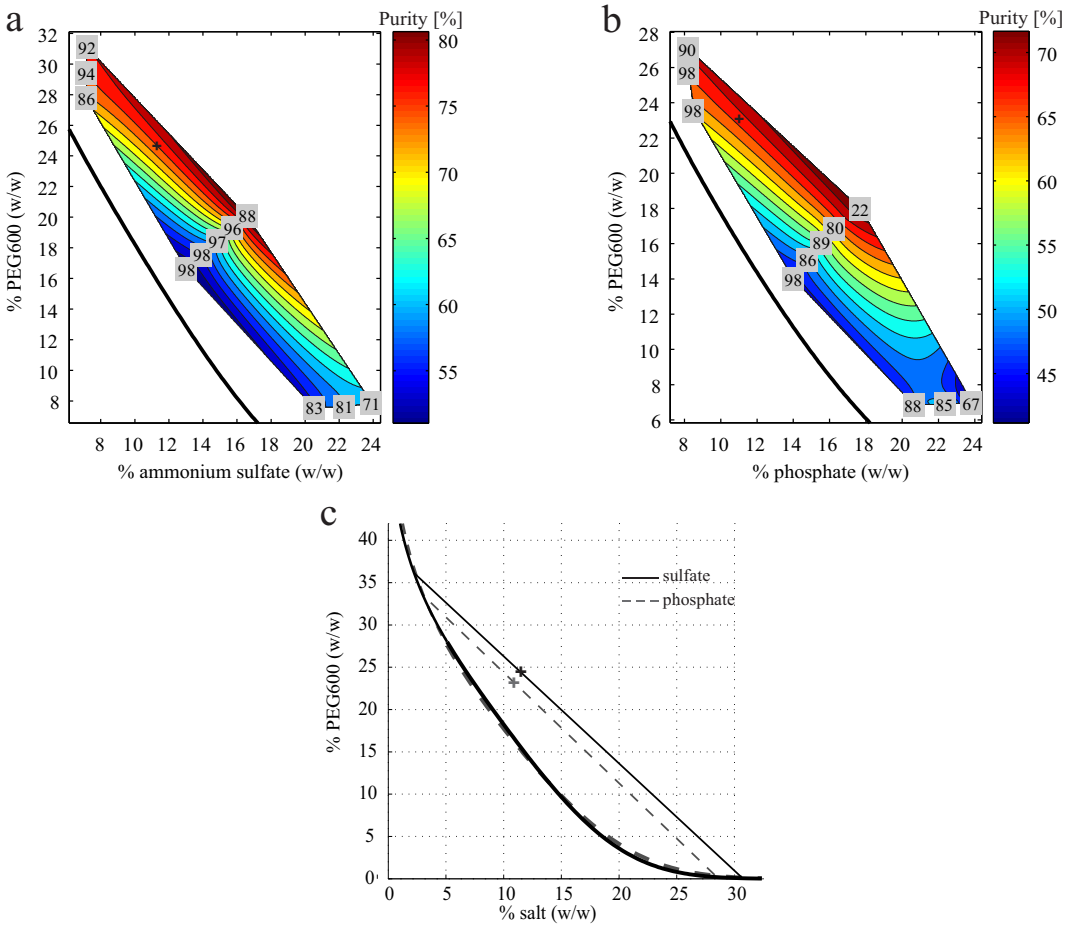

Figure 2: Avidin purity and yield obtained in PEG600/ammonium sulfate (a) and PEG600/phosphate (b) two-phase systems with 3% NaCl at pH 7. The obtained yields given in percentage are located at the corresponding systems. c: binodal curves and tie-lines of the ATPSs selected for scale-up.

3.1.2 System selection and scale-up

In order to demonstrate comparability between high-throughput experimentation and common laboratory experiments, suitable ATPSs were selected for scale-up by factor 230. The chosen system compositions are displayed in Figure 2c. For the selection, the largest tie-line lengths were considered at which no AV precipitation was observed in the screening results. Therefore, higher PEG and salt concentrations were chosen for the sulfate system compared to concentrations used in the phosphate systems. Furthermore, it was decided to calculate system points at an approximately pr=3 within the linearly interpolated AV Y- and Pu-plane. The intention behind this decision was to evaluate if interpolated data from the high-throughput screening could be used to sufficiently predict the performance of any other ATPS in the experimental landscape. It must be pointed out that for this reason, not necessarily the truly optimal systems were selected in regards of the best process performance.

In analogy to the robotic experiments, PPAS was added to homogeneously mixed stock solutions resulting in 150 mL of 11.5% PEG600, 24.4% ammonium sulfate, 3% NaCl and 10.9% PEG600, 23.2% potassium phosphate, 3% NaCl, respectively. Phase separation was reached by gravitation. It took about 75 minutes until a clear bottom phase and distinct interphase was obtained. For comparison, phase separation in a reference system using water instead of PPAS was completed already after 10

minutes. The obtained purities and yields in the bottom phases calculated for each protein are given in Figure 3.

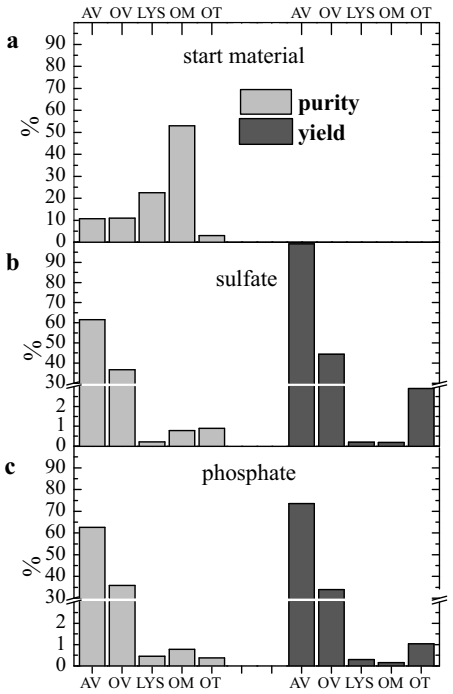

Figure 3: Purities and yields obtained in the bottom phases of 150 mL ATPSs composed of 24.4% PEG600, 11.5% ammonium sulfate (b) and 23.2% PEG600, 10.9% potassium phosphate (c) each with 3% NaCl at pH 7. a: Mass fractions of the proteins AV, OV, LYS, OM, and OT in the starting material PPAS.

An increase in the Pu of AV from 11% to >61% was achieved in both ATPSs. Furthermore, the contents of LYS and OM in the bottom phase were reduced by $\geq$ 99.7% and the OT concentration was reduced by $\geq$ 97%. The OV concentration in the bottom phase was reduced by 57% in the sulfate system compared to 66% in the phosphate system. The Y of AV obtained in the sulfate ATPS was 96%, and thus, significantly higher than in the phosphate system (73%). For comparison, the AV Y and Pu predicted by interpolation of the screening data for these systems were Y=94% and Pu=76%, and Y=91% and Pu=67% in the sulfate and phosphate systems, respectively. The results in the 150 mL scale thus confirmed the overall potential of ATPS for the purification of AV which was observed in high-throughput experiments. However, a precise prediction of Pu and Y by interpolation of the screening data was not achieved.

3.2 High-throughput precipitation screenings

The ATPS screenings had demonstrated distribution behaviors which significantly differed between AV and the impurities. It was therefore assumed that a selective separation by precipitation using one of the phase-forming component could also be achieved. A selective precipitation would either increase the purity of AV in solution (if impurities precipitated) or would lead to an additional product concentration (if AV precipitated). Since OV was not reduced as efficiently as the other impurities via aqueous two-phase extraction, the screening was in particular aiming at the identification of solubility differences between AV and OV. In the precipitation screenings the same polymers and salts were

applied as used for the ATPSs. Thus, an integration of precipitation to the ATPS extraction step would be possible without the need for an additional buffer exchange step.

3.2.1 Precipitation screenings using salt solutions

Precipitation experiments using phosphate and sulfate solutions were conducted in 96-well filter plates. The concentrations of each protein determined in the supernatant at different salt concentrations are shown in Figure 4a. The precipitation results applying only salt as precipitant were compared to precipitations with additional 1% PEG1000 in the salt solutions as approximately occurring in ATPS bottom phases. Here the motivation was to evaluate if PEG would act synergystically or adversely to salt as precipitant.

The salt concentration at which protein precipitation occurred increased in the order of OM < OT < OV < LYS < AV. AV precipitated at phosphate and phosphate/sulfate + PEG concentrations ≥27.5%. When using sulfate only, AV precipitated at 30% sulfate. The impurities, on the contrary, were less stable in sulfate and sulfate+PEG solutions than in phosphate solutions. The addition of 1% of PEG1000 to the sulfate solutions decreased the AV and LYS solubility by approx. 2.5%, whereas addition of PEG led to 1.5%–2.5% higher salt concentrations required to obtain precipitation of OV and OT. Regarding the stability of OM, no clear influence of the addition of PEG1000 was observed. Among all proteins, OV showed the highest sensitivity against a change in the precipitant. Addition of PEG1000 did not result in better separation of OV from AV, and therefore, further experiments, e.g. by adding PEG600 were not performed. The screening results demonstrated that by using 27.5% ammonium sulfate it would be able to reduce the content of the impurity OV by 95% whereas more than 90% of AV remained in solution.

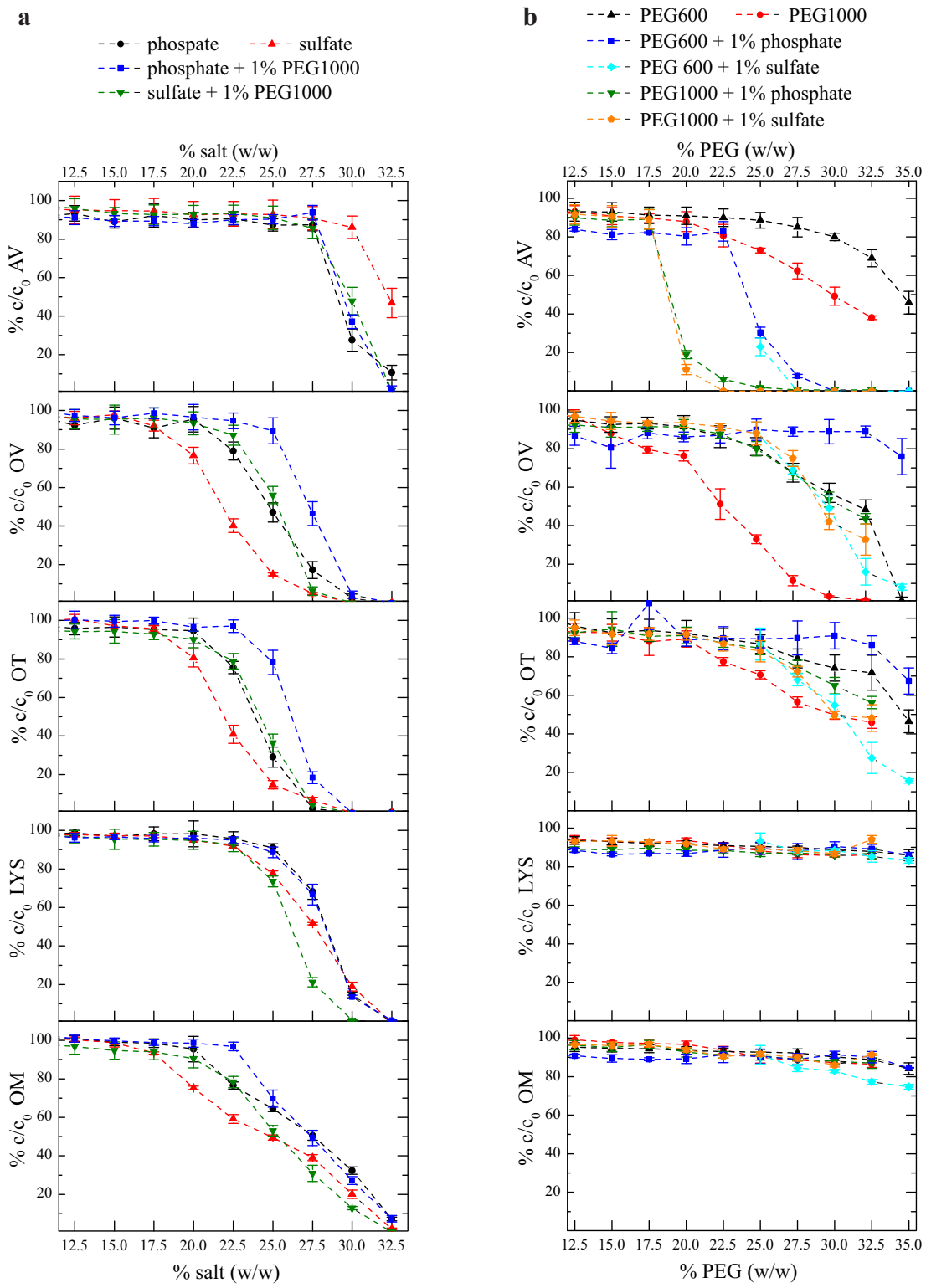

Figure 4: Protein concentration in the supernatant after incubation at different precipitant concentrations (a: salt; b: PEG) related to the start concentration c_0 in PPAS.

3.2.2 Precipitation screenings using PEG solutions

It was assumed that AV would be least stable in PEG solutions compared to the impurities, because of the lowest distribution values obtained in the ATPS screenings. Therefore, precipitation experiments were performed using PEG600 and PEG1000 solutions with concentrations of up to 35% and with and without 1% salt. The resulting protein precipitation curves are presented in Figure 4b.

The PEG concentrations at which precipitation occurred increased in the order of AV<OV≈OT, whereas LYS and OM did not precipitate at concentrations of up to 35% PEG. With an increasing PEG MW lower precipitant concentrations were required to achieve precipitation of AV, OV, and OT regardless whether salt was added or not. For both PEG MW, the addition of 1% salt resulted in precipitation of AV at significant lower PEG concentrations in the order of: PEG600 < PEG1000 < PEG600+salt < PEG1000+salt. The latter clearly showed the synergistic behavior of small amounts of salt in the case of AV, whereas the solubility of OV and OT were increased when adding phosphate.

Regarding a separation of AV from the impurities, the precipitation using either PEG600 with mass fractions of >30% or PEG1000 with mass fractions >25% at presence of 1% phosphate was shown the most promising, because at these conditions differences in the solubility were largest.

3.2.3 Resolubilization and scale-up

Following the screening for precipitation conditions, successful resolubilization of AV with preservation of its activity had to be demonstrated. As precipitation conditions 32.5% PEG600/1% phosphate were chosen followed by resolubilization in a 22.5% ammonium sulfate solution.

First, precipitation, removal of filtrate, and subsequent resolubilization of the precipitate was performed in 96-well filter plates. In order to confirm the screening results, the experiment was then performed in a larger scale with a start volume of 40 mL and a resolubilization in 5 mL. The mass fractions and yields of AV and the impurities obtained with precipitation and subsequent resolubilization in both scales are shown in Figure 5b,c and compared to the starting material PPAS (Figure 5a).

The yield of AV in the scale-up experiment was 89% and hence 23% higher compared to the value determined in the screening. The Pu value obtained in scale-up experiments (60%) was significantly lower compared to the Pu determined in the screening experiments (86%). This reduction of the purity was mainly due to the higher yield of OV in scale-up experiments. Furthermore, the yields of OT and OM were slightly increased in scale-up experiments compared to the values obtained in micro scale, whereas the opposite trend was observed for LYS. By this experimental set up, a volume reduction of factor eight was achieved.

3.3 Integration of precipitation step and ATP separation

In order to increase integration, a combination of precipitation with subsequent resolubilization and separation in an ATPS was evaluated. With this strategy, a reduction was aimed of the process volume and a large fraction of the impurities already in the precipitation step and the ATPS volume and the required amount of salt.

The precipitate obtained from the precipitation using 32.5% PEG600 + 1% phosphate in 40 mL was resolubilized with 5 mL of a homogeneously mixed ATPS composed of 11.5% ammonium sulfate, 24.5% PEG600, and 3% NaCl. Phase separation was achieved by centrifugation. While both phases became clear, some precipitate was seen in the interphase. The determined purities and yields in the

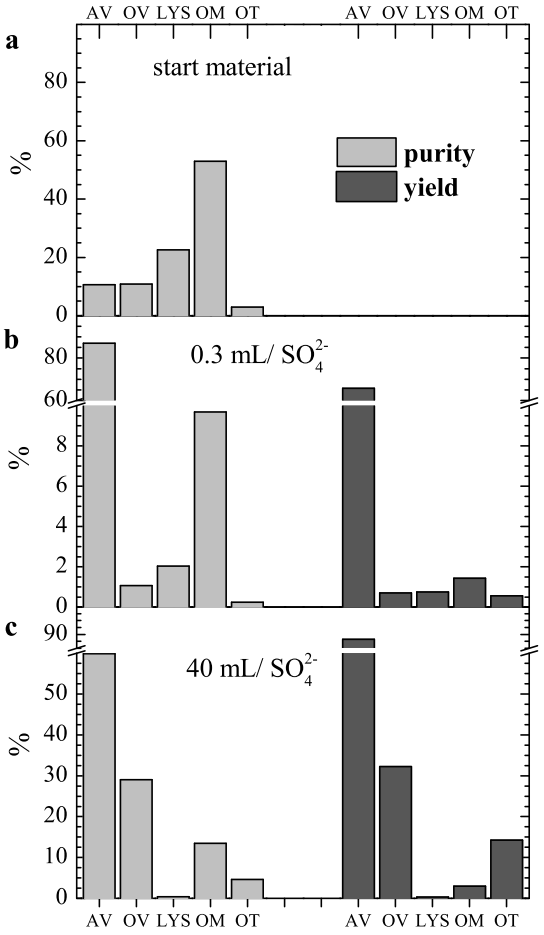

Figure 5: a: Protein mass fraction of proteins in PPAS. Results of precipitation by 32.5% PEG600 + 1%PO4 and resolubilization using 22.5% ammonium sulfate performed in micro-titer plates (b) and 50 mL centrifuge tubes (c).

obtained 1.5 mL of bottom phase are given in Figure 6.

With the combination of precipitation and aqueous two-phase extraction an AV purity of 77.6% and a total yield of 91% in the bottom phase were achieved. The LYS concentration was lower than the LLOQ which equals a reduction of the concentration by 99.97%. The OM concentration was reduced by >99.92% resulting in a mass fraction of 0.42%. The determined OT mass fraction was 0.23% which equaled a reduction of the concentration by 99.44%. With a final mass fraction of 21.7% (reduction of concentration by 80.7%), OV was the mainly remaining impurity in the bottom phase. The process of precipitation and separation in an ATPS resulted for the bottom phase in a volume reduction by factor 26.7.

3.4 High-throughput column chromatography experiments

In order to further increase AV purity while at the same time extracting AV from the ATPS bottom phase and reducing the fraction of the remaining impurity OV, a subsequent separation step was investigated. Due to their salt tolerance hydrophobic interaction chromatography (HIC) and mixed-

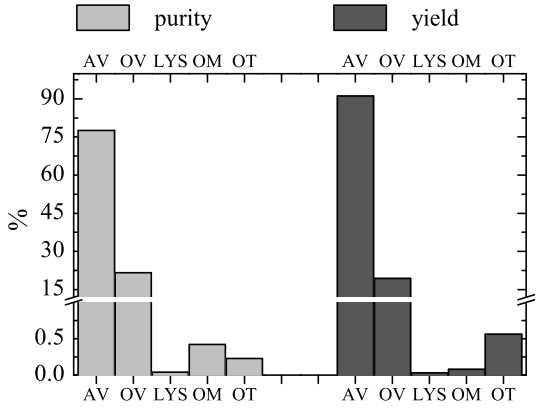

Figure 6: Yields and purities in the bottom phase obtained by PEG600-precipitation with resolubilization in an ATPS composed of 24.4% PEG600/11.5% ammonium sulfate/3% NaCl, pH 7.

mode chromatography (MM) were the favorable types of chromatography to be tested, as the salt-rich bottom phase of an aqueous two-phase separation could possibly be loaded directly on HIC or MM columns without requiring a buffer exchange.

Four HIC resins of different hydrophobicity and ligand type and four different MM resins were tested using 200 μL columns. These eight resins were loaded with eight column volumes of the bottom phases of ATPSs generated in lab-scale (see 3.1.2): a) the bottom phase composed of 11.5% PEG600, 24.4% ammonium sulfate, 3% NaCl (experiment 1), and b) the bottom phase composed of 10.9% PEG600, 23.2% potassium phosphate and 3% NaCl (experiment 2). Gradient elution was applied with decreasing sulfate concentrations from 1.9 M to 0.01 M over 60 CV in experiment 1 and decreasing phosphate concentrations from 1.8 M to 0.01 M over 60 CV in experiment 2. The salt gradient was hence performed beginning with approximately the loading salt concentrations.

Table III Overview of HIC and MM resin screening using ATPS bottom phases. 'Partly' indicates a beginning breakthrough and significant elution within wash step and the first 3 CV of the salt gradient. Separations of OV from AV providing baseline separation or minimal peak overlap were indicated with 'yes'.

Resin	Salt phase	Binding of AV	Separation of OV
HEA HyperCel™	phosphate	yes	yes
	sulfate	yes	yes
PPA HyperCel™	phosphate	yes	yes
	sulfate	yes	yes
MEP HyperCel™	phosphate	partly	yes
	sulfate	yes	yes
Capto MMC	phosphate	partly	no
	sulfate	yes	no
Fractogel® EMD Propyl S	phosphate	–	–
	sulfate	–	–
Toyopearl Phenyl-650S	phosphate	–	–
	sulfate	–	–
Butyl Sepharose HP	phosphate	–	–
	sulfate	partly	yes
Octyl Sepharose 4 FF	phosphate	partly	no
	sulfate	partly	yes

The screening results are summarized in Table III addressing two questions: Is a 100% binding of AV to the respective resins given, and can a separation of AV and OV be achieved in a linear salt gradient?

When using HIC resins, AV bound only partly, whereas complete binding was observed for six of the eight cases when using MM resins. For each resin type, the elution profile obtained with the phosphate system was similar compared to the profile obtained with the sulfate system. AV and OV eluted however at lower phosphate concentrations compared to the sulfate elution concentrations. A separation of AV from OV was achieved in the phosphate and sulfate systems on the three MM resins HEA HyperCel™, PPA HyperCel™, and MEP HyperCel™. No separation between AV and OV was achieved using the Capto MMC resin. Here, AV eluted in the regeneration step when using sulfate whereas AV coeluted with OV when using phosphate. As an example for the chromatograms obtained from the robotic experiments three chromatograms of MM chromatography experiments using a sulfate elution gradient are shown in Figure 7.

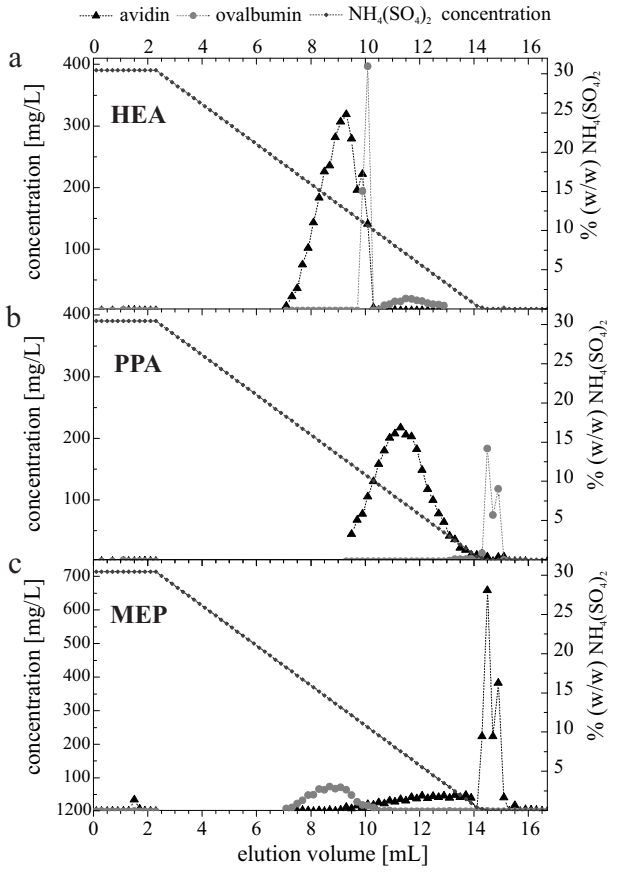

Figure 7: Elution profile of robotic chromatography experiments using mixed-mode adsorber PPAHyperCel™(a), HEAHyperCel™(b) and MEPHyperCel™(c) loaded with a bottom phase of a PEG600/sulfate/3% ATPS. The abrupt change in protein concentration at 9.5 – 10 mL elution volume in the case of HEAHyperCel™is assumably caused by an incorrect buffer composition and was therefore considered as outlier.

AV and OV eluted from the HEA column at higher salt concentrations compared to the concentrations required for the PPA resin. When using MEP, AV and OV changed elution order and AV was mainly eluted in the regeneration step applying 0.1 M NaOH. With the PPA resin, a total AV yield of 87% for the sulfate system and 85% for the phosphate system was achieved. The determined AV purities were 97.5% and 95.8%, respectively. Regarding the calculation of the purities one has to note that the concentration of impurities OV, LYS, and OT in the AV elution pool each was lower than the LLOQ. The LLOQ was hence used to calculate the minimal purity achieved and therefore the determined purity values are in fact the lowest possible values for the given AV concentration and might actually be higher.

3.5 Combined process

With the generated screening data and additional scale-up experiments shown in previous sections, following process conditions as well as the sequence of process steps is suggested for the purification of AV. First, precipitation using 30–32% PEG 600 and 1% potassium phosphate salt leading to a significant volume reduction (up to factor 8 was tested) and concentration of AV, and furthermore, a reduction of the LYS-, OM-, and OT-content. In an integrated process approach, a homogeneously mixed solution 24.4% PEG600/11.5% ammonium sulfate/3% NaCl is used for resolubilization and for the extraction of AV, which distributes in the salt-rich bottom phase. Besides further volume reduction (factor 3) AV purity is increased to 77% and LYS, OM, and OT are reduced close or below the quantification level. The mainly remaining impurity OV can be separated via mixed-mode chromatography applied in HIC-mode by loading the ATPS bottom phase directly on the column without the need for any conditioning steps. The reached AV purities and yields for the single steps ATPS and PEG precipitation as well as the combination of process steps are displayed in Figure 8. It must be noted that the determined total AV yield for the combined ATPS and precipitation process was higher than the yield determined for the precipitation step only. This is assumed to be due to the variance inherent in the experimental and analytical procedures.

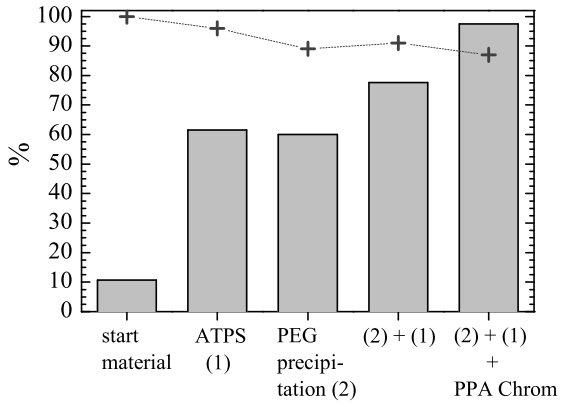

Figure 8: Overview of obtained AV purities (bars) and AV yields (+). Starting material: PPAS; ATPS: 24.4% PEG600/11.5% ammonium sulfate/3% NaCl; precipitation: 32% PEG 600 + 1% phosphate; PPA Chrom: mixed-mode chromatography using PPAHyperCel™ equilibrated with 1.8 M ammonium sulfate and loaded with ATPS bottom phase.

4 Discussion

4.1 Avidin purification using aqueous two-phase systems

In order to identify a suitable ATPS for the purification of AV from its four major impurities LYS, OM, OT, and OV, three steps were undertaken. First, a crude screening of three different PEG/sulfate/NaCl combinations was performed complemented by data from an earlier screening study [34]. In a second screening step the two most promising systems were analyzed with variation of the phase ratio. Third, based on the screening results two ATPSs were selected for performance testing in a larger scale. The use of ATPSs of either PEG600/sulfate/3% NaCl or PEG600/phosphate/3% NaCl was demonstrated to be very effective for a reduction of the impurity fraction. The AV purity was increased from 11% in the PPAS to greater than 61% in only one extraction step. In particular, the large difference of the distribution ratios obtained in the PEG600/sulfate/3% NaCl system (D_{LYS} = 264 and D_{AV} = 0.004) yielded a reduction of the LYS concentration by ≥99.7%. This is a considerably high value when taking into account that a preparative separation of these proteins using common ion-exchange chromatography has been reported to be challenging due to their similar isoelectric points of AV (pI=10) and LYS (pI=10.7) [26, 27]. If considering multiple extraction steps, the impurities LYS, OM and OT could probably be fully separated from the AV-rich bottom phase.

Varying the phase ratio

A volume reduction of the phase enriched with the protein of interest lowers the volume to be processed in subsequent steps and is of great economic interest. The phase ratio was therefore taken into account in the second step of the screening procedure for optimal ATPS conditions. The following relation of yield (Y) in the bottom phase, distribution ratio (D), and phase ratio (pr) derives by combining equations 2, 3 and 5.

$$Y = \frac{1}{1 + D \cdot pr}. \tag{7}$$

From this relationship, which is visualized in figure 9a, it becomes clear that the optimization of the purity and yield poses a Pareto optimization problem. With increasing phase ratio the yield in the bottom phase must decrease providing that the partitioning coefficient is constant and no precipitation occurs. Protein yields and purities strongly depend on the phase ratio, in particular for moderate D-values between 0.02 – 50. In the case of AV, the D-values were ≤0.01 for pr ≥1.0. Thus, corresponding to equation 7 a Y ≥90% could be achieved with phase ratios up to 10. In this study, however, we deliberately chose a phase ratio of only three for subsequent experiments in larger scale in order to investigate the potential of predicting experimental outcomes from the interpolated data of the high throughput experiments ($0.25<pr<4$). With this in mind, further process optimization will be carried out using higher phase ratios of up to ten in order to achieve an even further decrease of the OV fraction in the bottom phase. With a distribution value of about 0.4 in the sulfate system, the calculated yield of OV using equation 7 at a phase ratio of pr =3 is 38%. This is in good accordance with the experimentally obtained yield of OV in the large scale experiment (43%). If the phase ratio would be increased up to 10, for example, only 20% of OV would be yielded in the bottom phase.

Scale comparability

The large scale experiments were performed at pr=3. The expected purities of AV estimated with the

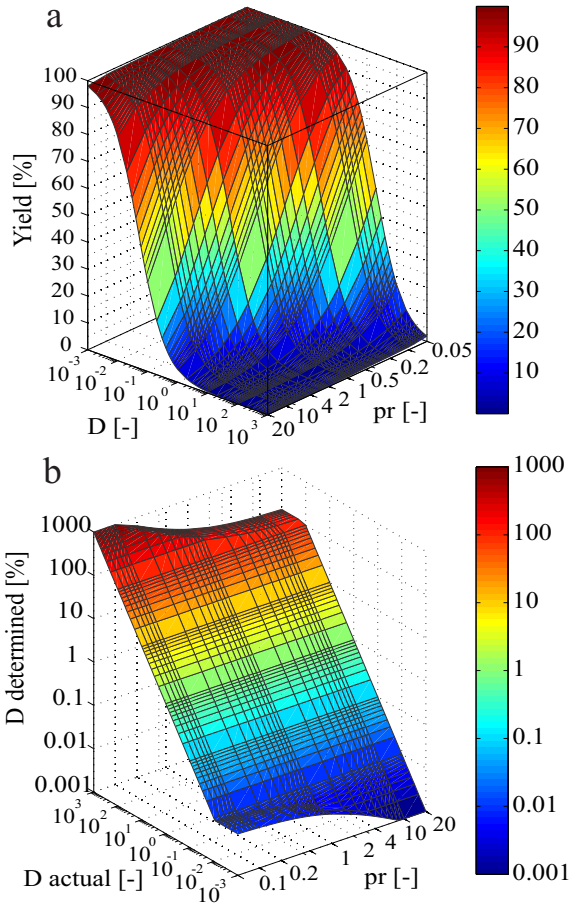

Figure 9: a): Visualization of the correlation of phase ratio, distribution ratio and theoretical yield in the bottom phase (equation 7). b) Experimentally determined distribution ratio (z-axis) of AV for a LLOQ of 0.5 mg/L and a start concentration of 0.27 g/L plotted versus the actual distribution ratio and different phase ratio. With extreme distribution ratios and phase ratios protein concentrations can be lower the analytical LLOQ, and hence, the determined distribution ratio differs from the actual distribution ratio.

high-throughput data shown in figure 2a and b for this *pr* were found to be 14% (11.5% PEG600/24.4% sulfate/3% NaCl) and 5% (10.9% PEG600/23.2% phosphate/3% NaCl) higher than the experimentally derived values in the large scale experiments. The lower purity values are mostly explained by the lower D-value for OV determined in the larger scale (0.4) compared to the value obtained in miniaturized PEG/sulfate systems ($\sim 0.55 - 0.6$). This deviation could be caused by the different phase separation times and vessel materials used in the high-throughput experiment compared to the experiments performed in 250 mL glass bottles. In a future study, a further scale-up to pilot or production scale should be performed in order to evaluate the quantitative prediction of manufacturing data from high-throughput and lab-scale aqueous two-phase experiments.

The yield of AV predicted by interpolation of the screening data for the sulfate system were in accordance with the experimental value obtained in the large scale experiment (94% and 96%, respectively), whereas the yield values deviated significantly for the phosphate system (91% vs. 73%). This deviation can not be explained by the experimental error of the high-throughput experiments (6%). Partly precipitation of AV was indicated by the recovery value of soluble protein of only 74%. The results from the precipitation studies (section 3.2) supports this hypothesis by showing precipitation of AV at phosphate concentrations $\geq$27.5% compared to sulfate concentrations $\geq$30.0% after 3 hours.

Concerning a robust extraction process this observation favors the sulfate system over the phosphate system.

Apparent distribution values

The correlation presented in Figure 9a can obviously be only used if the D-value has been determined correctly. In the case of extreme D-values and/or protein precipitation, however, this can be challenging. When protein concentrations in the start material or in one of the two phases are below the LLOQ of the chosen analytical method, the D-value is determined only on the basis of the concentration determined in the second phase, and hence, true and apparent D-values differ. This was observed in the case of LYS, for example, as can be seen in Figure 1j. The determined 'apparent' D-values increased with a decrease in phase ratio and hence with an increase in the LYS concentration in the top phase, whereas the concentration in the bottom phase was below the LLOQ of 0.3 mg/L.

The impact of concentrations below the LLOQ on the apparent D-value is presented as an example for AV in Figure 9b. For a varying phase ratio the true D-value of AV was compared with the apparent D-value, which would be experimentally determined if protein concentrations below the LLOQ were determined. In this example, an AV concentration of 0.27 g/L, a LLOQ for AV of 0.5 mg/L and a fivefold dilution of the samples prior analysis were taken into account. It becomes clear that at true D-values <0.02 and >100 the analytical LLOQ leads to apparent D-values based on experimental data differing greatly from the true D-value. For example, at phase ratios of 0.25 and a true D-value of 0.005 the apparent D-value would be determined to 0.032. Furthermore, the relative error of an analytical method is commonly increasing for very low and very large concentrations resulting in higher uncertainties for the determination of extreme D-values. The increased error for very low concentrations was therefore averaged by performing the experiments in quadruplicates. It can be concluded that, if very strong affinity to one phase is determined, the analytical sensitivity and used phase ratio should be taken into account when comparing results only based on distribution values.

Effect of precipitation

Besides the analytical sensitivity, protein precipitation affects the determination of the distribution value. Huddleston and co-workers [43] pointed out in their work that the distribution value can change due to increased salting-out in PEG/salt systems when varying the phase ratio. In the case of OV and for the longest tie-line lengths (figure 1i), the determined D-value increased with reduced recovery values. The OV concentration in the upper phase remained constant, whereas the concentration in the bottom phase was significantly reduced, probably caused by the high sulfate concentration in the bottom phase and precipitation of OV therein. The evaluation of distribution characteristics of a protein in ATPSs should therefore be related to the overall mass balance of soluble protein found in the top and bottom phase. For this reason, the results of the precipitation screenings will be discussed in the following section not only in regards to a possible purification of AV but also in context of the partitioning and stability in ATPSs.

4.2 Screening for precipitation coupled with ATPS

It has been described in literature that protein hydrophobicity determined by solubility analysis in salt and polymer phases correlates to certain extent with the partitioning in aqueous two-phase systems [41, 44, 45]. The ATPS screening had revealed very large differences in the distribution of AV and the impurities OV, LYS, OM, and OT. Therefore, the solubility of AV and its impurities was

investigated by mixing PPAS and different levels of PEG and salt concentrations assuming that AV could be further purified by selective precipitation.

Precipitation in salt solutions

AV was the most stable protein in solutions of increasing salt concentrations. AV was less stable in phosphate solutions compared to sulfate solutions, whereas the opposite was observed for the stability of the impurities OV, OM, and OT. Precipitation of AV occurred in sulfate solutions only at concentrations >30%, whereas OV, which distributed similarly to AV into the bottom phase, precipitated already at sulfate concentrations ≥18%. This result leads to the conclusion that in PEG/sulfate ATPSs low recovery values should be obtained for OV, because the sulfate concentrations in the bottom phase were above 22%. However, the experimentally determined recoveries of OV in PEG/sulfate systems were >80%. A possible explanation for this discrepancy lies in the different processing time in the aqueous two-phase and the precipitation experiments. While samples from an ATPS were diluted for analysis approximately 10 minutes after phase separation, sample withdrawal and dilution of precipitation samples was performed after three hours.

The concentration at which a protein began to precipitate determined in this study followed the order OT<OV<LYS. This result is not in accordance with the order given in Andrews *et al.* [41] where protein precipitation was investigated in single-component systems: LYS<OV<OT. This might be due to large differences in protein concentrations and possible protein-protein interactions occurring in PPAS.

The addition of PEG1000 to the sulfate solutions resulted in an increased solubility of OV. Therefore, the high recovery values of OV in ATPSs could also be due to the 0 – 2% PEG present in the ATPS bottom phases. The stabilizing effect of PEG on proteins in salt solutions might be due to interaction of PEG molecules with nonpolar patches on the protein surface, and thus, sterically stabilizing the proteins. Remarkably, when adding 1% PEG1000 the solubility of all impurities increased, whereas AV precipitated at lower sulfate concentrations. When using phosphate, however, an effect of the additional PEG was not observed. Further studies would be required in order to give a sound explanation for this observation which was out of scope for this work.

Precipitation in PEG solutions

With regard to an AV purification process, the high solubility of AV in sulfate solutions could be used to precipitate OV before performing a PEG/sulfate two-phase separation. However, we aimed at a precipitation of AV in order to achieve an increased AV concentration by resolubilization in low liquid volume. In solutions with PEG600 and PEG1000 concentrations of up to 35%, AV, OV, and OT could be precipitated, whereas LYS and OM remained soluble. This observation is in accordance with the results obtained in the ATPS experiments, in which high recoveries and the highest distribution ratios of all investigated proteins had been obtained for LYS and OM. Furthermore, the low solubility of AV in PEG solutions compared to the stability in highly concentrated salt solutions explains the propensity to partition into the salt-rich bottom phase.

Lower concentrations of PEG1000 compared to PEG600 were required in order to obtain precipitation. This is in accordance with a decreasing slope of the precipitation curve with decreasing PEG molecular weight as described in literature [46, 47].

The addition of 1% salt (either phosphate or sulfate) to the PEG solutions was demonstrated to have a large effect on the solubilities of AV. AV precipitated at an approximately 5% lower PEG

content when adding potassium phosphate. In contrast, the solubilities of OT and particularly of OV were increased (approximately 7.5% higher PEG content). When adding sulfate to PEG600, however, no difference in the solubilities of OV and OT were observed when compared to using only PEG600. The results of the PEG600/1% phosphate experiments in small scale (confirmed in a repeated experiment), were not in accordance with the corresponding large scale experiment. Therefore, the effect of additional phosphate on the solubility of OV and OT might not be as strong as indicated from the small-scale precipitation curves shown in figure 4.

An increase in protein stability in PEG solutions with the addition of salt has been previously reported for albumin and chymotrypsin in PEG4000 solutions by Miekka and Ingham [48]. They observed that at acidic pH, the addition of ammonium sulfate led to decreasing stability whereas at a pH>pI the addition of ammonium sulfate resulted in increased solubility. The precipitation data we have presented match this observation, because at pH 7 OV has a negative net charge, whereas AV is positively charged. Furthermore, in their work, Miekka and Ingham correlate the different stabilities in solutions of PEG either with or without the presence of salt with the extent of self-associated protein. They showed that with the presence of salt, the degree of self-association is reduced leading to stabilization in PEG solutions, whereas solubility is decreased when no additional salt is present. Self-association for OV has been reported [49] supporting this hypothesis. It would however require a dedicated study in order to fully explain the observed stability differences between AV, OV, and OT by considering self-association, inter-monomer interactions of AV [50], and molecular properties such as the distribution of hydrophobic patches and charges.

The precipitation screenings demonstrated that solutions of both PEG1000 + 1% phosphate as well as PEG600 + 1% phosphate could be used to selectively precipitate AV. The successful resolubilization of AV precipitate in both a 100 mM phosphate buffer and a highly-concentrated sulfate solution provides the possibility of using either cation-exchange chromatography or hydrophobic interaction chromatography as subsequent step. The screening results were in adequate accordance with the results in 40 mL scale. The mass fractions of OV and OT, however, were considerably larger when using centrifugal tubes compared to the results obtained in micro-titer plates. This might be due to the different vessel material used, the mixing procedure and different methods for solid-liquid separation.

The precipitation screening provided useful information on the solubilities of all contaminants in potential top and bottom phases, respectively. Oelmeier *et al* [8] have proposed an ATPS selection scheme on the basis of the solubilities of the target protein determined in precipitation screenings. The suitability of this approach is supported by the presented data. Nevertheless, it was demonstrated that, if the stability of proteins in bottom phases of ATPSs is estimated based on precipitation screenings, in some cases thepresence of PEG remaining in the bottom phase need be taken into account. A correlation of partitioning behavior and solubility as presented in some publications [41, 44, 45] could not been stated for all proteins present in PPAS. Possible protein-protein interactions occurring in the mixture of PPAS could explain different results compared to literature data that were obtained in experiments using only a single, already highly purified protein.

4.3 Processing of ATPS bottom phase using mixed-mode chromatography

After precipitation and ATPS separation OV remained in solution with a fraction of about 22%. Due to the low pI of 4.5, a separation of OV from AV by using ion-exchange chromatography was assumed to be feasible at neutral or basic pH-values. However, this would require the reduction of the salt concentration in the bottom phase by using diafiltration, for example. A seamless process would be

beneficial and hydrophobic interaction or mixed-mode chromatography were therefore considered as potential techniques.

When using HIC resins, neither with a bottom phase composed of phosphate (1.9 M) nor sulfate (1.8 M) sufficient binding of AV was obtained or AV eluted completely in the first fractions of the salt gradient. A separation from OV was only achieved when using butyl and octyl phases providing the greatest hydrophobicity. For these resins, HIC might be optimized to function as flow-through step. This, however, would not provide the advantage of a further concentration of AV. When using MM resins, AV and OV were adsorbed. The chosen salt gradient provided a separation of OV from AV on the HEA and PPA HyperCel™ resins and step yields ≥95%. However, the AV mass recovered in the elution pool was about 25% lower for the experiments using phosphate compared to when using sulfate. The discrepancy to the determined yield for the phosphate systems could be due to further precipitation in the application indicating low stability of AV in the phosphate systems. AV eluted from the PPA resin at lower salt concentration when compared to the elution from the HEA column. This is probably due to the more hydrophobic ligand of PPA compared to HEA and is in accordance with studies using different proteins [51]. The reversed elution order of AV and OV when using MEP in comparison to PPA and HEA underlines the different binding mechanism for proteins on the MEP resin [52, 53].

With both HEA and PPA, a good separation of AV and OV was achieved. When regarding potential follow-up processing steps, the use of the PPA is favored over the use of HEA, because the target protein elutes at lower salt concentrations compared to the elution from the HEA resin. In the case of PPA, the concentrations of every impurity in the elution pool were below the LLOQ. Therefore, the purity of ≥97%, which was determined in the case of PPA and the sulfate system, in fact manifests the lowest limit of achieved purity. Further optimization of the elution (e.g. including elution by pH-shift), investigations on the binding capacity and confirmation in large scale is required to fully evaluate the potential of the developed process.

4.4 Process performance

The presented process scheme of precipitation, resolubilization and separation in an ATPS and subsequent MM chromatography poses an effective alternative AV purification process. By means of HTPD, an integrated process was developed which is simple and scalable. With regard to the high purity achieved and the currently high market costs of avidin the process layout is economically promising. The proposed process further features the effective separation of AV and LYS which, due to their similar isoelectric points, is difficult to achieve with common ion-exchange chromatography. It further does not require time-intensive and hence expansive buffer exchange or dialysis steps. However, further optimization should take recycling of PEG and salts into account in order to reduce environmental burden.

The presented process is considered favorable over existing processes published in literature. For comparison, Piskarev et al. [28] presented a purification process consisting of a two-step ammonium sulfate precipitation, ethanol precipitation, sodium acetate precipitation and a final cation-exchange HPLC step using a concave gradient that yielded 75%-85% AV and a purity of approximately 99% (SDS-PAGE). Rothemund and co-workers [29] demonstrated the separation of avidin from egg white using an electrophoresis instrument (Gradiflow) and obtained a yield of 60%-65% and an AV fraction containing approximately 0.1% of lysozyme. In contrast to the newly developed process, a preparative application of this technique seems questionable. In 2002, Rao *et al.* [54] patented an interesting

approach by using 4'-hydroxy azobenzene-2-carboxylic acid (HABA) in order to selectively elute AV from an weak cation-exchange resin. HABA was subsequently separated from AV at low pH and dialysis. The yield percentage was not stated whereas the purity was estimated 98% via SDS PAGE. The use of dialysis, however, appears not favorable for a large-scale process.

Regarding the purification of AV, future studies should focus on even further volume reduction in the precipitation step and an investigation of the capacity of the ATPS extraction process. Furthermore, when having larger amounts of sample material at hand, large-scale PPA column experiments as polishing step should be performed to evaluate potential column loads and impurity reduction factors.

The presented integrated process consisting of precipitation, aqueous two-phase extraction and mixed-mode chromatography (or HIC) possibly can be used as a platform for the development of purification processes for other molecules. In particular, the process is suitable for molecules being either very hydrophilic (AV, polymer precipitation, ATPS, MM/HIC) or very hydrophobic (salt precipitation, ATPS + back extraction into bottom phase, MM/HIC).

5 Conclusion

We have developed a purification process for the extraction of AV from an egg white solution by exclusively using high-throughput techniques. The HTPD-approach included the three unit operations precipitation, aqueous two-phase extraction and chromatography as well as a fast tandem RP HPLC assay for protein quantification. ATPSs composed of PEG600/ammonium sulfate/3%NaCl and PEG600/potassium phosphate/3%NaCl were demonstrated to be suitable for the separation of OM, LYS and OT with an AV step yield of 96% and step purity of 61%. The precipitation of AV employing high PEG600 concentrations and additional 1% salt was successfully integrated prior to the extraction via an ATPS resulting in a volume reduction and an AV purity of 77%. The results obtained were mainly confirmed in 230-fold and 133-fold scaled-up experiments. Furthermore, the screening provided information on the sensitivity of a possible process against changes in the phase ratios and the concentrations of PEG and salt. HIC and MM resins were evaluated by dynamic salt gradient elution experiments. Two resins, PPA and HEA Hypercel™ provided a separation of OV from AV with step yields >95% resulting in total AV yield of about 85% and final purities >95%. The obtained data build the basis for further studies addressing the binding capacity and optimization of elution conditions.

The high-throughput process development methodology presented in this work underlines the benefits of such an approach: First, only little sample material was needed: 65 μL per ATPS, 30 μL per precipitation experiment, 1.6 mL per chromatography experiment. Secondly, the presented study (robotic experiments, HPLC analysis and data evaluation) required in total only a about three weeks work for one person (not including the development time for the high-throughput methods). One must keep in mind that ATPS and precipitation screenings were performed in quadruplicates in order to increase data quality. Therefore, the experimental time needed could have been reduced even further. Such short development times, however, can only be reached, if a toolbox of high-throughput routines for ATPS, precipitation, chromatography screenings, and analytics is already well-established in the lab as well as combined with automated data handling methods. Once developed and implemented, HTPD was shown to enable the fast and simple evaluation of process alternatives that could be used for economic considerations as well as initial facility fit simulations.

Acknowledgement

We gratefully thank Marie-Luise Schwab for performing parts of the experimental work and Stefan A. Oelmeier for useful discussions and the review of the manuscript.

References

[1] J. Thiemann, J. Jankowski, J. Rykl, S. Kurzawski, T. Pohl, B. Wittmann-Liebold, H. Schlüter, Principle and applications of the protein-purification-parameter screening system., *J. Chrom. A* 1043 (1) (2004) 73–80.

[2] M. Bensch, P. Schulze Wierling, E. von Lieres, J. Hubbuch, High Throughput Screening of Chromatographic Phases for Rapid Process Development., *Chem. Eng. Technol* 28 (11) (2005) 1274–1284.

[3] J. L. Coffman, J. F. Kramarczyk, B. D. Kelley, High-throughput screening of chromatographic separations: I. Method development and column modeling., *Biotechnol. Bioeng* 100 (4) (2008) 605–618.

[4] 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 - Miniaturization, Automation and Parallelization of Breakthrough and Elution Studies., *Chem. Eng. Technol* 31 (6) (2008) 893–903.

[5] K. Treier, S. Hansen, C. Richter, P. Diederich, J. Hubbuch, P. Lester, High-throughput methods for miniaturization and automation of monoclonal antibody purification processes., *Biotechnol. Prog* 28 (3) (2012) 723–732.

[6] M. Wiendahl, C. Völker, I. Husemann, J. Krarup, A. Staby, S. Scholl, J. Hubbuch, A novel method to evaluate protein solubility using a high throughput screening approach., *Chem. Eng. Sci.* 64 (17) (2009) 3778–3788.

[7] C. Knevelman, J. Davies, L. Allen, N. J. Titchener-Hooker, High-throughput screening techniques for rapid peg-based precipitation of igg4 mab from clarified cell culture supernatant., *Biotechnol. Prog.* 26 (3) (2010) 697–705.

[8] S. A. Oelmeier, C. Ladd-Effio, J. Hubbuch, High throughput screening based selection of phases for aqueous two-phase system-centrifugal partitioning chromatography of monoclonal antibodies., *J Chrom. A* 1252 (2012) 104–114.

[9] S. A. Oelmeier, F. Dismer, J. Hubbuch, Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation., *Biotechnol. Bioeng.* 108 (1) (2011) 69–81.

[10] K. Rege, M. Pepsin, B. Falcon, L. Steele, M. Heng, High-throughput process development for recombinant protein purification., *Biotechnol. Bioeng.* 93 (4) (2006) 618–630.

[11] M. D. Melamed, N. M. Green, Avidin., *Biochem. J.* 89 (1963) 591–599.

[12] R. J. McMahon (Ed.), *Avidin-Biotin Interactions - Methods and Applications.*, Humana Press, Totowa, NJ, 2008.

[13] M. Wilchek, E. A. Bayer, Biotin-binding proteins: Overview and prospects, in: M. Wilchek, E. A. Bayer (Eds.), *Avidin-Biotin Technology.*, Vol. 184 of Method Enzymol, Academic Press, 1990, pp. 49 – 51.

[14] E. P. Diamandis, T. K. Christopoulos, The biotin-(strept)avidin system: principles and applications in biotechnology., *Clin. Chem.* 37 (5) (1991) 625–636.

[15] H. Sakahara, T. Saga, Avidin-biotin system for delivery of diagnostic agents., *Adv. Drug Delivery Rev.* 37 (1-3) (1999) 89–101.

[16] R. J. Boado, Y. Zhang, Y. Zhang, C.-F. Xia, Y. Wang, W. M. Pardridge, Genetic engineering, expression, and activity of a chimeric monoclonal antibody-avidin fusion protein for receptor-mediated delivery of biotinylated drugs in humans., *Bioconjugate Chem* 19 (3) (2008) 731–739.

[17] P. Savi, J. P. Herault, P. Duchaussoy, L. Millet, P. Schaeffer, M. Petitou, F. Bono, J. M. Herbert, Reversible biotinylated oligosaccharides: a new approach for a better management of anticoagulant therapy., *J. Thromb. Haemost.* 6 (10) (2008) 1697–1706.

[18] I. Paty, M. Trellu, J.-M. Destors, P. Cortez, E. Boëlle, G. Sanderink, Reversibility of the anti-FXa activity of idrabiotaparinux (biotinylated idraparinux) by intravenous avidin infusion., *J. Thromb. Haemost.* 8 (4) (2010) 722–729.

[19] Y. Mine, Recent advances in the understanding of egg white protein functionality., *Trends Food Sci. Technol.* 6 (1995) 225–232.

[20] A. C. Awade, On hen egg fractionation: applications of liquid chromatography to the isolation and the purification of hen egg white and egg yolk proteins., *Z. Lebensm.Unters. For.* 202 (1) (1996) 1–14.

[21] P. Cuatrecasas, M. Wilchek, C. B. Anfinsen, Selective Enzyme Purification By Affinity Chromatography., *Proc. Natl. Acad. Sci. USA* 61 (2) (1968) 636–643.

[22] G. Heney, G. A. Orr, The Purification of Avidin and Its Derivatives on 2-Iminobiotin-6-aminohexyl-Sepharose 4B., *Anal. Biochem.* 114 (1981) 92–96.

[23] C. Guérin-Dubiard, M. Pasco, A. Hietanen, A. Quiros del Bosque, F. Nau, T. Croguennec, Hen egg white fractionation by ion-exchange chromatography., *J. Chrom. A* 1090 (1-2) (2005) 58–67.

[24] T. Durance, Separation, purification, and thermal stability of lysozyme and avidin from chicken egg white., in: J. S. Sim, S. Nakai (Eds.), *Egg uses and processing technologies. New developments*, Cab Int, Wallingford, UK, 1994, pp. 77–85.

[25] G. Gatti, M. Bolognesi, A. Coda, F. Chiolerio, E. Filippini, M. Malcovati, Crystallization of hen egg-white avidin in a tetragonal form., *J. Mol. Biol.* 178 (3) (1984) 787–789.

[26] T. D. Durance, S. Nakai, Simultaneous Isolation of Avidin and Lysozyme., *J. Food Sci.* 53 (4) (1988) 1096–1101.

[27] E. Li-Chan, S. Nakai, J. Sim, D. B. Bragg, K. V. Lo, Lysozyme Separation from Egg White by Cation Exchange Column Chromatography., *J. Food Sci.* 51 (4) (1986) 1032–1036.

[28] V. E. Piskarev, A. M. Shuster, A. G. Gabibov, A. G. Rabinkov, A novel preparative method for the isolation of avidin and riboflavin-binding glycoprotein from chicken egg-white by the use of high-performance liquid chromatography., *Biochem. J.* 265 (1) (1990) 301–304.

[29] D. L. Rothemund, T. M. Thomas, D. B. Rylatt, Purification of the basic protein avidin using Gradiflow technology., *Protein Express. Purif* 26 (1) (2002) 149–152.

[30] B. A. Andrews, S. Nielsen, J. A. Asenjo, Partitioning and purification of monoclonal antibodies in aqueous two-phase systems., *Bioseparation* 6 (5) (1996) 303–313.

[31] A. M. Azevedo, P. A. J. Rosa, I. F. Ferreira, M. R. Aires-Barros, Chromatography-free recovery of biopharmaceuticals through aqueous two-phase processing., *Trends Biotechnol.* 27 (4) (2009) 240–247.

[32] P. A. J. Rosa, A. M. Azevedo, S. Sommerfeld, W. Bäcker, M. R. Aires-Barros, Aqueous two-phase extraction as a platform in the biomanufacturing industry: economical and environmental sustainability., *Biotechnol Adv* 29 (6) (2011) 559–567.

[33] J. A. Asenjo, B. A. Andrews, Aqueous two-phase systems for protein separation: phase separation and applications., *J. Chrom. A* 1238 (September 2011) (2012) 1–10.

[34] P. Diederich, S. Amrhein, F. Hämmerling, J. Hubbuch, Evaluation of PEG/phosphate aqueous two-phase systems for the purification of the chicken egg white protein avidin by using high-throughput techniques., *Chem. Eng. Sci.* 104 (2013) 945–956.

[35] J. C. Merchuk, B. A. Andrews, J. A. Asenjo, Aqueous two-phase systems for protein separation: Studies on phase inversion., *J. Chrom. B* 711 (1-2) (1998) 285–293.

[36] J. G. Huddleston, H. D. Willauer, R. D. Rogers, Phase Diagram Data for Several PEG + Salt Aqueous Biphasic Systems at 25 °C., *J. Chem. Eng. Data* (2003) 1230–1236.

[37] S. L. Mistry, A. Kaul, J. Merchuk, J. A. Asenjo, Mathematical modelling and computer simulation of aqueous two-phase continuous protein extraction., *J Chrom. A* 741 (2) (1996) 151–163.

[38] L. A. Ferreira, J. A. Teixeira, L. M. Mikheeva, A. Chait, B. Y. Zaslavsky, Effect of salt additives on partition of nonionic solutes in aqueous PEG-sodium sulfate two-phase system., *J. Chrom. A* 1218 (31) (2011) 5031–5039.

[39] J. Huddleston, A. Veide, K. Köhler, J. Flanagan, S. O. Enfors, A. Lyddiatt, The molecular basis of partitioning in aqueous two-phase systems., *Trends Biotechnol* 9 (11) (1991) 381–388.

[40] B. Zaslavsky, *Aqueous Two-Phase Partitioning, Physical Chemistry and Bioanalytical Applications.*, Marcel Dekker, New York, 1995.

[41] B. A. Andrews, A. S. Schmidt, J. A. Asenjo, Correlation for the partition behavior of proteins in aqueous two-phase systems: effect of surface hydrophobicity and charge., *Biotechnol. Bioeng.* 90 (3) (2005) 380–390.

[42] J. G. Huddleston, K. W. Ottomar, D. M. Ngonyani, A. Lyddiatt, Influence of system and molecular parameters upon fractionation of intracellular proteins from Saccharomyces by aqueous two-phase partition., *Enzyme Microb. Technoly* 13 (1) (1991) 24–32.

[43] J. G. Huddleston, R. Wang, J. A. Flanagan, S. O. Brien, A. Lyddiatt, Variation of protein partition coefficients with volume ratio in poly (ethylene glycol) -salt aqueous two-phase systems., *J. Chrom. A* 668 (1994) 3–11.

[44] J. Asenjo, A. Schmidt, H. F., A. B.A., Model for predicting the partition behaviour of proteins in aqueous two phase systems., *J Chrom. A* 668 (1994) 47–54.

[45] J. Huddleston, J. C. Abelaira, R. Wang, A. Lyddiatt, Protein partition between the different phases comprising poly(ethylene glycol)-salt aqueous two-phase systems, hydrophobic interaction chromatography and precipitation: a generic description in terms of salting-out effects., *J. Chrom. B* 680 (1-2) (1996) 31–41.

[46] W. Hönig, M. R. Kula, Selectivity of protein precipitation with polyethylene glycol fractions of various molecular weights., *Anal. Biochem.* 72 (1976) 502–512.

[47] D. H. Atha, K. C. Inghamg, Mechanism of Precipitation of Proteins by Polyethylene Glycols., *J. Biol. Chem.* 256 (23) (1981) 12108–12117.

[48] S. I. Miekka, K. C. Ingham, Influence of self-association of proteins on their precipitation by poly(ethylene glycol)., *Arch. Biochem. Biophys.* 191 (2) (1978) 525–536.

[49] N. Muramatsu, A. P. Minton, Hidden self-association of proteins., *J. Mol. Recognit.* 1 (4) (1989) 166–171.

[50] M. Wilchek, E. A. Bayer, O. Livnah, Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction., *Immunol. Lett.* 103 (1) (2006) 27–32.

[51] V. Brenac Brochier, A. Schapman, P. Santambien, L. Britsch, Fast purification process optimization using mixed-mode chromatography sorbents in pre-packed mini-columns., *J. Chrom. A* 1177 (2) (2008) 226–233.

[52] S. C. Burton, D. R. Harding, Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aqueous buffers., *J. Chrom. A* 814 (1-2) (1998) 71–81.

[53] G. Zhao, X.-Y. Dong, Y. Sun, Ligands for mixed-mode protein chromatography: Principles, characteristics and design., *J. Biotechnol.* 144 (1) (2009) 3–11.

[54] M. Rao, M. Gupta, I. Roy, Process for the isolation and purification of the glycoprotein avidin. *European Patent Office* EP 1 505 882 B1, Filed Oct 14 2003.

MANUSCRIPT 5

High-Throughput Column Chromatography performed on Liquid Handling Stations

—

Process Characterization and Error Analysis

Patrick Diederich, Jürgen Hubbuch

Institute of Engineering in Life Sciences, Section IV:
Biomolecular Separation Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

submitted to: *Preparative Chromatography for Separation of Proteins*
Editors: A. Staby, S. Ahuja and A. S. Rathore

Abstract

Automated and parallelized chromatography on liquid handling workstations using miniaturized columns has become a widely-accepted experimental tool for purification process development of biopharmaceuticals. Established knowledge on limitations and pitfalls is however limited. As a consequence, some uncertainty regarding the data quality remains with the experimenters and eventually the regulatory bodies. In order to increase confidence in the data obtained using this technology, we have conducted a thorough investigation of specific parameters and error sources and their influence on the precision of the experimental outcome. The study comprised investigations on the influence of pipetting precision, absorption measurements in micro-titer plates, peak fractionation, flow patterns, and salt step heights in gradient elution experiments. Separate and combined effects were qualitatively and quantitatively investigated using both experiments and simulations based on a mechanistic model.

The results demonstrate that with a sufficient number of fractions collected per peak a significant improvement in precision can be obtained despite of a low analytical precision. The flow interruptions, which are required to perform gradient elution experiments, were demonstrated to strongly affect peak shapes. Therefore, the resulting deviations of actual and nominal residence time should be taken into account if data are used to predict scale-up performance of common laboratory experiments or to calibrate chromatography models. The interplay of fractionation and salt step heights resulted in elution profiles that vary strongly when judged visually. In contrast to the visual impression, it was shown that the influence of different salt step heights can however be neglected up to step heights of 44 mM when using a peak fitting routine. Altogether, the new insights into the technology allow a more confident choice of experimental parameters and provides better understanding of the observed variances.

Keywords: high-throughput column chromatography (HTCC), RoboColumn, fractionation, error analysis, error simulation

1 Introduction

Liquid chromatography is indispensable for the purification of therapeutic proteins. However, the development of chromatographic processes can be a challenging task. There is a variety of chromatography-based separation principles that can be applied and combined, and for each of them dozens of process parameters need to be optimized. Consequently, the required experimental effort for process development and process optimization can be huge. This is in particular true if only little is known on the molecular properties of target protein and contaminants and heuristic rules do not apply. Even if process development scientists can employ a purification platform process or support the development by using mechanistic modeling, the number of experiments, the time needed for model parameter estimation and process characterization, and the material consumption can still be considerably high.

In response to these challenges, miniaturized chromatography formats have been developed within the last decade for the parallel and automated experimentation on liquid handling stations (LHSs). This has enabled high-throughput experimentation (HTE) within liquid chromatography. Batch chromatography for resin screening and isotherm determination can be performed using 96-well filter plates and 384-well micro-plates [1–6]. Packed bed chromatography can be conducted in pipette tips filled with resin and miniaturized chromatography columns providing a fast way to obtain information on dynamic binding capacities and optimal binding and elution conditions [7–12]. Furthermore, HTE chromatography has been used for the initial determination of parameters required for mechanistic modeling and simulation of chromatographic processes [1, 13–16]. Particularly in early industrial process development, in which process development work is mainly driven by short timelines, HTE in the form of screening procedures have been implemented. Here, the advantages of low sample volume and high throughput will outweigh the larger experimental errors which are often inherent in HTE. However, if the intention is to use μL-scale HTE column chromatography (in the following referred to as HTCC) for more than screening purposes, i.e. as scale-down model for detailed process optimization, process characterization, and calibration of chromatography models, two categories of questions related to the quality of data arise. The questions in the first category deal with scale, i.e. 'What chromatography parameters can in fact be investigated in μL-scale compared to the Liter-scale?' 'How is scale 'comparability' defined by development departments and regulatory authorities?' The questions in the second category deal with data quality, i.e. 'What are the main error sources in HTCC?' 'Which effects are caused specifically by using a LHS and do not occur when using common chromatographic systems?' Or, 'To what extent can the accuracy be influenced by optimizing LHS settings and method parameters?' In this work, we have addressed questions of the second category, focusing on HTCC at a scale of 200 μL column volume (CV).

The study aims to generate novel insights into the performance of HTCC and develop a better understanding of the limitations while incorporating results of recently published studies on the uncertainty in HTE batch chromatography [6, 17]. We have studied the influence of LHS settings related to pipetting accuracy, working in micro-titer plates (MTPs), fractionation, fluid flow regime, and step gradient elution by employing a combination of experimental data and simulations. In combination with different levels of analytical errors, we aim to identify the main sources of experimental noise and to quantitatively determine the correlation to the resulting data quality regarding peak characteristics such as retention times or peak width. The results also illustrate which HTCC settings require special attention whereas others could be left out from extensive and time-consuming optimization without compromising the resulting data quality. A detailed knowledge on expected error ranges will help

approaching issues from the above-mentioned first-category on scale comparability and on quality of model parameters estimated based on data from HTCC.

1.1 High-throughput column chromatography – method review

In order to facilitate a better understanding of the presented objectives and results, this section will briefly review the experimental setup of HTCC while focusing on chromatography operated using salt gradient elution. The typical work-flow in HTCC is displayed in Figure 1.

The chromatography columns are fixed on the HTCC carrier which is located above a plate shuttle. The plate shuttle is a part of the plate storage and delivery system. On this shuttle, MTPs are moved with defined intervals underneath the columns which enables collection of column eluate droplets and fractionation hereof. The liquid flow on typically eight columns operated in parallel is realized using an eight-channel pipetting arm which aspirates and dispenses the required process liquids such as buffers, protein solutions, regeneration and cleaning solutions. Most commonly, the collected fractions are analyzed photometrically with a UV-VIS plate reader that is operated at-line fully integrated in the LHS. Further information including a description of the miniaturized columns used for HTCC are given in [9, 18]. The photometric data are automatedly exported for subsequent data evaluation. The resulting 'chromatogram' in fact is a scatter or bar plot with the width of the fraction volume. Peak fitting functions can be used to generate a 'continuous' absorption signal vs. the elution volume.

When considering the usage of HTCC three major system characteristics must be kept in mind. First, the plunger size determines the maximal volume that can be dispensed without flow interruption. The most common size of the plunger in current format is 1 mL. Therefore, a sequence of dispense steps interrupted by aspiration steps must be performed in order to process the total volume required for dynamic chromatography experiments. Second, the minimal possible volumetric flow rate is given by the fixed plunger volume and the minimal step size for the plunger movement e.g. 0.83 μL/s for 1 mL plunger which equals a residence time of 4 min when using 200 μL columns. Third, only one liquid at a time can be dispensed into a column and therefore inline gradient mixing is not possible. Hence, a gradient must be mimicked by a multiple step elution (step gradient). Preferably, this is done with buffers previously prepared by the LHS [12].

1.2 High-throughput column chromatography – error sources

It has been demonstrated that chromatograms resulting from HTCC in 200 μL scale are comparable to chromatograms obtained with 2.5 mL or 1 mL used on an laboratory LC system [9, 12, 16]. However, HTCC features apparently higher noise levels when compared to the precision obtained on a laboratory scale LC system [9, 16]. In particular, the variations in mass balances are relatively high. For example, a range of 86%–104% (HPLC analysis) and 83%–104% (photometric analysis) was stated in Hansen *et al.* [19] (four replicates). Treier *et al.* [12] stated a mass balances of 111% with a standard deviation (STD) of 10% (photometric analysis; eight replicates). No publications were found regarding a quantitative investigation of HTCC in order to identify the main contributors to the observed variances.

A cause-and-effect diagram is shown in Figure 2 providing an overview of error sources potentially affecting the experimental outcomes. From this overview, LHS settings were identified which are specific for HTCC when compared to column chromatography using standard laboratory LC systems. These settings were chosen for further investigation. Error sources which are similarly important to

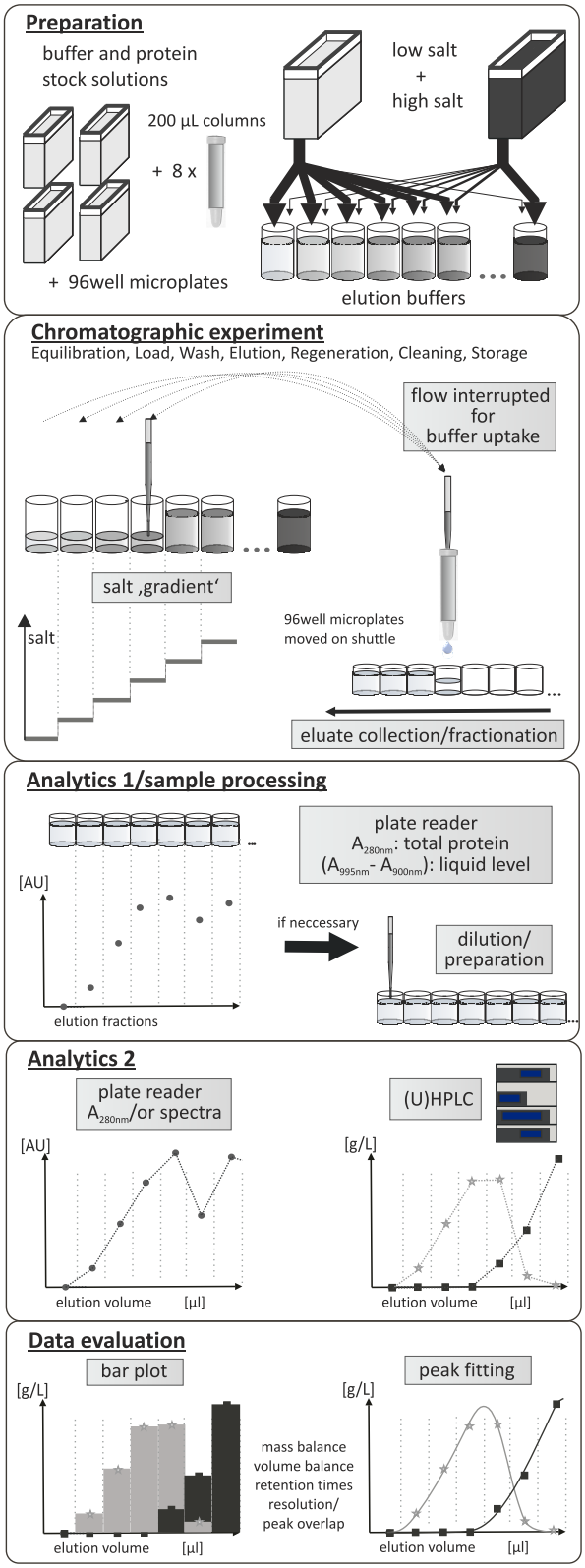

Figure 1: Sequence of HTCC process steps in case of gradient elution experiments.

the experimental results when performing laboratory LC experiments such as false weighing of buffer components and proteins were not considered for further analysis.

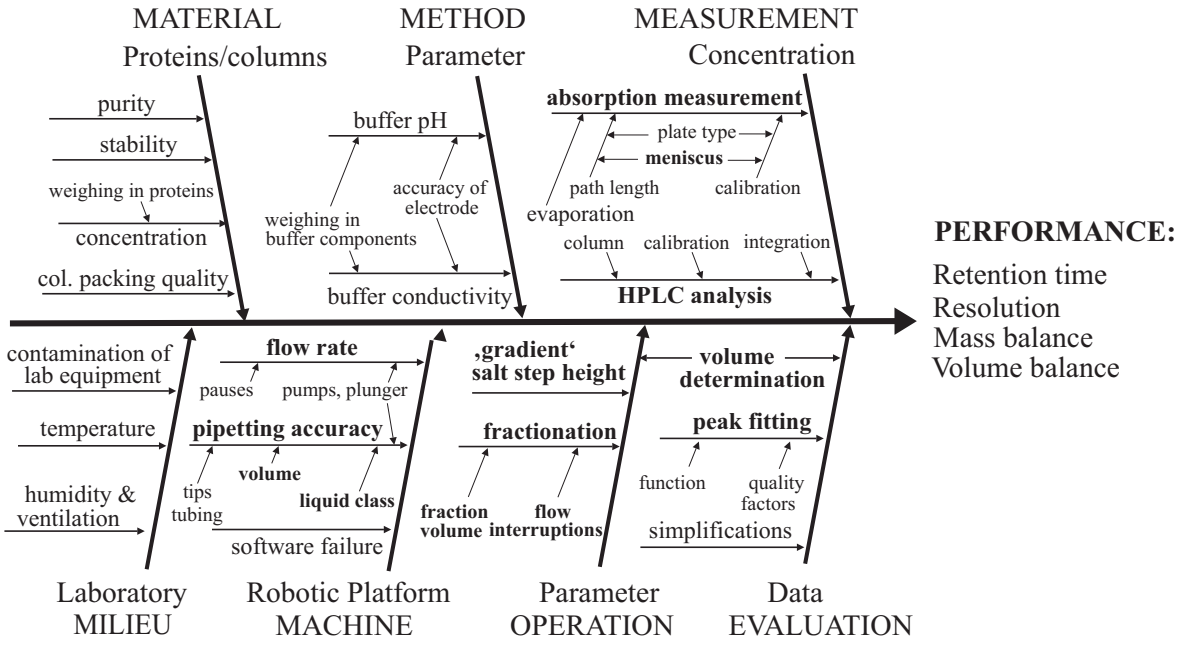

Figure 2: Fishbone diagram on potential parameters affecting the quality of HTCC experiments on a LHS. The process parameters typed in bold represent robot-specific parameters chosen for detailed investigation.

It must be pointed out that parameters such as fraction volume and salt step height are parameters that needs to be set in advance. These are self-evidently well controlled, whereas precision of pipetting can only be maximized to a certain extent by optimization of pipetting parameters. In the investigation, the following process parameters were included and sorted by increasing degree of complexity and cumulation of possible error sources:

- pipetting
- UV-VIS absorption measurements in 96-well MTP
- fractionation volume (number of fractions)
- discontinuous flow and needle movement in and out of the column
- salt step height and flow interruptions in salt gradient elution experiments.

2 Material and methods

2.1 Materials

2.1.1 Protein Solutions

Chicken egg white lysozyme (prod.-no. 62970) was purchased from Sigma-Aldrich Co. (Taufkirchen, DE). Purified avidin was provided by an industrial partner. Protein stock solutions were prepared in 100 mM sodium phosphate solution at pH 4.5, sterile filtered (0.22 μm, Cellulose Acetate syringe filter, Pall, Dreieich, DE) and stored at 7 °C until use. Each category of experiments was performed using the same protein solution.

2.1.2 Chemicals and stock solutions

Purified water was generated using an Arium®pro UV-system (Sartorius Stedim Biotech, Göttingen, DE). NaH_2PO_4, Na_2HPO_4 and NaCl of analysis grade were purchased from Merck KGaA (Darmstadt, DE). Ethanol and NaOH were supplied from (Sigma-Aldrich Co., Taufkirchen, DE) and diluted with purified water prior the use for column regeneration (0.1 M NaOH) and column storage (20% v/v EthOH with 100 mM NaCl).

2.2 Liquid handling station

HTCC was performed on a Freedom Evo®200 station (Tecan, Crailsheim, DE), equipped with an 8-channel liquid handling arm (non-coated fixed tips), a centric gripper, a rotational shaker (Te-shake, Tecan Crailsheim, DE), a plate delivery and storage devices (Te-Stack, Tecan, Crailsheim, DE), and a chromatography unit (Te-Chrom, Tecan Crailsheim, DE). An Infinite200 spectrophotometer (Tecan, Crailsheim, DE) was integrated into the platform for automated absorption measurements in a microplate format. For comparison of pipetting accuracy low-volume tips coated with fluorinated ethylene propylene (FEP) were used.

The pipetting accuracy of the liquid handling arm was tested by pipetting a nominal liquid volume onto an analytical balance (WXTS205DU, Mettler-Toledo, Greifensee, CH) providing a readability of 0.1 μg. The actual pipetted volume was calculated by dividing the measured mass by the density of the corresponding liquid. The density was determined using a 5 mL pycnometer (BLAUBRAND®, VWR International GmbH, Darmstadt, DE). The deviations between pipetted volume and nominal volume were correlated by a linear function. This correlation and additional settings on pipetting speed are referred to as 'Liquid Class' in the control software.

2.3 Software and data processing

The liquid handling station was controlled using Evoware 2.4 SP3. The software both imported pipetting values and exported photometric data employing Excel®-files (Microsoft, Redmond, WA, USA). Matlab®R2011b (The Mathworks, Natick, ME, USA) was used to further process exported absorption values and to provide automated data evaluation. This included the calculation of chromatograms based on UV 280 nm, the performance of peak fitting, the calculation of retention times and peak widths, the calculation of liquid volumes in each well, providing the selection of the samples for subsequent HPLC analysis and the calculation of dilution values (if the absorption exceeded an absorption of 2.0 AU). Matlab®was also used to perform chromatography simulations and in order to perform error simulations on robotic peak fractionation. Furthermore, the used chromatography model and simulation procedures were implemented in Matlab®environment [20]. Chromeleon®(6.80 SR10) was used to control the UltiMate®3000 RSLC x2 Dual system.

2.4 High-throughput column chromatography

Both isocratic and gradient HTCC experiments were performed using 200 μL RoboColumns (1 cm bed height) filled with SP Sepharose FF resin purchased from Atoll GmbH (Weingarten, DE). For description of the general procedures of HTCC it is referred to section 1.1. The separation of the proteins lysozyme (14.3 kDa) and avidin (≈68 kDa) at pH 4.5 was chosen as experimental system.

The fractions from both, isocratic and gradient elution experiments were photometrically analyzed at wavelengths 280 nm, 900 nm, and 990 nm using the integrated plate reader prior to HPLC analysis.

2.4.1 Isocratic elution

Isocratic elution experiments were performed using a flow rate of 90 cm/h (5 μL/s) which equals a residence time of 40 s. The fraction volume was 75 μL. The default pipetting setting allowed a maximal volume during aspiration and dispension of 1 mL per step. Saying this, by using firmware commands in the control software, an aspiration volume of 3.2 mL was achieved (the tubing volume from needle to switching valve was approx. 4 mL). Thus, for isocratic elution, a continuous dispension of 2.7 mL (=15 CV) elution buffer was performed, interrupted only briefly after each plunger volume (approx. 2 sec). The excess in aspiration volume of 0.5 mL was needed as a decrease in the salt concentration of buffer aspirated at the front end had been observed when aspirating such large volumes. This decrease in salt concentration indicates dilution during the aspiration process probably by the remaining thin film of system liquid within the tubing, which clearly only becomes significant when aspirating more than the nominal plunger volume of 1 mL. The total elution volume used during isocratic HTCC was 5.4 mL, which corresponds to 27 CV. Thus, the elution phase was interrupted only once for the aspiration of new elution buffer.

Isocratic experiments were performed using the same eight columns throughout all experiments in order to evaluate whether the inter-column variability significantly differed from the variance caused by the different flow patterns investigated. Two sets of four columns were used with a column load of 0.64 mg of lysozyme or 0.64 mg of avidin.

2.4.2 Gradient elution

The applied gradient lenghts ranged from 8 CV – 28 CV for an increase in NaCl concentration from 150 mM NaCl to 675 mM. The fraction volume of 75 μL and a step length of 150 μL were kept constant for all gradient experiments. In the case of a salt gradient length of 18 CV, the salt step height in the robotic experiments corresponded to 21.9 mM. Used flow rates were 54 cm/h (3 μL/s) and 144 cm/h (8 μL/s). Columns were loaded with 0.21 g lysozyme and 0.64 g avidin.

2.5 Lab-scale column chromatography

For comparison to HTCC, gradient chromatography experiments were performed on an Äkta Purifier10 laboratory LC system from GE healthcare (Uppsala, SE). Pre-packed mini-columns with a bed volume of 200 μL purchased from Atoll GmbH (Weingarten, DE) were used providing identical dimensions compared to the RoboColumns. For the separation of lysozyme from avidin, a salt gradient with an increasing NaCl concentration from 0.15 M – 0.675 M in 18 CV was applied. The eluate was collected in fractions of 75 μL. Column loads were equivalent to the loads used in gradient HTCC experiments.

2.6 Peak fitting

The discrete data points of each chromatographic experiment were fitted using an exponential-Gaussian hybrid function (EGH) described in [21]. This function features excellent fitting of also strongly asymmetric peaks. The EGH was compared to the commonly used exponentially-modified Gaussian function [22, 23] to be used in Monte-Carlo simulations and found being more robust in finding fit parameters. Both fit functions have the advantage compared to many other fit functions of requiring

only four data points for the determination of the function parameters. Details on the EGH and its parameters are given in appendix A.1 and in [21, 24].

2.7 Monte-Carlo simulations on peak fractionation with varying number of fractions

The effect of different error levels for the determination of the fraction volume (V), protein concentration (c) and pipetting error (Pip) were investigated using in-silico fractionation of two differently shaped, experimentally derived chromatography peaks. In Figure 3, the simulation procedure is schematically summarized for one peak shape. These peaks were initially fitted with the EGH function, thus, serving as reference peaks for normalization (step 1). Afterwards, these peaks were in-silico fractionated with fraction volumes ranging from 50 μL – 350 μL (step 2). The theoretical concentrations in each of these artificial fractions were then altered by adding experimental errors of different levels for the determination of the concentration (y-coordinate) and the fraction volume (x-coordinate) (step 3). These varied data points were then fitted again in order to calculate the retention time and the peak area (step 4). For each fraction size, the concentrations and the volumes were altered 100 times each resulting in 10,000 simulated peaks per fraction size. Finally, the 95% confidence intervals of the mean retention time and mean peak area of all simulated peaks were calculated (step 5) and plotted versus the fraction size normalized to the widths of the peaks, thus, indicating the precision in relation to a specific fractionation scheme.

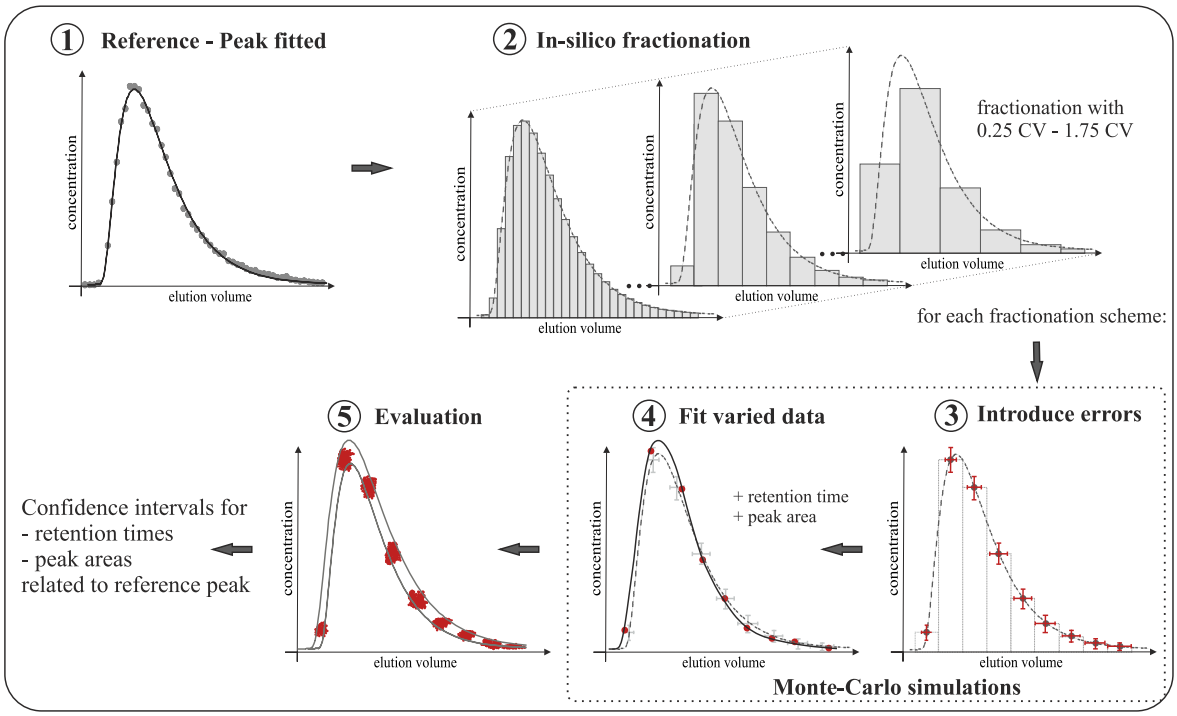

Figure 3: Illustration of the Monte-Carlo simulation procedure. Randomly distributed errors of different levels on fraction volume and concentration were added in 10,000 different combinations. See text for further explanation.

All errors were assumed to be standard normal distributed and were added to the initial volume and concentration value for each fraction ($y_{ref,frac,V}$ and $y_{ref,frac,c}$) resulting in a simulated data point defined by the modified volume $\widetilde{y}_V$ and concentration $\widetilde{y}_c$:

Table I: Error correlations used in the Monte-Carlo simulations. The path length in full area plates was assumed to be identical to the length in half area MTPs with half of the volume. The constants k was 75 μL in the case of fractions <150 μL (half area MTPs) and 150 μL for fractions ≥150 μL (full area MTPs).

Parameter	Error term	Error source	Notation
volume (V)	$s_{V,frac}$= $a/100$	calibration + detector	$V_{a\%}$
	$E_{syst,V}$= $0.075 \cdot \exp(-11c) \cdot \frac{V}{k}$	meniscus	$V_{syst,meniscus}$
concentration (c)	$s_{c,frac}$=$0.05 \cdot \exp(-8c) + 0.001 \cdot c^3 + p$	analytical error	
	$p = 0.004$		$c_{0.5\%}$
	$p = 0.019$		$c_{2\%}$
	$p = 0.099$		$c_{10\%}$
	$E_{syst,c}$= 0		
pipetting	$s_{pip10\mu L,frac}$=0.038	precision, 1 mL plunger	$Pip_{dil:V}$
	$s_{pip50\mu L,frac}$= 0.0064	precision, 1 mL plunger	$Pip_{dil:V}$

$$\widetilde{y}_V = y_{ref,frac,V} + E_{syst,V} + E_{random,V}, \tag{1}$$

with

$$E_{random,V} = y_{ref,frac,V} \cdot s_{V,frac}, \tag{2}$$

and

$$\widetilde{y}_c = y_{ref,frac,c} + E_{syst,c} + E_{random,c}, \tag{3}$$

with

$$E_{random,c} = y_{ref,frac,c} \cdot (randn \cdot s_{c,frac} + randn \cdot s_{pip,frac}). \tag{4}$$

The systematic deviation was given by the absolute values E_{syst}. The statistical errors E_{random} were calculated by multiplying a random number normally distributed around the mean zero with a function $s_{V,frac}$,$s_{c,frac}$, or $s_{pip,frac}$ representing the relative standard deviation (RSTD). An overview on the used error terms is given in Table I.

The error functions for the determination of concentration were selected to cover a wide range of analytical error levels and were based on data reported for photometric analysis in 96-well plates [19, 25]. The functions incorporated the experimental observation that the common random error increases for very small and for large concentrations. Systematic errors were not considered.

For the determination of the volume by photometric analysis, a RSTD of 2% was experimentally derived and used in all simulations unless stated differently. This value included the pipetting error during calibration and photometric precision with a RSTD of ∼0.5%. Additionally, a concentration-dependent, systematic deviation on the determined volume, $E_{syst,V}$, was calculated for each fraction representing the error caused by meniscus formation as will be presented for half area plates with 75 μL in section 3.2. This absolute error linearly depended on the liquid level and hence on the choice of half or full area plates.

The variance introduced by pipetting for dilution was added to fractions with concentrations >0.2 g/L. In this study, we chose to add an average variance for the dilution step on the concentration value rather than calculating the dilution error in dependence of the concentration, as the concentrations of the considered peak were relative low. For a 20-fold dilution (10 μL protein solution), concentration values were varied with a RSTD of 3.8%, whereas a RSTD of 0.64% was used for a four-fold dilution (50 μL protein solution).

For evaluation of the error impact, the 95% confidence intervals of the simulated peaks were plotted versus the fraction size normalized to the widths of the peaks, thus, indicating the precision in relation to a specific fractionation scheme.

2.8 Simulations of cation-exchange chromatography

2.8.1 Model description

A general rate model was used to simulate the mass transfer in the chromatography column including convection, dispersion, film diffusion and intra-particle diffusion. The steric mass-action model (SMA) introduced by Brooks and Cramer [26] was chosen as sorption model. The employed solver of the general rate model followed von Lieres and Andersson [20]. The Matlab®routine was adapted for the simulation of HTCC, i.e. incorporating a multiple step elution with variable step height and step volume (mimicking the salt gradient) and the corresponding flow interruptions for the uptake of each salt buffer. The corresponding model equations describing dispersive and convective transport in the bulk phase, film mass transfer and diffusion in the beads, and adsorption and desorption kinetics can be found in appendix A.2. For further information on the discretization and solution of the model equations we refer to von Lieres and Andersson [20].

2.8.2 Model parameters

The determination of the SMA-isotherm parameters was performed using the inverse method as presented by Osberghaus *et al.* [27]. As data for calibration, we used experimental data from elutions on the LHS with salt gradient slopes of 525 mM NaCl over 8 CV, 13 CV, 18 CV, 23 CV, and 28 CV. For both lysozyme and avidin, the characteristic charge, ν, the shielding factor σ and the ad- and desorption equilibrium constants k_{ad} and k_{des} were determined by minimizing the least-squares residuals to experimental data using the matlab function `lsqnonlin`. The isotherm parameters resulting from the fitting procedure are given in Table II.

Table II: SMA-isotherm parameters for lysozyme and avidin using SP Sepharose FF at pH 4.5

Parameter	Lysozyme	Avidin
ν	6.25	11.18
σ_{SMA}	10	45
k_{ad}	55	256
k_{des}	80	811

The used process parameters and column model parameters are summarized in Table III. Porosities were obtained in pulse experiments using a penetrating (acetone) and non-penetrating tracer (2 MDa Dextran). An ionic capacity Λ of 0.8 M as reported in literature for SP Sepharose FF was used [27]. Axial dispersion, film and intra-particle diffusion coefficients were retrieved from manufacturers, literature, or by empirical correlations.

2.9 Protein analytics

For rapid protein quantification based on absorption measurements, a Tecan plate reader (Magellan, Tecan, Crailsheim, DE) was used. For selective protein quantification, Reversed Phase (RP) HPLC was performed on a UHPLC system as described in [32]. The lower limit of quantification (LLOQ)

Table III: Process model parameters for HTCC.

Parameter	value	determination
process and column parameters		
column length, L	0.01 m	manufacturer information
column diameter, d_{col}	0.005 m	manufacturer information
linear flow rate, u	36-144 cm/h	–
column porosity, $\epsilon_{col}c$	0.33	experimental, moment analysis
particle porosity, ϵ_p	0.82	experimental, moment analysis
particle diameter, d_p	90 $\cdot 10^{-6}$ m	manufacturer information
axial dispersion coefficient, D_{ax}	6.8 $\cdot 10^{-8}$ m^2/s	estimated by correlation [28]
mass transfer parameters		
film mass transfer coefficient, k_f	NaCl: 1.9 $\cdot 10^{-4}$m/s	literature [29]
	lysozyme: 3.5$\cdot 10^{-5}$ m/s	estimated by correlation [30, 31]
	avidin: 2.4$\cdot 10^{-5}$ m/s	estimated by correlation [30, 31]
effective intra-particle diffusion coefficient, D_p	NaCL: 5.9$\cdot 10^{-10}$ m^2/s	literature [29]
	lysozyme: 5.75$\cdot 10^{-11}$ m^2/s	estimated from literature [31]
	avidin: 2.33 $\cdot 10^{-11}$ m^2/s	estimated from literature [31]

were estimated by 10 times the noise signal from blank injections. The average RSTD for protein concentrations >50 mg/L was less than 1%.

3 Results and discussion

The challenge of investigating the influence of each single factor lies in finding proper methods to distinguish between effects exclusively caused by the robotic method and the general effects of working in such small scale. The following result sections start with investigations on very basic effects and gain complexity with sequentially incorporating more aspects of gradient elution in HTCC.

3.1 Pipetting accuracy

Pipetting accuracy is probably the first error source to be considered when discussing the goodness of automated methods. In common control software the researcher can adjust many pipetting parameters such aspirate and dispense speed as well as experimentally derived correlations ('Liquid Classes') between plunger movement and pipetted liquid volume for each type of liquid. However, optimization of a large number of Liquid Classes can become time-consuming, especially if with each class a wide volume range needs to be covered accurately. Liquid Class parameters and data on pipetting accuracy can be found in literature [12, 17, 33, 34] demonstrating that an optimization of the Liquid Class is necessary in order to reduce the deviation between nominal and true pipetted volume below 1%.

In Figure 4, the RSTD is plotted as a function of the nominal pipetting volume when using optimized Liquid Classes for the pipetting of a phosphate buffer and a protein solution (1 g/L avidin). Two liquid handling set-ups were compared: first, using stainless steel (non-coated) tips and a 1 mL plunger volume, and second, using low volume stainless steel tips (FEP-coated) and a plunger volume of 0.25 mL. The pipetting precision decreased exponentially with a decreasing pipetted volume which is in concert with trends reported by Treier [12] and Osberghaus [17]. When using 1 mL plungers the precision decreased strongly from RSTD values <1% for volumes larger than ∼25 μL to a RSTD of ∼10% for pipetting less than 5 μL. With a reduced plunger volume, a RSTD of <1% was obtained for volumes larger ∼15 μL and the precision was reduced to 5% when pipetting 3 μL. The data clearly demonstrates that low volumes should be avoided for sample dilution, i.e. volumes lower than 15 μL.

The average pipetting accuracy of all channels was 100%±1% when using optimized Liquid Classes calibrated in three volume ranges (2 – 15 μL, 15 – 55 μL, 55 – 200 μL) and for each needle tested in

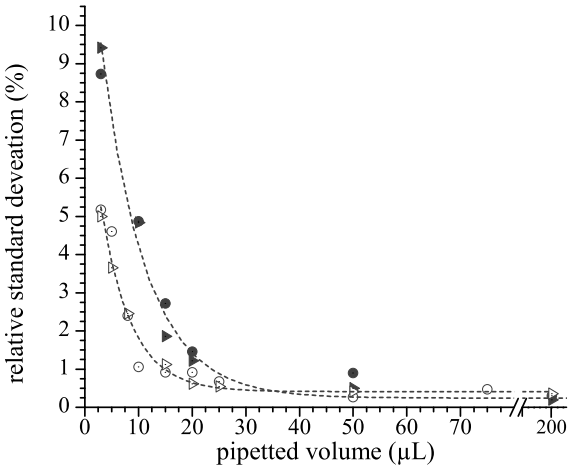

Figure 4: Relative standard deviation as a function of the pipetted volume. -●-: 50 mM phosphate buffer, pH 7, with fixed tip, stainless steel non-coated, 1 mL syringe; -○-: 50 mM phosphate buffer, pH 7, with low volume fixed tip, stainless steel FEP-coated, 0.25 mL syringe; -▶-: 1 g/L avidin with fixed tip, stainless steel non-coated, 1 mL syringe; -▷-: 1 g/L avidin with low-volume fixed tip, stainless steel FEP-coated, 0.25 mL syringe.

duplicate. In terms of accuracy, there was no significant difference obtained when comparing the two types of set-ups. The accuracy depends only on the goodness of the calibration, and therefore, only the random variance in pipetting was taken into account in this study.

3.2 Absorption measurement in micro-titer plates

Photometric analysis is probably the most suitable analytical technique for high-throughput screenings. It is very fast, cheap, label-free, non-invasive, and can be automated on LHSs in 96-well and 384-well plate format. Most commonly, the absorption at a wavelength of 280 nm is used for the quantification of the protein content. Furthermore, the liquid level within a well and thus the liquid volume can be linearly correlated to the absorption difference of wavelengths $\geq$ 970 nm and at 900 nm [12, 35].

The accurate determination of volume balances and of mass balances (if protein concentration is analyzed by absorption measurements) hence depends on the accuracy of the photometric measurement in MTPs and the quality of the calibration curves. The main differences of measurements in micro-plate wells compared to measurements using cuvettes are the variable path length. The dropwise collection of the elution volume during a HTCC experiment can lead to deviations in collected volume between fractions by one drop ($\sim$25 μL), and hence, the path length is not constant for all fractions. Furthermore, meniscus formation results in a decrease of the path length and, hence, in a decrease of the absorption value according to Lambert Beer's law. The curvature of the meniscus and hence the path length depend on liquid/protein characteristics and protein concentration, all having an influence on the liquid surface tension [36].

Determination of volume based on absorption difference

A typical calibration curve obtained when measuring the absorption difference $A_h = A_{990nm}$ - A_{900nm} as a function of the pipetted volume in half area and full area MTPs is given in Figure 5a for a constant protein concentration of 0.6 g/L lysozyme and for a 100 mM phosphate buffer solution, pH 7.0. The linear relationship as described in literature [12, 35] was confirmed by achieving coefficients of

determination for linear regression R^2>0.998 (8 replicates for each volume level were performed). For both microplate types, A_h-values were higher when using buffer compared to using protein solution with an offset being rather constant for all volumes. In Figure 5b, A_h obtained in half area and full area MTPs at volumes of 75 μL and 150 μL, respectively, is shown as a function of a varying lysozyme concentration. The decreasing absorption difference measured is due to a path length which decreases exponentially with increasing protein concentration until a rather constant value is obtained for lysozyme concentrations higher than approx. 0.15 g/L. The formation of a meniscus, being the cause of a decreasing path length with increasing protein concentration, thus clearly affects the use of a single calibration curve with constant protein concentration as shown in Figure 4a, and thus, the calculation of liquid volume in MTPs based on path length determination via $A_{990nm} - A_{900nm}$. As the meniscus can change with the composition of the protein sample, the accurate determination of the volume using an univariate calibration becomes practically not feasible.

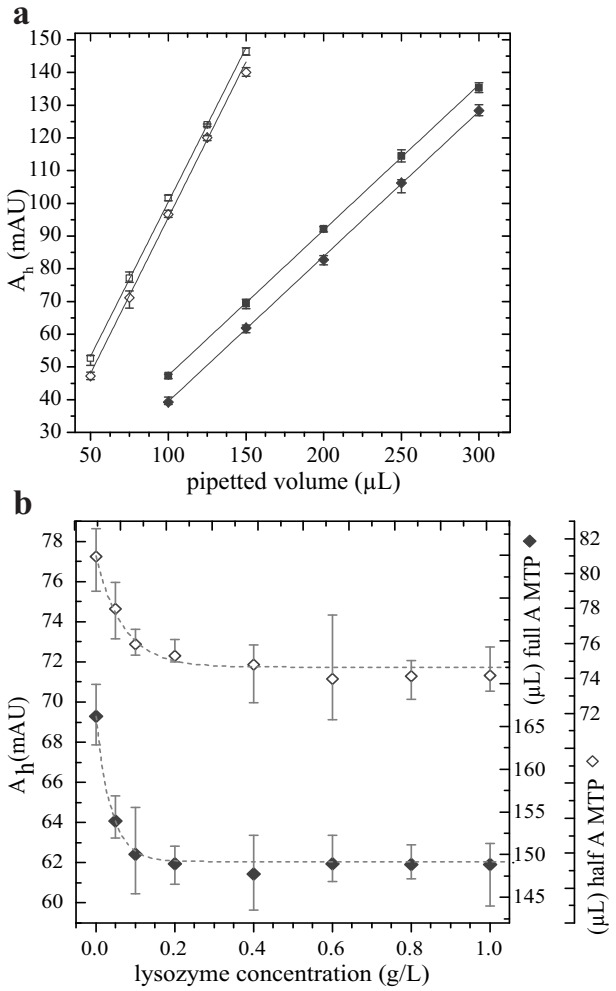

Figure 5: a: Absorption difference A_h=A_{990nm} - A_{900nm} as a function of pipetted volumes of 100 mM phosphate buffer (marker squares) and 0.6 g/L lysozme (marker diamonds) in half area MTPs (open marker) and full area MTPs (filled marker). b: Absorption difference A_h=A_{990nm} - A_{900nm} obtained in half area (-◇-) and full area MTPs (-♦-) at nominal volumes of 75 μL and 150 μL, respectively, as a function of the lysozyme concentration in 100 mM phosphate buffer, pH 7.0. The y-axis to the right is obtained relying on a calibration curve established by using a lysozyme solution of 0.6 g/L as displayed in Figure 4a.

The different path lengths for a 1 g/L lysozyme solution compared to the buffer blank resulted in deviations of predicted volumes of ∼6 μL and ∼16 μL in half area MTP wells and full area MTP wells, respectively. This equals an error of 8% and 11% for sample volumes of 75 μL and 150 μL. These values are in concert with data Treier *et al.* [12] reported as offsets between protein solutions and water (18 μL, full area MTP).

Determination of protein concentration based on UV 280 nm
When using common cuvette measurements the path length is constant and thus a fixed value when correlating A_{280nm} and protein concentration. In MTP measurements, however, the increased path length for samples with low protein concentrations leads to higher absorption values which is an effect not represented by a linear calibration curve for A_{280nm}. It may thus be advisable to correlate A_{280nm}/A_h to the protein concentration when establishing the calibration curve. Furthermore, we have observed that the constant value of a linear correlation between A_{280nm} and the protein concentration usually is higher than the absorption value of the buffer (data not shown). Therefore, one can conclude that the blank value should not be included in the calibration data set.

In conclusion, when using micro-plate absorption measurements for the determination of protein concentrations and fraction volumes, potential meniscus effects must be considered when setting up the calibration curves. Deviations from 100% in the total volume balance of a HTCC experiment can possibly be explained by meniscus effects rather than false pipetting. The procentual error caused by meniscus formation however decreases proportionally with increasing sample volume. Thus, the error can be reduced by choosing larger sample volumes and fractions.

3.3 Effect of fractionation and number of fractions

The number of collected fractions is an important parameter to be set when performing HTCC. The fractionation using low volumes, on the one hand, results in more information on the peak shape and provides a higher resolution which, in turn, provides the possibility of performing fractionation and peak cutting comparable to a large-scale process. On the other hand, the analytical effort obviously increases with a larger number of fractions and the number of fractions might be limited by the minimum fraction volume required for subsequent analysis. The influence of the chosen fraction size regarding the determination of the peak area and the retention time was investigated for two different peak shapes, a rather symmetric peak (Figure 6a) and a asymmetric, strongly tailing peak (Figure 7a). By using Monte-Carlo simulations, different error levels for the determination of volume and protein concentration as well as pipetting errors from potential dilution of fractions were investigated using the procedure described and visualized in section 2.7 and Figure 3.

The effect of a varying number of fractions per peak width on the peak area and peak retention time is displayed in Figure 6 and Figure 7. For both investigated peak types, error scenarios are compared for errors introduced during determination of the fraction volume (V) (Figure 6b,c, and Figure 7b,c), the concentration (c) (Figure 6d,e, and Figure 7d,e), and dilution procedures (Pip) (Figure 6f,g, and Figure 7f,g).

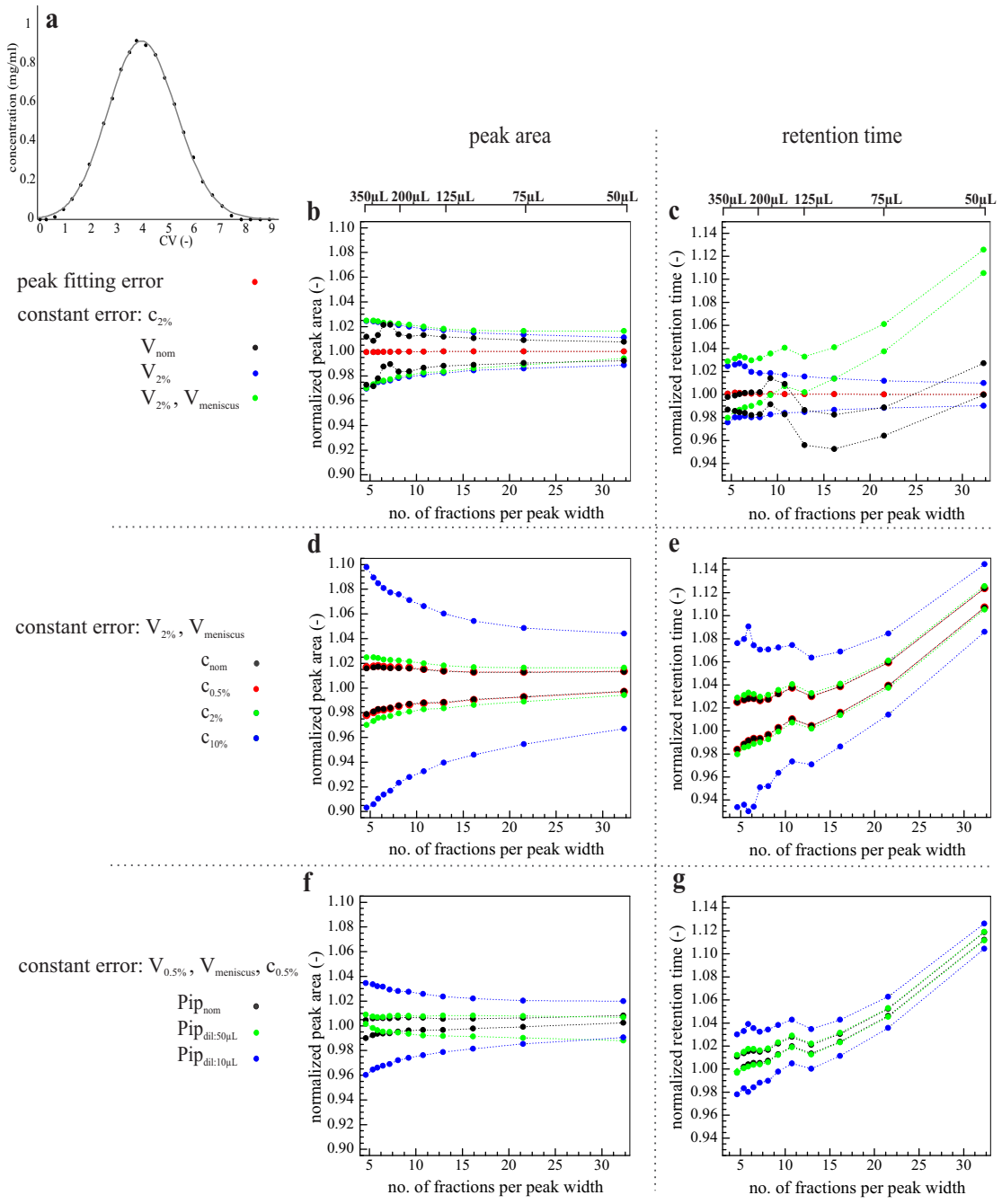

Figure 6: Effect of the number of fractions per peak width on the peak area (b,d,f) and peak retention time (c,e,g) determined after peak fitting for a symmetrical elution peak displayed by 95% confidence intervals. The confidence intervals were calculated for peak area and retention time of the simulated peaks related to the values of the start peak. Lower and upper boundaries of the intervals are plotted in same color. Introduced error categories were b,c: determination of volumes; d,e: determination of protein concentration; f,g: dilution procedures. Following errors introduced were set constant: b,c: $c_{2\%}$; d,e: $V_{2\%}, V_{meniscus}$; f,g: $V_{0.5\%}, V_{meniscus}, C_{0.5\%}$.

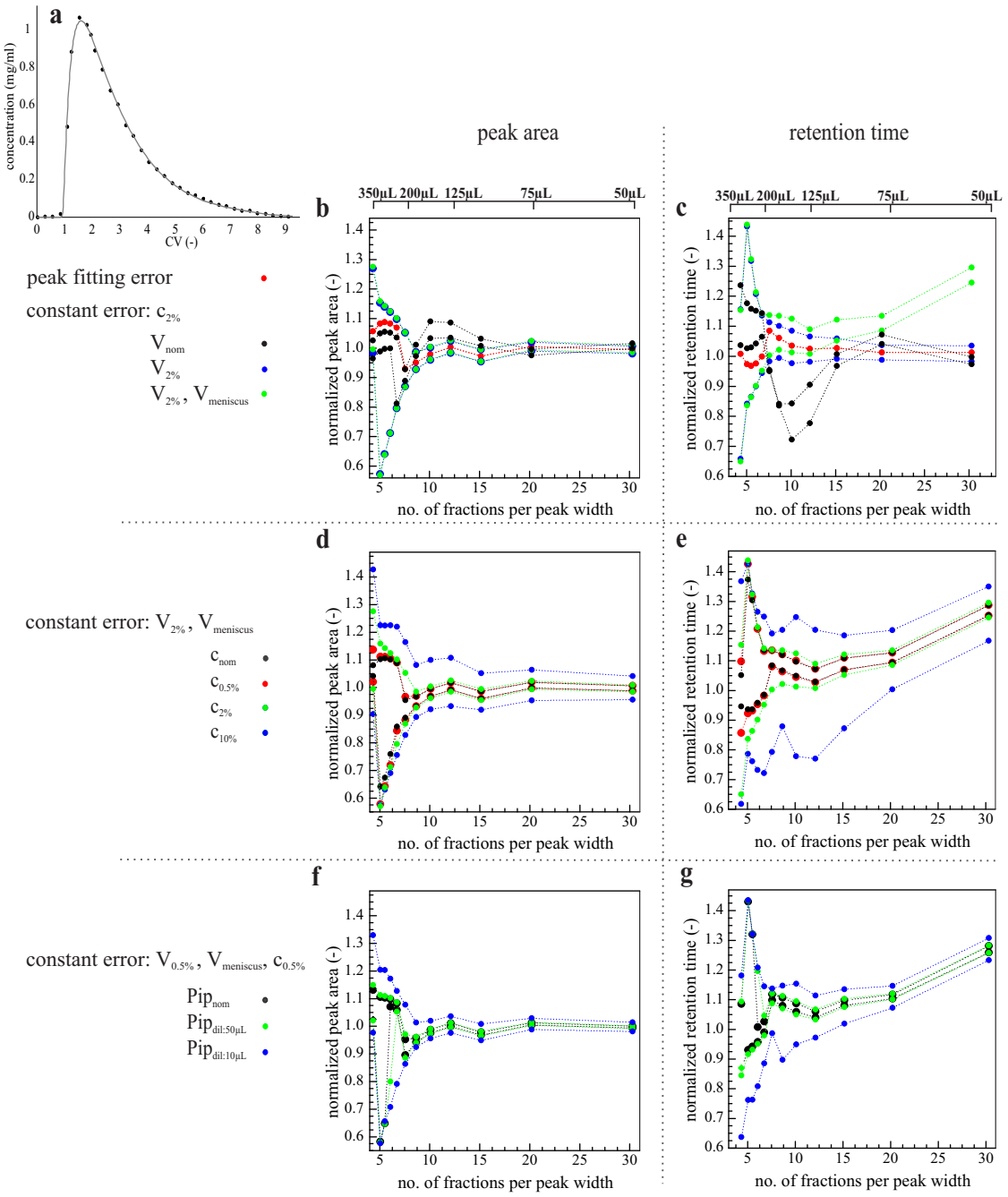

Figure 7: Effect of the number of fractions per peak width on the peak area (b,d,f) and peak retention time (c,e,g) determined after peak fitting for an asymmetrical elution peak displayed by 95% confidence intervals. The confidence intervals were calculated for peak area and retention time of the simulated peaks related to the values of the start peak. Lower and upper boundaries of the intervals are plotted in same color. Introduced error categories were b,c: determination of volumes; d,e: determination of protein concentration; f,g: dilution procedures. Following errors introduced were set constant: b,c: $c_{2\%}$; d,e: $V_{2\%}, V_{meniscus}$; f,g: $V_{0.5\%}, V_{meniscus}, C_{0.5\%}$.

3.3.1 Error by peak fitting

The error caused by peak fitting was relatively small (<0.2%) for the symmetrical peak (Figure 6b,c). For the asymmetrical peak, however, a low fraction number of <8 – 10 resulted in deviations up to 9% and 4% for the determined peak area and retention time, respectively (Figure 7,b,c). It becomes clear, that even if the applied analytics was free of errors, accurate results for asymmetrical peaks would only be achieved when using a large number of fractions per peak width (e.g. 20). The reason for a decreased accuracy when fitting asymmetrical peaks lies in the steep peak front of the asymmetrical peak. With changing fraction sizes, the fraction that covers the peak maximum is pointing either to the peak front or the peak tail. For which number of fractions the peak maximum shifts is furthermore influenced by the start of the fractionation. Nevertheless, the shown overall trends and the demonstrated sensitivity of the asymmetrical peak towards the number of fractions are expected to hold true regardless of a change in starting volumes or the elution peak width.

3.3.2 Effect of fraction number - general trends

For all investigated error levels, the confidence intervals decreased with an increasing number of fractions per peak width. This is as expected, because a peak shape is obviously better represented by more data points as well as the errors are averaged and partly cancel each other out when fitting the peak. From Figure 6 and 7 it can be seen that the decrease of the confidence interval approached a width which did not decrease further with increasing number of fractions. This limit was reached for fraction volumes of <100 μL and represented the approximately maximal feasible experimental precision at corresponding error levels. For a number of fractions ≥15, the width of the confidence intervals of asymmetric and symmetric peaks became similarly small, thus indicating that an effect of the peak shape can be neglected if a large number of data points per peak width are available to fit the peak. Precision of the determination of the retention time and peak area increased with increasing number of fractions. Accuracy was only increasing for the peak area, whereas for the retention time an increasing discrepancy to the nominal retention time with increasing number of fractions was observed as explained in the following section.

While the overall trend of increasing precision with increasing number of fractions was the same for both peak shapes, huge differences in the absolute confidence intervals as well as in the shape of the confidence interval bands were determined. The confidence intervals obtained for the asymmetrical peak were up to ±45% when using a fraction number of five, whereas the intervals obtained for the symmetrical peak shape were up to 10% for five fractions. For the asymmetrical peak, large variations in the confidence intervals were obtained even if the number of fractions was changed by only one (e.g. changing the number of fractions from five to four, Figure 7). These spikes in the confidence intervals result from the peak fitting algorithm, by which for the altered number of data points a significantly different peak maximum and hence retention time is identified.

3.3.3 Accuracy of retention times

In all simulations in which the meniscus formation of a liquid in a MTP was taken into account ($V_{meniscus}$) an increase in retention times with increasing number of fractions was obtained. This can be explained with the data given in section 3.2 and Figure 5. For the simulations, we chose a calibration scenario for the determination of the fraction volumes using a protein solution with a concentration >0.2 g/L. From Figure 5 it becomes clear that the volume with concentrations >∼0.2 g/L will be

determined correctly using this calibration, whereas the volume of fractions with concentrations below will be determined higher than the actual values. As shown in Figure 5, the offset of an additional, false volume for buffer fractions compared to protein fractions is up to about 6 μL and 16 μL for half-area MTPs and full-area MTPs, respectively. Randomized volumes up to these values were therefore added as absolute values to the fraction volume, and hence, the larger the number of fractions, the larger the absolute error and shifting towards later retention times. The widths of the confidence intervals and hence the precision, however, was not influenced by V_{men}. Obviously, the retention times would become shorter when calibrating using a buffer solution or might barely be influenced when a different volume-offset is present (protein-specific). It is therefore essential to investigate the calibration for the volume determination when using a high number of frations and aiming at an accurate determination of the retention times.

3.3.4 Effects of volume errors

The effect of errors introduced during the determination of fraction volumes is shown in Figure 6b,c, for a symmetric peak and in Figure 7b,c for an asymmetric peak. Under the assumption of a constant error of 2% introduced by the determination of protein concentration, $c_{2\%}$, three different scenarios were compared: a) assuming the preset constant nominal fraction volume, V_{nom}, b) volumes with a constant RSTD of 2%, $V_{2\%}$, and c) volumes with a constant RSTD of 2% combined with the effect caused by meniscus formation, $V_{meniscus}$.

As can be seen when comparing case b) and c) the addition of errors caused by meniscus formation had very little effect on the widths of the confidence intervals. However, the accuracy of the determination of retention times strongly decreased with an error of about +11% and +27% for the symmetric and asymmetric peak, respectively, due to the assumption of using a protein solution for the calibration of the volume determination as explained above. When omitting the determination of fraction volumes and instead using the nominal fraction volume (case a)), the precision was higher for a number of fractions $\sim$<10 and the accuracy of the determined retention time was higher for a number of fractions $\sim$>20. This is an interesting result demonstrating that an additional effort of measuring and calculating the true volumes (V_{men}) in fact can result in less accurate retention times. Omitting the volume determination would furthermore be beneficial as the experimental and evaluation method can be simplified. However, when not measuring the fraction volumes, significant deviations to the actual retention times were seen for a number of fractions of 12 – 22 in Figure 6c and for a number of fractions of 8 – 13 in Figure 7c. The reason for such deviations lies in the collection of the eluate in droplets. The fractions often differ in the number of collected droplets, which was taken into account for the reference and simulated peaks by alternating the fraction size by half of a drop volume (13 μL). Hence, the x-coordinate for the concentrations being fitted were different for the scenario using fixed nominal values compared to the scenario taking the droplets into account. For both considered peaks, fitting of the peak was thereby affected resulting in deviations of accuracy using V_{nom}.

3.3.5 Effects of concentration errors

The effect of errors introduced during the determination of protein concentration is shown in Figure 6d,e for a symmetric peak and in Figure 7d,e for an asymmetric peak. The errors applied for the determination of protein concentrations were set to 0% (c_{nom}), 0.5%, 2%, and 10% (see Table I) while keeping an error of the determination of fraction volumes ($V_{2\%}$ + $V_{meniscus}$) constant for all systems.

A RSTD of 0.5% in the concentration did not result in a noticeable change on the precision when compared to simulations without an concentration error. Such low error levels apply if precise analytical techniques, such as HPLC, are available for the determination of the concentration in each fraction. A decrease in analytical precision from 0.5% to 2% resulted in an absolute increase of the confidence interval of maximal 0.1 for the symmetric peak (Figure 6d). For the asymmetric peak, the effect was larger for a number of fractions <8 with an increase of the confidence interval by maximal 0.25 (Figure 7e). In the case of the symmetric peak and for a large analytical error of 10%, the determined confidence interval for the peak area and the retention time decreased with increasing number of fractions from approximately 0.2 to 0.08 and from approximately 0.15 to 0.06, respectively. In the case of the asymmetric peak and an analytical error of 10%, the determined confidence interval for the peak area and the retention time decreased with increasing number of fractions from approximately 0.55 to 0.1 and from approximately 0.75 to 0.18, respectively.

It becomes clear that the compensation of the analytical error by using a larger number of fractions is greater for a large error in the determination of the concentration compared to a rather high precision of the analytical technique (e.g. $c_{0.5\%}$). For an asymmetrical peak, using a larger number of fractions (e.g. >20) would therefore be more effective than optimizing the precision of the analytical method. In contrast, the precision for symmetrical peaks is mostly dominated by the precision of the analytical technique and the optimization thereof outweighs the benefit from using more fractions.

3.3.6 Effect of dilution errors

The effect of errors introduced during a dilution process is shown in Figure 6f,g for symmetric and in Figure 7f,g for asymmetric peak shapes. Under the assumption of error levels for the determination of fraction volumes, $V_{0.5\%}$,$V_{meniscus}$ and for the determination of protein concentrations, $c_{2\%}$, for all systems, three different scenarios for the pipetting during dilution were investigated: a) no pipetting error (Pip_{nom}), a dilution by pipetting 50 μL of protein solution, $Pip_{dil:50\mu L}$, and a dilution by pipetting 10 μL of protein solution, $Pip_{dil:10\mu L}$. For these two dilution volumes, a constant RSTD derived from the data presented in Figure 4 was used in the simulations.

A dilution using a large volume did not result in a significant decrease of precision, since the pipetting error is small, as discussed in section 3.1. With decreasing sample volume used for dilution and hence an increasing pipetting error, the effect on the precision for the determination of the peak area and retention time increases. This effect was seen from the simulation results for both the symmetric and asymmetric peak. In the case of the symmetric peak, the confidence intervals increased up to 0.055 (Figure 6f), whereas in the case of the asymmetric peak, the confidence intervals increased by up to 0.26. An increase in precision at a low number of fractions for the low pipetting error as seen in Figure 6f is assumed to be caused by the errors canceling each other out. Furthermore, for a larger number of fractions the mean of the confidence interval shifted from approx. 1.01 for (Pip_{nom}) to the expected value of 1.00 $Pip_{dil:50\mu L}$ (Figure 6f), demonstrating that in some cases accuracy might increase despite a reduction of the precision.

The fraction volume needs to be adequately large to enable the use larger volumes for dilution. The use of larger fraction volumes, however, is in contrast to the general conclusion that a higher number of fractions, i.e. smaller fraction volumes, should be used to improve precision.

The error caused by dilution depends also on the concentration in the eluate. With increasing concentrations the absolute error caused by dilution increases significantly, which results in higher

uncertainty of the peak fitting result compared to eluting peaks of lower concentrations.

In summary, it was demonstrated that even with very precise analytical tools the choice of a low fraction number can not be balanced out. If the analytical precision is known to be imprecise, however, a sufficient level of the overall precision on the mass balance and retention time still can be reached by increasing the number of fractions. With about $\geq$10 fractions per peak width, the experimental precision and the error caused by strong asymmetry can be significantly reduced. The maximal possible precision for the retention time and mass balance was estimated by the simulations using the low error levels $V_{0.5\%}$ and $c_{0.5\%}$ and can be reached by using a fraction number of 15 – 20. For long elution times and hence a large number of buffer fractions or low concentrated protein fractions, volume determination affected by meniscus formation can result in significantly inaccurate retention times. In that case, the use of the preset fraction volume appears favorable. It must be noted that the results for the asymmetric peak could be different when changing the asymmetry and peak width, because the peak fit strongly depends on the number of data points by which the peak front is represented.

3.4 Influence of flow regime

During HTCC, the pipetting needles leave the columns multiple times as explained in section 1.1, Figure 1. In contrast to typical laboratory LC systems, the fluid flow is thus interrupted for a period of time required for the uptake of new buffer solutions. This leads to longer residence times and potential reflux caused during withdrawal of the needle. In order to investigate the quantitative influence of robotic process steps on the chromatographic behavior in isocratic HTCC, the following fluid flow pattern for a step elution from a 200 μL RoboColumn were analyzed:

1. 'continuous' flow
2. 'flow interruptions' every 150 μL for 19 sec, keeping needles in columns
3. 'needle' movement out and into the column every 150 μL[1]
4. 'combination' of flow interruption and needle movement.[2]

In Figure 8, the retention times, peak widths and asymmetries determined after peak fitting with the EGH function are displayed including mean values and 95% and 99% confidence intervals.

No effect of the chosen flow pattern on either retention time, peak width or peak asymmetry was observed when using the protein lysozyme. The differences between the flow patterns were statistically not larger than the differences between the columns. The elution profiles of avidin, however, showed an effect when comparing the 'continuous' flow with the other patterns, which all featured flow interruptions. In the case of avidin, the mean retention time and the peak width were both reduced by ≈20% compared to the three other flow patterns. The asymmetry was increased by factor 1.8. Between the flow patterns 'flow interruptions', 'needle', and 'combination' no differences were determined.

The difference between the two proteins indicates that the influence of changing from a continuous flow to a flow pattern incorporating flow interruptions depends on the diffusion rates specific for the

[1] this implied an interruption of the flow, approximately 14 sec
[2] the total time for buffer uptake performing all program steps took 40 sec

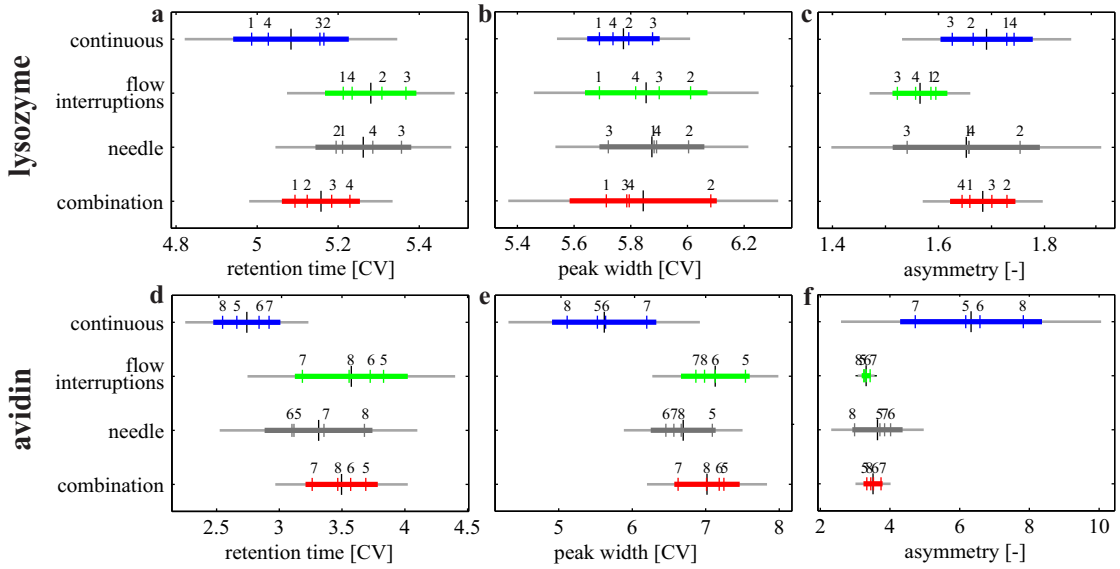

Figure 8: Influence of the used fluid flow pattern on peak retention time, peak width at 50% peak height and peak asymmetry. For each fluid flow pattern, the same four columns were used in parallel for each protein and are indicated with numbers. The determined values (|, colored), the mean (|, black), and the 95% and 99% confidence intervals widths (horizontal bars, colored and light gray, respectively) are displayed. A trend was determined, if the 95% confidence intervals overlayed vertically. The effect was considered significant, if also the 99% confidence intervals for each flow pattern overlayed vertically.

protein/resin system as well as on the flow rate applied. With changing residence times and flow rates the ratio of external mass transfer rate to intra-particle diffusion, expressed by the Biot number, changes, eventually resulting in a change in the rate-limiting step [37, 38]. The diffusion coefficient of avidin is about three times smaller than the value for lysozyme explaining among others the larger difference of the peak width and asymmetry between continuous flow and flow with interruptions for avidin compared to lysozyme.

Concentration changes after flow interruptions were not detected in the elution profiles of any case (data not shown). This indicates a 'smoothing' effect when collecting the eluate in fractions compared to the continuous signal from a flow-cell. The effect of flow interruptions, however, will be detectable when performing breakthrough curves. Wiendahl *et al.* [9] reported that the concentration during protein breakthrough dropped significantly when interrupting the flow for aspirating new protein solution, hence affecting the shape of the breakthrough curve.

The frequent movements of the pipetting needles out of the column and into the columns did not lead to a detectable change in retention time and peak shape. There was no liquid volume detected by visual detection being soaked from the column outlet into the columns by an underpressure generated when the needles were withdrawn out of the columns.

When comparing the results for each column with each other no statistically relevant differences were detected that would indicate deviations in packing quality or bed volume. Since identical buffers and protein solutions were used and each column was operated each time with the same pipetting channel, the systematic errors were considered being negligible compared to the overall variance. The mean mass and volume balances and their uncertainties for all 2 x 16 experiments were calculated in order to estimate the precision for yield calculations. For lysozyme, the average mass balance was 94%±1.5% (RSTD of 3.1%). The mean volume balance for lysozyme experiments was 101%±1.3%

(1.7% RSTD). In the case of avidin, the mean mass balance was 98%±2.1% (RSTD 4.1%) and the mean volume balance was 102%±0.4% (RSTD 0.8%). The obtained deviations were therefore close to the theoretical value for the peak area confidence interval (approx. 2%) determined by simulations of fractionation schemes and a corresponding 2% error level (see section 3.3).

3.5 Gradient elution experiments

When performing isocratic elution in HTCC mode, flow interruptions can be reduced to a minimum required for changing fraction collection plates. When performing gradient elutions, however, flow interruptions for buffer uptake can not be avoided when using current HTCC hardware set-ups. During HTCC, the application of salt-gradients is mimicked by sequential pipetting of buffers with varying salt concentrations or pH (Figure 1). The salt concentration c_{salt} for each buffer is calculated by equation 5 with n being the total number of salt steps:

$$c_{salt} = c_{gradient,start} + \frac{c_{gradient,end} - c_{gradient,start}}{n} \cdot \frac{2i-1}{2}, i = 1, 2, 3...n. \quad (5)$$

It is obvious that the smaller the applied salt steps (and thus the more steps), the closer the approximation to a gradient applied on a laboratory LC system. In a study published by Welsh *et al.* [39], the effect of three different salt step heights on the elution profile was qualitatively shown demonstrating a significant influence of the salt steps on the elution profile. Furthermore, with each flow interruption required the discrepancy between applied flow rate and average residence time increases. In Table IV, ratios between times of flow and interruptions due to buffer uptake were exemplarily calculated for 12 and 48 interruptions during elution. Significant differences in residence times occur depending on the chosen flow rate and frequency of flow interruptions.

Table IV: Relation of flow rate and residence times during elution with 12 and 48 interruptions in HTCC experiments. The duration of a flow interruption ('pause') for wash and buffer uptake was 30 s.

flow rate (μL/s)	no. of pauses	average residence time (s)	ratio to continuous flow
1	12	220	1.1
1	48	280	1.4
3	12	147	2.2
3	48	87	1.3
8	12	45	1.8
8	48	105	4.2

In the following sections, investigations of gradient elutions in HTCC are presented adressing:

1. the effect of different salt step heights mimicking a gradient
2. the combined influence salt steps, flow interruptions and flow rate on elution performance
3. the comparability of HTCC, simulation, and common laboratory LC experiments.

As it is not possible to investigate separately the effect of different salt step heights and flow interruptions for salt-gradient runs experimentally on a LHS, simulations were used for the comparison of various scenarios of salt step heights, flow interruptions, and fractionation. For the simulations, a general-rate model and the SMA isotherm model (see section 2.8 and the appendix) was applied.

3.5.1 Salt step height

In a first step, the separation of lysozyme from avidin was simulated using NaCl step heights between 11 mM and 66 mM (corresponding to buffer steps from 75 μL to 450 μL, respectively). The qualitative effect of the salt step height can be seen from the simulated chromatograms in Figure 9.

The application of buffer salt steps caused a deviation from the common peak profile. With increasing salt step height the deformation of the peaks became more pronounced. When using salt step heights below 11 mM, an effect of buffer steps was not visually detected. For salt step heights of 21.9 mM, the effect of using salt steps still appeared to be negligible. For salt step heights of 54.7 mM and 65.6 mM, the avidin peak eventually split into two main peaks. It became clear that peak fitting is required in order to enable determination of peak retention times. However, despite the large peak deformation, a similar qualitative conclusion on the separation efficiency of the system of resin, solvent, and gradient would be drawn regardless of the salt step height used.

Simulation results would obviously be different for other protein systems, flow rates, and binding kinetics. However, the observation based on the presented simulated elution profiles match the experimental data using a monoclonal antibody reported by Welsh *et al.* [39]. As general conclusion one can state that the suspected 'noise' often observed in data from robotic experiments is mostly caused by the application of salt steps. These peak deformation must therefore not be considered as noise but in fact reflect the elution profile of a multiple step elution. The reason why these unsmooth elution profiles are not always detected in the obtained pseudo-chromatograms lies in the chosen fractionation volumes (as will be shown in the next subsection), but it depends also on the step length and hence on the gradient slope.

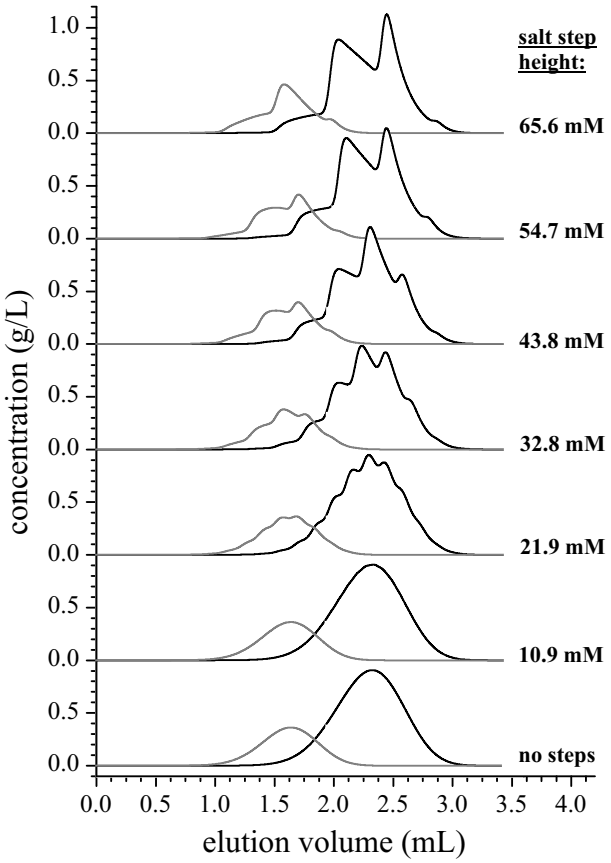

Figure 9: The separation of lysozyme (gray) from avidin (black) in robotic experiments simulated using different salt steps for elution. The flow rate was 54 cm/h (3μL/s).

3.5.2 Salt steps and flow interruptions

In a next step, simulations were performed taking the flow interruptions caused by each buffer uptake in gradient HTCC into account by setting the flow rate to zero for 30 s before every salt step. Furthermore, we expanded the simulations introducing in-silico fractionation with fractions of 75 μL, 15 μL, 225 μL, 300 μL, and 375 μL. In order to determine retention time, asymmetry, and peak width, for each of these fractions the theoretical average concentration was calculated and a peak fit was performed using these average values. It was furthermore assumed that additional pauses should not be introduced during a salt step, and thus, the number of breaks corresponded to the number of applied salt steps.

Effect of salt steps and flow interruptions - qualitative evaluation

Simulated elution profiles obtained for NaCl concentration step sizes of 10.9 mM and 43.8 mM, each with fractionation in 75 μL or 300 μL, are displayed for flow rates of 54 cm/h and 144 cm/h in Figure 10.

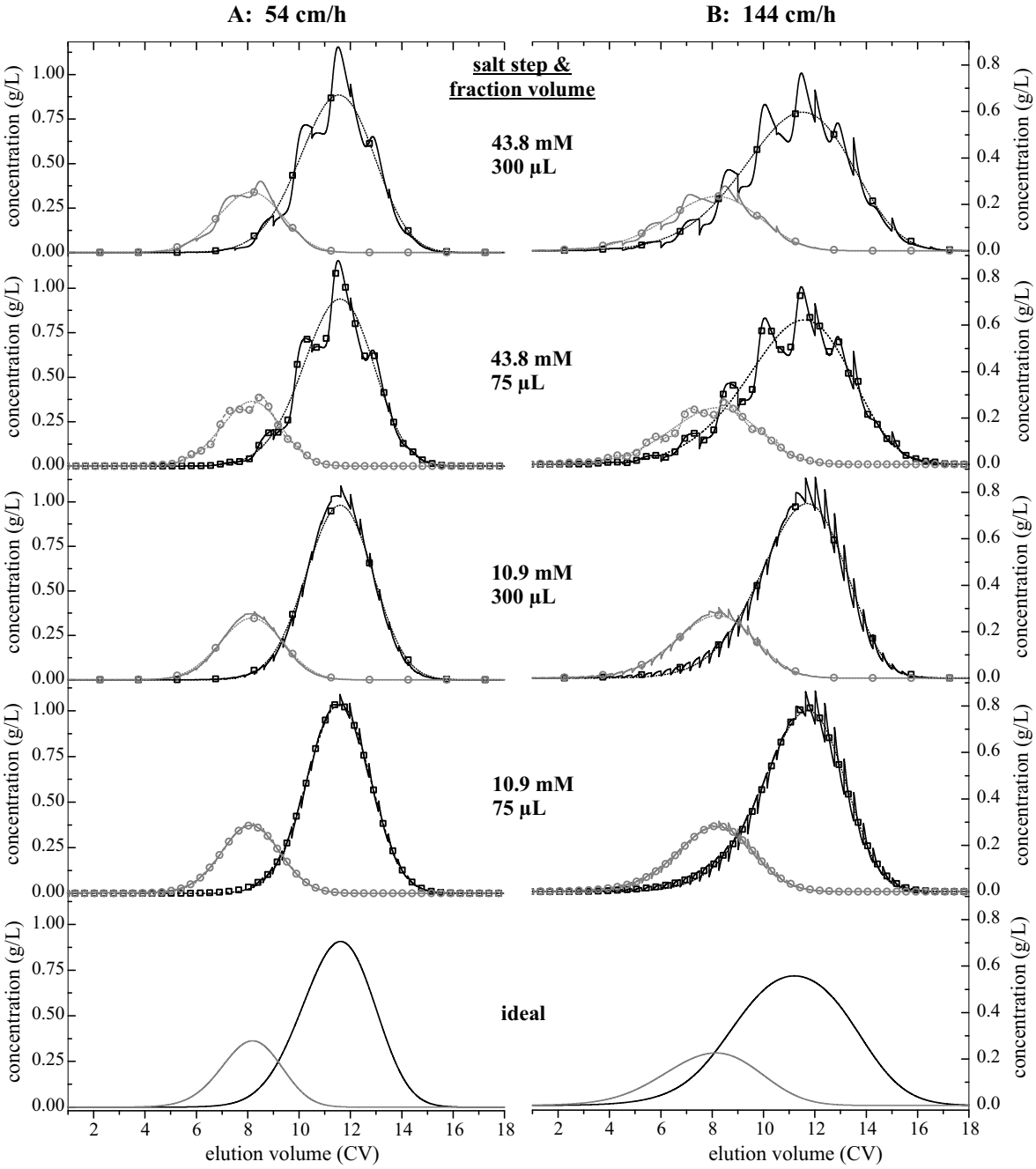

Figure 10: Effect of salt step height, fractionation, and flow rate on the elution profile for the separation of lysozyme (gray) from avidin (black). The gradient slope was +525 mM NaCl over 18 CV. Solid lines: simulated elution profiles for corresponding salt step height. Markers (circles: lysozyme, boxes: avidin): concentration values obtained in theoretical fractions of corresponding size (300 μL or 75 μL). Dashed lines: fit of the fraction concentrations using the EGH function. The step size of 11 mM and 43.8 mM corresponded to a step length of 75 μL and 300 μL, respectively. The 'ideal' chromatogram displays the simulation result assuming a linear gradient without flow interruptions.

Similar to the result shown previously (Figure 9), the influence of the salt step was large for the 43.8 mM salt step, whereas when using 11 mM steps the peak shapes were very similar to those of the reference ('ideal') gradient elution. The introduction of flow interruptions generated prompt changes in the concentration after each break.

The concentration in the bulk phase decreases at peak front after each stop in flow because protein pre-dominantly diffuses into the beads. At peak tail, the protein concentration in the bulk phase increases after each stop in flow due to protein diffusion out of the beads. The magnitude of this effect is more pronounced at high flow rates at which the ratio of convection to diffusion is larger. Compared to the effect of different salt steps, the effect of flow interruptions on the overall peak shape appeared smaller.

The influence of the fraction volume becomes very clear from the chromatograms shown in Figure 10. When using small fraction volumes the number of distinct data points is high enough to reflect the actual elution profile including the variations in concentration. That means, an 'unsmooth' elution peak will be observed potentially leading to a false conclusion of having large data noise levels. When using large fraction volumes the elution profile will be smoothed and the resulting common peak shape intuitively gives the impression of a more precise result. However, as demonstrated in section 3.3, a fraction number per peak lower than 8 – 10 may lead to a significant reduction in precision of the determined retention time and mass balance (area).

By performing a peak fit the resulting peaks were very similar regardless of the chosen fractionation or salt step height indicating the importance of fitting the peaks prior to evaluation.

Effect of salt steps and flow interruptions - quantitative evaluation

The quantitative deviation of the determined peak characteristics for different gradient patterns and fractionation schemes related to an simulated 'ideal' gradient experiment are presented in Figure 11 for lysozyme (Figure 11a,b) and avidin (Figure 11c,d,e). The effect on the determined peak resolution is shown in Figure 11f. The relative deviations were calculated for different fraction volumes and two flow rates, each grouped by a certain salt step height (and hence number of flow interruptions). In addition to simulations with flow interruptions, the theoretical scenario of performing salt steps without having flow interruptions are given for the smallest and largest fraction volume.

A clear difference between the output parameters retention times on one hand and peak asymmetry, peak width and peak resolution on the other hand can be seen in Figure 11 from the magnitude of deviations. The effect of different experimental combinations on the determined retention times was rather low (<4.5% for all cases), whereas the range of deviations in peak asymmetry and avidin peak widths was very large (-29% – +22%). The range of deviations in peak width for lysozyme was -21% – +11% (data not shown). In accordance with the large deviations in peak width, the determined resolution values increased significantly (+48%) compared to the 'ideally' performed gradient when using the highest number of flow interruptions, i.e. smallest salt step heights and smallest fractions.

The effect of only introducing salt steps can be seen from the simulations without flow interruptions. The effect on the quantitative result of the chromatographic output parameters was small (<8%) when applying the two small salt step heights. For a step height of 65.6 mM and a step height of 43.8 combined with the largest fraction volume, deviations of up to 20% were obtained for peak asymmetries (Figure 11b, d). For these larger salt step heights, the output parameters were more affected by the choice of the fraction volume compared to the salt step height.

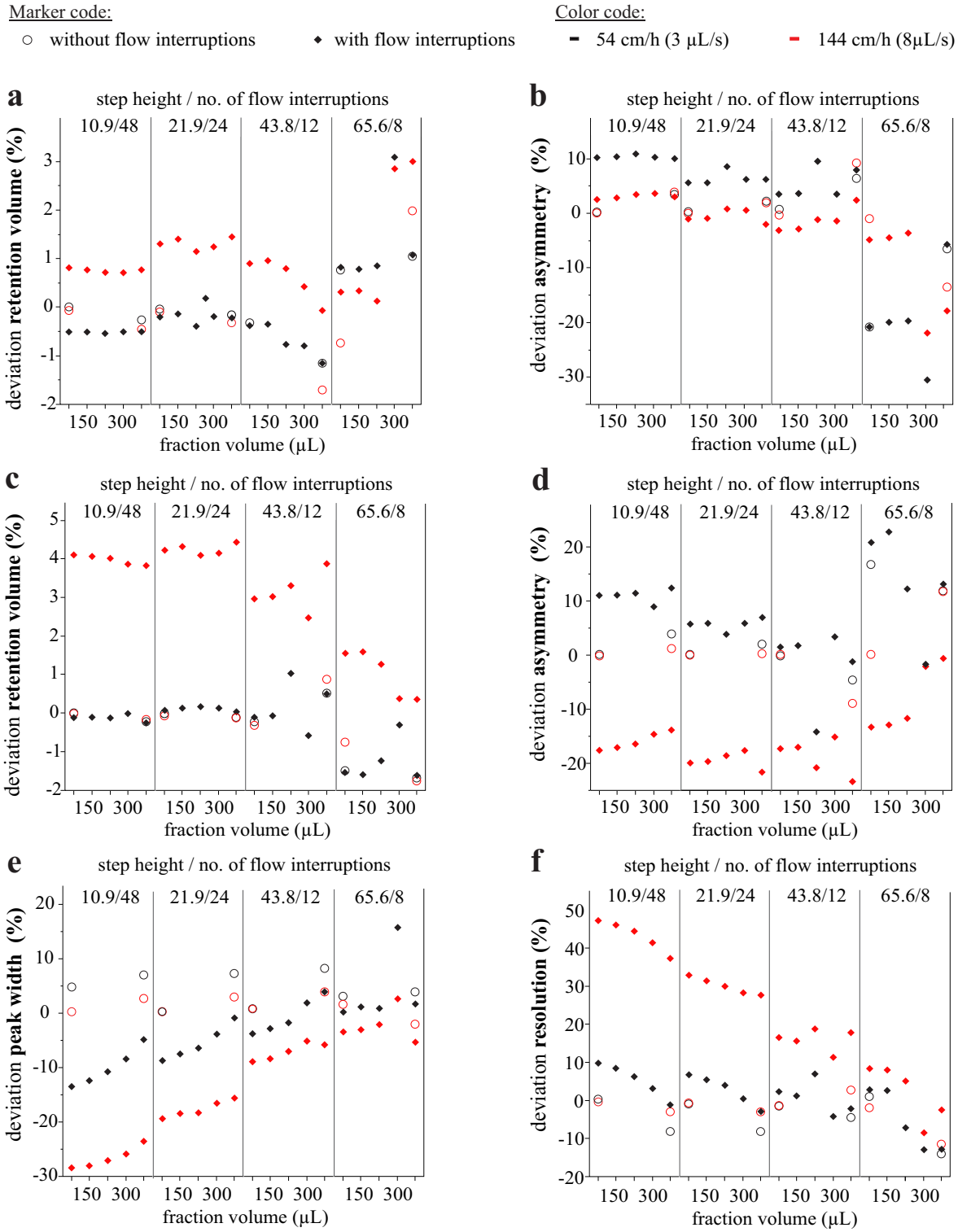

Figure 11: Relative deviations in the determined output parameters from the simulated 'ideal' gradient experiment for varying fraction volumes and flow rates, each grouped by a certain salt step height. The effect on the determined retention times and peak asymmetries are given for lysozyme (a+b) and avidin (c+d). The effect on the peak width is only displayed for avidin (e). The effect on the resolution is displayed in subplot f.

The effect of the flow interruptions can be seen when comparing the simulation results of identical salt step heights with and without flow interruptions as displayed for each fraction volume of 75 μL and 375 μL. It becomes clear that the effect of flow interruptions on the quantitative results is significantly larger than the variation of the salt step height, although qualitatively the chromatograms indicated the opposite (Figure 10).

For a combination of salt steps and flow interruptions, the overall results were closer to the result from an 'ideal' gradient experiment when using larger salt step heights compared to results using small salt steps. This is due to the fact that with increasing salt step heights a lower number of buffers and hence a lower number of flow interruptions is required which outweighs the peak deformation resulting from larger step heights.

For a variation of only the fraction volume at constant salt step height, peak widths and asymmetries for avidin increased with the chosen fraction volumes, and larger fraction volumes resulted in less deviation compared to the 'ideal' case. This has several reasons: First, a lower peak height of the determined peaks was obtained when fitting the peak with less fractions. As the peak widths were calculated at 0.1 peak height, the peak widths increased with increasing fraction volume, and thus, coming closer to the ideal case, in which peaks were broad. Second, the concentrations after a flow interruption at peak front were reduced and the concentrations after a flow interruption at peak tail were increased (see Figure 10) leading to a further increase of the peak height, and thus, the determined peak width. Lastly, with an increasing fraction volume and hence using less fractions for fitting the peak, the influence of flow interruptions decreased due to the generated smoothing effect.

The trends described above were very similar for both flow rates. However, for a flow rate of 144 cm/h the deviations in peak width and resolution were more than doubled compared to a flow rate of 54 cm/h. This is due to the larger relative changes in residence times (see Table IV) and hence larger differences in mass transfer rates as discussed in section 3.3. The abrupt changes in concentrations in the bulk phase after a flow interruption were larger at 144 cm/h compared to 54 cm/h, as can be seen in Figure 10. Thus, the peak shifted slightly to longer retention times, and at the same time the peak width decreased. From this result, higher comparability of experiments in HTCC mode to traditional laboratory LC experiments would be expected when using lower flow rates, because differences in the residence times become smaller. However, with further decreasing flow rates, the rate limiting step might change from the intra-particle diffusion to the external mass transfer as discussed by Łacki [38].

3.5.3 Comparability of simulation, HTCC, and laboratory LC results

In the previous sections it was shown in simulations that HTCC process parameters have both a quantitative and qualitative effect on the outcome of chromatography experiments. However, to investigate whether the determined effects and deviations from the 'ideal' case can also be detected in experimental data, results from HTCC experiments were compared for two flow rates with data generated in traditional laboratory LC experiments and simulation data (Figure 12 and Table V). For an evaluation of the variance between parallel HTCC runs, four elution profiles from HTCC experiments were compared. Furthermore, the comparison included experiments at flow rates of 54 cm/h and 144 cm/h. An artificial fractionation in fractions of 75 μL was applied for all simulated chromatograms matching the experimental procedure.

From Figure 12 it can be seen that the experimental chromatographic profiles largely match the simulated elution profiles, hence confirming the use of simulations as a suitable tool for the presented investigation of HTCC process parameters. The concentration profiles of the elution peaks show a

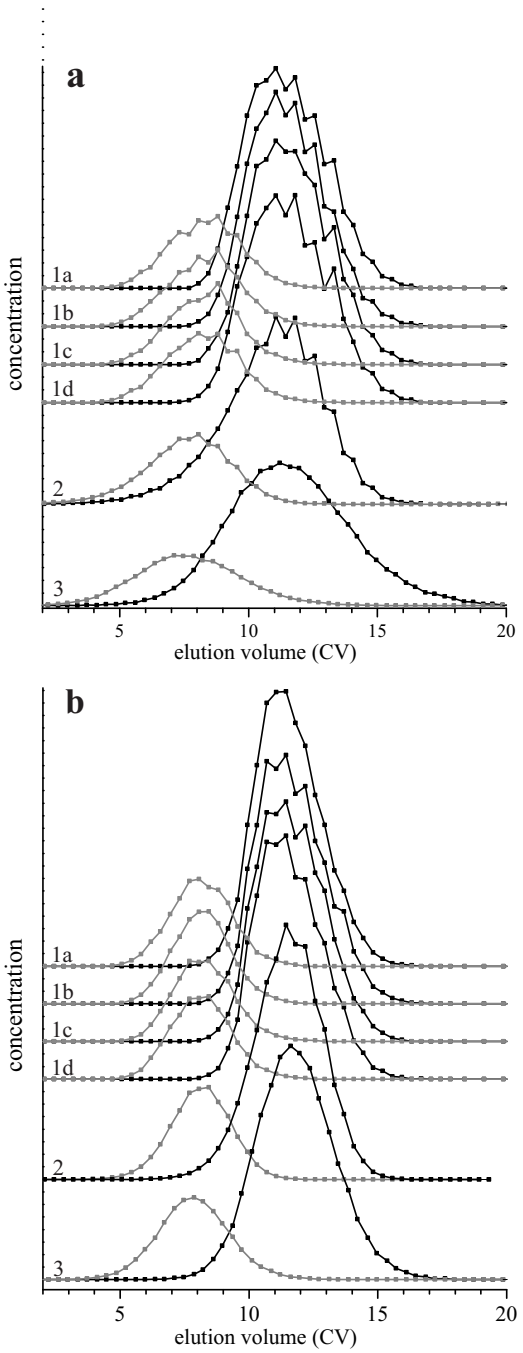

Figure 12: Chromatograms of the separation of avidin (black) from lysozyme (gray) at flow rates 144 cm/h (a) and 54 cm/h (b) on SP Sepharose FF (CV=200 μL). Separations were performed on a LHS in four parallel experiments (1a) – 1d)), salt step height: 21.9 mM, fraction volume: 75 μL, simulated including flow interruptions and salt steps (2), and performed on a laboratory LC system (ideal gradient and without flow interruptions, 3). All experiments were performed using columns with CV=200 μL.

pattern of changing concentrations every second value. This pattern was seen for all four HTCC experiments indicating these variations to be reproducible. The concentration pattern is furthermore in agreement with the predictions from the simulated elution profiles using the same fractionation scheme consisting of step heights of 21.9 mM (step lengths = 150 μL) and a fraction size of 75 μL.

Table V: Deviations of retention time and peak widths to the reference of a simulated 'ideal' robotic experiment without salt steps and flow interruptions ('pauses') for different experimental set-ups. The relative standard deviations (RSTD) were calculated from four parallel experiments on a LHS. The resolution values are given as absolute values. All separations were performed using the same gradient slope of 525 mM over 18 CV.

experimental mode	**protein**[1]	**retention volume**		**peak width**		**resolution**
		deviation (%)	RSTD (%)	deviation (%)	RSTD (%)	mean (-)
54 cm/h (3 μL/s)						
HTCC simulated 'ideal'	lys/avi	–	–	–	–	1.32
HTCC experimental 'steps+interruptions'	lys	-0.9	0.3	-1.5	2.7	–
HTCC simulated 'steps+interruptions'	lys	-0.2	–	-2.0	–	–
Laboratory LC experimental 'ideal'	lys	-4.3	–	14.4	–	–
HTCC experimental 'steps+interruptions'	avi	-3.1	0.5	-7.7	2.9	1.24
HTCC simulated 'steps+interruptions'	avi	0.1	–	-8.4	–	1.40
Laboratory LC experimental 'ideal'	avi	0.7	–	4.4	–	1.36
144 cm/h (8 μL/s)						
HTCC simulated 'ideal'	lys/avi	–	–	–	–	0.76
HTCC experimental 'steps+interruptions'	lys	2.5	0.8	-23.4	2.7	–
HTCC simulated 'steps+interruptions'	lys	1.5	–	-14.7	–	–
Laboratory LC experimental 'ideal'	lys	-7.4	–	6.0	–	–
HTCC experimental 'steps+interruptions'	avi	0.3	1.0	-31.3	7.5	0.98
HTCC simulated 'steps+interruptions'	avi	4.3	–	-19.3	–	1.00
Laboratory LC experimental 'ideal'	avi	0.3	–	2.3	–	0.86

[1] lys: lysozyme, avi: avidin

The elution profile from laboratory LC experiments (continuous flow and ideal gradient) shows the expected Gaussian peak shape. Small variations to the ideal peak profile reflect the analytical error for the determination of the concentration. The experimental data confirmed the earlier findings that variations in concentration are more pronounced at higher flow rates (Figure 10).

The differences of retention times and peak widths obtained in experimental and simulated HTCC and laboratory LC experiments related to the reference of an simulated 'ideal' robotic experiment (without salt steps and flow interruptions) are summarized in Table V. In average, the deviations for the different experimental modes regarding the determined retention times were significantly lower ($\sim$2.1%) than the deviations determined for the peak widths ($\sim$11.3%). The RSTDs determined from four HTCC experiments for retention times and peak widths were $\leq$1.0% and $\leq$2.9%, respectively, with the exception of 7.5% RSTD for the avidin peak width. With respect to this variation it can be concluded that there is no statistical effect of the experimental mode on the retention times. This conclusion is affirmed when taking potential deviations in the determined retention time due to different fractionation schemes into account as presented in Figure 6 and Figure 7. Furthermore, the differences between results from HTCC and results from laboratory LC experiments were larger than the differences between the two simulated HTCC modes ('ideal' and 'steps+interruptions').

A significant difference of HTCC to ideal gradient and laboratory LC experiments was however obtained for the peak widths. The peaks obtained from the laboratory LC experiments were broader compared to the peaks from HTCC experiments. This result is in agreement with the findings discussed in section 3.5.2, Figure 11. Furthermore, for a flow rate of 144 cm/h the peak widths determined from experimental HTCC were significantly smaller than the simulated values (Table V). This indicates that the peak width and hence peak resolution might in fact be even more affected by the HTCC experimental set-up than concluded from preceding simulation results.

4 Summary and Conclusion

In this study, we systematically investigated in which way robot-specific experimental parameters contribute to the experimental error of HTCC. Furthermore, differences between elution profiles obtained in HTCC and profiles obtained with laboratory LC systems were explained.

It was shown that the precision of pipetting (e.g. for dilution prior to analysis) increases exponentially with increasing pipetting volumes. Furthermore, the use of the MTPs was demonstrated to cause specific challenges. First, meniscus formation being different for solvent and varying protein concentrations affects the path length in absorption measurements, and thus, causes inaccuracies in the determination of concentration and volume. Calibration ranges need therefore to be chosen adequately and without forcing the calibration curve through zero (buffer blank). Secondly, when collecting eluate in MTPs the elution signal comprises a series of distinct absorption values rather than a continuous absorption signal and a chromatogram needs to be rebuild by fitting discrete absorption/concentration data points. The importance of using peak fitting functions to obtain suitable information on retention times and peak resolution as well as to increase precision was highlighted.

It was shown that a low number of collected elution fractions significantly contribute to imprecision of the retention time and peak area (mass balance) obtained from HTCC data, particularly when elution peaks are asymmetric. Remarkably, using very precise analytical techniques can not outweigh the increase of variance caused by of a low number of fractions. If the analytical precision is known to be imprecise, however, a sufficiently precise determination of mass balance and retention time can still be reached by increasing the number of fractions. About 10 fractions collected per peak width appear sufficient in order to significantly reduce the variance and the maximal possible precision is approached. A possible further gain of precision using more fractions might eventually not justify the increasing effort required for the analysis of more fractions. Furthermore, for low fraction volumes, the absolute error of the volume determination via absorption measurements increases. In this case, it becomes favorable to instead use the nominal fraction volumes.

It was explained that the 'smoothness' of an elution peak in HTCC gradient experiments is mainly depending on the choice of the salt step height. By choosing small fractions variations in the concentration profile will be reflected giving the impression of apparent noise in the data. In contrast, when collecting less fractions the common peak shape can be obtained. This is a fact which might be counter-intuitive, as one would expect better representation of the common peak shape when collecting more fractions rather than less. It was quantitatively shown that when using peak fitting functions the choice of the salt step height is not as important as the resulting number of flow interruptions. The effect of flow interruptions was furthermore demonstrated for isocratic elutions showing an effect only for the larger protein investigated. The peak widths were strongly influenced by the number of flow interruptions, whereas the determination of the retention times was shown to be more affected by the intrinsic variation caused by the fractionation scheme and analytical errors. Using larger salt steps (i.e. less flow interruptions) increased comparability to common laboratory LC experiment. However, larger deviations in retention times and asymmetry were seen in this case study for salt step heights >43.8 mM. Therefore, a suitable compromise for the salt step heights would be in a range of 22 mM - 44 mM. Future work could expand this study by taking other mass transfer scenarios into account, e.g. by varying the diffusion rates (different proteins and resins), flow rate and column dimensions.

There is no doubt that for screening purposes the use of HTCC fulfills the requirements on accuracy and precision. However, the results of this study also demonstrated that robotic process parameters should be chosen thoughtful on a case by case basis. Variabilities in the determined retention times

and peak widths need to be taken into account, if data are used to calibrate mechanistic models or if the technology is considered for detailed process development and robustness studies. The changes in residence times depending on the flow interruptions need to be considered in order to evaluate comparability to larger scales. The presented investigation of HTCC-specific parameters provides not only comprehensive knowledge on potential error sources and experimental variances, but it can also be considered as guidance for planning HTCC experiments according to the desired quality.

Acknowledgement

We thank Anna Osberghaus for the support regarding chromatography simulations as well as Sven Amrhein for his contributions to the Monte-Carlo simulations. Further, we thank Katharina Drechsel for carrying out parts of the isocratic elution experiments.

Appendix

A.1 Peak Fitting

The exponential Gaussian hybrid function (EGH) was used for fitting HPLC concentration data in order to determine peak width, asymmetry, and retention volume. The EGH represents a combination of the Gaussian function and an truncated exponential function. Details on this function can be found in [21, 24]. The fit function has four fit parameters: retention time t_R, the peak height H, the STD of the corresponding Gaussian function $\tilde{\sigma}$, and the time constant of the corresponding exponential function τ. It is piecewise defined as

$$f_{EGH}(t) = \begin{cases} H \exp\left(\frac{-(t-t_R)^2}{2\tilde{\sigma}^2+\tau(t-t_R)}\right), & 2\tilde{\sigma}^2+\tau(t-t_R) > 0 \\ 0, & 2\tilde{\sigma}^2+\tau(t-t_R) \leq 0 \end{cases} \tag{6}$$

The fit parameters were determined using the matlab *fit* function based on non-linear least squares regression. A maximum of 10^{12} iterations were performed with a maximum tolerance of 10^{-15}. The peak width was determined at 10% of the peak height. The peak asymmetry was determined at 50% of the peak height. Peak resolution was defined by

$$R = \frac{t_2 - t_1}{2(\sigma_2 + \sigma_1)} \tag{7}$$

with $t_{1,2}$ being the retention times of peak 1 and peak 2 and σ_1 and σ_1 being the standard deviations of peak 1 and peak 2 estimated via Moment analysis.

A.2 Chromatography model

The used model was based on following assumptions:

- ideally-packed bed
- spherically symmetrical and equally-sized beads

- incompressible fluid
- plug-flow regime and homogeneous radial distribution of components in the column
- no convective transport within the bead pores
- instantaneous local equilibrium between the bead macro pore surface and fluid inside the macro pores
- isothermal process
- diffusional mass transfer being independent of the concentration of other components
- dead volumes at column inlet and outlet are negligible.

The concentrations for component i in the interstitial phase $c_i(t, z)$ are a function of time t and axial position $z \in [0, L]$ with L being the column length. The concentration in the intra-particle fluid phase $c_{p,i}(t, z, r)$ and stationary phase $q_{p,i}(t, z, r)$ are functions of time and both axial and radial position with $r \in [0, r_p]$.

Mass transfer

Equation 8 describes the concentration change, $\partial c_i/\partial t$, depending on time t and location z in the interstitial liquid phase for components $1, 2, ...i$. The right-hand side terms describe the convective transport in the bulk phase, the dispersive transport and the diffusion through the stagnant film to the particle surface:

$$\frac{\partial c_i}{\partial t} = -u_{int}\frac{\partial c_i}{\partial z} + D_{ax}\frac{\partial^2 c_i}{\partial z^2} - \frac{1-\epsilon_{col}}{\epsilon_{col}}\frac{3}{r_p}k_{f,i}(c_i - c_{p,i}). \tag{8}$$

The interstitial velocity, u_{int}, is the quotient of the linear velocity, u, divided by the column porosity, ϵ_{col}. D_{ax} is the axial dispersion coefficient, r_p is the particle radius, $k_{f,i}$ are referred to as the film mass transfer coefficients of component i, and $c_{p,i}$ are the component's concentrations at r=r_p.

The mass balance for the bead is given by the diffusion terms through the film and within the particle:

$$\frac{\partial c_{p,i}}{\partial t} = D_{p,i}\left(\frac{\partial^2 c_{p,i}}{\partial r^2} + \frac{2}{r}\frac{\partial c_{p,i}}{\partial r}\right) - \frac{1-\epsilon_p}{\epsilon_p}\frac{\partial q_i}{\partial t}, \tag{9}$$

in which $D_{p,i}$ denotes the effective intra-particle diffusion coefficients, ϵ_p refers to as the particle porosity, and q_i is the stationary phase concentration.

The system of partial differential-algebraic equations were discretized in the spatial domain with 200 axial nodes and 25 radial nodes employing a finite volume method (see [20] for detailed information).

Adsorption and desorption

The used Steric Mass-Action (SMA)-model for ion-chromatography [26] describes the ad- and desorption kinetics by the rate coefficients k_{ad} and k_{des} for adsorption and desorption, respectively, the characteristic charge, ν, a shielding factor σ_{SMA}, the counter ions available for binding, $\bar{q}_0$, and the ionic capacity, Λ. For each component $i \geq 1$, the concentration on the surface, q_i, changes over time as expressed by:

$$\frac{\partial q_i}{\partial t} = k_{ad,i} c_{p,i} \bar{q}_0^{\nu_i} - k_{des,i} q_i c_{p,0}^{\nu_i} \tag{10}$$

From the condition of electro-neutrality the ionic capacity of the adsorbent must equal the sum of the available counter ions for binding and the binding sites occupied by the bound species:

$$\Lambda = \bar{q}_0 + \sum_{i=1}^{n} q_i(\nu_i + \sigma_{SMA,i}) \tag{11}$$

where the total concentration of salt ions, q_0, is related to the available counter-ions by

$$\bar{q}_0 = q_0 - \sum_{i=1}^{n} q_i \sigma_{SMA,i}. \tag{12}$$

Boundary conditions

The boundary conditions were chosen according to [20]. The concentrations at the column inlet ($z = 0$) and outlet ($z = L$) were:

$$c_i(t,0)u = c_{in,i}(t)u + D_{ax}\frac{\partial c_i}{\partial z}(t,0) \tag{13}$$

$$\frac{\partial c_i}{\partial z}(t,L) = 0 \tag{14}$$

The diffusive transport into the beads was assumed to equal the flux through the stagnant film at the particle surface ($r = r_p$) and hence the concentration change can be written as

$$\frac{\partial c_{p,i}}{\partial r^2}(.,.,rp) = \frac{k_{f,i}(c_i - c_{p,i})}{\epsilon_p D_{p,i}}. \tag{15}$$

The concentration in the bead center ($r = 0$) is

$$\frac{\partial c_{p,i}}{\partial r^2}(.,.,0) = 0. \tag{16}$$

Model calibration

The model was calibrated using the inverse method as described in detail in [27]. Five chromatographic separations of lysozyme from avidin in 200 μL columns filled with SP Sepharose FF were performed on the LHS. The salt gradients slopes of 525 mM NaCl over 8 CV, 13 CV, 18 CV, 23 CV and 28 CV were performed at pH 4.5, 22°C and with a flow rate of 3 μL/s. Fractions were collected in 75 μL volumes at a salt step width of 150 μL for each gradient. The salt step heights were calculated using equation 5. The protein concentrations were determined by the RP-HPLC method described in section 2.9.

The resulting experimental chromatograms and the corresponding simulated chromatograms obtained when using the estimated SMA-parameters (Table III) are displayed in Figure 13. Taking all five chromatograms into account the simulated results were in good agreement with the experimental results and therefore, the quality of the calibration was considered to be sufficient for the purpose of this study.

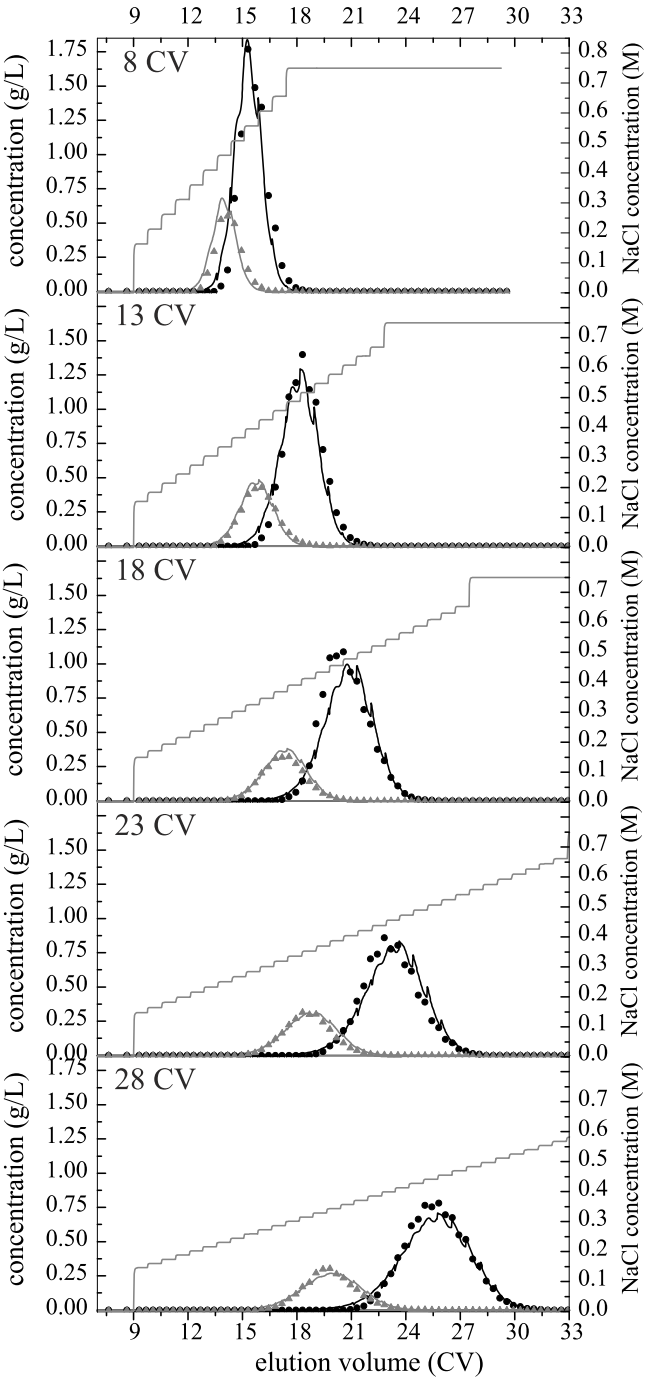

Figure 13: Separation of lysozyme (gray) and avidin (black) by applying different salt gradients over 525 mM in 200 μL RoboColumns operated on a LHS. Displayed are the experimental values (dots) and the simulated concentration profiles (solid lines).

References

[1] J. L. Coffman, J. F. Kramarczyk, B. D. Kelley, High-throughput screening of chromatographic separations: I. Method development and column modeling., *Biotechnol. Bioeng.* 100 (4) (2008) 605–18.

[2] 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 (5) (2008) 950–63.

[3] J. F. Kramarczyk, B. D. Kelley, J. L. Coffman, High-throughput screening of chromatographic separations: II. Hydrophobic interaction., *Biotechnol. Bioeng.* 100 (4) (2008) 707–20.

[4] T. Bergander, K. Nilsson-Välimaa, K. Oberg, K. M. Lacki, High-throughput process development: determination of dynamic binding capacity using microtiter filter plates filled with chromatography resin., *Biotechnol. Progr.* 24 (3) (2008) 632–9.

[5] L. Connell-Crowley, E. a. Larimore, R. Gillespie, Using high throughput screening to define virus clearance by chromatography resins., *Biotechnol. Bioeng.* 110 (7) (2013) 1984–94.

[6] J. Kittelmann, J. Hubbuch, Robust high-throughput batch screening method in 384-well format with optical in-line resin quantification, *J. Chrom. B* 988 (2015) 98–105.

[7] P. Schulze Wierling, R. Bogumil, E. H. J. Knieps-Grünhagen, High-Throughput Screening of Packed-Bed Chromatography Coupled With SELDI-TOF MS Analysis: Monoclonal Antibodies Versus Host Cell Protein., *Biotechnol. Bioeng.* 98 (2) (2007) 440–450.

[8] M. D. Wenger, P. DePhillips, C. E. Price, D. G. Bracewell, An automated microscale chromatographic purification of virus-like particles as a strategy for process development., *Biotechnol. Appl. Biochem.* 47 (2) (2007) 131–139.

[9] 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 - Miniaturization, Automation and Parallelization of Breakthrough and Elution Studies., *Chem. Eng. Technol.* 31 (6) (2008) 893–903.

[10] S. Chhatre, D. G. Bracewell, N. J. Titchener-Hooker, A microscale approach for predicting the performance of chromatography columns used to recover therapeutic polyclonal antibodies., *J. Chrom. A* 1216 (45) (2009) 7806–15.

[11] A. Susanto, K. Treier, E. Knieps-Grünhagen, E. von Lieres, J. Hubbuch, High Throughput Screening for the Design and Optimization of Chromatographic Processes: Automated Optimization of Chromatographic Phase Systems, *Chem. Eng. Technol.* 32 (1) (2009) 140–154.

[12] K. Treier, S. Hansen, C. Richter, P. Diederich, J. Hubbuch, P. Lester, High-throughput methods for miniaturization and automation of monoclonal antibody purification processes., *Biotechnol. Progr.* 28 (3) (2012) 723–32.

[13] K. Rege, N. Tugcu, S. M. Cramer, Predicting Column Performance in Displacement Chromatography from High Throughput Screening Batch Experiments., *Sep. Sci. Technol.* 38 (7) (2003) 1499–1517.

[14] A. Susanto, E. Knieps-Grünhagen, E. von Lieres, J. Hubbuch, High Throughput Screening for the Design and Optimization of Chromatographic Processes: Assessment of Model Parameter Determination from High Throughput Compatible Data., *Chem. Eng. Technol.* 31 (12) (2008) 1846–1855.

[15] A. Osberghaus, K. Drechsel, S. Hansen, S. 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.

[16] W. R. Keller, S. T. Evans, G. Ferreira, D. Robbins, S. M. Cramer, Use of minicolumns for linear isotherm parameter estimation and prediction of benchtop column performance, *J. Chrom. A* 1418 (2015) 94–102.

[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 (8) (2012) 1456–1464.

[18] L. Britsch, T. Schroeder, J. Friedle, Automated, High-Throughput Chromatographic Separation of Biological Compounds., *American Biotechnology Laboratory* 26 (6) (2008) 20–23.

[19] 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 (11) (2011) 2661–9.

[20] E. von Lieres, J. Andersson, A fast and accurate solver for the general rate model of column liquid chromatography., *Comp. Chem. Eng.* 34 (8) (2010) 1180–1191.

[21] K. Lan, J. W. Jorgenson, A hybrid of exponential and gaussian functions as a simple model of asymmetric chromatographic peaks., *J. Chrom. A* 915 (1-2) (2001) 1–13.

[22] M. S. Jeansonne, J. P. Foley, Review of the Exponentially Modified Gaussian (EMG) Function Since 1983., *J. Chrom. Sci.* 29 (6).

[23] P. J. Naish, S. Hartwell, Exponentially modified gaussian function - a good model for chromatographic peaks in isocratic hplc?, *Chromatographia* 26 (1).

[24] J. Li, Comparison of the capability of peak functions in describing real chromatographic peaks., *J. Chrom. A* 952 (1-2) (2002) 63–70.

[25] S. Kreusch, S. Schwedler, B. Tautkus, G. A. Cumme, A. Horn, UV measurements in microplates suitable for high-throughput protein determination., *Anal. Biochem.* 313 (2) (2003) 208–15.

[26] C. A. Brooks, S. M. Cramer, Steric mass-action ion exchange: Displacement profiles and induced salt gradients., *AIChE Journal* 38 (12).

[27] 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.

[28] H. Schmidt-Traub (Ed.), *Preparative Chromatography.*, 1st Edition, Wiley-VCH, Weinheim, 2005.

[29] C. A. Orellana, C. Shene, J. A. Asenjo, Mathematical modeling of elution curves for a protein mixture in ion exchange chromatography applied to high protein concentration., *Biotechnol. Bioeng.* 104 (3) (2009) 572–81.

[30] T. Kataoka, H. Yoshida, K. Ueyama, Mass transfer in laminar region between liquid and packing material surface in the packed bed., *J. Chem. Eng. Japan* 5 (2).

[31] M. Schröder, E. von Lieres, J. Hubbuch, Direct quantification of intraparticle protein diffusion in chromatographic media., *J Phys. Chem. B* 110 (3) (2006) 1429–36.

[32] P. Diederich, S. Amrhein, F. Hämmerling, J. Hubbuch, Evaluation of PEG/phosphate aqueous two-phase systems for the purification of the chicken egg white protein avidin by using high-throughput techniques., *Chem. Eng. Sci.* 104 (0) (2013) 945 – 956.

[33] I. H. Xie, M. H. Wang, R. Carpenter, H. Y. Wu, Automated calibration of TECAN genesis liquid handling workstation utilizing an online balance and density meter. (2004).

[34] S. A. Oelmeier, F. Dismer, J. Hubbuch, Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation., *Biotechnol. Bioeng.* 108 (1) (2011) 69–81.

[35] E. L. McGown, D. G. Hafeman, Multichannel pipettor performance verified by measuring pathlength of reagent dispensed into a microplate., *Anal. Biochem.* 258 (1) (1998) 155–7.

[36] M. G. Cottingham, C. D. Bain, D. J. T. Vaux, Rapid method for measurement of surface tension in multiwell plates., *Laboratory investigation; a journal of technical methods and pathology* 84 (4) (2004) 523–9.

[37] Y. Tao, G. Carta, Rapid monoclonal antibody adsorption on dextran-grafted agarose media for ion-exchange chromatography., *J. Chrom. A* 1211 (1-2) (2008) 70–9.

[38] K. M. Lacki, High-throughput process development of chromatography steps: Advantages and limitations of different formats used., *Biotechnol. J.* 7 (10) (2012) 1192–1202.

[39] J. P. Welsh, M. G. Petroff, P. Rowicki, H. Bao, T. Linden, D. J. Roush, J. M. Pollard, A practical strategy for using miniature chromatography columns in a standardized high-throughput workflow for purification development of monoclonal antibodies, *Biotechnol. Progr.* 30 (3) (2014) 626–635.

4 Conclusion

This work dealed with three different aspects of High-Throughput Process Development. As essential element of HTPD, analytical methods matching the requirements of high-throughput analytics regarding speed, sample volume, and automation were developed in the first part of this work. With (Ultra)HPLC technology being the 'Gold-standard' for protein concentration and purity analysis, the presented studies have focussed on alternative utilization of HPLC equipment and control programs, in order to reduce the typically long analysis times. Parallel operation of two columns paired with interlaced injection was found an effective way to cut-down analysis times for applications using isocratic elution. For the quantification of antibody aggregate/monomer content, the analysis time was reduced to below two minutes, while accuracy and precision were identical to the traditional method using single injections on one column. For the avidin egg white protein system used in this work, the application of two columns in tandem mode was successfully implemented for the separation of AV, OV, OM, OT, and LYS. Besides providing short analysis times and high precision, robust automated peak integration was achieved with the highly developed HPLC control software. The latter is as an important factor for a high-thoughput method. For comparison, a detailed evaluation of a high-throughput analytical technique (LabChip®) was conducted revealing major disadvantages when it comes to method robustness, precision and automation of data evaluation. However, there is no doubt that with technological improvements of such new instruments, control software, and improved data quality, the partly existing analytical bottleneck in HTPD will be further opened. With the used methods, HPLC technology was demonstrated to be a qualified element in the HTA toolbox. Nevertheless, the time-consuming development for HPLC methods, which is required whenever the analytical task changes, needs to be considered when choosing the analytical technique.

In the second part of this work, high-throughput methods were used to develop an alternative process for the separation of AV from egg white. A sequence of process steps were investigated in high-throughput mode in order to identify a simple and scalable process for large-scale manufacturing of highly pure AV. Extraction using aqueous two-phase systems was found to be very effective regarding the separation of LYS from AV in comparison to traditionally used ion-exchange chromatography. With ATPS as core element of the suggested process, suitable process conditions for a selective precipitation of AV before, and a volume reduction and final purification of AV after the extraction step could be identified by using the HTPD-approach. While being out of scope for this work, further studies, however, need to evaluate existing methods for re-cycling of the used polyethylene glycol and salts, in order to further reduce costs and the environmental burden.

During development work and the application of both HTE and HTA, it became clear that automated data processing, evaluation and data storage are important elements which also need to be addressed when implementing HTPD. Only by adaquate methods for data evaluation (ideally applicable for various development tasks) included in the experimental workflow, the real benefits of HTPD can be exploited. In conclusion from the practical experiences gained during this work and

from discussions with industrial researchers, the author considers efficient HTPD as a network of several elements (Figure 3). While in biopharmaceutical sciences the focus is often on the experimental and analytical methods only, future work should also strive for more standardized solutions for data transfer and communication between control software of existing HTE/HTA equipment as well as automated sorting, pre-processing and final data evaluation.

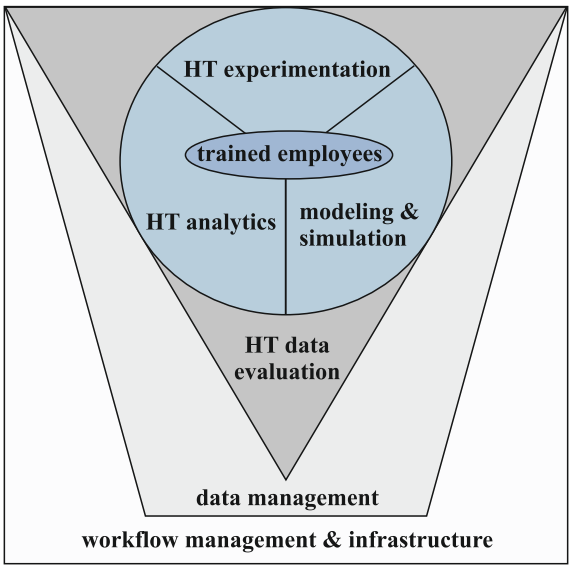

Figure 3: The core elements in current HTPD methodology. In many publications on HTPD applications, already the combination of High-throughput (HT) experimentation and HT analytics is termed HTPD (Łacki 2014) while in others modeling and simulation techniques (Hanke 2014) or advanced data analysis (Hubbuch 2012) is included. However, for efficient HTPD both automation and standardization of data evaluation and the linkage to easy-accessible data management systems are required. In particular, in larger organizations, in which experimentation and analytics are performed by different groups, coordination of the workflow and the obtained data becomes important.

The last study of this work contributed to a better understanding of the apparent noise and reduced precision obtained when using high-throughput column chromatography, HTCC. Correlations of decreasing precision were quantitatively determined for decreasing pipetting volumes, a lower number of fractions collected during elution and by taking the precision of methods for the determination of volume and concentration into account. For the first time, simulations of gradient elution HTCC were performed taking all method-specific parameters into account, in order to determine their influence on the experimental outcome. By this approach, it could be shown that the selected salt step height when mimicking a gradient qualitatively contributes strongly to the elution profile. However, the effect is rather small with regards to quantitative results such as the determined retention time or peak width. Furthermore, simulations allowed the investigation of the interplay of peak fractionation, salt step height, and flow interruptions, which are inherent for gradient elutions in HTCC. The conflict of increasing discrepancy between actual and nominal residence time when using higher flow rates and choosing smaller salt step heights (= many flow interruptions), in order to achieve a better approximation of a gradient, was elucidated. The introduction of flow interruptions influenced the retention time and peak width in the case of avidin for both isocratic and gradient elutions, whereas an effect when using lysozyme was not determined. Although it became clear that a general recommendation on the optimal HTCC method parameters can not be stated, the presented study provides valuable information which will serve other users in the selection of HTCC method parameters. Furthermore, researchers aiming to use HTCC data for calibrating chromatography models

are provided with data on experimental variances for various conditions which might help to improve model robustness validation.

With the knowledge of method-specific effects, future work must address remaining questions regarding the scale-differences to traditional lab-scale and production scale (e.g. wall effects, packing (in-)homogeneity, liquid flow distribution in mini columns). Most-importantly, the limitations of scaling-down via residence time or by linear flow rate require the identification of other scaling factors more suitable for predicting the scale-up from HTCC. With the advanced knowledge on method-specific effects presented in this work and scale-specific effects obtained in future studies, HTCC can eventually be used not only for screening purposes but also for detailed process optimization and characterization.

5 Glossary

Abbreviations

ATPS	Aqueous two-phase system
AV	Avidin
CV	Column volume
DAD	Diode array detector
DNA	Deoxyribonucleic acid
EGH	Exponential-Gaussian hybrid function
ELISA	Enzyme linked immunosorbent assay
FEP	Fluorinated ethylene propylene
FPLC	Fast preparative liquid chromatography
HABA	4'-hydroxy-azobenzene-2-carboxylic acid
HCP	Host cell protein
HIC	Hydrophobic interaction chromatography
HPLC	High performance liquid chromatography
HS	High-Sensitivity
HT	High-throughput
HTA	High-throughput analytics
HTCC	High-throughput column chromatography
HT CGE	High-throughput capillary gel electrophoresis
HTE	High-throughput experimentation
HTPD	High-throughput process development
HTS	High-throughput screening
ID	Inside diameter
LHS	Liquid handling station
LLOQ	Lower limit of quantification
LYS	Lysozyme
mAb	Monoclonal antibody
MC	Monte-Carlo
MM	Mixed mode
MTP	Microtiter plate
MW	Molecular weight
MWCO	Molecular weight cut-off
OM	Ovomucoid
OT	Ovotransferrin
OV	Ovalbumin
PEG	Polyethylene glycol
PI	Isoelectric point
Pip	Pipetting
PI-SEC	Parallel interlaced size exclusion chromatography
POI	Protein of interest
PPAS	Pre-purified avidin solution
RNA	Ribonucleic acid
RP	Reversed phase
RSD	Relative standard deviation
SMA	Steric mass-action
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
ST	Standard
UHPLC	Ultra high performance liquid chromatography
UV-VIS	Ultraviolet-visible radiation

Symbols

Symbol	Description	Unit
a	Peak area	mAU·s
A	Absorption	mAU
c	Concentration	g/L or M
c_p	Intra-particle concentration	M
d	Diameter	m
D	Distribution ratio	–
D_{ax}	Axial dispersion coefficient	m^2/s
D_p	Effective intra-particle diffusion coefficient	m^2/s
E_c	Deviation of concentration	μg/mL
E_V	Deviation of volume	μL
H	Peak heigth	mg/L
k	Sorption rate	–
k_f	Film mass transfer coefficient	m/s
L	Length	m
pr	Phase ratio	–
Pu	Purity	%
q	Adsorbed protein concentration	M
q_0	Total concentration of salt ion	M
$\bar{q}_0$	Concentration of available counter-ion	M
r	Radial position	m
R	Resolution	–
REC	Recovery	%
s	Relative standard deviation	–
t	Time	s
t_R	Retention time	s
u	Linear velocity	cm/h
u_{int}	Interstitial velocity	m/s
V	Volume	mL
x_{PEG}	Mass fraction PEG	kg/kg
x_{salt}	Mass fraction salt	kg/kg
y_V	Elution volume of fraction	μL
y_c	Concentration in elution fraction	μg/mL
$\tilde{y}_V$	y_V modified by error term	μL
$\tilde{y}_c$	y_c modified by error term	μg/mL
Y	Yield	%
z	Axial position	m

Greek symbols

Symbol	Description	Unit
ϵ	Porosity	–
ϵ_A	Molar extinction coefficient	L/mmol·cm
λ	Wavelength	nm
Λ	Ionic capacity	M
ν	Characteristic charge	–
σ	Standard deviation	s
$\tilde{\sigma}$	Standard deviation Gaussian function	μL^2
σ_{SMA}	Shielding factor	–
τ	Time constant of exponential decay	s

Indices

0	Relating to start condition
$+Biotin$	After addition of biotin
$+HABA$	After addition of HABA
$active$	Biologically active
ad	Adsorption
AV	Relating to avidin
bp	Bottom phase
c	Concentration
col	Column
des	Desorption
dil	Diluted
$dimer$	Antibody dimer
$frac$	Eluate fraction (in one well)
$HABA - AV$	Avidin complex with HABA
h	Liquid Level
$hold$	Holding
i	Component i
in	Inlet
inf	Information
lag	Lag
LYS	Relating to lysozyme
M	Mixing point
$mean$	Mean value
$meniscus$	Meniscus of liquid in a well
$monomer$	Antibody monomer
nom	Nominal
OM	Relating to ovomucoid
OT	Relating to ovotransferrin
OV	Relating to ovalbumin
p	Particle
PEG	Relating to polyethylene glycol solution
pip	Pipetting
$phase$	Phase of an ATPS
$random$	Random
ref	Reference value
$salt$	Relating to salt
$syst$	Systematic
$system$	Experimental system
tp	Top phase
V	Volume

6 Bibliography

Airenne, K. J., Marjoma, V. S., Kulomaa, M. S., 1999. Recombinant avidin and avidin - fusion proteins. *Biomol. Eng.* 16, 87–92.

Albertsson, P.-A., 1986. *Partition of Cell Particles and Macromolecules*, 3rd Edition. Wiley-Interscience, New York.

Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G., Kell, D. B., 2003. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* 21 (6), 692–6.

Altenhöner, U., Meurer, M., Strube, J., Schmidt-Traub, H., 1997. Parameter estimation for the simulation of liquid chromatography. *J. Chrom. A* 769 (1), 59–69.

Andrews, B. A., Nielsen, S., Asenjo, J. A., Jan. 1996. Partitioning and purification of monoclonal antibodies in aqueous two-phase systems. *Bioseparation* 6 (5), 303–313.

Andrews, B. A., Schmidt, A. S., Asenjo, J. A., 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: effect of surface hydrophobicity and charge. *Biotechnol. Bioeng.* 90, 380–390.

Asenjo, J. A., Andrews, B. A., 2012. Aqueous two-phase systems for protein separation: phase separation and applications. *J. Chrom. A* 1238, 1–10.

Asenjo, J. A., Schmidt, A. S., Hachem, F., Andrews, B. A., 1994. Model for predicting the partition behaviour of proteins in aqueous two phase systems. *J. Chrom. A* 668, 47–54.

Atha, D. H., Inghamg, K. C., 1981. Mechanism of Precipitation of Proteins by Polyethylene Glycols. *J. Biol. Chem.* 256 (23), 12108–12117.

Aubert, J. H., Tirrell, M., Aug. 1983. Flow Rate Dependence of Elution Volumes in Size Exclusion Chromatography: A Review. *J. Liquid Chrom. Related Technol.* 6 (9), 219–249.

Awade, A. C., 1996. On hen egg fractionation: applications of liquid chromatography to the isolation and the purification of hen egg white and egg yolk proteins. *Z. Lebensm. Untersuchung Forsch.* 202, 1–14.

Azevedo, A. M., Rosa, P. A. J., Ferreira, I. F., Aires-Barros, M. R., 2009. Chromatography-free recovery of biopharmaceuticals through aqueous two-phase processing. *Trends Biotechnol.* 27, 240–247.

Baskir, J. N., Hatton, T. A., Suter, U. W., 1989. Protein partitioning in two-phase aqueous polymer systems. 1989 *Biotechnol. Bioeng.* 34, 541–548 34, 541–548.

Bayer, E., Ehrlich-Rogozinski, S., Wilchek, M., 1996. Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. *Electrophoresis* 17, 1319–1324.

Benavides, J., Rito-Palomares, M., 2008. Practical experiences from the development of aqueous two-phase processes for the recovery of high value biological products. *J. Chem. Technol. Biot.* 142, 133–142.

Bensch, M., Schulze Wierling, P., von Lieres, E., Hubbuch, J., Nov. 2005. High Throughput Screening of Chromatographic Phases for Rapid Process Development. *Chem. Eng. Technol.* 28 (11), 1274–1284.

Bensch, M., Selbach, B., Hubbuch, J., 2007. High throughput screening techniques in downstream processing: Preparation, characterization and optimization of aqueous two-phase systems. *Chem. Eng. Sci.* 62 (7), 2011–2021.

Berek, D., 2010. Size exclusion chromatography–a blessing and a curse of science and technology of synthetic polymers. *J. Sep. Sci.* 33 (3), 315–35.

Berg, A., 2008. *Development of High-Throughput Screening Methods for the automated optimization of inclusion body protein refolding processes*. Ph.d., Karlsruhe Institute of Technology (KIT).

Berg, A., Schuetz, M., Dismer, F., Hubbuch, J., 2013. Automated measurement of apparent protein solubility to rapidly assess complex parameter interactions. *Food Bioprod. Proc.* 92 (2), 133–142.

Bergander, T., Nilsson-Välimaa, K., Öberg, K., Łacki, K. M., 2008. High-Throughput Process Development: Determination of Dynamic Binding Capacity Using Microtiter Filter Plates Filled with Chromatography Resin. *Biotechnol. Progr.* 24 (3), 632–639.

Bernardi, S., Gétaz, D., Forrer, N., Morbidelli, M., 2013. Modeling of mixed-mode chromatography of peptides. *J. Chrom. A* 1283, 46–52.

Bevan, P., Ryder, H., Shaw, I., 1995. Identifying small-molecule lead compounds: The screening approach to drug discovery. *Trends Biotechnol.* 13 (3), 115–121.

Bhambure, R., Kumar, K., Rathore, A. S., 2011. High-throughput process development for biopharmaceutical drug substances. *Trends Biotechnol.* 29 (3), 127–35.

Bielefeld-Sevigny, M., 2009. AlphaLISA immunoassay platform – The "nowash" high throughput alternative to ELISA. *Assay Drug Dev. Techn.* 7 (1), 90–92.

Bioscrience, T., 2010. Analysis of human immunoglobulins by size exclusion chromatography. Tosoh Bioscience LLC, *Application note A13L07A*.

Boado, R. J., Zhang, Y., Zhang, Y., Xia, C.-F., Wang, Y., Pardridge, W. M., 2008. Genetic engineering, expression, and activity of a chimeric monoclonal antibody-avidin fusion protein for receptor-mediated delivery of biotinylated drugs in humans. *Bioconjugate Chem.* 19, 731–739.

Bosma, J. C., Wesselingh, J. A., 2000. Partitioning and diffusion of large molecules in fibrous structures. *J. Chrom. B* 743 (1-2), 169–80.

Brenac Brochier, V., Schapman, A., Santambien, P., Britsch, L., Jan. 2008. Fast purification process optimization using mixed-mode chromatography sorbents in pre-packed mini-columns. *J. Chrom. A* 1177 (2), 226–233.

Britsch, L., Schroeder, T., Friedle, J., 2008. Automated, High-Throughput Chromatographic Separation of Biological Compounds. *American Biotechnology Laboratory* 26 (6), 20–23.

Broach, J., Thorner, J., 1996. High-throughput screening for drug discovery. *Nature* 384 (6604 Suppl), 14–16.

Brooks, C. A., Cramer, S. M., 1992. Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal* 38 (12).

Burton, S. C., Harding, D. R., Jul. 1998. Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aqueous buffers. *J. Chrom. A* 814 (1-2), 71–81.

Caliper LS, 2009. *HT Protein Express LabChip Kit, Version 2 LabChip GXII*. Caliper LS, user Guide Rev 4.

Campbell, L., Raikos, V., Euston, S. R., 2003. Review Modification of functional properties of egg-white proteins. *Food/Nahrung* 47 (6), 369–376.

Carta, G., Jungbauer, A. (Eds.), 2010. *Preparative Chromatography: Process Development and Scale-Up*, 1st Edition. WILEY-VCH Verlag, Weinheim.

Chandler, M., Zydney, A., 2004. High throughput screening for membrane process development. *J. Membr. Sci.* 237 (1–2), 181–188.

Chase, H. A., 1994. Purification of proteins by adsorption chromatography in expanded beds. *Trends Biotechnol.* 12, 296–303.

Chen, J.-P., 1992. Partitioning and Separation of a-Lactalbumin and fl-Lactoglobulin in PEG / Potassium Phosphate Aqueous Two-Phase Systems. *J. Ferment. Bioeng.* 73, 140–147.

Chen, X., Flynn, G. C., 2009. A high throughput dimer screening assay for monoclonal antibodies using chemical cross-linking and microchip electrophoresis. *J. Chrom. B* 877, 3012–3018.

Cheng, W., Hollis, D., 1987. Flow-rate effect on elution volume in size-exclusion chromatography. *J. Chrom. A* 408, 9–19.

Chhatre, S., Bracewell, D. G., Titchener-Hooker, N. J., 2009. A microscale approach for predicting the performance of chromatography columns used to recover therapeutic polyclonal antibodies. *J. Chrom. A* 1216 (45), 7806–15.

Chhatre, S., Francis, R., Bracewell, D. G., Titchener-Hooker, N. J., 2010. An automated packed protein G micro-pipette tip assay for rapid quantification of polyclonal antibodies in ovine serum. *J. Chrom. B.* 878 (30), 3067–3075.

Chhatre, S., Konstantinidis, S., Ji, Y., Edwards-Parton, S., Zhou, Y., Titchener-Hooker, N. J., 2011. The simplex algorithm for the rapid identification of operating conditions during early bioprocess development: case studies in FAb precipitation and multimodal chromatography. *Biotechnol. Bioeng.* 108 (9), 2162–70.

Chhatre, S., Titchener-Hooker, N. J., 2009. Microscale methods for high-throughput chromatography development in the pharmaceutical industry. *J Chem. Technol. Biot.* 84 (7), 927–940.

Chhatre, S., Titchener-hooker, N. J., 2011. Comprehensive Biotechnology, second ed. Edition. Vol. 1. Elsevier.

Chilamkurthi, S., Sevillano, D. M., Albers, L. H. G., Sahoo, M. R., Verheijen, P. J. T., van der Wielen, L. A. M., den Hollander, J. L., Ottens, M., 2014. Thermodynamic description of peptide adsorption on mixed-mode resins. *J. Chrom. A* 1341, 41–9.

Coffman, J. L., Kramarczyk, J. F., Kelley, B. D., 2008. High-throughput screening of chromatographic separations: I. Method development and column modeling. *Biotechnol. Bioeng.* 100 (4), 605–18.

Connell-Crowley, L., Larimore, E. a., Gillespie, R., 2013. Using high throughput screening to define virus clearance by chromatography resins. *Biotechnol. Bioeng.* 110 (7), 1984–94.

Cottingham, M. G., Bain, C. D., Vaux, D. J. T., 2004. Rapid method for measurement of surface tension in multiwell plates. *Laboratory investigation; a journal of technical methods and pathology* 84 (4), 523–9.

Cuatrecasas, P., Wilchek, M., 1968. Single-step purification of avidin from egg white by affinity chromatography on biocytin-sepharose columns. *Biochem. Biophys. Res. Commun.* 33 (2), 1967–1968.

Cuatrecasas, P., Wilchek, M., Anfinsen, C. B., 1968. Selective Enzyme Purification By Affinity Chromatography. *Roc. Natl. Acad. Sci. USA* 61 (2), 636–643.

da Silva, L. H. M., Coimbra, J. S. R., Jose, A., Meirelles, D. A., 1997. Equilibrium Phase Behavior of Poly (ethylene glycol) + Potassium Phosphate + Water Two-Phase Systems at Various pH and Temperatures. *J. Chem. Eng. Data* 9568, 398–401.

de Neuville, C. B., Tarafder, A., Morbidelli, M., 2013. Distributed pore model for bio-molecule chromatography. *J. Chrom. A* 1298, 26–34.

de Oliveira, F. C., dos Reis Coimbra, J. S., da Silva, L. H. M., Rojas, E. E. G., do Carmo Hespanhol da Silva, M., 2009. Ovomucoid partitioning in aqueous two-phase systems. *Biochem Eng. J.* 47, 55–60.

de Sousa, R. D. C. S., Coimbra, J. S. D. R., da Silva, L. H. M., da Silva, M. D. C. H., Rojas, E. E. G., Vicente, A. A. A., 2009. Thermodynamic studies of partitioning behavior of lysozyme and conalbumin in aqueous two-phase systems. *J. Chrom. B* 877, 2579–2584.

Degerman, M., Westerberg, K., Nilsson, B., 2009. A Model-Based Approach to Determine the Design Space of Preparative Chromatography. *Chem. Eng. Technol.* 32 (8), 1195–1202.

Delange, R. J., 1970. Egg White Avidin: I. Amino Acid Composition; Sequence of the Amino- and Carboxyl-Terminal Cyanogen Bromide Peptides. *J. Biol. Chem.* 245 (5), 907–916.

Desert, C., Guérin-Dubiard, C., Nau, F., Jan, G., Val, F., Mallard, J., 2001. Comparison of Different Electrophoretic Separations of Hen Egg White Proteins. *J. Agric. Food Chem.* 49 (10), 4553–4561.

Di Marco, V. B., Bombi, G. G., 2001. Mathematical functions for the representation of chromatographic peaks. *J. Chrom. A* 931 (1-2), 1–30.

Diamandis, E. P., Christopoulos, T. K., May 1991. The biotin-(strept)avidin system: principles and applications in biotechnology. *Clinical Chemistry* 37 (5), 625–636.

Diederich, P., Amrhein, S., Hämmerling, F., Hubbuch, J., 2013. Evaluation of PEG/phosphate aqueous two-phase systems for the purification of the chicken egg white protein avidin by using high-throughput techniques. *Chem. Eng. Sci.* 104, 945–956.

Dismer, F., 2009. *Characterization of Transport and Adsorption Mechanism in Chromatographic Media*. Ph.d., Karlsruhe Institute of Technology (KIT).

Donovan, J. W., Ross, K. D., 1973. The Stability of Avidin Produced by Binding of Biotin. A Differential Scanning Calorimetric Study of Denaturation by Heat. *Biochemistry* 12 (3), 512–517.

Drakeman, D. L., 2014. Benchmarking biotech and pharmaceutical product development. *Nat. Biotechnol.* 32 (7), 621–5.

Durance, D. T., 1994. Separation, purification and thermal stability of lysozyme and avidin from chicken egg white. In: Sim, J. S., Nakai, S. (Eds.), *Egg Uses and Processing Technologies. New Developments.* International CAB, Wallingford, UK, pp. 77–93.

Durance, T. D., Nakai, S., 1988a. Purification of Avidin by Cation Exchange, Gel Filtration, Metal Chelate Interaction and Hydrophobic Interaction Chromatography. *Can. Instit. Food Sci. Technol. J.* 21 (3), 279–286.

Durance, T. D., Nakai, S., 1988b. Simultaneous Isolation of Avidin and Lysozyme. *J. Food Sci.* 53, 1096–1102.

Eakin, R. E., McKinley, W. A., Williams, J., 1940. Egg-white injury in chicken and its relationship to a deficiency of vitamin H (biotin). *Science* 92, 224–225.

Eakin, R. E., Snell, E. E., Williams, J., 1941. The concentration and assay of avidin, the injury-producing protein in raw egg white. *J. Biol. Chem.* 140, 535–543.

Ed Bayer Group, . *Avidin-biotin publications*. Accessed November 2014.
URL `http://wws.weizmann.ac.il/Biological_Chemistry/scientist/Bayer/pubAvidin`

EFPIA, 2014. The Pharmaceutical Industry in Figures. Tech. rep., European Federation of Pharmaceutical Industries and Association.

Elvin, J. G., Couston, R. G., van der Walle, C. F., 2013. Therapeutic antibodies: market considerations, disease targets and bioprocessing. *Int. J. Pharm.* 440 (1), 83–98.

Fahrner, R. L., Knudsen, H. L., Basey, C. D., Galan, W., Feuerhelm, D., Vanderlaan, M., Blank, G. S., 2001. Industrial Purification of Pharmaceutical Antibodies: Development, Operation, and Validation of Chromatography Processes. *Biotechnol. Gen. Eng.* 18 (1), 301–327.

Farnan, D., Moreno, G., Stults, J., Becker, A., Tremintin, G., van Gils, M., 2009. Interlaced size exclusion liquid chromatography of monoclonal antibodies. *J. Chrom. A* 1216 (51), 8904–9.

Farnan, D., Rea, J. C., Lou, Y., Parikh, R., Moreno, G. T., Nov 1, 2010. High-Throughput Multi-Product Liquid Chromatography for Characterization of Monoclonal Antibodies. *Biopharm International*.

Ferreira, L. A., Teixeira, J. A., Mikheeva, L. M., Chait, A., Zaslavsky, B. Y., Aug. 2011. Effect of salt additives on partition of nonionic solutes in aqueous PEG-sodium sulfate two-phase system. *J. Chrom. A* 1218 (31), 5031–5039.

Food and Drug Administration, 2009. Guidance for Industry - Q8(R2) Pharmaceutical Development.
URL `http://www.fda.gov/downloads/Drugs/Guidances/ucm073507.pdf`

Franzreb, M., Siemann-Herzberg, M., Hobley, T. J., Thomas, O. R. T., 2006. Protein purification using magnetic adsorbent particles. *Appl. Microbiol. Biotechnol.* 70 (5), 505–16.

Fraser, S., Cameron, M., O'Connor, E., Schwickart, M., Tanen, M., Ware, M., 2014. Next generation ligand binding assays-review of emerging real-time measurement technologies. *AAPS journal* 16 (5), 914–24.

Gatti, G., Bolognesi, M., Coda, A., Chiolerio, F., Filippini, E., Malcovati, M., Sep. 1984. Crystallization of hen egg-white avidin in a tetragonal form. *J. Molecular Biol.* 178 (3), 787–789.

Geng, F., Huang, Q., Wu, X., Ren, G., Shan, Y., Jin, G., Ma, M., Aug. 2012. Co-purification of chicken egg white proteins using polyethylene glycol precipitation and anion-exchange chromatography. *Sep. Purif. Technol.* 96, 75–80.

Gerontas, S., Shapiro, M. S., Bracewell, D. G., 2013. Chromatography modelling to describe protein adsorption at bead level. *J. Chrom. A* 1284, 44–52.

Golshan-shirazi, S., Guiochon, G., 1994. Modeling of preparative liquid chromatography. *J. Chrom. A* 658, 149–171.

Gotmar, G., Fornstedt, T., Guiochon, G., 1999. Peak tailing and mass transfer kinetics in linear chromatography - Dependence on the column length and the linear velocity of the mobile phase. *J. Chrom. A* 831, 17–35.

Green, N. M., 1963a. Avidin 1. The use of ^{14}C Biotin for kinetic studies and for assay. *Biochm. J.* 89 (1959), 585–591.

Green, N. M., 1963b. Avidin 3. The nature of the biotin-binding site. *Biochm. J.* 89, 599–609.

Green, N. M., 1963c. Avidin 4. Stability at extremes of pH and dissociation into subunits by guanidin hydrochloride. *Biochm. J.* 89, 609–620.

Green, N. M., 1965. A spectrophotometric assay for avidin and biotin based on binding of dyes by avidin. *Biochemical J.* 94 (23-24).

Green, N. M., 1970. *Spectrophotometric determination of avidin and biotin. Methods in Enzymology*. Academic Press, New York, NY.

Green, N. M., 1975. Avidin. In: C.B. Anfinsen, J. T. E., Richards, F. M. (Eds.), *Advances in Protein Chemistry*. Vol. 29. Academic Press, pp. 85–133.

Green, N. M., Toms, E. J., 1970. Purification and Crystallization of Avidin. *Biochm. J.* 118, 67–70.

Gu, T., Tsai, G.-J., Tsao, G. T., 1993. Modeling of Nonlinear Multicomponent Chromatography. *Adv. Biochem. Eng. Biot.* 49, 46–69.

Guérin-Dubiard, C., Pasco, M., Hietanen, A., Quiros del Bosque, A., Nau, F., Croguennec, T., 2005. Hen egg white fractionation by ion-exchange chromatography. *J. Chrom. A* 1090, 58–67.

Guiochon, G., Beaver, L. A., 2011. Separation science is the key to successful biopharmaceuticals. *J. Chrom. A* 1218 (49), 8836–8858.

Haber, C., 2006. Microfluidics in commercial applications; an industry perspective. *LabChip* 6 (9), 1118–1121.

Hanke, A. T., Ottens, M., 2014. Purifying biopharmaceuticals: knowledge-based chromatographic process development. *Trends Biotechnol.* 32 (4), 210–20.

Hansen, K. S., 2013. *High Throughput Analytics for Protein Purification Process Development and Process Control*. Ph.d., Karlsruhe Institute of Technology (KIT).

Hansen, S. K., Skibsted, E., Staby, A., Hubbuch, J., 2011. A label-free methodology for selective protein quantification by means of absorption measurements. *Biotechnol. Bioeng.* 108 (11), 2661–2669.

Harenberg, J., 2009. Development of idraparinux and idrabiotaparinux for anticoagulant therapy. *Thrombosis and Haemostasis* 102 (5), 811–5.

Hatti-Kaul, R., 2000. *Aqueous Two-Phase Systems - Methods and Protocols*. Humana Press, Totowa, New Jersey.

Heney, G., Orr, G. A., 1981. The Purification of Avidin and Its Derivatives on 2-Iminobiotin-6-aminohexyl-Sepharose 4B. *Anal. Biochem.* 114, 92–96.

Herper, M., 2013. The Cost Of Creating A New Drug Now $5 Billion, Pushing Big Pharma To Change. forbes magazine. `\http://www.forbes.com/sites/matthewherper/2013/08/11/how-the-staggering-cost-of-inventing-new-drugs-is-shaping-the-future-of-medicine/`, accessed Nov 10, 2014.

Hiller, Y., Gershoni, J. M., Bayer, E. A., Wilchek, M., 1987. Biotin binding to avidin Oligosaccharide side chain not required for ligand association. *Biochm. J.* 248, 167–171.

Hönig, W., Kula, M. R., May 1976. Selectivity of protein precipitation with polyethylene glycol fractions of various molecular weights. *Anal. Biochem.* 72, 502–512.

Hood, E. E., Witcher, D. R., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R. J., Hernan, R., Kappel, W. K., Ritland, D., Li, C. P., Howard, J. A., 1997. Commercial production of avidin from transgenic maize : characterization of transformant, production, processing, extraction and purification. *Mol. Breeding* 3, 291–306.

Hook, D., 1996. Ultra high-throughput screening – a journey into Nanoland with Gulliver and Alice. *Drug Discovery Today* 1 (7), 267–268.

Huang, T. S., DeLange, R. J., 1971. Egg White Avidin II. Isolation, composition, and amino acid sequences of the tryptic peptides. *J. Biol. Chem.* 246 (3), 686–697.

Hubbuch, J., 2012. Editorial: *High-throughput process development. Biotechnol. J.* 7 (10), 1185–1185.

Hubbuch, J., Thömmes, J., Kula, M.-R., 2005. Biochemical Engineering Aspects of Expanded Bed Adsorption. *Adv. Biochem. Eng. Biot.* 92, 101–123.

Huddleston, J., Abelaira, J. C., Wang, R., Lyddiatt, A., 1996. Protein partition between the different phases comprising poly(ethylene glycol)-salt aqueous two-phase systems, hydrophobic interaction chromatography and precipitation: a generic description in terms of salting-out effects. *J. Chrom. B* 680, 31–41.

Huddleston, J., Veide, A., Köhler, K., Flanagan, J., Enfors, S. O., Lyddiatt, A., 1991a. The molecular basis of partitioning in aqueous two-phase systems. *Trends Biotechnol.* 9, 381–388.

Huddleston, J. G., Ottomar, K. W., Ngonyani, D. M., Lyddiatt, A., 1991b. Influence of system and molecular parameters upon fractionation of intracellular proteins from Saccharomyces by aqueous two-phase partition. *Enzyme Microb. Technol.* 13, 24–32.

Huddleston, J. G., Wang, R., Flanagan, J. A., Brien, S. O., Lyddiatt, A., 1994. Variation of protein partition coefficients with volume ratio in poly (ethylene glycol) -salt aqueous two-phase systems. *J. Chrom. A* 668, 3–11.

Huddleston, J. G., Willauer, H. D., Rogers, R. D., 2003. Phase Diagram Data for Several PEG + Salt Aqueous Biphasic Systems at 25 Â° C. *J. Chem. Eng. Data* 48, 1230–1236.

Hui, R., Edwards, A., 2003. High-throughput protein crystallization. *J. Struct. Biol.* 142 (1), 154–161.

Hytönen, V. P., Laitinen, O. H., Airenne, T. T., Kidron, H., Meltola, N. J., Porkka, E. J., Hörhä, J., Paldanius, T., Määttä, J. a. E., Nordlund, H. R., Johnson, M. S., Salminen, T. a., Airenne, K. J., Ylä-Herttuala, S., Kulomaa, M. S., 2004. Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli. *Biochemical J.* 384, 385–90.

IMS Institute for Healthcare Informatics, Nov 2013. The Global Use of Medicines: Outlook through 2017. URL `http://www.imshealth.com`

Jackson, N., Liddell, J., Lye, G., 2006. An automated microscale technique for the quantitative and parallel analysis of microfiltration operations. *J. Membr. Sci.* 276 (1-2), 31–41.

Jeansonne, M. S., Foley, J. P., 1989. Measurement of statistical moments of resolved and overlapping chromatographic peaks. *J. Chrom. A* 461 (088), 149–163.

Jeansonne, M. S., Foley, J. P., 1991. Review of the Exponentially Modified Gaussian (EMG) Function Since 1983. *J. Chrom. Sci.* 29 (6).

Kaliszan, R., Baczek, T., 2010. QSAR in Chromatography: Quantitative Structure-Retention Relationships (QSRRs). In: Puzyn, T., Leszczynski, J., Cronin, M. T. (Eds.), *Recent Advances in QSAR Studies*, 1st Edition. Springer Science, Heidelberg London New York, pp. 223–251.

Kataoka, T., Yoshida, H., Ueyama, K., 1972. Mass transfer in laminar region between liquid and packing material surface in the packed bed. *J. Chem. Eng. Japan* 5 (2).

Kazemi, A. S., Latulippe, D. R., 2014. Stirred well filtration (swf) – a high-throughput technique for downstream bio-processing. *J. Membr. Sci.* 470 (0), 30 –39.

Keller, W. R., Evans, S. T., Ferreira, G., Robbins, D., Cramer, S. M., 2015. Use of minicolumns for linear isotherm parameter estimation and prediction of benchtop column performance. *J. Chrom. A* 1418, 94–102.

Kelley, B., Sep 2009. Industrialization of mab production technology: the bioprocessing industry at a crossroads. *mAbs* 1 (5), 443–52.

Kelley, B. D., Switzer, M., Bastek, P., Kramarczyk, J. F., Molnar, K., Yu, T., Coffman, J., 2008. High-throughput screening of chromatographic separations: IV. Ion-exchange. *Biotechnol. Bioeng.* 100 (5), 950–63.

Kinch, M. S., 2014. The rise (and decline?) of biotechnology. *Drug Discovery Today* 19 (11).

Kinch, M. S., Haynesworth, A., Kinch, S. L., Hoyer, D., 2014. An overview of FDA-approved new molecular entities: 1827–2013. *Drug Discovery Today* 19 (8), 1033–1039.

King, R., Miller-Stein, C., Magiera, D., Brann, J., Jan 2002. Description and validation of a staggered parallel high performance liquid chromatography system for good laboratory practice level quantitative analysis by liquid chromatography/tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM* 16 (1), 43–52.

Kittelmann, J., Hubbuch, J., 2015. Robust high-throughput batch screening method in 384-well format with optical in-line resin quantification. *J. Chrom. B* (submitted).

Kittelmann, J., Hubbuch, J., Ottens, M., 2015. Robust high-throughput batch screening method in 384-well format with optical in-line resin quantification. *J. Chrom. B (submitted).*

Knevelman, C., Davies, J., Allen, L., Titchener-Hooker, N. J., 2009. High-throughput screening techniques for rapid PEG-based precipitation of IgG4 mAb from clarified cell culture supernatant. *Biotechnol. Progr.* 26 (3), 697–705.

Knevelman, C., Davies, J., Allen, L., Titchener-Hooker, N. J., 2010. High-throughput screening techniques for rapid peg-based precipitation of igg4 mab from clarified cell culture supernatant. *Biotechnol. Progr.* 26 (3), 697–705.

Kobayashi, H., Sakahara, H., Endo, K., Yao, Z.-s., Toyama, S., Konishi, J., 1995. Repeating the Avidin Chase Markedly Improved the Biodistribution of Radiolabelled Biotinylated Antibodies and Promoted the Excretion of Additional Background Radioactivity. *Eur J Canc* 31 (10).

Konstantinidis, S., Chhatre, S., Velayudhan, A., Heldin, E., Titchener-Hooker, N., 2012. The hybrid experimental simplex algorithm – an alternative method for 'sweet spot' identification in early bioprocess development: case studies in ion exchange chromatography. *Analytica Chimica Acta* 743, 19–32.

Kontoravdi, C., Samsatli, N. J., Shah, N., 2013. Development and design of bio-pharmaceutical processes. *Curr. Opin. Chem. Eng.* 2 (4), 435–441.

Koseki, T., Kitabatake, N., Doi, E., 1988. Conformational changes in ovalbumin at acid pH. *J. Biochem.* 103, 425–430.

Kramarczyk, J. F., Kelley, B. D., Coffman, J. L., 2008. High-throughput screening of chromatographic separations: II. Hydrophobic interaction. *Biotechnol. Bioeng.* 100 (4), 707–20.

Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N., Scholtz, J. M., Apr. 2012. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. *Biophysical J.* 102 (8), 1907–1915.

Kreusch, S., Schwedler, S., Tautkus, B., Cumme, G. A., Horn, A., 2003. UV measurements in microplates suitable for high-throughput protein determination. *Anal. Biochem.* 313 (2), 208–15.

Kumar, V., Sharma, V. K., Kalonia, D. S., Jan. 2009. Effect of polyols on polyethylene glycol (PEG)-induced precipitation of proteins: Impact on solubility, stability and conformation. *Int. J. Pharm* 366 (1-2), 38–43.

Łacki, K. M., 2012. High-throughput process development of chromatography steps: Advantages and limitations of different formats used. *Biotechnol. J.* 7 (10), 1192–1202.

Łacki, K. M., 2014. High throughput process development in biomanufacturing. *Curr. Opin. Chem. Eng.* 6, 25–32.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *Nature* 227 (5259), 680–685.

Laitinen, O. H., Nordlund, H. R., Hytönen, V. P., Kulomaa, M. S., 2007. Brave new (strept)avidins in biotechnology. *Trends Biotechnol.* 25 (6), 269–77.

Lan, K., Jorgenson, J. W., 2001. A hybrid of exponential and gaussian functions as a simple model of asymmetric chromatographic peaks. *J. Chrom. A* 915 (1-2), 1–13.

Landy, M., Dicken, D. M., Bicking, M. M., Mitchell, W. R., 1942. Use of Avidin in Studies on Biotin Requirement of Microorganisms. *Exp. Biol. Med.* 49 (3), 441–444.

Lang, K. M., Kittelmann, J., Dürr, C., Osberghaus, A., Hubbuch, J., 2015. A comprehensive molecular dynamics approach to proteinretention modeling in ion exchange chromatography. *J. Chrom. A*.

Latour, R. A., 2008. Molecular simulation of protein-surface interactions: Benefits, problems, solutions, and future directions. *Biointerphases* 3 (3), 1–12.

Leader, B., Baca, Q. J., Golan, D. E., 2008. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discovery* 7, 21–39.

Li, J., 2002. Comparison of the capability of peak functions in describing real chromatographic peaks. *J. Chrom. A* 952 (1-2), 63–70.

Li, M., 2015. Bioanalytical laboratory automation development: why should we and how could we collaborate? *Bioanalysis* 7 (2), 153–155.

Li-Chan, E., Nakai, S., 1989. Biochemical basis for the properties of egg white. *CRC Critical Rev. Poultry Biol.* 2, 21–58.

Li-Chan, E., Nakai, S., Sim, J., Bragg, D. B., Lo, K. V., 1986. Lysozyme Separation from Egg White by Cation Exchange Column Chromatography. *J. Food Sci.* 51, 1032–1036.

Lienqueo, M. E., Mahn, A., Asenjo, J. A., 2002. Mathematical correlations for predicting protein retention times in hydrophobic interaction chromatography. *J. Chrom. A* 978, 71–79.

Livnah, O., Bayert, E. A., Wilchek, M., Sussman, J. L., 1993. Three-dimensional structures of avidin and the avidin-biotin complex. *Biochemistry* 90, 5076–5080.

Love, J. C., Love, K. R., Barone, P. W., 2013. Enabling global access to high-quality biopharmaceuticals. *Curr. Opin. Chem. Eng.* 2 (4), 383–390.

Lye, G. J., Ayazi-Shamlou, P., Baganz, F., Dalby, P. A., Woodley, J. M., 2003. Accelerated design of bioconversion processes using automated microscale processing techniques. *Trends Biotechnol.* 21 (1), 29–37.

Macko, T., Berek, D., May 2001. Pressure Effects in Hplc: Influence of Pressure and Pressure Changes on Peak Shape, Base Line, and Retention Volume in Hplc Separations. *J. Liquid Chrom. Related Technol.* 24 (9), 1275–1293.

Matsudomi, N., Yamamuraand, Y., Kobayashi, K., 1986. Heat-induced Aggregation between Ovalbumin and Lysozyme. *Agricultural Biological Chem.* 50, 1389–1395.

Mazza, C. B., Rege, K., Breneman, C. M., Sukumar, N., Dordick, J. S., Cramer, S. M., 2002. High-throughput screening and quantitative structure-efficacy relationship models of potential displacer molecules for ion-exchange systems. *Biotechnol. Bioeng.* 80 (1), 60–72.

McCormick, D. B., 1965. Specific Purification of avidin by column chromatography on biotin-cellulose. *Anal. Biochem.* 13, 194–198.

McGown, E. L., Hafeman, D. G., 1998. Multichannel pipettor performance verified by measuring pathlength of reagent dispensed into a microplate. *Anal. Biochem.* 258 (1), 155–7.

McLeod, F., Arnold, F., 2006. Computer Assisted Design of Column Switching Instrument Methods for Reduced Chromatography Run Times. `http://www.dionex.com/en-us/webdocs/35679-LPN\%201795-02\_Computer\%20assisted.pdf`.

McMahon, R. J. (Ed.), 2008. *Avidin-biotin interactions: methods and applications*, 1st Edition. Humana Press, Totowa, NJ.

Melamed, M., Green, N., 1963a. Avidin. *Biochemistry* 89, 591–599.

Melamed, M. D., 1967. Electrophoretic properties of ovomucoid. *Biochem. J.* 103 (3), 805–810.

Melamed, M. D., Green, N. M., 1963b. Avidin 2. Purification and composition. *Biochm. J.* 89 (1959), 591–599.

Merchuk, J. C., Andrews, B. A., Asenjo, J. A., 1998. Aqueous two-phase systems for protein separation: Studies on phase inversion. *J. Chrom. B* 711, 285–293.

Mestre-ferrandiz, J., Sussex, J., Towse, A., 2012. The R&D cost of a new medicine. Tech. rep., Office of Health Economics, London.
URL http://www.ohe.org

Miekka, S. I., Ingham, K. C., Dec. 1978. Influence of self-association of proteins on their precipitation by poly(ethylene glycol). *Archives of biochemistry and biophysics* 191 (2), 525–536.

Mine, Y., 1995. Recent advances in the understanding of egg white protein functionality. *Trends Food Sci. Technol.* 6, 225–235.

Mistry, S., Kaul, A., Merchuk, J., Asenjo, J., Aug. 1996. Mathematical modelling and computer simulation of aqueous two-phase continuous protein extraction. *J. Chrom. A* 741 (2), 151–163.

Mollerup, J. M., 2007. The thermodynamic principles of ligand binding in chromatography and biology. *J. Biotechnol.* 132 (2), 187–95.

Moors, E. H. M., Cohen, A. F., Schellekens, H., 2014. Towards a sustainable system of drug development. *Drug Discov. Today* 19 (11), 1–10.

Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., Greyson, D., 2011. The cost of drug development: a systematic review. *Health Policy* 100 (1), 4–17.

Muramatsu, N., Minton, A. P., Apr. 1989. Hidden self-association of proteins. *J. Molecular Recognition* 1 (4), 166–171.

Muthukumar, S., Rathore, A. S., 2013. High throughput process development (HTPD) platform for membrane chromatography. *J. Membr. Sci.* 442, 245–253.

Naish, P. J., Hartwell, S., 1998. Exponentially modified gaussian function - a good model for chromatographic peaks in isocratic hplc? *Chromatographia* 26 (1).

Nau, F., Guérin-Dubiard, C., Croguennec, M., 2007. Avidin-biotin technology. In: Huopalahti, R., López-Fandiño, R., Anton, M., Schade, R. (Eds.), *Bioactive Egg Compounds*. Springer, Berlin Heidelberg, pp. 287–292.

Nelson, M., Dolan, J., 2004. Parallel chromatography - double your money. *LC GC North America* 22 (4), 338–343.

Nfor, B. K., Hylkema, N. N., Wiedhaup, K. R., Verhaert, P. D. E. M., van der Wielen, L. A. M., Ottens, M., 2011. High-throughput protein precipitation and hydrophobic interaction chromatography: salt effects and thermodynamic interrelation. *J. Chrom. A* 1218 (49), 8958–73.

Nfor, B. K., Noverraz, M., Chilamkurthi, S., Verhaert, P. D. E. M., van der Wielen, L. A. M., Ottens, M., 2010. High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *J. Chrom. A* 1217 (44), 6829–50.

Nfor, B. K., Verhaert, P. D. E. M., van der Wielen, L. A. M., Hubbuch, J., Ottens, M., 2009. Rational and systematic protein purification process development: the next generation. *Trends Biotechnol.* 27 (12), 673–679.

Oelmeier, S., Dismer, F., Hubbuch, J., Jan. 2011. Application of an Aqueous Two-Phase Systems High-Throughput Screening Method to Evaluate mAb HCP Separation. *Biotechnol. Bioeng.* 108 (1), 69–81.

Oelmeier, S. A., 2012. *Aqueous two phase extraction of proteins: From molecular understanding to process development*. Ph.d., Karlsruhe Institute of Technology (KIT).

Oelmeier, S. A., Ladd-Effio, C., Hubbuch, J., 2012a. High throughput screening based selection of phases for aqueous two-phase system-centrifugal partitioning chromatography of monoclonal antibodies. *J. Chrom. A* 1252 (2010).

Oelmeier, S. A., Ladd-Effio, C., Hubbuch, J., 2012b. High throughput screening based selection of phases for aqueous two-phase system-centrifugal partitioning chromatography of monoclonal antibodies. *J. Chrom. A* 1252, 104–114.

Omana, D. A., Wang, J., Wu, J., 2010. Co-extraction of egg white proteins using ion-exchange chromatography from ovomucin-removed egg whites. *J. Chrom. B* 878, 1771–1776.

Orellana, C. A., Shene, C., Asenjo, J. A., 2009. Mathematical modeling of elution curves for a protein mixture in ion exchange chromatography applied to high protein concentration. *Biotechnol. Bioeng.* 104 (3), 572–81.

Osberghaus, A., Baumann, P., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012a. Detection, Quantification, and Propagation of Uncertainty in High-Throughput Experimentation by Monte Carlo Methods. *Chem. Eng. Technol.* 35 (8), 1456–1464.

Osberghaus, A., Drechsel, K., Hansen, S., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012b. Model-integrated process development demonstrated on the optimization of a robotic cation exchange step. *Chem. Eng. Sci.* 76, 129–139.

Osberghaus, A., Drechsel, K., Hansen, S., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012c. Model-integrated process development demonstrated on the optimization of a robotic cation exchange step. *Chem. Eng. Sci.* 76, 129–139.

Osberghaus, A., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012d. Determination of parameters for the steric mass action model-A comparison between two approaches. *J. Chrom. A* 1233, 54–65.

Paty, I., Trellu, M., Destors, J.-M., Cortez, P., Boëlle, E., Sanderink, G., Apr. 2010. Reversibility of the anti-FXa activity of idrabiotaparinux (biotinylated idraparinux) by intravenous avidin infusion. *J. Thrombosis Haemostasis* 8 (4), 722–729.

Peters, F. T., Maurer, H. H., 2005. Bioanalytical method validation and its implications for forensic and clinical toxicology - a review. In: *Validation in Chemical Measurement*. Springer, pp. 1–9.

Pina, A. S., Lowe, C. R., Roque, A. C., 2014. Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnol. Advances* 32 (2), 366–81.

Piskarev, V. E., Shuster, A. M., Gabibov, A. G., Rabinkov, A. G., 1990. A novel preparative method for the isolation of avidin and riboflavin-binding glycoprotein from chicken egg-white by the use of high-performance liquid chromatography. *Biochem. J.* 265 (1), 301–304.

Popovici, S., Schoenmakers, P., Dec 2005. Fast size-exclusion chromatography–theoretical and practical considerations. *J. Chrom. A* 1099 (1-2), 92–102.

Rao, M., Gupta, M., Roy, I., . Process for the isolation and purification of the glycoprotein avidin. *European Patent Office* EP 1 505 882 B1, Filed Oct 14 2003.

Rayat, A. C., Lye, G. J., Micheletti, M., 2014. A novel microscale crossflow device for the rapid evaluation of microfiltration processes. *J. Membr. Sci.* 452, 284–293.

Rege, K., Heng, M., 2010. Miniaturized parallel screens to identify chromatographic steps required for recombinant protein purification. *Nat. Protoc.* 5 (3), 408–17.

Rege, K., Ladiwala, A., Cramer, S. M., 2005. Multidimensional high-throughput screening of displacers. *Anal. Chem.* 77 (21), 6818–6827.

Rege, K., Pepsin, M., Falcon, B., Steele, L., Heng, M., Mar. 2006. High-throughput process development for recombinant protein purification. *Biotechnol. Bioeng.* 93 (4), 618–630.

Rege, K., Tugcu, N., Cramer, S. M., 2003. Predicting Column Performance in Displacement Chromatography from High Throughput Screening Batch Experiments. *Sep Purif. Technol.* 38 (7), 1499–1517.

Rey, G., Wendeler, M. W., 2012. Full automation and validation of a flexible ELISA platform for host cell protein and protein A impurity detection in biopharmaceuticals. *J. Pharm. Biomed. Anal.* 70, 580–586.

Ricker, R. D., Sandoval, L. A., Aug 1996. Fast, reproducible size-exclusion chromatography of biological macromolecules. *J. Chrom. A* 743 (1), 43–50.

Rosa, P. A. J., Azevedo, A. M., Sommerfeld, S., Bäcker, W., Aires-Barros, M. R., 2011. Aqueous two-phase extraction as a platform in the biomanufacturing industry: economical and environmental sustainability. *Biotechnol Adv* 29 (6), 559–567.

Rosenberg, A. S., Jan 2006. Effects of protein aggregates: an immunologic perspective. *The AAPS journal* 8 (3), E501–7.

Rothemund, D. L., Thomas, T. M., Rylatt, D. B., 2002. Purification of the basic protein avidin using Gradiflow technology. *Protein Expression Purif.* 26 (1), 149–52.

Safarik, I., Safarikova, M., 2004. Magnetic techniques for the isolation and purification of proteins and peptides. *Biomagnetic Research Technol.* 2 (1), 7.

Sakahara, H., Saga, T., 1999. Avidin-biotin system for delivery of diagnostic agents. *Adv. Drug Delivery Rev.* 37, 89–101.

Savi, P., Herault, J. P., Duchaussoy, P., Millet, L., Schaeffer, P., Petitou, M., Bono, F., Herbert, J. M., Oct. 2008. Reversible biotinylated oligosaccharides: a new approach for a better management of anticoagulant therapy. *J. Thrombosis Haemostasis* 6 (10), 1697–1706.

Scannell, J. W., Blanckley, A., Boldon, H., Warrington, B., 2012. Diagnosing the decline in pharmaceutical R D efficiency. *Nat. Rev. Drug Discovery* 11 (3), 191–200.

Schenk, J., Balazs, K., Jungo, C., Urfer, J., Wegmann, C., Zocchi, A., Marison, I. W., Stockar, U. V., Bioscience, B., 2008. Influence of Specific Growth Rate on Specific Productivity and Glycosylation of a Recombinant Avidin Produced by a Pichia pastoris Mut R Strain. *Biotechnol. Bioeng.* 99, 368–377.

Schmidt, A. S., Andrews, B. A., Asenjo, J. A., 1996. Correlations for the partition behavior of proteins in aqueous two-phase systems: effect of overall protein concentration. *Biotechnol. Bioeng.* 50, 617–626.

Schmidt-Traub, H. (Ed.), 2005. *Preparative Chromatography*, 1st Edition. Wiley-VCH, Weinheim.

Schröder, M., von Lieres, E., Hubbuch, J., 2006. Direct quantification of intraparticle protein diffusion in chromatographic media. *J Phys. Chem. B* 110 (3), 1429–36.

Seidel-Morgenstern, A., Kessler, L. C., Kaspereit, M., 2008. New developments in simulated moving bed chromatography. *Chem. Eng. Technol.* 31 (6), 826–837.

Seidel-Morgenstern, A., Schmidt-Traub, H., Michel, M., Epping, A., Jupke, A., 2012. *Modeling and Model Parameters*. In: Schmidt-Traub, H., Seidel-Morgenstern, A., M, S. (Eds.), *Preparative Chromatography*, 2nd Edition. WILEY-VCH, Weinheim, Ch. 6, pp. 321–424.

Shukla, A. A., Hubbard, B., Tressel, T., Guhan, S., Low, D., 2007. Downstream processing of monoclonal antibodies–application of platform approaches. *J. Chrom. B* 848 (1), 28–39.

Shulgin, I. L., Ruckenstein, E., Apr. 2006. Preferential hydration and solubility of proteins in aqueous solutions of polyethylene glycol. *Biophys. Chem* 120 (3), 188–198.

Sittampalam, G. S., Kahl, S. D., Janzen, W. P., 1997. High-throughput screening: advances in assay technologies. *Curr. Opin. Chem. Biol.* 1, 384–391.

Smith, C., 2005. Striving for purity: advances in protein purification. *Nat. Methods* 2 (1), 71–77.

Statista, 2014. Top-selling biotech drugs based on revenue in 2013. Accessed Nov 25, 2014.
URL `http://www.statista.com/statistics/299138/top-selling-biotech-drugs-based-on-revenue/`

Stevens, R. C., 2000a. Design of high-throughput methods of protein production for structural biology. *Structure* 8 (9), 177–185.

Stevens, R. C., 2000b. High-throughput protein crystallization. *Curr. Opin. Struct. Biol.* 10 (5), 558–563.

Straathof, A. J. J., 2011. Comprehensive Biotechnology, second ed. Edition. Vol. 1. Elsevier.

Su, C.-K., Chiang, B. H., 2006. Partitioning and purification of lysozyme from chicken egg white using aqueous two-phase system. *Process Biochem.* 41, 257–263.

Sullivan, M., Nicholls, J., 1942. Nutritional dermatoses in the rat: V. signs and symptoms resulting from a diet containing unheated dried egg white as the source of protein. *Arch. Dermatol. Syph.* 45 (2), 295–314.

Susanto, A., Knieps-Grünhagen, E., von Lieres, E., Hubbuch, J., 2008. High throughput screening for the design and optimization of chromatographic processes: Assessment of model parameter determination from high throughput compatible data. *Chem. Eng. Technol.* 31 (12), 1846–1855.

Susanto, A., Treier, K., Knieps-Grünhagen, E., von Lieres, E., Hubbuch, J., 2009. High Throughput Screening for the Design and Optimization of Chromatographic Processes: Automated Optimization of Chromatographic Phase Systems. *Chem. Eng. Technol.* 32 (1), 140–154.

Sutherland, I. A., 2007. Recent progress on the industrial scale-up of counter-current chromatography. *J. Chrom. A* 1151 (1–2), 6 –3.

Swartz, M., Apr 2005. Uplc: An introduction and review. *J. Liquid Chrom. Related Technol.* 28 (7), 1253–1263.

Tao, Y., Carta, G., 2008. Rapid monoclonal antibody adsorption on dextran-grafted agarose media for ion-exchange chromatography. *J. Chrom. A* 1211 (1-2), 70–9.

Thiemann, J., Jankowski, J., Rykl, J., Kurzawski, S., Pohl, T., Wittmann-Liebold, B., Schlüter, H., 2004. Principle and applications of the protein-purification-parameter screening system. *J. Chrom. A* 1043 (1), 73–80.

Tia, S., Herr, A. E., 2009. On-chip technologies for multidimensional separations. *Lab on a Chip* 9 (17), 2524–36.

To, B. C., Lenhoff, A. M., 2007. Hydrophobic interaction chromatography of proteins: ii. solution thermodynamic properties as a determinant of retention. *J. Chrom. A* 1141 (2), 235–243.

Treier, K., Berg, A., Diederich, P., Lang, K., Osberghaus, A., Dismer, F., Hubbuch, J., 2012a. Examination of a genetic algorithm for the application in high-throughput downstream process development. *Biotechnol. J.* 7 (10), 1203–1215.

Treier, K., Hansen, S., Richter, C., Diederich, P., Hubbuch, J., Lester, P., 2012b. High-throughput methods for miniaturization and automation of monoclonal antibody purification processes. *Biotechnol. Progr.* 28 (3), 723–32.

van Deemter, J. J., Zuiderweg, F. J., Klinkenberg, A., 1956. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chem. Eng. Sci.* 5 (6), 271 – 289.

Vazquez-Rey, M., Lang, D. a., Jul 2011. Aggregates in monoclonal antibody manufacturing processes. *Biotechnol. Bioeng.* 108 (7), 1494–508.

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., Bignon, C., 2004. High-throughput automated refolding screening of inclusion bodies. *Prot. Sci.* 13 (10), 2782–2792.

Vlachakis, D., Bencurova, E., Papangelopoulos, N., Kossida, S., 2014. Current State-of-the-Art Molecular Dynamics Methods and Applications. *Adv. Prot. Chem. Struct. Biol.* 94, 269–313.

von Lieres, E., Andersson, J., 2010. A fast and accurate solver for the general rate model of column liquid chromatography. *Comp. Chem. Eng.* 34 (8), 1180–1191.

Wahler, D., Reymond, J.-L., 2001. High-throughput screening for biocatalysts. *Curr. Opin. Biotechnol.* 12 (6), 535–544.

Walsh, G., 2010. Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today* 15 (17–18), 773–80.

Walter, H., Brooks, D. E., Fisher, D., 1985. *Partitioning in aqueous two-phase systems : theory, methods, uses, and applications to biotechnology*. Academic Press, Orlando.

Warner, R. C., 1954. . In: Neurath, H., Bailey, K. (Eds.), *The proteins*, vol 2 part Edition. Academic Press, New York, Ch. 17, p. 440.

Welsh, J. P., Petroff, M. G., Rowicki, P., Bao, H., Linden, T., Roush, D. J., Pollard, J. M., 2014. A practical strategy for using miniature chromatography columns in a standardized high-throughput workflow for purification development of monoclonal antibodies. *Biotechnol. Progr.* 30 (3), 626–635.

Wenger, M. D., DePhillips, P., Price, C. E., Bracewell, D. G., 2007. An automated microscale chromatographic purification of virus-like particles as a strategy for process development. *Biotechnol. Appl. Biochem.* 47 (2), 131–139.

Wiendahl, M., Schulze Wierling, P., Nielsen, J., Fomsgaard Christensen, D., Krarup, J., Staby, A., Hubbuch, J., Jun. 2008. High Throughput Screening for the Design and Optimization of Chromatographic Processes - Miniaturization, Automation and Parallelization of Breakthrough and Elution Studies. *Chem. Eng. Tech.* 31 (6), 893–903.

Wiendahl, M., Völker, C., Husemann, I., Krarup, J., Staby, A., Scholl, S., Hubbuch, J., 2009. A novel method to evaluate protein solubility using a high throughput screening approach. *Chem. Eng. Sci.* 64 (17), 3778–3788.

Wierling, P. S., Bogumil, R., Knieps-Grünhagen, E., Hubbuch, J., 2007. High-Throughput Screening of Packed-Bed Chromatography Coupled With SELDI-TOF MS Analysis : Monoclonal Antibodies Versus Host Cell Protein. *Biotechnol. Bioeng.* 98 (2), 440–450.

Wilchek, M., Bayer, E. A., 1988. The Avidin-Biotin Complex in Bioanalytical Applications. *Anal. Biochem.* 171, 1–32.

Wilchek, M., Bayer, E. A., 1989. Avidin-biotin technology ten years on: Has it lived up to its expectations? *Trends Biochem. Sci.* 14 (10), 408–412.

Wilchek, M., Bayer, E. A., 1990a. Avidin-Biotin Technology. *Methods in Enzymology* 184, 14–45.

Wilchek, M., Bayer, E. A. (Eds.), 1990b. Biotin-binding proteins: Overview and prospects. Vol. 184 of Methods in Enzymology. Academic Press.

Wilchek, M., Bayer, E. a., Livnah, O., 2006. Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunol. Letters* 103 (1), 27–32.

Xie, I. H., Wang, M. H., Carpenter, R., Wu, H. Y., 2004. Automated calibration of TECAN genesis liquid handling workstation utilizing an online balance and density meter.

Zaslavsky, B. Y., 1994. *Aqueous Two-Phase Systems - Physical Chemistry and Bioanalytical Applications*. Marcel Dekker, Inc., New York, NY.

Zavaleta, C. L., Phillips, W. T., Soundararajan, A., Goins, B. A., 2007. Use of avidin-biotin-liposome system for enhanced peritoneal drug delivery in an ovarian cancer model. *Intern. J. Pharmaceutics* 337 (1–2), 316–28.

Zhang, J., Burman, S., Gunturi, S., Foley, J. P., 2010. Method development and validation of capillary sodium dodecyl sulfate gel electrophoresis for the characterization of a monoclonal antibody. *J Pharm. Biomed. Anal.* 53 (5), 1236–1243.

Zhao, G., Dong, X.-Y., Sun, Y., Oct. 2009. Ligands for mixed-mode protein chromatography: Principles, characteristics and design. *J. Biotechnol.* 144 (1), 3–11.