CHARACTERIZATION, MANIPULATION, AND PREDICTION OF PROTEIN AGGREGATION IN MODEL SYSTEMS

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. Lara Galm

aus Öhringen

Referent:	Prof. Dr. Jürgen Hubbuch
Korreferent:	Prof. Dr.-Ing. Matthias Kind
Tag der mündlichen Prüfung:	18. Dezember 2015

Danksagung

Von ganzem Herzen danken möchte ich:

- Prof. Dr. Jürgen Hubbuch, meinem Doktorvater. Jürgen, ich danke dir für die Möglichkeit, meine Arbeit in deiner Arbeitsgruppe durchführen zu können. Danke für die gestalterische Freiheit, für deine Unterstützung und für deine Offenheit.

- Prof. Dr.-Ing. Matthias Kind für die freundliche Übernahme des Korreferats.

- Dem Bundesministerium für Bildung und Forschung (BMBF) für die finanzielle Unterstützung meiner Arbeit im Rahmen des Verbundprojekts "Proteinaggregation bei der Herstellung moderner Biopharmazeutika".

- Meinen Kollegen am MAB für die tolle Zeit und die freundschaftliche Atmosphäre. Für gemeinsame sportliche Unternehmungen und den einen oder anderen Grillabend. Danke auch an die "erste Generation" an Doktoranden am MAB.

- -129 im Speziellen. Das hat einfach gepasst. Und es war mir ein Vergnügen für den einen oder anderen Running Gag verantwortlich gewesen zu sein und euch damit den Tag zu versüßen.

- Kai, Juliane, Heike, Katha und Sven für die erfolgreiche fachliche Zusammenarbeit.

- Josi. Du warst einfach eine Wahnsinns-Diplomandin.

- Katha, Sven und Carsten für morgendliche Tennis-Sessions und für die entspannten Arbeitstage danach.

- Katha, Heike, Sven, Frank, Therese und Carsten für die schönen Stunden, die wir privat miteinander verbracht haben.

- Meinen Eltern, Ute und Helmut. Danke für eure Liebe, für eine Heimat, für Wurzeln und Flügel. Für eure Unterstützung, für euer Vorbild und für alles, was ich hier nicht in Worte fassen kann. Ihr seid einfach die Besten!

- Meinen Schwestern, Almuth und Lisa. Danke für eure Entscheidungshilfen, fürs Zuhören und fürs Teilen. Ihr seid immer da. Ihr seid immer an meiner Seite. Ihr seid vor allem so viel mehr als nur Schwestern. Ich liebe euch und ich bin so dankbar, dass es euch gibt.

- Jutta und Hans. Für euer Vertrauen, eure Unterstützung, für guten Wein und gutes Essen. Für eure Lebensweisheiten und euren Erfahrungsschatz.

- Johanna, meiner großen Karlsruher Schwester, und Jochen. Was für eine glückliche Fügung, die uns in der Mühle zusammengeführt hat. Danke für die vielen, vielen Abende mit Konzerten, Kochen und Kreativem. Für Slowenien. Und für grandioses Frühstück. Danke für die Ablenkung und die andere Welt, in die man mit euch eintaucht. Danke euch beiden für einfach alles.

• Katha. Für dein immer offenes Ohr und fürs Kopf waschen. Fürs Pläneschmieden und "über den Sinn des Lebens philosophieren". Auf die Gegenwart und die Zukunft! Und auf die ereignisreichen Jahre, die jetzt auf uns zukommen!

• Sven. Wie schon Kathy sagte: "Danke. Einfach für alles." Für deine Feinfühligkeit, deine Geselligkeit, deine Kochkünste und ganz viel mehr. Ich habe selten einen so empathischen, loyalen und aufmerksamen Menschen wie dich kennengelernt.

• Caro und Uli. Für die lange Zeit und die lange Freundschaft, die uns verbindet.

• Dem Frauenzirkel. Katha, Rabea, Inga und Tashima. Schön wars. Und zwar wirklich jedes Mal. Ihr habt meine Zeit in Karlsruhe noch viel besser gemacht.

• Meiner Skiurlaubs-Truppe. Ihr wisst es nicht, aber unser erster Urlaub hat den Wendepunkt gebracht. Danke!

• Sebi, Rike, Ramina, Steffen B., Steffen S., Jochen, Jens und Albert für eine tolle Studienzeit in Karlsruhe!

• Allen meinen Freunden im Allgemeinen. Und den Menschen, die ich namentlich nicht explizit erwähnt habe, die aber auch –wissentlich oder unwissentlich- zum Erfolg dieser Arbeit beigetragen haben.

…und sobald du die Antwort hast,
ändert das Leben die Frage.

-unbekannter Verfasser-

Contents

Abstract

The general purpose of this thesis is the characterization, manipulation, and prediction of protein aggregation behavior for selected model proteins. Considering the characterization of protein aggregation behavior, main influencing factors favoring protein aggregation are to be identified regarding solution conditions of the surrounding aqueous solution (e.g. pH value, salt concentration) and regarding physicochemical properties of the model proteins themselves (e.g. net charge, hydrophobicity). Based on these experimental results, concepts are developed for manipulation of protein aggregation behavior, i.e. for selective control of the protein aggregate or phase state and for suppression of non-native aggregation. Finally, based on selected parameters that significantly impact protein aggregation behavior, approaches are developed to predict protein aggregation propensity. This thesis further aims to increase the basic knowledge about protein aggregation as mechanistic understanding is still limited. Generally, protein aggregation has been described as desirable or undesirable as well as controlled or accidental event during purification, formulation, and storage of protein therapeutics. On the one hand protein aggregation (especially crystallization and precipitation) is an acknowledged process step in biopharmaceutical industries either for purification or formulation purposes. On the other hand, conformational instabilities, e.g. partial unfolding, and colloidal instabilities both likewise favor protein aggregation and might not always be mutually exclusive, but accompany each other. However, aggregation due to conformational instabilities might result in irreversible, non-native aggregation and might lead to reduced product activity of the protein therapeutic and to immunogenic reactions after application. Thus, in either case it is essential to ensure that the protein remains in its native conformation and that accidental protein aggregation is reduced to a minimum. Therefore, characterization of the main influencing factors favoring protein aggregation, characterization of the protein conformation during aggregation (native or non-native) as well as their manipulation and selective control are of increasing importance to biopharmaceutical research and development and might aid to identify parameters for prediction of protein aggregation propensity.

For characterization of solution conditions favoring aggregation an approach to generate protein phase diagrams in high throughput was successfully implemented on an automated liquid handling station. Protein solutions were prepared in microbatch experiments in 24 µL scale. Thus, the approach enabled to generate protein phase diagrams with minimized time and product consumption (2 hours and 150 mg protein for 96 microbatch experiments). The method allowed for realization of an extensive and systematic study of protein phase behavior. The developed approach was applied to five different proteins, namely lysozyme from chicken egg white (lysozyme), human lysozyme, glucose oxidase, glucose isomerase, and α-lactalbumin,

covering a wide range of protein sizes and isoelectric points. Solution conditions were chosen to meet typical conditions arising during biopharmaceutical processes, i.e. high salt concentrations (0-2.5 M sodium chloride, 0-1.5 M ammonium sulfate) and extreme pH values (pH 3 to pH 9). Thus, the screened conditions provide valuable information for process design in biopharmaceutical industries either for selection of conditions to avoid protein aggregation or for selection of conditions with the desired aggregate or phase state (e.g. crystallization or precipitation). For the investigated proteins it was possible to induce all known phase transitions, which are crystallization, precipitation, skin formation, gelation, and liquid-liquid phase separation. The investigated salts thereby showed various effects on the protein phase behavior that could not be generalized with regard to the salt type. But, very high salt concentrations almost always induced aggregation. Extremely acidic pH values (pH 3) induced specific phase transitions such as skin formation and gelation. As gelation and skin formation are assumed to be related to partial or complete protein denaturation and thus to non-native aggregation their occurrence is unfavorable for biopharmaceutical processes. Generally, the wide screening of protein phase behavior experimentally supported a pH dependency of short-range attractive forces -most likely arisen due to hydrophobic interactions between the protein monomers- which was not described in literature earlier.

At some points of a biopharmaceutical process, extreme salt concentrations (e.g. during ion exchange or hydrophobic interaction chromatography) or extremely acidic pH values (e.g. during low pH viral inactivation) cannot be circumvented. However, the first part of this thesis could show that high salt concentrations generally increase protein aggregation propensity and that extremely acidic pH values increase the risk of non-native aggregation. Thus, in these cases strategies to completely prevent aggregation, to prevent non-native aggregation or to selectively control aggregate states are needed. Special additive classes, namely osmolytes, and polyethylene glycol (PEG) are known to influence protein aggregation as they are supposed to stabilize the native conformational state of proteins or to reduce attractive protein-protein interactions. Though, their mode of action is not completely clear yet, and their impact on protein solubility and on the type of aggregate state to evolve is even more uncertain as predicting their impact on protein conformation. In this thesis the impact of two osmolytes, namely glycerol and glycine, and of PEG 1000 on lysozyme phase behavior, crystal size and morphology, and on lysozyme solubility was investigated by comparing it to two different reference systems (lysozyme at pH 3 and pH 5, using sodium chloride as precipitant). Therefore, the automated high throughput approach for generation of binary protein phase diagrams in microbatch experiments, consisting of protein and precipitant as solution components, was extended to generation of ternary protein phase diagrams, consisting of protein, precipitant, and the respective additive as solution components. Protein conformation of the reference systems and the impact of glycerol, glycine, and PEG 1000 on conformational stability were investigated

using Fourier transform infrared (FT-IR) spectroscopy. Basically, it could be shown that the impact of glycerol, glycine, and PEG 1000 on lysozyme phase behavior, crystal size and morphology, and on lysozyme solubility were strongly pH dependent. The pH dependency could be explained by differences in the conformational stability of lysozyme at pH 3 and pH 5 in presence of sodium chloride. This correlation between the pH dependent additive impact and the protein's conformational stability was not described in literature earlier. For conditions where conformational stability of lysozyme was low in the binary reference system, addition of glycerol and PEG 1000 increased the conformational stability, whereas addition of glycine further decreased conformational stability. PEG 1000 in particular even completely inhibited non-native aggregation, whereas glycine even promoted non-native aggregation. However, stabilization or destabilization of the native conformational state did not prevent aggregation and did not affect lysozyme solubility and crystal size and morphology. But, phase states could be successfully selectively controlled, meaning that for example former precipitated phase states could be transferred to crystalline ones and vice versa. For conditions where the conformational stability of lysozyme was high in the binary reference system no impact of the additives on lysozyme's native conformation could be observed. Though, the additives successfully prevented aggregation and exhibited a significant impact on lysozyme crystal size and morphology, and on lysozyme solubility. In summary, all of the three additives investigated showed characteristics that might be of interest for biopharmaceutical processes: aggregate or phase states could be selectively controlled and non-native aggregation could be prevented. Aggregation could be successfully manipulated with or without effects on protein solubility, crystal size and morphology.

The experimental results collected during characterization and manipulation of protein aggregation led to the conclusion that protein surface hydrophobicity, or more precisely short-range forces due to protein hydrophobicity, as well as conformational stability are of major importance for protein aggregation. Those parameters thus were used to develop a predictive approach for protein aggregation propensity based on experimental and *in silico* methods. As protein surface hydrophobicity still is hard to assess both experimentally and *in silico*, a novel non-invasive high throughput compatible approach for determination of protein surface hydrophobicity based on surface tension of protein solutions was developed during this work. The presented methodology showed to outclass other widely used spectrophotometric methods with hydrophobicity sensitive dyes as it gave reasonable results without restrictions on pH, protein species, and dye solubility. A number of model proteins, namely lysozyme, human lysozyme, bovine serum albumin, and α-lactalbumin, were investigated using this novel non-invasive approach and a pH dependent hydrophobicity ranking could be developed. This pH dependent hydrophobicity ranking further supported the observation from the first part of this thesis on a pH dependent hydrophobic character of proteins and resulting pH dependent short-

range hydrophobic forces. Conformational stability, which should in addition to protein hydrophobicity be used to develop a predictive approach for protein aggregation propensity, generally refers to the tendency of proteins to perform profound conformational changes (i.e. partial unfolding or denaturation). However, the previous results showed that most aggregate states in this thesis were native, i.e. that protein monomers involved in aggregation processes only differ slightly in structure. Thus, conformational flexibility was investigated in the following rather than conformational stability as conformational flexibility describes only minor structural fluctuations between conformational states with similar energy. Conformational flexibility of lysozyme, human lysozyme, and α-lactalbumin could be assessed by analysis of the root-mean-square-fluctuation (RMSF) of the C_α, C, and N backbone atoms of the proteins using an *in silico* approach by means of molecular dynamics (MD) simulations. The predictive power of protein surface hydrophobicity and conformational flexibility for the aggregation behavior of lysozyme, human lysozyme, and α-lactalbumin at pH 3, 5, 7, and 9 without precipitants and with precipitants, namely sodium chloride and ammonium sulfate, was tested. Therefore, the correlation between protein surface hydrophobicity, determined experimentally, conformational flexibility, determined *in silico*, and aggregation propensity was probed. Protein aggregation propensity hereby was extracted from protein phase diagrams. Both parameters could be identified to be main determinants for aggregation propensity and to be suitable parameters for prediction of aggregation propensity, albeit with some limitations. Neither for sodium chloride nor for ammonium sulfate a clear correlation between surface hydrophobicity or protein flexibility and aggregation propensity as function of pH within one protein species could be found. Aggregation propensity could only be correlated to surface hydrophobicity or flexibility at fixed pH values by comparison of different protein species. For sodium chloride as precipitant, correlation between surface hydrophobicity or protein flexibility and aggregation propensity revealed a pH dependency. For ammonium sulfate as precipitant protein surface hydrophobicity was the key driver for aggregation propensity for all pH values investigated. Furthermore, it could be shown that protein surface hydrophobicity might indicate the aggregate state to evolve as hydrophilicity could be related to crystallization and high hydrophobicity to gelation.

All of the presented approaches were developed and validated using selected model proteins (lysozyme, human lysozyme, glucose oxidase, glucose isomerase, bovine serum albumin, and α-lactalbumin), but can easily be transferred to any protein of interest, such as monoclonal antibodies, and will give valuable information for biopharmaceutical purification and formulation processes.

Abstract in German - Zusammenfassung

Die Zielsetzung der vorliegenden Dissertation ist die Charakterisierung, Manipulation und Vorhersage des Aggregationsverhaltens von ausgewählten Modellproteinen. Unter dem Aspekt der Charakterisierung des Aggregationsverhaltens sollen zuerst die Haupteinflussfaktoren identifiziert werden, die Proteinaggregation begünstigen und zwar einerseits bezogen auf Eigenschaften der umgebenden wässrigen Lösung (z.B. pH, Salzgehalt) und andererseits bezogen auf die physikochemischen Eigenschaften der Proteine selbst (z.B. Nettoladung, Hydrophobizität). Die experimentellen Ergebnisse dienen als Grundlage für die Entwicklung von Ansätzen zur gezielten Beeinflussung des Aggregationsverhaltens, der Aggregationsneigung und zur Unterdrückung nicht-nativer Aggregation. Abschließend sollen basierend auf ausgewählten Parametern, die das Aggregationsverhalten von Proteinen maßgeblich beeinflussen, Ansätze entwickelt werden um das Aggregationsverhalten und die Aggregationsneigung von Proteinen vorherzusagen.

Im Allgemeinen beschreibt Proteinaggregation die Zusammenlagerung nativer und nicht-nativer Proteinmonomere zu Proteinmultimeren. Sowohl während des Aufreinigungsprozesses, der Formulierung, und auch während der Lagerung von Biopharmazeutika kann es zu Proteinaggregation kommen. Proteinkristallisation und –präzipitation, zwei spezielle Formen der Proteinaggregation, werden auch gezielt als Aufreinigungs- und Formulierungsschritte eingesetzt. Allerdings wird Proteinaggregation sowohl durch Instabilitäten der Proteinkonformation (z.B. partielle Entfaltung, Denaturierung) als auch durch kolloidale Instabilitäten (aufgrund attraktiver Protein-Protein Wechselwirkungen) ausgelöst. Aggregation aufgrund von Instabilitäten der Proteinkonformation resultiert hierbei meist in irreversibler, nicht-nativer Aggregation, die widerum zu einer verminderten Produktaktivität und zu immunogenen Reaktionen nach Verabreichung des Proteintherapeutikums führen kann. Es ist daher in jedem Fall unumgänglich sicherzustellen, dass, falls Proteinaggregation auftritt, das Protein in seiner nativen Konformation verbleibt und dass zufällige, unkontrollierte Proteinaggregation auf ein Minimum reduziert wird. Daher ist die oben beschriebene Zielsetzung dieser Dissertation zur Identifizierung der Haupteinflussfaktoren von Proteinaggregation, zur gezielten Manipulation des Aggregationsverhaltens und der Unterdrückung nicht-nativer Aggregation, sowie zur Vorhersage von Proteinaggregation von großer Bedeutung für die biopharmazeutische Industrie.

Für die Charakterisierung des Aggregationsverhaltens von Proteinen und zur Identifizierung der Haupteinflussfaktoren wurde das Phasenverhalten von Proteinen untersucht. Hierfür wurde ein automatisiertes Hochdurchsatzverfahren entwickelt, um die Protein-Phasendiagramme in Microbatch-Experimenten im 24 µL-Maßstab zu erzeugen. Mithilfe dieses

Hochdurchsatzverfahrens im 24 µL-Maßstab konnten Protein-Phasendiagramme mit minimiertem zeitlichen Aufwand und mit minimiertem Produktverbrauch erzeugt werden (120 min und 150 mg Protein für 96 unterschiedliche Bedingungen). Dies ermöglichte die Realisierung einer sehr umfangreichen, systematischen Studie zum Protein-Phasenverhalten, bzw. zum Protein-Aggregationsverhalten. Das Hochdurchsatzverfahren wurde verwendet, um binäre Phasendiagramme fünf verschiedener Proteine zu erstellen (Lysozym aus Hühnereiweiß, humanes Lysozym, Glucose Oxidase, Glucose Isomerase und α-Lactalbumin). Diese Proteine wurden so ausgewählt, dass ein weites Spektrum an Molekulargewichten und isoelektrischen Punkten abgedeckt wird. Der Ausdruck „binäre Phasendiagramme“ wird in dieser Dissertationsschrift verwendet, wenn die untersuchten Lösungen aus dem Protein und einem Präzipitanten (hier Natriumchlorid, Ammoniumsulfat, Polyethylenglykol (PEG) 300 oder Polyethylenglykol (PEG) 1000) als Lösungskomponenten bestanden. Die einzelnen Phasendiagramme wurden in Abhängigkeit der Protein- und Präzipitantkonzentration bei festem pH-Wert und konstanter Temperatur bestimmt. Die Lösungseigenschaften der umgebenden wässrigen Lösungen wurden so ausgewählt, dass sie Eigenschaften abbilden, die auch typischerweise in der biopharmazeutischen Industrie auftreten, z.B. hohe Salzkonzentrationen (0-2.5 M Natriumchlorid, 0-1.5 M Ammoniumsulfat) und extreme pH-Werte (pH Bereich 3 bis 9). Für die untersuchten Proteine konnten alle bekannten Phasenübergänge (Kristallisation, Präzipitation, Hautbildung, Gelierung und Phasenseparation) beobachtet werden. Die untersuchten Salze zeigten hierbei verschiedenartige Einflüsse auf das Phasen- oder Aggregationsverhalten der Proteine, das bezüglich der Salzart nicht generalisiert werden konnte. Es zeigte sich jedoch, dass bei hohen Salzkonzentrationen fast immer Proteinaggregation induziert wurde, dass eine steigende Salzkonzentration die Aggregationsneigung also deutlich erhöht. Bei extrem sauren pH-Werten (pH 3) wurden sehr spezielle Phasenübergänge, wie Gelierung und Hautbildung, beobachtet. Es wird im Allgemeinen angenommen, dass Gelierung und Hautbildung eng mit partieller Entfaltung und Denaturierung von Proteinen zusammenhängen und dass sie somit nicht-native Aggregationsformen bzw. Phasenzustände darstellen, deren Auftreten für die biopharmazeutische Industrie sehr unvorteilhaft ist. Grundsätzlich ergab diese umfassende Studie, dass die kurzreichweitigen, attraktiven Kräfte zwischen Proteinen, die vor allem durch hydrophobe Wechselwirkungen hervorgerufen werden, eine Abhängigkeit vom pH-Wert aufweisen. Dies wurde in der Literatur zuvor noch nicht erwähnt und stellt somit eine neuartige Erkenntnis da.

Der zweite Teil der vorliegenden Arbeit beschäftigte sich mit der gezielten Beeinflussung des Aggregationsverhaltens von Proteinen. An manchen Stellen eines biopharmazeutischen Prozesses sind sehr hohe Salzkonzentrationen (z.B. bei der Ionenaustausch- oder Hydrophoben-Interaktions-Chromatographie) oder extrem saure pH-Werte (z.B. bei der Virus-Inaktivierung) unumgänglich. Allerdings konnte im ersten Teil dieser Arbeit gezeigt werden, dass hohe

Salzkonzentrationen die Aggregationsneigung deutlich steigern und dass extrem saure pH-Werte die Gefahr von nicht-nativer Aggregation erhöhen. In diesen Fällen werden Möglichkeiten benötigt Proteinaggregation komplett zu verhindern, nicht-native Aggregation zu unterdrücken oder zumindest die Aggregationsform, bzw. den Phasenzustand, gezielt zu steuern. Osmolyte, eine bestimmte Klasse von Additiven, und PEG beeinflussen bekanntermaßen das Aggregationsverhalten von Proteinen, vermutlich durch die Stabilisierung der nativen Konformation der Proteine oder durch Reduktion attraktiver Protein-Protein-Wechselwirkungen. Die genaue Wirkweise dieser Additive ist allerdings noch nicht komplett aufgeklärt und Vorhersagen über ihren Einfluss auf die Proteinlöslichkeit oder auf entstehende Aggregationsformen, bzw. Phasenzustände, sind sogar noch unsicherer als Vorhersagen über ihre Wirkung auf die Proteinkonformation. In dieser Dissertation wurde der Einfluss zweier Osmolyte (Glycerol und Glycin) und der Einfluss von PEG 1000 auf das Phasenverhalten, die Kristallgröße und –morphologie, und auf die Löslichkeit von Lysozym aus Hühnereiweiß untersucht und mit zwei Referenzsystemen verglichen (Lysozym bei pH 3 und pH 5 mit Natriumchlorid als Präzipitant). Hierfür wurde das automatisierte Hochdurchsatzverfahren zur Erstellung binärer Protein-Phasendiagramme in Microbatch-Experimenten, das im ersten Teil dieser Arbeit entwickelt wurde, erweitert für die Erstellung ternärer Phasendiagramme, bestehend aus Protein, Präzipitant und dem jeweiligem Additiv als Lösungskomponenten. Die Proteinkonformation der Referenzsysteme und der Einfluss von Glycerol, Glycin und PEG 1000 auf die Stabilität der Proteinkonformation wurde mithilfe der Fourier-Transformations-Infrarot (FT-IR) Spektroskopie untersucht. Es konnte gezeigt werden, dass der Einfluss von Glycerol, Glycin und PEG 1000 auf das Phasenverhalten von Lysozym, die Kristallgröße und –morphologie, sowie auf die Löslichkeit eine starke Abhängigkeit vom pH-Wert aufweist. Diese pH-Abhängigkeit konnte erklärt werden auf Grundlage von Unterschieden der Stabilität der Proteinkonformation in den Referenzsystemen. Diese Korrelation zwischen pH-abhängigem Additiveinfluss und pH-abhängiger Stabilität der Proteinkonformation wurde in der Literatur bis dahin noch nicht beschrieben. Bei pH-Werten mit niedriger Stabilität der Proteinkonformation in den binären Referenzsystemen erhöhte die Zugabe von Glycerol und PEG 1000 die Stabilität der Proteinkonformation, wohingegen die Zugabe von Glycin die Stabilität der Proteinkonformation sogar noch weiter verringerte. Glycerol konnte das Auftreten von nicht-nativer Aggregation verringern und die Zugabe von PEG 1000 konnte nicht-native Aggregation sogar komplett unterdrücken. Glycin hingegen begünstigte nicht-native Aggregation im Vergleich zum binären Referenzsystem. Die Stabilisierung oder Destabilisierung der nativen Proteinkonformation konnte Aggregation im Allgemeinen im Vergleich zum Referenzsystem jedoch nicht verzögern und beeinflusste weder die Lysozymlöslichkeit noch die Kristallgröße und –morphologie. Die Aggregationsform konnte jedoch erfolgreich selektiv gesteuert werden, was bedeuted, dass z.B. im Referenzsystem ursprünglich präzipitierte Zustände in kristalline Zustände überführt werden

konnten und umgekehrt. Bei pH-Werten mit hoher Stabilität der Proteinkonformation in den binären Referenzsystemen konnte kein Einfluss der Additive auf die Proteinkonformation beobachtet werden. Allerdings konnte in diesen Fällen die Proteinaggregation im Vergleich zum Referenzsystem verzögert und ein signifikanter Einfluss auf die Kristallgröße und –morphologie, sowie auf die Lysozymlöslichkeit beobachtet werden. Zusammenfassend ist festzustellen, dass alle drei Additive Wirkweisen ausbildeten, die für biopharmazeutische Prozesse von Interesse sind: Aggregationsformen, bzw. Phasenzustände, konnten selektiv gesteuert und nicht-native Aggregation verhindert werden. Das Aggregationsverhalten konnte erfolgreich gezielt beeinflusst werden, mit oder ohne Einfluss auf die Proteinlöslichkeit, Kristallgröße und Kristallmorphologie.

Die experimentellen Ergebnisse die unter den Aspekten der Charakterisierung und Manipulation von Proteinaggregation gesammelt wurden, lassen die Schlussfolgerung zu, dass die Protein-Oberflächenhydrophobizität, oder genauer die kurzreichweitigen Kräfte die durch die Proteinhydrophobizität hervorgerufen werden, und die Stabilität der Proteinkonformation die Parameter sind, die Proteinaggregation maßgeblich beeinflussen. Diese Parameter wurden daher verwendet, um einen Ansatz zu entwickeln die Aggregationsneigung von Proteinen vorherzusagen, basierend auf experimentellen und *in silico* Methoden. Da die Protein-Oberflächenhydrophobizität immer noch schwer zugänglich ist, sowohl experimentell als auch *in silico*, wurde ein neuartiges, nicht-invasives Hochdurchsatzverfahren entwickelt, mit dem die Protein-Oberflächenhydrophobizität basierend auf der Oberflächenspannung von Proteinlösungen bestimmt werden konnte. Es konnte gezeigt werden, dass diese Methodik traditionellen spektrophotometrischen Methoden mit hydrophobizitätssensitiven Farbstoffen deutlich überlegen ist, da sie sinnvolle Ergebnisse liefert ohne Einschränkungen bezüglich pH-Wert, Proteinspezies, und Löslichkeit des Farbstoffs in wässriger Lösung. Die Oberflächenhydrophobizität von Lysozym aus Hühnereiweiß, humanem Lysozym, Rinderserumalbumin und α-Lactalbumin wurde mithilfe dieses neu entwickelten, nicht-invasiven Hochdurchsatzverfahrens bestimmt und es konnte ein pH-abhängiges Hydrophobizitäts-Ranking erstellt werden. Die Stabilität der Proteinkonformation, die ebenfalls für die Vorhersage der Aggregationsneigung verwendet werden sollte, beschreibt im Allgemeinen die Tendenz von Proteinen deutliche Strukturänderungen zu vollziehen (partielle Entfaltung, komplette Denaturierung). Allerdings zeigten die vorangegangenen Untersuchungen, dass die meisten in dieser Arbeit beobachteten Aggregate nativ waren, d.h. die Proteinmonomere, die in die Aggregationsprozesse involviert waren veränderten sich (wenn überhaupt) nur sehr gering in ihrer Struktur. Deshalb wurde im Folgenden die Proteinflexibilität statt der Stabilität der Proteinkonformation untersucht. Die Proteinflexibilität beschreibt die geringfügige Fluktuation der nativen Proteinkonformation zwischen Zuständen ähnlicher Energie. Die Proteinflexibilität von Lysozym aus Hühnereiweiß, humanem Lysozym und α-

Lactalbumin konnte über eine Analyse der mittleren quadratischen Fluktuation (RMSF) der Position der C_α, C und N- Atome des Proteinrückgrats erfasst werden. Die mittlere quadratische Fluktuation wurde über einen *in silico* Ansatz mittels Moleküldynamik-Simulationen bestimmt. Die Vorhersagekraft der Protein-Oberflächenhydrophobizität und der Proteinflexibilität für das Aggregationsverhalten von Lysozym aus Hühnereiweiß, humanem Lysozym und α-Lactalbumin ohne und mit zugesetzten Präzipitanten (Natriumchlorid und Ammoniumsulfat) wurde für pH 3, 5, 7 und 9 überprüft. Hierfür wurde die Korrelation zwischen Protein-Oberflächenhydrophobizität, Proteinflexibilität und Aggregationsneigung untersucht. Die Aggregationsneigung wurde aus Protein-Phasendiagrammen extrahiert. Beide Parameter konnten als Haupt-Einflussfaktoren für die Aggregationsneigung und als für die Vorhersage der Aggregationsneigung geeignet bestätigt werden, wenn auch mit Einschränkungen. Weder bei Verwendung von Natriumchlorid, noch bei Verwendung von Ammoniumsulfat als Präzipitant konnte ein eindeutiger Zusammenhang zwischen der Protein-Oberflächenhydrophobizität oder der Proteinflexibilität und der Aggregationsneigung als Funktion des pH-Wertes innerhalb einer Proteinart gefunden werden. Die Aggregationsneigung konnte also nur bei festen pH-Werten durch den Vergleich unterschiedlicher Proteinarten mit der Protein-Oberflächenhydrophobizität oder der Proteinflexibilität korreliert werden. Für Natriumchlorid als Präzipitant zeigte sich für die Korrelation zwischen Protein-Oberflächenhydrophobizität, Proteinflexibilität und Aggregationsneigung eine pH-Wert-Abhängigkeit. Für Ammoniumsulfat als Präzipitant war die Protein-Oberflächenhydrophobizität für alle untersuchten Bedingungen die treibende Kraft für die Aggregationsneigung. Desweiteren konnte gezeigt werden, dass die Protein-Oberflächenhydrophobizität einen starken Einfluss auf die entstehende Aggregatform hat. Lysozym aus Hühnereiweiß, ein Protein mit hydrophiler Oberfläche, kristallisierte bereitwillig. Humanes Lysozym und α-Lactalbumin wiesen bei pH 3 eine hohe Oberflächenhydrophobizität auf, was Gelierung begünstigte.

Alle in dieser Arbeit präsentierten Methoden und Ansätze wurden entwickelt und validiert unter Verwendung von Modellproteinen, können aber direkt auf jedes beliebige Protein, z.B. monoklonale Antikörper, übertragen werden.

1. Introduction

Protein therapeutics have played an outstanding role in the past decades and represent a rapidly growing proportion of marketed drugs (Johnson-Léger et al., 2006) with therapeutic and diagnostic antibodies being the fastest growing area of biopharmaceutical applications (Daugherty and Mrsny, 2006). Protein therapeutics are used for example in the treatment of a number of inflammatory diseases, immune disorders and forms of cancer. For many years the treatment of these diseases was limited to small molecule compounds whose use is associated with unwanted side effects including renal toxicity, diabetes and osteoporosis. Besides reduced unwanted side effects for some indications protein drugs are the only effective therapy. (Johnson-Léger et al., 2006) Though, purification and formulation process steps of protein therapeutics are challenging. Purification steps with high salt concentrations or extreme pH values and high concentrated protein formulations are amongst others hard to realize, as the proteins tend to aggregate, i.e. protein monomers assemble to protein multimers. This assembly in some cases is undesirable because of the concern that protein aggregates may lead to immunogenic reactions and reduced product activity (Cromwell et al., 2006). On the other hand, in some cases protein aggregation is an acknowledged process step in biopharmaceutical industries either for purification or formulation purposes (Scopes, 1994). Thus, protein aggregation might be an undesirable as well as an desirable event during purification and formulation of protein therapeutics and identification of aggregation conditions as well as its manipulation and control is of increasing importance to pharmaceutical research and development (Mahler et al., 2009).

In the first subsection of this introduction (Section 1.1) protein aggregation will be defined in detail and acknowledged classification schemes and descriptions of aggregation mechanisms will be shown. As proteins are large three-dimensional molecules with complex hydrophilic, hydrophobic, charged, and uncharged properties that cause interactions between protein monomers and clearly influence aggregation propensity, protein properties and their impact on aggregation propensity will be discussed in Section 1.2. Methods to visualize and characterize protein aggregation by means of protein phase states and protein conformation are described in Section 1.3 and Section 1.4.

1.1 Protein aggregation

1.1.1 Definition of protein aggregation

Protein aggregation describes the assembly of protein monomers to protein multimers. Thus, aggregation characterizes the formation of protein crystals, amorphous precipitates and also incorporates cross-linked structures like protein gels. Conformational and colloidal instabilities of proteins both likewise favor protein aggregation. The widespread opinion exits that aggregation processes, especially amorphous precipitation, are usually associated with a conformational change, i.e. partial unfolding, of the protein monomers (Fink, 1998). Partially unfolded protein monomers are especially prone to aggregation (Chi et al., 2003) as in most cases partial unfolding increases the hydrophobicity of the protein monomers. Aggregation of conformationally unstable protein monomers results in a loss of the protein's native state and thus in non-native aggregation. Protein aggregation due to colloidal instabilities describes the assembly of native protein monomers to aggregates and occurs as result of attractive intermolecular interactions (Chi et al., 2003). Crystallization for example is assumed to be a multimerization of native protein molecules (McPherson, 2004; Philo and Arakawa, 2009), but also native precipitation has been observed (Matheus et al., 2009). Attractive intermolecular interactions occur due to the surface characteristics of protein monomers, as described in Section 1.2.2. Major contributions to interactions between protein molecules in aqueous solution are electrostatic and hydrophobic interactions. Conformational and colloidal instabilities might not always be mutually exclusive, but accompany each other (Chi et al., 2003; Mahler et al., 2009).

1.1.2 Classification of protein aggregation

According to the comprehensive review of Mahler et al. (Mahler et al., 2009) protein aggregation might be classified into the following categories by

- type of bond:
  non-covalent aggregates (bound by electrostatic forces) vs. covalent aggregates (bound by disulfide bridges)
- reversibility:
  reversible vs. irreversible aggregates
- size:
  small soluble oligomers (dimers, trimers, tetramers, etc.) vs. large oligomers (≥ 10-mer

oligomers) vs. aggregates in the diameter range between 20 nm and 1 µm vs. insoluble aggregates in the diameter range between 1 µm and 25 µm vs. visible aggregates

- protein conformation:
  native vs. non-native aggregates

1.1.3 Mechanisms of protein aggregation

Published reports on aggregation pathways of pharmaceutically relevant proteins are still limited. Additionally, there is no single protein aggregation pathway but a variety of pathways, which differ between proteins and result in different end states. The aggregation pathway as well is effected by environmental conditions, such as temperature, pH, protein concentration itself, and by the type of applied stress e.g. through shaking, stirring, pumping, freezing, and thawing. Also, as described above, the initial state of a protein monomer prone to aggregation may differ, i.e. aggregation might occur due to conformational or colloidal instabilities or due to a combination of both. (Mahler et al., 2009) A very extensive description of possible protein aggregation pathways was presented by Philo and Arakawa (Philo and Arakawa, 2009). Native (Mechanism 1), non-native (Mechanism 2), chemically induced (Mechanism 3), nucleation controlled (Mechanism 4), and surface induced (Mechanism 5) aggregation pathways were described and are schematically illustrated in Fig. 1. According to Philo and Arakawa understanding the mechanism of aggregation is advantageous in process and formulation development.

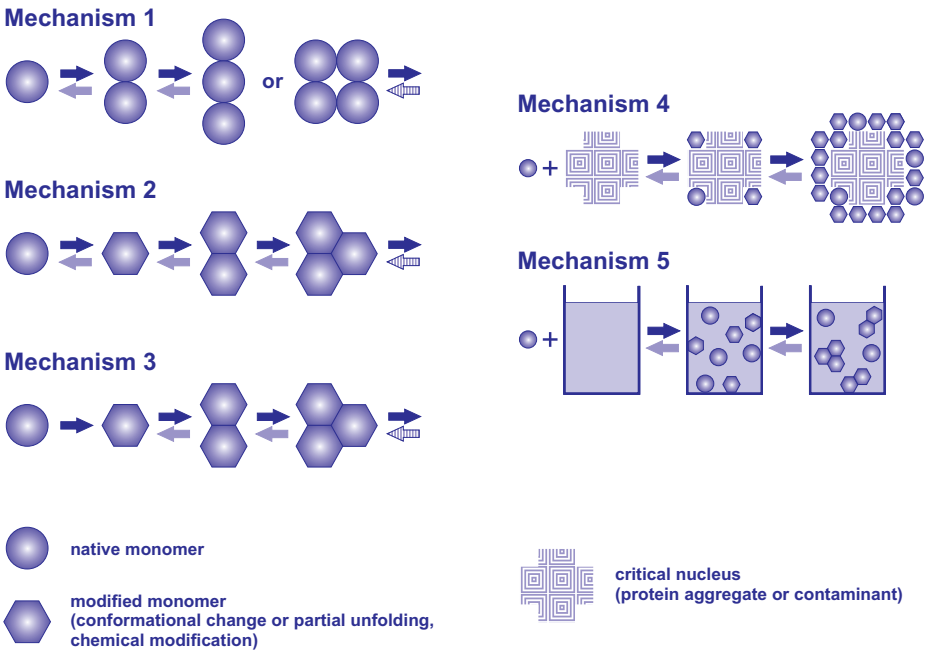

Fig. 1. Schematic illustration of five common aggregation mechanisms, adapted from (Philo and Arakawa, 2009).

1.1.4 Protein aggregation in biopharmaceutical industries

Protein aggregation can occur in every step of a biopharmaceutical process and also during storage of biopharmaceutics. During purification the proteins face solution conditions that are unfavorable for their conformational and colloidal stability, e.g. extreme pH changes, high ionic strengths or agitation stress due to stirring, pumping, and shaking (Chi et al., 2003; Cromwell et al., 2006; Mahler et al., 2009). During formulation protein solutions are usually concentrated and thus protein concentration itself might induce aggregation. The most problematic class of aggregates for biopharmaceutical applications are non-native ones. The presence of such non-native aggregates is considered to be undesirable as immunogenic reactions might be caused and product activity might be influenced. Thus, non-native protein aggregation needs to be inhibited to ensure patient safety. However, protein aggregation on the other hand is an acknowledged process step in biopharmaceutical industries either for purification or formulation purposes (Scopes, 1994). Matheus et al. (Matheus et al., 2009) applied precipitation and subsequent resolubilisation as intermediate process step for the production of highly concentrated drug formulations. Crystalline drug formulations have shown significant benefits in the delivery of protein therapeutics to achieve high-concentration, high-stability, low-viscosity, and controlled-release formulations (Basu et al., 2004; Jen and Merkle, 2001). Crystalline insulin formulations as well as insulin formulations associated with precipitation at injection site are market approved (Basu et al., 2004; Brange and Vølund, 1999; Vajo et al., 2001) and crystalline antibody formulations are studied, too (Yang et al., 2003).

Thus, aggregation in biopharmaceutical industries might be a controlled or an accidental as well as a wanted or an unwanted event. In either case it is essential to ensure that the target protein remains in its native conformation, that biological activity is preserved and that accidental protein aggregation is reduced to a minimum. Therefore, identification of solution conditions inducing protein aggregation, characterization of the nature of the protein aggregate (native or non-native) as well as its manipulation and selective control is of increasing importance to pharmaceutical research and development.

1.2 Impact of protein properties on protein aggregation

1.2.1 Protein conformation

Protein structure can be grouped into four conformational stages. The tertiary structure (third conformational stage and generally referred to as the proteins' native state) of a glucose oxidase monomer and the quaternary structure (fourth conformational stage) of dimeric glucose

oxidase, protein data bank (PDB) entry 1GPE, are illustrated in Fig. 2(a.) and Fig. 2(b.), respectively.

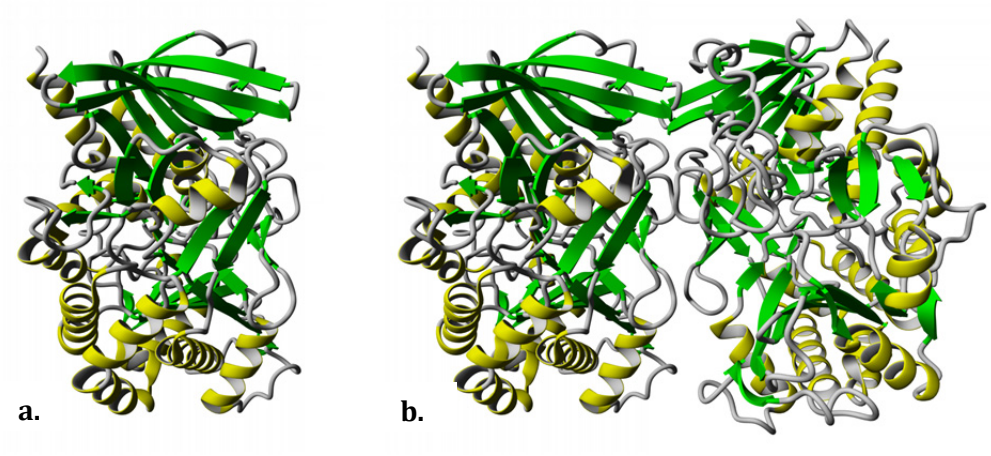

Fig. 2. Glucose oxidase in its tertiary structure as monomer (a.) and in its quaternary structure as dimer (PDB entry 1GPE) (b.), both depicted in ribbon style. Secondary structure elements are colored, i.e. α–helices are depicted in yellow, β–sheets are depicted in green, random coil structures are depicted in grey. Molecular graphics were created using Yasara (www.yasara.org).

Basically, proteins are polymers composed of amino acids that are joined by peptide bonds. The sequence of these joined amino acids gives the primary structure of a protein and determines how the protein organizes into higher-level structures during folding. Glucose oxidase as dimer (Fig. 2(b.)) for example is comprised of 1174 amino acids, 587 amino acids per monomer (Kiess et al., 1998). The secondary structure arises due to regular hydrogen-bonding between N-H and C=O groups in the polypeptide backbone of the primary structure. This regular hydrogen-bonding causes the linear primary structure to adopt α–helix, β-sheet and random coil structures. These secondary structure elements are depicted in yellow, green, and grey in Fig. 2. In aqueous solution and at physiological conditions the protein folds into a three-dimensional tertiary structure which is shown in Fig. 2(a.) for a glucose oxidase monomer. Some proteins like glucose oxidase further assemble to higher protein complexes and build a quaternary structure, originating from the association of like protein monomers in their tertiary structure. (Petsko and Ringe, 2004a) Glucose oxidase naturally occurs in a dimeric form which is shown in Fig. 2(b.). For a protein to gain its function, it must usually be able to form a stable tertiary structure under physiological conditions. This stable tertiary structure is referred to as the protein's native conformational state. Non-native protein conformations cover incorrectly folded protein molecules as well as partially unfolded or completely denatured ones. Partially unfolded protein molecules as well as completely denatured ones are especially prone to aggregation as in most cases their hydrophobicity is increased (Chi et al., 2003). Conformational changes resulting in

non-native protein conformations can for example be identified by investigation of the secondary structure using Fourier transform infrared (FT-IR) spectroscopy (see Section 1.4).

1.2.2 Physicochemical properties of the protein surface

The protein organized in its native tertiary structure possesses a specific molecular surface. This molecular surface will be referred to as protein surface in the following. For a detailed definition of the molecular surface and other surface definitions (e.g. van der Waals surface, solvent accessible surface) see for example (Lijnzaad, 2007). The molecular surface of one monomer of dimeric glucose oxidase (PDB entry 1GPE) is depicted in Fig. 3.

The characteristics of the protein surface are of great importance for protein function as well as for protein-protein interactions. These characteristics are determined by the properties of the amino acids, or amino acid residues respectively, located at the protein surface. There are three major categories of amino acid residues according to Voet et al. (Voet et al., 1999): hydrophobic nonpolar, uncharged polar and charged polar residues. According to Petsko and Ringe (Petsko and Ringe, 2004a) there is also a group of amino acid residues with amphipathic (both polar and nonpolar) character. Some amino acid residues cannot clearly be assigned to only one of these major categories and thus occur in different categories in different publications. Charged polar and hydrophobic nonpolar residues are of major importance for the characteristics of the protein surface and therefore will be described more detailed in the following. By way of illustration, fractions of the protein surface with charged or hydrophobic properties are highlighted in different colors in Fig. 3.

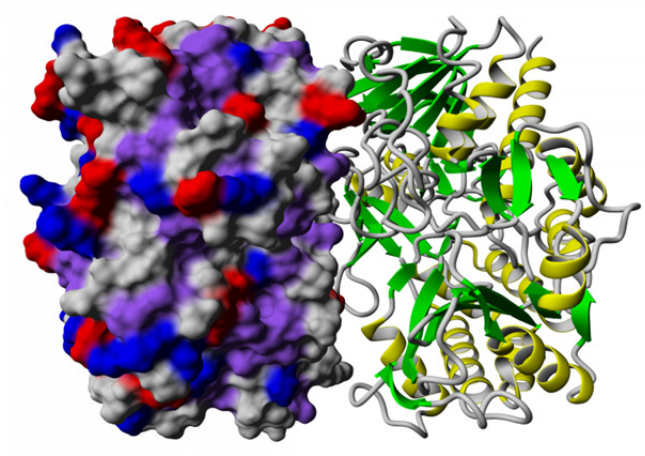

Fig. 3. Glucose oxidase (1GPE) in its quaternary structure, depicted in ribbon style. For the monomer on the left-hand side the molecular surface is depicted additionally. Molecular surface areas with a negative charge at pH 7.4 are colored blue, those with a positive charge at pH 7.4 are colored red. Hydrophobic molecular surface areas are colored purple. Molecular graphics were created using Yasara (www.yasara.org).

The charge of amino acid residues is strongly dependent on the pH of the surrounding aqueous solution and the pK_a value of the amino acid residues. For $pH>pK_a$ the titratable group is deprotonated. For $pH<pK_a$ the titratable group is protonated. For the acidic amino acids aspartic acid (Asp), glutamic acid (Glu), and histidine (His) this means that for $pH>pK_a$ they are negatively charged. For $pH<pK_a$ they are uncharged. For the basic amino acids arginine (Arg), cysteine (Cys), lysine (Lys), and tyrosine (Tyr) this means that for $pH>pK_a$ they are uncharged and for $pH<pK_a$ they are positively charged. The charge state of acidic and basic amino acids as function of pH is schematically shown for Glu ($pK_a \approx 4$) and Lys ($pK_a \approx 10.5$) in Fig. 4.

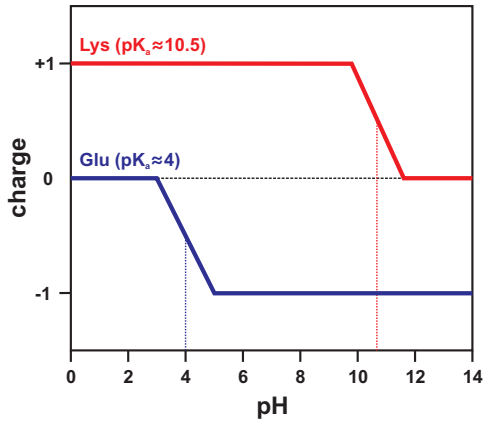

Fig. 4. Charge states of the acidic and basic amino acids glutamic acid (Glu) and lysine (Lys) as function of pH.

The net charge of a protein then is the sum of all individual charges of the amino acid residues at a specified pH value. It is possible to estimate it *in silico* using the amino acid sequence (Josuran, 2015; Putnam, 2013). However, it has to be considered that the protein in aqueous solution adopts its globular tertiary or quaternary structure and thus the net charge of the protein surface is of major relevance, as the characteristics of the protein surface are crucial for protein function as well as for protein-protein interactions. Thus, calculation of a protein's net charge based on its three-dimensional structure, e.g. on crystal structures obtained from the protein data bank (PDB) (Berman et al., 2000), is most relevant. One example for this purpose is H++ (Anandakrishnan et al., 2013). Furthermore, there are various methods to experimentally determine a protein's net charge in an aqueous solution with a specified pH value, for example by measuring the electrophoretic mobility which is related to the protein's zeta potential and net charge (Lehermayr et al., 2010; Winzor, 2004). Electrostatic forces between like-charged proteins are generally repulsive. Though, unlike-charged surface areas might attract each other. The pH value at which a protein's net charge is zero is called isoelectric point (pI) and is of great importance in biopharmaceutical industries as electrostatic repulsion between like protein molecules typically is at its minimum at the isoelectric point. Therefore, it is a rule of thumb to define the formulation's pH not too close to the isoelectric point as this increases the risk of

precipitation (Cleland et al., 1993).

Hydrophobic nonpolar amino acid residues are mainly found in the interior of a protein's native globular tertiary state as the contact of nonpolar residues with water is unfavorable. Thus, for minimization of the proportion of solvent exposed hydrophobic amino acid residues a hydrophobic core develops during protein folding. However, hydrophobic amino acids (Val, Ile, Leu, Met, Trp, Cyt) are also found at the protein surface as is shown in Fig. 3. Hydrophobic protein surface areas generate attractive protein-protein interactions as the fusion of these hydrophobic surface areas reduces the overall surface exposed hydrophobic area. Thus, protein surface hydrophobicity is of fundamental importance for protein aggregation propensity (Chennamsetty et al., 2009; Kumar et al., 2011; Zhang and Topp, 2012). Within the last decades a huge research effort has been spent on the development of experimental and *in silico* approaches for assessing protein surface hydrophobicity but it is still difficult to quantify and existing approaches suffer from some serious drawbacks. *In silico* methods apply hydrophobicity scales of single amino acids to quantify protein surface hydrophobicity. The outcome of the calculation is strongly dependent on the selected hydrophobicity scale (Biswas et al., 2003) and cannot consider pH dependency. Hydrophobicity sensitive dyes like Nile Red, 1-anilinonaphthalene-8-sulphonate (ANS) and bromophenol blue (BPB) (Bertsch et al., 2003; Cardamone and Puri, 1992; Sackett and Wolff, 1987) are widely used to experimentally estimate protein hydrophobicity (Hawe et al., 2008). These dyes are either hardly soluble in aqueous solution or charged. Thus, either organic solvents need to be added that might denature protein molecules or an interaction between the dye and charged protein surface areas instead of the hydrophobic ones cannot be excluded. A stalagmometric approach that could avoid these drawbacks will be presented in Section 3.3.

1.2.3 Protein flexibility

Protein structures that emerge from X-ray crystallography and NMR seem rigid and static. In reality they are highly flexible and in constant motion between different conformational states with similar energy. (Petsko and Ringe, 2004b; Teague, 2003) This conformational flexibility of proteins is required for biological function, ligand binding, metabolism, transport and much more (Carlson and McCammon, 2000; Frauenfelder et al., 1988; Teague, 2003). However, protein conformational flexibility could also be related to aggregation propensity. Sousa (Sousa, 1995) twenty years ago postulated that increasing conformational flexibility will increasingly favor amorphous precipitation over crystallization. Valerio et al. (Valerio et al., 2005) ten years ago extended this statement. Using molecular dynamics (MD) simulations they could show that protein flexibility is the main determinant of protein aggregation propensity. In the present thesis molecular dynamics simulations were used to calculate the root-mean-square-fluctuation

(RMSF) of C_α, C, and N backbone atoms of proteins in order to use protein flexibility as parameter to predict protein aggregation propensity (see Section 3.4).

1.3 Basics of protein phase behavior

1.3.1 Protein solubility

A protein will stay in solution up to a certain concentration (Asherie, 2004). This concentration describes the solubility or the solubility line, respectively. The solubility line separates the undersaturated and the supersaturated region. A transgression of the solubility line is the fundamental prerequisite for any phase transition and thus for protein aggregation. In the supersaturated region protein crystallization, protein precipitation, and other phase transitions can occur. (Talreja et al., 2010) However, for the development of these phase transitions supersaturation needs to be high enough (Luft and DeTitta, 2009). The solubility line describes an equilibrium state, i.e. crystals that evolve in the supersaturated region grow up to a certain size while supersaturation in the supernatant decreases until this corresponding supernatant is saturated. At this point crystal growth stops and crystal and supernatant composition remain constant. Thus, solubility lines can be determined by probing the relationship between a crystalline and the corresponding saturated solution. Typically, this is conducted by dissolving crystals in an undersaturated solution until saturation is reached. The protein concentration in the equilibrated supernatant determines the point on the solubility line for the given solution composition. (Asherie, 2004; Talreja et al., 2010) Using this method it takes days or weeks to determine one point on the solubility line and there are some difficulties involved e.g. finding the appropriate initial proportion of crystals to solution. Though, this method has been used to measure the solubility of more than 40 proteins. (Asherie, 2004) It is also possible to expose the crystals to a supersaturated solution. Equilibrium then is reached by crystal growth and possibly by formation of new crystals. (Asherie, 2004; Talreja et al., 2010) Regardless of starting with an undersaturated or a supersaturated solution the protein concentration in the equilibrated supernatant should give the same solubility value. Besides these methods that apply exposure of protein crystals to either under- or supersaturated solutions, microbatch experiments are also suitable to determine protein solubility lines. Here, the protein solution is either undersaturated, saturated, or supersaturated at the beginning of the experiment. If supersaturation is high enough, crystallization occurs and crystals will grow until equilibrium, i.e. saturation, is reached (Boistelle and Astier, 1988; Luft and DeTitta, 2009). In this work solubility lines were determined from the protein concentration in the supernatant of crystalline suspensions after 40 days.

1.3.2 Protein crystallization

As described above, protein crystals are of special interest for experimental determination of protein solubility lines. Moreover, protein crystals are of special interest in biopharmaceutical industries. Crystallization as purification step is thought to be highly effective as it is supposed to be highly selective (Judge et al., 1998). Crystallization for formulation purposes has yet been shown to be particularly well suited for delayed release formulations (Jen and Merkle, 2001). But, as mentioned above, crystals will not grow out of all supersaturated solutions. A certain energy or nucleation barrier ΔG* needs to be overcome as can be seen in Fig. 5. (Faber and Hobley, 2006)

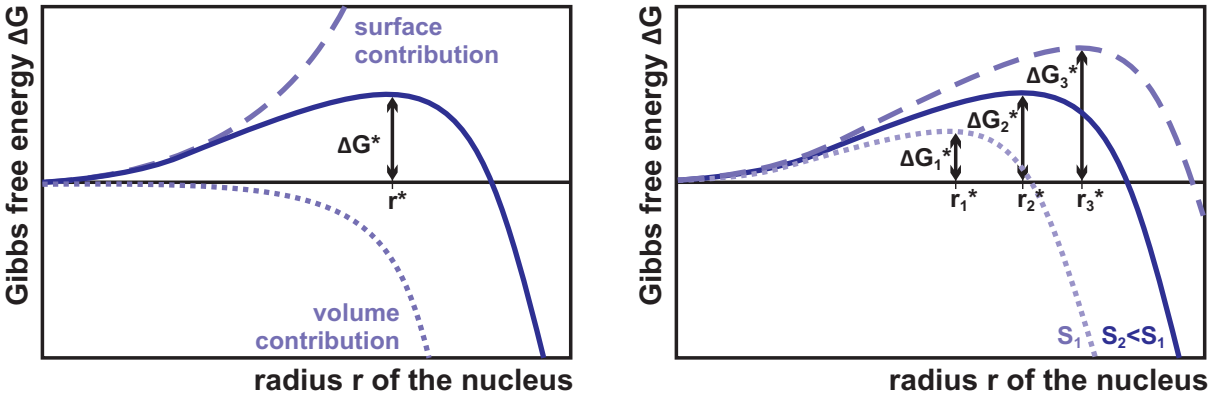

Fig. 5. Gibbs free energy upon formation of a spherical nucleus of the radius r. This figure was adapted from García-Ruiz (García-Ruiz, 2003).

The nucleation barrier ΔG* can be derived theoretically. The following description on nucleation of protein crystals is based on the classical nucleation theory that was developed 1926 for the condensation of a drop from its vapor by Volmer and Weber (Volmer and Weber, 1926). Its application to protein crystallization was pioneered by Feher and Kam (Feher and Kam, 1985). ΔG(r) describes the Gibbs free energy upon formation of a spherical nucleus of the radius r. This Gibbs free energy consists of a surface term (first summand in Eq. (1)) and a volume term (second summand in Eq. (1)) as illustrated in Fig. 5 and can according to García-Ruiz and Faber and Hobley (Faber and Hobley, 2006; García-Ruiz, 2003) be described by

$$\Delta G(r) = 4\pi r^2 \gamma - \frac{4}{3}\pi r^3 \frac{kT}{\Omega} lnS \qquad (1)$$

γ is the interfacial free energy between a crystal nucleus and the bulk solution, r the radius of the nucleus and kT/Ω lnS the free energy difference between a protein molecule in solution and incorporated into the crystal. Ω hereby describes the molar volume occupied by one protein molecule, S the solution supersaturation, k the Boltzmann constant and T the absolute

temperature. The radius at which Eq. (1) exhibits a maximum is the radius of the critical nucleus (r*) and thus is defined by

$$\left(\frac{\partial \Delta G(r)}{\partial r}\right)_{S,T} = 0 \qquad (2)$$

resulting in

$$r^* = \frac{2\Omega\gamma}{kTlnS} \qquad (3)$$

Substituting r* into Eq. (1) gives the nucleation barrier ΔG*. r* and ΔG* vary inversely with supersaturation, meaning the higher the supersaturation S the lower r* and ΔG* which is schematically illustrated in the right-hand section of Fig. 5.

The supersaturation S can be used to differentiate between the metastable and the labile zones in the generic c_{prot}/c_{precip}-phase diagram (see Fig. 6(c.)). The solubility line describes the saturated state of the protein solution, i.e. S=1. r* tends to infinity as S tends to 1 and nucleation is thermodynamically not possible. In the metastable zone supersaturation is not high enough, i.e. the volume term of ΔG(r) (second summand in Eq. (1)) does not exceed the contribution from the surface term (first summand in Eq. (1)). Consequently, no nucleation can proceed, but existing crystals can continue growing. On the borderline between metastable and labile zone the volume and the surface term counterbalance each other. In the labile zone supersaturation is high enough for nucleation to proceed and once the critical nucleus size r* is overcome, the crystal continues to grow. (Faber and Hobley, 2006) The nucleation rate J can be expressed as (García-Ruiz, 2003)

$$J = \kappa_0 \exp\left(-\frac{\Delta G^*}{kT}\right) \qquad (4)$$

The pre-exponential term κ_0 is difficult to derive theoretically. It is related to the kinetics of monomer attachment to the nucleus and it depends, for instance, on the viscosity and density of the solution. To give one example Vekilov et al. (Vekilov et al., 1996) expressed κ_0 as

$$\kappa_0 = \nu Z n \qquad (5)$$

where ν is the rate of attachment of monomers to the nucleus, Z is the Zeldovich factor and n is the number density of protein molecules in the solution.

García-Ruiz (García-Ruiz, 2003) and Faber and Hobley (Faber and Hobley, 2006) describe the

difference between crystallization and precipitation based on this nucleation theory. According to García-Ruiz precipitation occurs for very high supersaturations. Then the critical nucleus size r* becomes smaller than the smallest structural unit, i.e. the size of one protein monomer, and the nucleation barrier disappears. Faber and Hobley as well describe precipitation to occur in the absence of a nucleation barrier.

1.3.3 Protein phase states and phase diagrams

Information on protein phase behavior is important for protein purification and formulation process design. The importance is related both to the avoidance of undesired phase transitions during processing and to the application of phase transitions for purification purposes. Once the solubility line is exceeded, a phase transition (i.e. aggregation) can occur resulting in crystallization, precipitation, gelation, liquid-liquid phase separation or skin formation (Ahamed et al., 2007). A protein phase diagram is a graphical display of these possible phase transitions or phase or aggregate states of a protein. Hence, a protein phase diagram will provide information on the phase behavior of a protein, i.e. on whether crystallization, precipitation, gelation, liquid-liquid phase separation or skin formation occur under the given solution composition. Phase diagrams can be determined and visualized subject to various parameters, e.g. protein concentration (c_{prot}), precipitant concentration (c_{precip}), temperature (T) and pH value. Usually two of these parameters are varied with the other parameters held constant. Phase diagrams derived from thermodynamic approaches usually show protein concentration as a function of temperature (see for example (Asherie et al., 1996; Dumetz et al., 2008; Haas and Drenth, 1998)). One example therefore is shown in Fig. 6(a.). Attempts have been made to describe and visualize protein phase diagrams as a function of a protein interaction parameter such as the second osmotic virial coefficient (B_{22}) (Ahamed et al., 2007). One example therefore is shown in Fig. 6(b.). In correspondence to the typical physiological, laboratory, and industrial conditions, protein phase transitions are typically considered under constant temperature and pressure (Vekilov, 2012). Thus, protein phase diagrams then are illustrated as function of precipitant concentration (c_{precip}) as it is shown in Fig. 6(c.). Anyway, the general shape of these c_{prot}/T-, c_{prot}/B_{22}- and c_{prot}/c_{precip}-phase diagrams is very similar. There are three main phases or regions in a protein phase diagram: the one-phase region of the dilute solution, a two-phase region and the one-phase region of the solid crystal (Vekilov, 2012). The first one describes the undersaturated region of a protein solution where no phase transitions occur and is separated from the two-phase region by the solubility line. This region is indicated with a checked pattern in Fig. 6. Not much is known about the solidus line which separates the two-phase region from the pure solid crystal region, which is why the solidus line is usually drawn as horizontal line (Vekilov, 2012) as can be seen in Fig. 6(b.). The phase transitions of interest occur in the two-

phase region. These are crystallization, precipitation, gelation and liquid-liquid phase separation. No skin formation is considered in the two-phase region. Skin formation is thought to be a layer of denatured protein and the model phase diagrams only account for phase transitions with the protein in its native state. It is not known if all proteins can be described by these model phase diagrams, i.e. it is not known if all proteins adopt all of the phase states which are described in the generic phase diagrams. There is also uncertainty about the definition of phase transitions and about the classification of different phase states into the different sections in the phase diagram. In some publications only crystallization and liquid-liquid phase separation are defined as true phase transitions. According to Asherie (Asherie, 2004), the formation of amorphous clusters as occurring for example during precipitation, is not a phase transition as no macroscopic phase forms and thus no aggregation phase boundary exists. Faber and Hobley (Faber and Hobley, 2006) as well refuse to consider unordered amorphous precipitate as true solid phase. The term precipitate generally underlies many different definitions. Muschol and Rosenberger (Muschol and Rosenberger, 1997) and Dumetz et al. (Dumetz et al., 2008) for example classified two different forms of precipitates. The formation of cloudy precipitates is assigned to liquid-liquid phase separation whereas the formation of amorphous precipitate is assigned to an incomplete gelation which occurs only for high supersaturations. The difference between crystallization and precipitation based on the nucleation theory was described in Section 1.3.2.

1.3.4 Experimental methods to determine protein phase diagrams

Protein phase diagrams can be determined by different experimental approaches. The most popular ones are via vapor diffusion or microbatch experiments. A schematic depiction of these two processes is shown in Fig. 6(d.). Vapor diffusion experiments are conducted by using an undersaturated protein solution in the form of a hanging or a sitting drop. The drop equilibrates against a reservoir that contains a higher concentration of precipitant than the protein solution (Chayen and Saridakis, 2008; Luft and DeTitta, 2009). During equilibration the drop volume decreases, leading to an increase of protein and precipitant concentration in the drop as illustrated in Fig. 6(d.). If the supersaturation in the drop is high enough phase transitions like crystallization can occur (Luft and DeTitta, 2009). In vapor diffusion experiments conditions in the protein solution change throughout the equilibration process (Chayen and Saridakis, 2008). Therefore protein and precipitant concentrations at a phase transition cannot be determined exactly. As opposed to vapor diffusion experiments, the solution conditions will remain constant when performing microbatch experiments (Chayen and Saridakis, 2008). Here, the solution is either undersaturated, saturated, or supersaturated at the beginning of the experiment. If supersaturation is high enough, crystallization occurs and crystals will grow until equilibrium,

i.e. saturation, is reached (Boistelle and Astier, 1988; Luft and DeTitta, 2009). Microbatch experiments give a more accurate description of solution compositions resulting in a phase transition, as the overall conditions in the protein solution do not change throughout the experiment. They also enable subsequent determination of protein solubility by measurement of the protein concentration in the supernatant of crystalline suspensions.

In the present work binary and ternary phase diagrams were determined via microbatch experiments. Binary phase diagrams result from aqueous solutions with protein and precipitant as solution components. Ternary phase diagrams additionally contained glycerol, glycine or polyethylene glycol (PEG) 1000 as third solution component. The experimental setup for automated high throughput generation of protein phase diagrams and the resulting binary phase diagrams are described in detail in Section 3.1. Ternary phase diagrams are presented in Section 3.2.

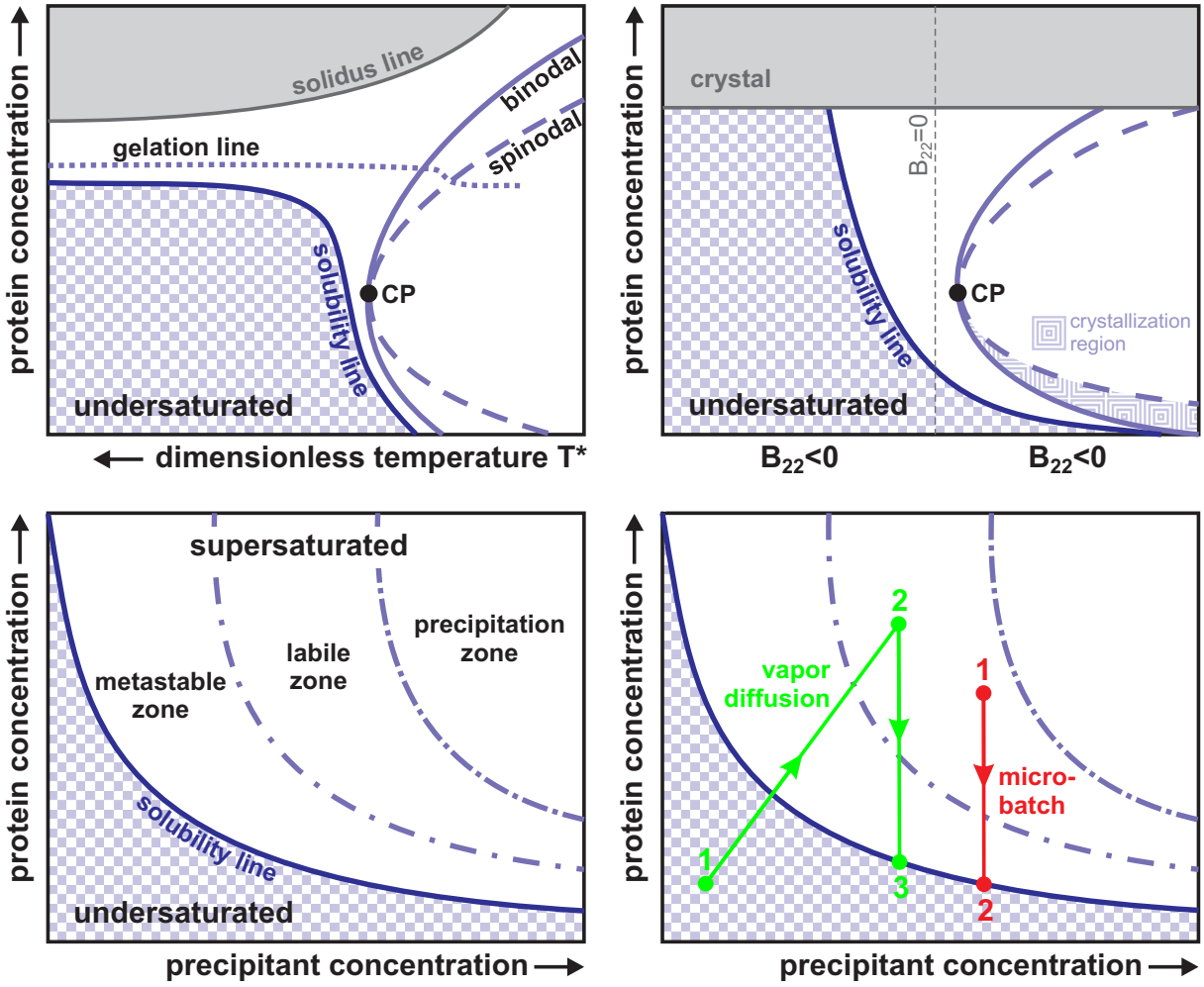

Fig. 6. Protein phase diagrams visualized subject to various parameters. The one-phase region of the dilute solution is indicated with a checked pattern. The one-phase region of the solid crystal is colored grey. The two-phase region is not colored, but different sub-regions within are separated by binodal, spinodal, and gelation lines. Figures were adapted from (Dumetz et al., 2008) (a.), (Ahamed et al., 2007) (b.), (Asherie, 2004) (c.), (Chayen, 2004) (d.).

1.4 Investigation of protein conformation during protein aggregation

Protein conformation can be studied using Fourier transform infrared (FT-IR) spectroscopy. For a detailed description on the principle of FT-IR spectroscopy in general and applied to proteins, on sampling techniques (transmission, attenuated total reflectance (ATR)), and data processing see for example (Barth, 2007). The main advantage of FT-IR spectroscopy with respect to protein studies is that proteins can be studied in any physical state, i.e. in aqueous solution, frozen, dried, or even as an insoluble aggregate. Thus, FT-IR spectroscopy is best suited for investigation of protein conformation in soluble, crystalline, precipitated, and other phase states.

In a FT-IR spectrum, there are nine characteristic absorption bands (amide A, amide B, amide I-VII) that arise from the peptide linkages, i.e. that are associated with the protein backbone (Barth, 2007; Dong et al., 1995). The most important ones are the amide I and the amide II absorption bands as they are correlated to the secondary structure of the protein backbone while hardly affected by the nature of amino acid side chains. As the correlation between the amide I absorption band and secondary structure is more straightforward than for the amide II absorption band, studies on protein secondary structure and conformational changes as well as on aggregation are almost exclusively conducted using the amide I band. (Barth, 2007) The amide I band itself arises as superposition from overlapping secondary structure component bands such as α–helix, β-sheet, β-turns, and random coil structures as illustrated in Fig. 7.

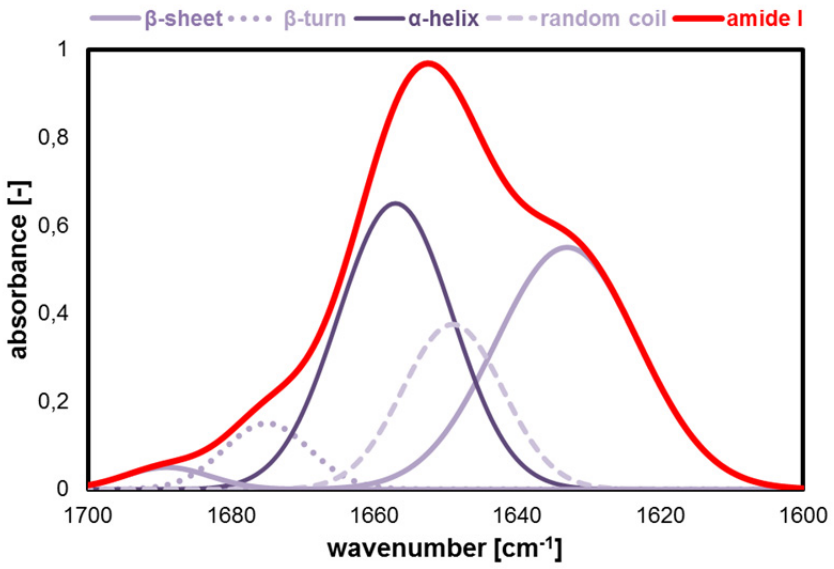

Fig. 7. Illustration of the amide I band as superposition from overlapping secondary structure component bands.

These secondary structure component bands are in close proximity to one another and instrumentally unresolvable. Mathematical methods are necessary to resolve the individual component bands. There are several methods of which Fourier self-deconvolution (Byler and Susi, 1986) and second derivative analysis (Susi and Byler, 1983) are the most popular ones. For different proteins the mean band position for different secondary structure component bands varies between extremes. In aqueous solution bands between 1658 and 1660 cm^{-1} could be assigned to α-helices. Bands between 1623 and 1643 cm^{-1} are assigned to β–sheets by different authors. Bands located between 1662 and 1689 cm^{-1} are assigned to β–turn structures. The characteristic band for disordered or random coil structures could be assigned to a band located between 1640 and 1651 cm^{-1}. (Dong et al., 1990; Dong et al., 1995; Jackson and Mantsch, 1995) Additionally Jackson and Mantsch correlated bands between 1610 and 1628 cm^{-1} and between 1675 and 1695 cm^{-1} to antiparallel β–sheets and aggregated strands, although there is no experimental evidence that it is possible for globular proteins to distinguish between the bands of parallel and antiparallel β-sheets (Khurana and Fink, 2000; Susi and Byler, 1987). Jackson and Mantsch further assigned bands between 1660 and 1670 cm^{-1} to 3_{10}-helix structures, although this structure is rarely found in proteins (Kong and Yu, 2007). This structure as well as other rare or minor structures might interfere with the described band assignments. Thus, there is no simple correlation between infrared spectra and secondary structural components and caution has to be exercised in the interpretation of it. The correlation between amide I band positions and secondary structure is concisely shown in Tab. 1.

Tab. 1. Correlation between amid I band positions and secondary structure elements.

Secondary structure element	Mean band position in aqueous solution [cm^{-1}]	Extremes of band position in aqueous solution [cm^{-1}]
Aggregated strands		1610-1628[2]
β-sheet	1624[3] 1627[3] 1632[3] 1635[1] 1638[3] 1642[3]	1623-1643[3] 1625-1640[2]
Disordered	1648[1] 1650[3]	1640-1648[2] 1649-1651[3]
α-helix	1656[3] 1658[1]	1648-1660[2] 1654-1658[3]
β-turns	1666[3] 1668[1] 1672[3] 1677[1] 1680[3] 1688[3]	1665-1689[3]
β-sheet/antiparallel β-sheet/ aggregated strands	1689[1]	1675-1695[2]

[1](Dong et al., 1995)
[2](Jackson and Mantsch, 1995)
[3](Dong et al., 1990)

FT-IR spectroscopy has been widely used for protein studies such as for investigation of protein conformation, to follow protein aggregation, and to study protein flexibility to name just a few. Selected examples of the application of FT-IR spectroscopy for protein studies will be presented in the following. The most interesting application of FT-IR spectroscopy for this work is the investigation of protein aggregation. Dong et al. (Dong et al., 1995) reviewed studies on thermally induced (cold- and heat-induced) protein aggregation and found that formation of β–sheets, regardless of the initial secondary structural composition of the native proteins, often coincides with the appearance of unordered secondary structure during aggregation. Lyophilization-induced protein aggregation (drying, dehydration) as well produced a strong band near 1620 cm^{-1} and an associated weak band near 1685 cm^{-1} which are indicative of intermolecular β–sheets. Thus, these bands can be used to monitor and quantify aggregation. Fink (Fink, 1998) could show that salting-out aggregates (precipitates) retained native conformation. Matheus et al. (Matheus et al., 2009) extended the observations of Fink, as they observed that ammonium sulfate, sodium citrate and PEG 4000 precipitated and re-dissolved IgG1 molecules retained their native secondary structure as well as binding activity and stability. They could additionally show that in the precipitated suspension secondary structure was maintained using PEG 4000. Only using ethanol they detected unacceptable changes in the secondary structure of IgG1. Their results could support that intermediate precipitation is a promising technique to manufacture liquid, highly concentrated IgG1 antibody formulations without causing risks for the patient or influencing product activity.

In this work, FT-IR spectroscopy was as well applied to study protein conformation during aggregation and to identify non-native aggregates. Furthermore, it was used to evaluate conformational stability. Results are shown in Section 3.2.

2. Research proposal

Protein aggregation can occur through different mechanisms at any step of a biopharmaceutical production process, e.g. initiated by changes in pH, ionic strength, protein concentration, by shearing forces or by interactions with facility equipment (Cromwell et al., 2006; Philo and Arakawa, 2009). On the one hand it is an acknowledged process step in biopharmaceutical industries either for purification or formulation purposes (Scopes, 1994a). On the other hand accidental and non-native protein aggregation need to be inhibited to ensure patient safety and to maintain product activity. Up to now, little is known about the solution conditions and protein properties that are of major importance for protein aggregation propensity and phase behavior. Interestingly, original classifications of proteins depended on their aggregation propensity as a function of salt concentration. Globulins were defined as proteins with high aggregation propensity in solutions with a low salt concentration whereas albumins showed low aggregation propensity in these solutions. (Scopes, 1994a) However, extremely high salt concentrations and extremes of pH in most cases provoke aggregation or even denaturation regardless of the protein species. Differing aggregation propensities of different protein species are supposed to be dictated by protein surface hydrophobicity especially at conditions where electrostatic repulsion between like proteins is low, i.e. at pH values near the isoelectric point and at low to medium salt concentrations. To expand these general observations, the present work focusses on three subsections dealing with characterization, manipulation, and prediction of protein aggregation. The first subsection will serve to characterize protein aggregation, meaning to assess aggregation propensity of model proteins at different pH values and using different precipitants, to identify preferred phase transitions, e.g. crystallization, precipitation or gelation, and to investigate if surface characteristics like surface hydrophobicity of the model proteins can be related to either aggregation propensity or preferred phase transitions. Aggregation will be induced by addition of sodium chloride, ammonium sulfate, polyethylene glycol (PEG) 300 or polyethylene glycol (PEG) 1000 at pH 3, 5, 7, or 9. The impact of the precipitants on the model proteins will be illustrated in binary phase diagrams. Binary as they will result from protein and precipitant as solution components. These binary protein phase diagrams reflect aggregation propensity as well as the type of phase transition. Data hereof can be used to determine solution properties or suitable precipitants for realization of aggregation processes for purification or formulation. The binary phase diagrams can also be used to readout solution compositions to inhibit protein aggregation. Additionally, it will be examined if a correlation between surface hydrophobicity of the model proteins and aggregation propensity or preferred type of phase transitions exists. Therefore, a novel non-invasive high-throughput stalagmometric approach for determination of protein surface hydrophobicity based on surface tension will be developed, as

protein surface hydrophobicity up to now is experimentally hard to assess and current methods suffer from serious drawbacks. The second subsection will deal with manipulation of protein aggregation propensity as in some cases solution composition is not freely selectable, e.g. in cases where specified pH values or ionic strengths result due to virus inactivation or chromatographic process steps. Then manipulation of protein aggregation can increase processability of the protein solutions especially if aggregation without manipulation would be irreversible or non-native. One chance to manipulate protein aggregation at fixed solution composition is using stabilizing additives. Some uncharged sugars, polyols, and free amino acids are known to stabilize the proteins' native state and to reduce aggregation propensity (Timasheff, 2002; Yancey, 2001). PEG of low molecular weight is thought to reduce attractive protein-protein interactions resulting in a reduced aggregation propensity as well. Systems with differing phase states will be chosen from the variety of binary phase diagrams. Ternary phase diagrams with the same protein and precipitant concentrations will be generated under addition of constant concentrations of additives and compared to the binary phase diagrams. This comparison will reveal the potential of different additives to manipulate, i.e. to prevent or to selectively control, protein aggregation. In this subsection about manipulation of protein aggregation a special focus will additionally be placed on the analysis of protein aggregation regarding native and non-native aggregation processes. Fourier transform infrared (FT-IR) spectroscopy will be used to investigate the secondary structure of the proteins during aggregation. Non-native aggregation forms can be identified and the potential of additives to prevent non-native aggregation, i.e. to stabilize the proteins native conformational state, will be tested. The last section of this publication will present an attempt to predict aggregation propensity. Protein conformational flexibility as well as surface hydrophobicity and surface charge are supposed to be the main mechanistic determinants of aggregation propensity (Lauer et al., 2012; Valerio et al., 2005) and protein flexibility is even thought to decide on formation of precipitation or crystallization (Sousa, 1995). Physicochemical (surface hydrophobicity and surface charge) and dynamic (conformational flexibility) properties of model proteins in aqueous solution at pH 3, 5, 7, and 9 will be correlated to the phase behavior at pH 3, 5, 7, and 9 with sodium chloride and ammonium sulfate as precipitants. Surface hydrophobicity of the model proteins will be determined using the novel non-invasive high-throughput stalagmometric approach. Electrostatic properties of the model proteins will be experimentally determined using the zeta potential. Dynamic properties, i.e. conformational flexibility, of the model proteins will be assessed using an *in silico* approach by means of molecular dynamics simulations that capture the mobility of the protein backbone. These three parameters will be correlated to aggregation propensity to extend and confirm the above assumptions.

In summary, this publication aims to deepen the basic understanding of protein aggregation mechanisms and protein aggregation propensity. Approaches will be presented to detect

solution conditions and protein characteristics aiming in an increased aggregation propensity. Strategies to manipulate aggregation mechanisms, i.e. to prevent non-native protein aggregation or to selectively control the kind of phase transition, will be probed using model proteins. The correlation between protein phase behavior and physicochemical surface properties or dynamic properties of the model proteins might give a powerful tool for prediction of protein aggregation propensity.

In any of the experimental methods sample volume will be minimized to save product consumption and high throughput methods, whenever possible, will be used to reduce time consumption. As the development and validation of these studies are though protein consuming, model proteins will be used. The developed experimental methods can easily be transferred to proteins of interest, such as monoclonal antibodies.

3. Publications and Manuscripts

3.1 Determination of protein phase diagrams by microbatch experiments: Exploring the influence of precipitants and pH

Lara Galm, Kai Baumgartner, Juliane Nötzold, Heike Sigloch, Josefine Morgenstern, Kristina Schleining, Susanna Suhm, Stefan A. Oelmeier, Jürgen Hubbuch

This publication presents the systematic generation and extensive study of protein phase diagrams of model proteins. An experimental method was set up for automated generation of protein phase diagrams in microbatch experiments with minimized protein and time consumption. Phase diagrams were generated for four proteins, namely lysozyme from chicken egg white, human lysozyme, glucose isomerase, and glucose oxidase. These proteins were selected to cover a broad range in size and isoelectric points. Their phase behavior was investigated at four different pH values – 3, 5, 7, 9 – using four different precipitants – sodium chloride, ammonium sulfate, PEG 300 and 1000.

My contribution to this work was the calibration of the liquid handling for PEG solutions of different PEG molecular weights and for different PEG concentrations and the generation and evaluation of overall 32 phase diagrams for lysozyme from chicken egg white and human lysozyme. I could show that for accurate and reproducible liquid handling particular liquid classes need to be used dependent on PEG molecular weight. Evaluation of the protein phase diagrams revealed that despite the high structural similarity of lysozyme from chicken egg white and human lysozyme they develop very different phase behavior. The most common phase transition of lysozyme from chicken egg white was crystallization whereas human lysozyme mainly precipitated. This finding could be explained by differing protein surface properties arising due to the dissimilar primary structure of the two proteins. This explanation as well gave rise to the assumption of an extraordinary importance of hydrophobic forces between proteins for protein phase behavior.

published in: ***International Journal of Pharmaceutics***. 2015; 479(1): 28-40

3.2 Manipulation of lysozyme phase behavior by additives as function of conformational stability

Lara Galm, Josefine Morgenstern, Jürgen Hubbuch

In this work the potency of different additives, namely glycerol, PEG 1000, and glycine, to prevent protein aggregation and selectively manipulate protein phase behavior, solubility, and crystal size and morphology was investigated. Lysozyme from chicken egg white was used as a model protein. Furthermore, this paper aims to find a link between the pH

dependent mode of action of additives and the conformational stability of lysozyme. To the best of our knowledge up to now there are no investigations on the pH dependent mode of action of additives as a function of conformational stability of proteins. The impact of the additives on conformational stability and their potency to stabilize or destabilize the protein's native state was evaluated using Fourier transform infrared spectroscopy. The impact of the additives on lysozyme phase behavior was evaluated by comparison of binary and ternary phase diagrams, consisting of lysozyme and sodium chloride as solution components or of lysozyme, sodium chloride, and the respective additive as solution components. Phase diagrams were generated according to Section 3.1. The microbatch experiments allowed to experimentally determine lysozyme solubility in cases where crystallization occurred. Again, comparison between the binary and ternary systems gives the additive impact on lysozyme solubility lines. The impact of glycerol, PEG 1000, and glycine on lysozyme crystal size and morphology was evaluated visually based on microscopic images of the microbatch experiments.

submitted to: ***International Journal of Pharmaceutics***.

3.3 Non-invasive high throughput approach for protein hydrophobicity determination based on surface tension

Lara Galm, Sven Amrhein, Katharina Christin Bauer, Jürgen Hubbuch

In this paper a high resolution stalagmometric method for surface tension determination was established on a liquid handling station, which can cope with accuracy as well as high throughput requirements. Surface tensions could be derived with low sample consumption and high reproducibility within a reasonable time. This stalagmometric method was applied as a novel non-invasive approach to determine protein hydrophobicity on base of the proteins' surface tension increments. Based on these surface tension increments a pH dependent hydrophobicity ranking of lysozyme from chicken egg white, human lysozyme, BSA, and α-lactalbumin was developed.

My contribution to this work was the implementation and optimization of the stalagmometric method for proteins and its application for lysozyme from chicken egg white and human lysozyme. Application of the stalagmometric method for proteins and comparison to literature values and a spectrophotometric reference method proved the applicability of the stalagmometric method for determination of protein surface hydrophobicity. Furthermore, I realized a dye based spectrophotometric method for determination of protein surface hydrophobicity in aqueous solution. I validated this spectrophotometric method using BSA and further on applied it for lysozyme from chicken egg white, human lysozyme, and α-lactalbumin. During realization of the spectrophotometric method I could show that the dye substance bromophenol blue

sodium salt (BPB Na) can be submitted in sufficient concentration in aqueous solution, i.e. without addition of ethanol or organic solvents. Validation of the spectrophotometric method showed that proteins can produce significant shifts of the absorption maximum of BPB Na for pH 5, 7, and 9, which could be used for evaluation of protein surface hydrophobicity. The extent of the shift could be shown to be pH dependent and to be extremely protein specific. This finding as well revealed the limits of the spectrophotometric method. Limitations of the spectrophotometric method could inter alia be related to protein size. Furthermore, a hydrophilic character is not as easy ascertainable as using the stalagmometric method.

accepted by: ***Biotechnology and bioengineering.*** doi: 10.1002/bit.25677

3.4 Predictive power of protein surface characteristics and conformational flexibility for protein aggregation propensity

Lara Galm, Sven Amrhein, Jürgen Hubbuch

This manuscript aims to evaluate the predictive power of conformational flexibility as well as electrostatic and hydrophobic properties of proteins for protein aggregation propensity. Conformational flexibility was determined by an *in silico* approach, whereas protein surface properties were determined experimentally. Conformational flexibility was assessed by analysis of the root-mean-square-fluctuation (RMSF) of the C_α, C, and N backbone atoms of the proteins using molecular dynamics (MD) simulations. To catch electrostatic surface characteristics, the zeta potential was determined by measuring the electrophoretic mobility. The hydrophobic character was determined as described in Section 3.3. To correlate conformational flexibility and protein surface characteristics to protein aggregation propensity protein phase states were investigated as function of pH, protein concentration, and as function of sodium chloride and ammonium sulfate concentration. α-lactalbumin, human lysozyme, and lysozyme from chicken egg white were used as model proteins.

This work was divided into an experimental and an *in silico* part. I conducted the experimental work, meaning the experimental determination of protein surface characteristics, generation of protein phase diagrams and determination of aggregation lines. The investigation based solely on experimental data already exhibited that aggregation propensity of different proteins in salt-free solutions correlates with their hydrophobic character and that electrostatic surface properties are of minor importance. Protein surface hydrophobicity –determined in salt-free solutions- even proved to be applicable for prediction of aggregation propensity under addition of ammonium sulfate. Furthermore, protein surface hydrophobicity could be shown to be related to the preferred phase state of the investigated proteins. Proteins with a hydrophobic character

tended to precipitate whereas lysozyme from chicken egg white, a protein with hydrophilic character, easily crystallized.

in preparation

3.1 Determination of protein phase diagrams by microbatch experiments: Exploring the influence of precipitants and pH

Kai Baumgartner[1], Lara Galm[1], Juliane Nötzold[1], Heike Sigloch, Josefine Morgenstern, Kristina Schleining, Susanna Suhm, Stefan A. Oelmeier, Jürgen Hubbuch*

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

** Corresponding author. Tel.: +49 721 608 47526; fax: +49 721 608 46240. E-mail address: juergen.hubbuch@kit.edu.*

[1] Contributed equally.

published in: ***International Journal of Pharmaceutics***. 2015; 479(1): 28-40

Abstract

Knowledge of protein phase behavior is essential for downstream process design in the biopharmaceutical industry. Proteins can either be soluble, crystalline or precipitated. Additionally liquid-liquid phase separation, gelation and skin formation can occur. A method to generate phase diagrams in high throughput on an automated liquid handling station in microbatch scale was developed. For lysozyme from chicken egg white, human lysozyme, glucose oxidase and glucose isomerase phase diagrams were generated at four different pH values – pH 3, 5, 7 and 9. Sodium chloride, ammonium sulfate, polyethylene glycol 300 and polyethylene glycol 1000 were used as precipitants. Crystallizing conditions could be found for lysozyme from chicken egg white using sodium chloride, for human lysozyme using sodium chloride or ammonium sulfate and glucose isomerase using ammonium sulfate. PEG caused destabilization of human lysozyme and glucose oxidase solutions or a balance of stabilizing and destabilizing effects for glucose isomerase near the isoelectric point. This work presents a systematic generation and extensive study of phase diagrams of proteins. Thus, it adds to the general understanding of protein behavior in liquid formulation and presents a convenient methodology applicable to any protein solution.

Keywords: *Protein phase diagram, High throughput, Automated imaging, Protein phase behavior, Microbatch crystallization*

1. Introduction

Information on protein phase behavior is important for protein purification and formulation process design. The importance is related both to the avoidance of undesired phase transitions during processing and to the application of phase transitions for purification purposes. Further, protein phase behavior is important where structure determination is conducted via diffraction studies.

A protein phase diagram is a graphical display of the possible phase states of a protein. Hence, a protein phase diagram will provide information on whether crystallization, precipitation, liquid-liquid phase separation or gelation occur under the given conditions (Ahamed et al., 2007; Asherie, 2004). It is generally accepted, that the solubility line divides an undersaturated zone where the protein is soluble, from a supersaturated zone where the protein is potentially either crystalline or precipitated (Asherie, 2004; Faber and Hobley, 2006). Phase diagrams can be

determined subject to various parameters, e.g. protein concentration, precipitant concentration, temperature and pH. Different precipitants have various influences on protein phase behavior (Dumetz et al., 2009; McPherson, 2001), pH-shifts can easily induce supersaturation (Faber and Hobley, 2006). Temperature as well has a significant but unpredictable impact on the phase behavior as could be shown by Lin et al. (Lin et al., 2008) for lysozyme, ribonuclease A, ribonuclease S, trypsin, concanavalin A, chymotrypsinogen A, papain, catalase and proteinase K. Commonly, protein phase behavior is visualized as a function of protein and precipitant concentration with all other parameters held constant (Ahamed et al., 2007; Asherie, 2004). The most popular methods to generate protein phase diagrams are via microbatch or vapor diffusion experiments. The latter is conducted by using an undersaturated protein solution in the form of a hanging or a sitting drop. The drop equilibrates against a reservoir that contains a higher concentration of precipitant than the protein solution (Chayen and Saridakis, 2008; Luft and DeTitta, 2009). During equilibration the drop volume decreases, leading to an increase of protein and precipitant concentration in the drop. If the supersaturation in the drop is high enough phase transitions like crystallization can occur (Luft and DeTitta, 2009). In phase diagrams generated by vapor diffusion experiments, the initial protein concentration is plotted against the initial precipitant concentration (Lin et al., 2008). However, the conditions in the protein solution change throughout the equilibration process (Chayen and Saridakis, 2008). Therefore protein and precipitant concentrations at a phase transition cannot be determined exactly. As opposed to vapor diffusion experiments, the solution conditions will remain constant when performing microbatch experiments (Chayen and Saridakis, 2008). Here, the solution is either undersaturated, saturated or supersaturated at the beginning of the experiment. If supersaturation is high enough, crystallization occurs and crystals will grow until the liquid-solid equilibrium, i.e. saturation, is reached (Boistelle and Astier, 1988; Luft and DeTitta, 2009). Hence, microbatch experiments not only give a more accurate description of phase states with known protein and precipitant concentration. They also enable determination of solubility by measurement of the protein concentration in the supernatant. Despite these benefits microbatch data can hardly be found in literature.

In addition, crystallization conditions are found mainly by trial-and-error studies using screening methods (McPherson, 2001; McPherson and Cudney, 2006) with sparse-matrix screens being the most popular ones (Chayen and Saridakis, 2008). Until now a systematic approach to map phase diagrams is still missing.

In the present work an experimental method was set up for automated generation of protein phase diagrams with minimized protein consumption. Microbatch experiments were chosen due to the above mentioned advantages. They were conducted in high throughput mode using an automated liquid handling station. Phase diagrams were generated for the four proteins lysozyme from chicken egg white, human lysozyme, glucose isomerase and glucose oxidase.

These proteins were selected to cover a broad range in size and isoelectric points. Their phase behavior was investigated at four different pH values – 3, 5, 7, 9 – using four different precipitants – sodium chloride, ammonium sulfate, polyethylene glycol (PEG) 300 and 1000. These precipitants were chosen based on findings in literature (Kalisz et al., 1990; Lu et al., 2003; McPherson, 2001; McPherson and Cudney, 2006; Sleutel et al., 2009).

The presented work thus demonstrates the use of high throughput experimentation for the generation of phase diagrams. The methodology can be applied to any protein solution and was used herein to generate phase diagrams of model proteins to an unmatched extent.

2. Materials and methods

2.1. Materials

The used buffer substances were citric acid (Merck, Darmstadt, Germany) and sodium citrate (Sigma-Aldrich, St. Louis, MO, USA) for pH 3, sodium acetate (Sigma-Aldrich, St. Louis, MO, USA) and acetic acid (Merck, Darmstadt, Germany) for pH 5, MOPSO (AppliChem, Darmstadt, Germany) for pH 7 and Bis-Tris propane (Molekula, Dorset, UK) for pH 9. Sodium chloride was obtained from Merck (Darmstadt, Germany), ammonium sulfate was from VWR (Radnor, PA, USA) and PEG 300 as well as PEG 1000 were purchased from Sigma-Aldrich (St. Louis, MO, USA). Hydrochloric acid and sodium hydroxide for pH adjustment were obtained from Merck (Darmstadt, Germany). pH adjustment was performed using a five-point calibrated pH-meter (HI-3220, Hanna Instruments, Woonsocket, RI, USA) equipped with a SenTix® 62 pH electrode (Xylem Inc., White Plains, NY, USA) or an InLab® Semi-Micro pH electrode (Mettler Toledo, Greifensee, Switzerland) dependent on the application. All buffers were filtered through 0.2 µm cellulose acetate filters, precipitant solutions through 0.45 µm cellulose acetate filters (Sartorius, Goettingen, Germany).

Lysozyme from chicken egg white (PDB 1LYZ, pI 11.4) (HR7-110) and glucose isomerase from Streptomyces rubiginosus (PDB 3KBS, pI 4.78) (HR7-100) were purchased from Hampton Research (Aliso Viejo, CA, USA). Human lysozyme recombinantly expressed in rice (PDB 1LZ1, pI 11.0) (L1167) and glucose oxidase from Aspergillus niger (PDB 1CF3, pI 4.55) (49180) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All isoelectric points were calculated using H++ 3.0 (Anandakrishnan et al., 2012) for the respective PDB code with protonation at pH 7 and the following parameters: salinity 0.15 M, internal dielectric constant 10, external dielectric constant 80.

Protein solutions were filtered through 0.2 µm syringe filters with PTFE membranes (VWR, Radnor, PA, USA). Size exclusion chromatography was conducted using a HiTrap Desalting Column (GE Healthcare, Uppsala, Sweden) on an AEKTAprime™ plus system (GE Healthcare, Uppsala, Sweden). A subsequent protein concentration step was performed using Vivaspin

centrifugal concentrators (Sartorius, Goettingen, Germany) with PES membranes and molecular weight cutoffs of 3 kDa for lysozyme from chicken egg white and human lysozyme, 30 kDa for glucose oxidase and glucose isomerase.

Protein phase diagrams were prepared on MRC Under Oil 96 Well Crystallization Plates (Swissci, Neuheim, Switzerland) in microbatch experiments with a Freedom EVO® 100 (Tecan, Maennedorf, Switzerland) automated liquid handling station. Calibration of pipetting for liquid handling of buffers, precipitant solutions and protein solutions were generated using a WXTS205DU analytical balance (Mettler-Toledo, Greifensee, Switzerland). The MRC Under Oil 96 Crystallization Plates were covered with HDclear™ sealing tape (ShurTech Brands, Avon, OH, USA) to prevent evaporation. A Rock Imager 54 (Formulatrix, Waltham, MA, USA) was used as an automated imaging system for protein crystallization. Protein concentration measurements were conducted using a NanoDrop2000c UV-Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

2.2. Methods

2.2.1. Preparation of stock solutions

To set up the buffers and precipitant stock solutions containing sodium chloride, ammonium sulfate, PEG 300 or PEG 1000, all substances were weighed in and dissolved in ultrapure water to 90 % of the final buffer volume. pH was adjusted with the appropriate titrant with an accuracy of ±0.05 pH units. After pH adjustment the buffers were brought to their final volume using ultrapure water. All buffers were filtered through 0.2 µm cellulose acetate filters. Precipitant stock solutions were filtered through 0.45 µm cellulose acetate filters. Buffers and precipitant stock solutions were used at the earliest one day after preparation and after repeated pH verification. Buffer capacity was 100 mM for all buffers. Precipitant stock solutions contained 2.5 or 5 M sodium chloride, 1.5 or 3 M ammonium sulfate and 30 %(w/V) PEG 300 or PEG 1000 additionally to the corresponding buffer substances.

To set up the protein stock solutions, protein was weighed in and dissolved in the appropriate buffer yielding a concentration of 80 mg/mL for lysozyme from chicken egg white, human lysozyme and glucose oxidase. The crystal suspension of glucose isomerase was diluted with appropriate buffer to redissolve the suspension. The protein solutions then were filtered through 0.2 µm syringe filters to remove particulates and desalted using size exclusion chromatography. Protein concentration was adjusted to 43.5 ± 1 mg/mL via centrifugal concentrators. A volume of 1 mL of protein stock solution was required for generation of one phase diagram.

2.2.2. pH stability in high concentrated salt solutions

The stability of pH in salt precipitant dilution series was evaluated for the different buffers at a scale of 1 mL. Sodium chloride concentration was varied between 0 and 5 M and ammonium

sulfate concentration between 0 and 3 M using the respective buffers including intermediate (2.5 M sodium chloride or 1.5 M ammonium sulfate) and high salt concentration buffers. pH values for the different dilution steps were measured using a five point calibrated pH meter.

2.2.3. Accuracy of automated liquid handling

Calibration of liquid handling is essential for reproducible generation of protein phase diagrams. The method to determine so called liquid classes, including important variables such as plunger volume, aspiration and dispense speed, was described earlier by Oelmeier et al. (Oelmeier et al., 2011). The challenge of pipetting different solutions arises due to different viscosities and densities resulting in possibly large errors in automated pipetting. Liquid classes were created for all buffers, precipitant solutions and protein solutions. The calibrated volume range was 7-200 µL.

To create liquid classes the following pipetting parameters were individually adjusted: the aspiration, the dispensing, and the breakoff speeds. Additionally three different air gaps in the pipetting tips which influence the pipetting behavior were adapted. The factors and offsets of regression lines, with which the otherwise incorrect pipetted volumes were corrected automatically, were adjusted individually.

2.2.4. Generation of phase diagrams

Protein phase diagrams were generated on 96 well crystallization plates in microbatch experiments with varying protein and precipitant concentration. The scheme of the following description is shown in Fig. 1. Protein stock solutions were adjusted to 43.5 ± 1 mg/mL. Eight protein dilution steps between 5 and 43.5 mg/mL for lysozyme from chicken egg white and human lysozyme, glucose oxidase and glucose isomerase were created by mixing of buffer and protein stock solution on a sample plate. Exceptions from this procedure were made for pH values near the isoelectric point of the proteins. Stock solution of glucose oxidase at pH 5 could only be concentrated up to 21.75 ± 1 mg/mL. The purchased crystal suspension of glucose isomerase could not be redissolved in pH 3. At pH 5 glucose isomerase stock solution could only be concentrated up to 30 ± 1 mg/mL. Twelve uniform precipitant dilution steps were created by mixing of buffer and 2.5 M or 2.5 and 5 M sodium chloride stock solution, buffer and 1.5 M or 1.5 and 3 M ammonium sulfate stock solution. For PEG as precipitant buffer and 30 %(w/V) PEG 300 stock solution or buffer and 30 %(w/V) PEG 1000 stock solution were mixed.

The protein phase diagrams were generated by adding 12 µL of diluted protein solution to 12 µL of diluted precipitant solution on the crystallization plate. Protein concentration was varied per row and precipitant concentration was varied per column. The resulting protein concentration on the crystallization plate ranged between 2.5 and 21.75 mg/mL for lysozyme from chicken egg white, human lysozyme, glucose oxidase and glucose isomerase besides the mentioned

exceptions. The corresponding precipitant concentration on the crystallization plate ranged between 0 and 2.5 M for sodium chloride, between 0 and 1.5 M for ammonium sulfate and between 0 and 15 %(w/V) for PEG 300 and PEG 1000. Crystallization plates were then centrifuged for 1 min at 1000 rpm to remove air bubbles and covered using optically clear and UV compatible sealing tape.

It has to be noted that microbatch crystallization experiments are normally conducted with a volume of 10 µL and using paraffin oil to cover the solution in order to avoid evaporation (Chayen and Saridakis, 2008; D'Arcy et al., 2003; Dumetz et al., 2008). However, D'Arcy et al. (D'Arcy et al., 2003) observed that paraffin oil influences crystallization probability. For that reason phase diagrams in this work were determined without paraffin oil, but with a sealing tape to avoid evaporation without influencing phase behavior.

The sealed plates were stored in the Rock Imager for 40 days at 20 °C.

2.2.5. Automated imaging

The Rock Imager performed the automated imaging of the crystallization plates. The imaging schedule was as follows: imaging every two hours during the first two days, imaging every six hours from third to ninth day and imaging once a day from tenth to fortieth day. For every well five focus levels were investigated and superimposed to one image. Resolution was 1.2 megapixels. The images were examined after 40 days. Six possible phase states were classified: clear solution, crystallization, precipitation, skin formation, gelation and phase separation. To distinguish between salt and protein crystals, the solution drops were exposed to UV light, exploiting the fluorescence of the aromatic amino acid tryptophan to identify protein crystals. Liquid-liquid phase separation was verified by birefringence of polarized light at 90° (Zeelen, 2009).

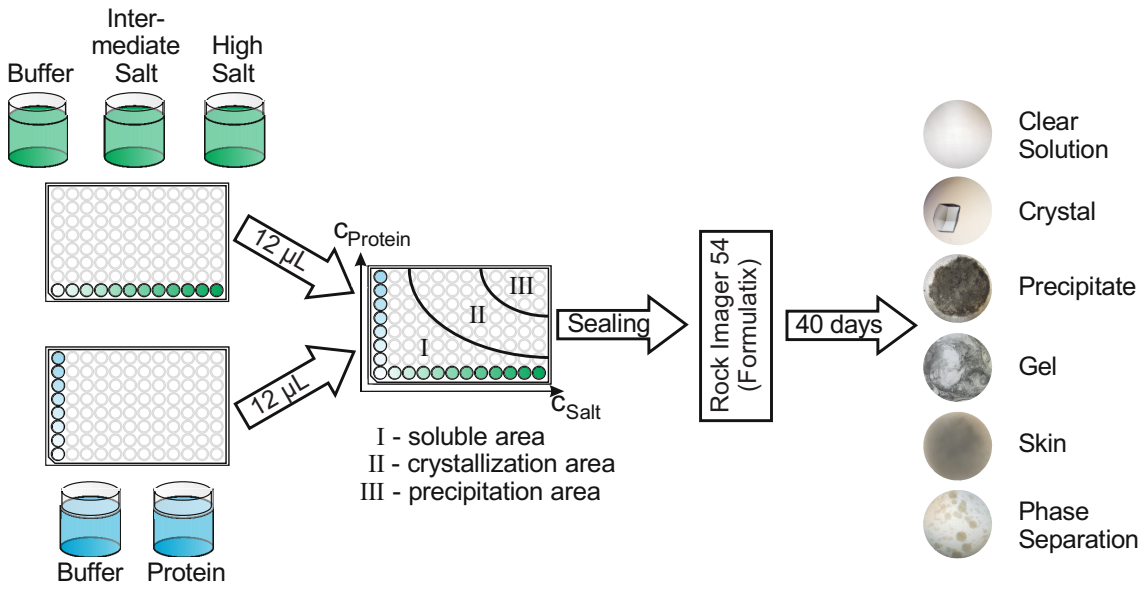

Fig. 1. Schematic illustration of the experimental setup and evaluation of protein phase states for the generation of protein phase diagrams. The soluble (I), crystallization (II) and precipitation (III) area are depicted in the phase diagram.

3. Results

3.1. pH stability in high concentrated salt solutions

The issue of pH stability is of high importance to avoid undesired pH deviations in microbatch scale due to used precipitants. Salts are known to shift the pH value (Debye and Hückel, 1923). To monitor this shift pH values of different buffers in the applied salt concentration ranges at different dilution steps were measured. The maximum deviation was found to be ±1 pH units. By the use of intermediate salt concentration buffers this deviation could be reduced to ±0.2 pH units.

3.2. Accuracy of automated liquid handling

As described in Section 2.2 different parameters were varied to ensure that correct volumes were pipetted. This provides the reproducible generation of phase diagrams. To reach this high accuracy the volume range was divided into subclasses. The subclasses ranged from 7 to 13 µL, 13.1 to 25 µL, 25.1 to 100 µL and from 100.1 to 200 µL. Within each of these subclasses the pipetting parameters were kept constant. In addition to division into subclasses, liquid handling was optimized for pipetting into air and into liquid.

To determine the quality of the liquid classes the errors of the pipetted volumes in each subclass were calculated according to Eq. (1) (Oelmeier et al., 2011).

$$r_{av} = \frac{100}{n} \sum \frac{x_{max} - x_{min}}{\bar{x}} \qquad (1)$$

r_{av} is the average range. x_{max} is the maximum and x_{min} is the minimum volume in percent of the mean volume $\bar{x}$ pipetted. n is the number of repetitions. For the buffers and precipitant solutions the r_{av} was always below 2 % – except for few outliers for the PEG solutions. For protein solutions r_{av} was below 3 %.

3.3. Phase diagrams

As stated in Section 2.2 64 protein phase diagrams were determined. Selected phase diagrams of each protein are described in detail in the following. All determined phase diagrams are shown in the supplementary online material.

3.3.1. Performance

A low volume high throughput method was established to obtain information about the phase behavior of proteins in salt or PEG environment. For the generation of one phase diagram for one protein and one precipitant in microbatch scale at a certain pH value, a total protein mass of less than 150 mg was needed. The total time consumption – for weighing the protein, dissolving it, conducting a buffer change to remove unwanted salts, concentrating the protein solution via centrifugal concentrators to the desired stock solution concentration, pipetting the microbatch plate on the robotic platform, sealing the plate and storing the microbatch plate in the Rock Imager – was around two hours.

3.3.2. Lysozyme from chicken egg white

For lysozyme from chicken egg white phase transitions occurred for sodium chloride and ammonium sulfate as precipitants. Crystallization, precipitation and skin formation could be observed. Crystallization only occurred for sodium chloride as precipitant. Precipitation and skin formation occurred for both precipitants.

3.3.2.1. Sodium chloride as precipitant

For lysozyme from chicken egg white and sodium chloride as precipitant phase transitions occurred over the whole investigated pH range. At pH 3 soluble, crystalline and precipitated phase states were observed. In some cases precipitate was accompanied by skin formation. The phase diagram is illustrated in Fig. 2(A). A soluble region occurred at low protein and precipitant concentrations. Increasing concentrations of both constituents let to crystallization and further increase to simultaneous crystallization and precipitation. At a high protein concentration of for example 21.75 mg/mL a low sodium chloride concentration of 0.45 M was needed to induce crystallization. With increasing sodium chloride concentration the protein concentration needed for induction of crystallization decreased. Three-dimensional crystals, microcrystals and needle-shaped crystals could be observed. Crystal morphology depended on protein and salt concentration. At sodium chloride concentrations of 2.05 M and above and protein

concentrations between 5.25 and 21.75 mg/mL skin formation appeared. It always co-occurred with precipitation.

The phase behavior of lysozyme from chicken egg white at pH 5 with sodium chloride as precipitant was similar to the phase behavior with the same precipitant at pH 3. The phase diagram is illustrated in Fig. 2(B). A soluble and a crystalline area were observed, as well as precipitation. However, the soluble area was wider than found at pH 3 and the crystallization area was shifted to higher salt and protein concentrations. Precipitation only occurred for the highest sodium chloride and protein concentration. The morphology of the crystals was similar to pH 3, although no microcrystals occurred at pH 5. Crystals for the same salt and protein concentration were bigger at pH 5 than at pH 3. No skin formation appeared at pH 5 with sodium chloride as precipitant.

At pH 7 sodium chloride induced both crystallization and precipitation. The latter occurred only for the two highest salt and protein concentrations. At pH 9 sodium chloride induced crystallization. No precipitation or skin formation was observed. At pH 7 and pH 9 no clear crystallization area could be defined because it was interspersed by phase states where the protein stayed soluble. However, there were more crystalline phase states at pH 9 than at pH 7. At both pH values isolated, three-dimensional crystals were observed. The crystals were much bigger than at pH 3 and pH 5. Crystal size generally increased with decreasing distance to the isoelectric point.

3.3.2.2 Ammonium sulfate as precipitant

For lysozyme from chicken egg white and ammonium sulfate as precipitant phase transitions occurred only at pH 3 and pH 5. The soluble area of the phase diagram was directly adjacent to the area where skin formation appeared. No crystalline phase states occurred previous to skin formation. Skin formation always co-occurred with precipitation. At pH 3 skin formation appeared for precipitant concentrations between 1.23 and 1.5 M and protein concentrations above 8 mg/mL.

At pH 5 skin formation occurred at 1.5 M ammonium sulfate and the two highest protein concentrations. For pH 7 and pH 9 no phase transition occurred using ammonium sulfate as precipitant.

3.3.2.3. PEG 300 and PEG 1000 as precipitant

No phase transitions of lysozyme from chicken egg white could be observed using PEG 300 or PEG 1000 as precipitant. The solutions stayed soluble over the whole investigated PEG and protein concentration range at each examined pH value.

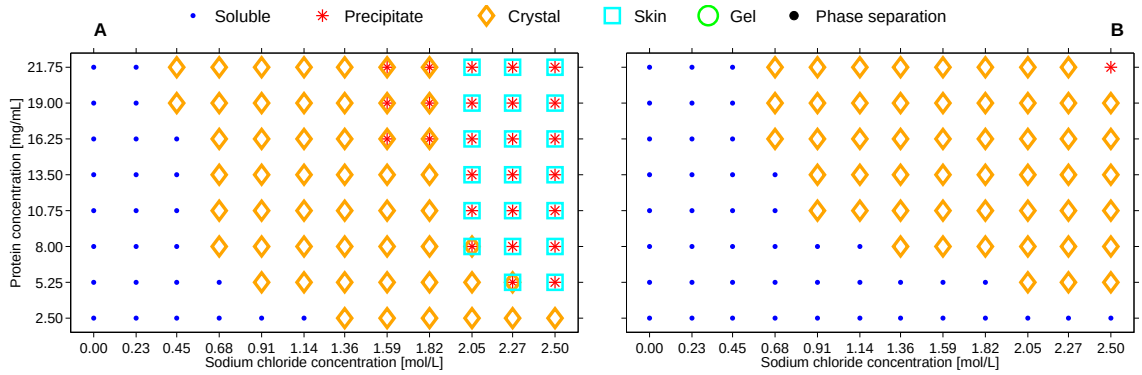

Fig. 2. Phase diagram of lysozyme from chicken egg white using sodium chloride as precipitant at pH 3 (A) and pH 5 (B).

3.3.3. Human lysozyme

For human lysozyme crystallization, gelation and precipitation were observed. Precipitation occurred predominantly and crystallization always occurred in combination with precipitation or gelation.

3.3.3.1. Sodium chloride as precipitant

Using sodium chloride as precipitant, soluble, crystalline, gelled and precipitated phase states occurred. Crystals always coexisted with gel or precipitate. For human lysozyme at pH 3 (Fig. 3(A)) phase transitions occurred for sodium chloride concentrations higher than 1.14 M. The soluble area was followed by an area where crystallization, gelation and precipitation occurred. The area where gelation appeared in combination with crystallization and/or precipitation was located in a sodium chloride concentration range between 1.36 and 1.82 M. Gelation alone appeared for 1.36 – 1.59 M sodium chloride and 8 – 10.75 mg/mL human lysozyme as well as for 1.82 – 2.05 M sodium chloride and 2.5 - 5.25 mg/mL human lysozyme. The area where gelation appeared was followed by an area where precipitation occurred. In two cases crystals and precipitated states coexisted. For human lysozyme at pH 5 using sodium chloride as precipitant (Fig. 4(A)) the precipitation area started at a sodium chloride concentration of 0.45 M and protein concentrations above 13.5 mg/mL. For sodium chloride concentrations from 2.05 M and protein concentrations between 10.75 and 21.75 mg/mL crystallization and precipitation occurred simultaneously.

At pH 7 and pH 9 only a soluble and a precipitated area could be identified, no crystals occurred. At pH 7 the precipitation started at 0.45 M sodium chloride and protein concentrations above 8 mg/mL. At pH 9 it started at 1.59 M sodium chloride and 21.75 mg/mL human lysozyme.

3.3.3.2. Ammonium sulfate as precipitant

Using ammonium sulfate as precipitant, soluble, crystalline, gelled and precipitated phase states occurred. Crystals always coexisted with precipitate. Gelation only occurred at pH 3. For human

lysozyme at pH 3 using ammonium sulfate as precipitant (Fig. 3(B)) gelation always co-occurred with precipitation. The combined gelation and precipitation area seems to represent a transfer region to a pure precipitate area. Gelation in combination with precipitation occurred for ammonium sulfate concentrations between 0.95 and 1.36 M and protein concentrations between 8 and 21.75 mg/mL. Precipitation without coexisting gelation occurred at ammonium sulfate concentrations of 1.36 M and 1.5 M and protein concentrations below 13.5 mg/mL or 19 mg/mL, respectively. Crystallization in combination with precipitation at pH 3 occurred for ammonium sulfate concentrations of 1.36 M and 1.5 M with human lysozyme concentrations of 21.75 mg/mL and 19 - 21.75 mg/mL, respectively.

At pH 5 (Fig. 4(B)) a pure precipitation area and an area with precipitation and crystallization in combination could be identified. Precipitation started at 0.55 M ammonium sulfate and 19 mg/mL human lysozyme. Crystals coexisted with precipitate from 0.95 M ammonium sulfate at 21.75 mg/mL human lysozyme. For ammonium sulfate concentrations above 0.95 M the protein concentration needed to induce crystallization decreased. At 1.5 M ammonium sulfate 8 mg/mL human lysozyme were needed to induce crystallization in addition to precipitation. At pH 7 precipitation started at 0.4 M ammonium sulfate and 19 mg/mL human lysozyme. Crystals coexisted with precipitate at and above 1.09 M ammonium sulfate and 19 mg/mL human lysozyme. At 1.5 M ammonium sulfate 8 mg/mL human lysozyme were needed to induce crystallization out of the precipitate.

At pH 9 precipitation started at 0.4 M ammonium sulfate and 16.25 mg/mL human lysozyme. Crystals coexisted with precipitate from 1.36 M ammonium sulfate and 21.75 mg/mL human lysozyme. At 1.5 M ammonium sulfate 19 mg/mL human lysozyme were needed to induce crystallization emerging from precipitate.

With decreasing distance to the isoelectric point and while using ammonium sulfate as precipitant the combined precipitated and crystalline area decreased continuously in its size. Only the phase behavior at pH 3 made an exception to this.

3.3.3.3. PEG 300 as precipitant

In the investigated pH range no phase transition of human lysozyme occurred when using PEG 300 as precipitant.

3.3.3.4. PEG 1000 as precipitant

At pH 3 and pH 5 no phase transition of human lysozyme occurred when using PEG 1000 as precipitant. At pH 7 and pH 9 precipitation occurred. Precipitation started at 2.73 %(w/V) PEG 1000 and for human lysozyme concentrations from 19 mg/mL in both cases. The effect of PEG on the phase behavior of human lysozyme was thus found to be dependent both on polymer size and pH.

Although phase behavior of human lysozyme at pH 7 and pH 9 with PEG 1000 as precipitant looks very similar there is a difference when considering precipitation kinetics. At pH 7 no spontaneous precipitation occurred. Precipitation evolved over time. At pH 9 spontaneous precipitation immediately after pipetting occurred for 9.55 %(w/V) PEG 1000 and 21.75 mg/mL human lysozyme and from 10.91 %(w/V) PEG 1000 for human lysozyme concentrations of 16.25 mg/mL and above.

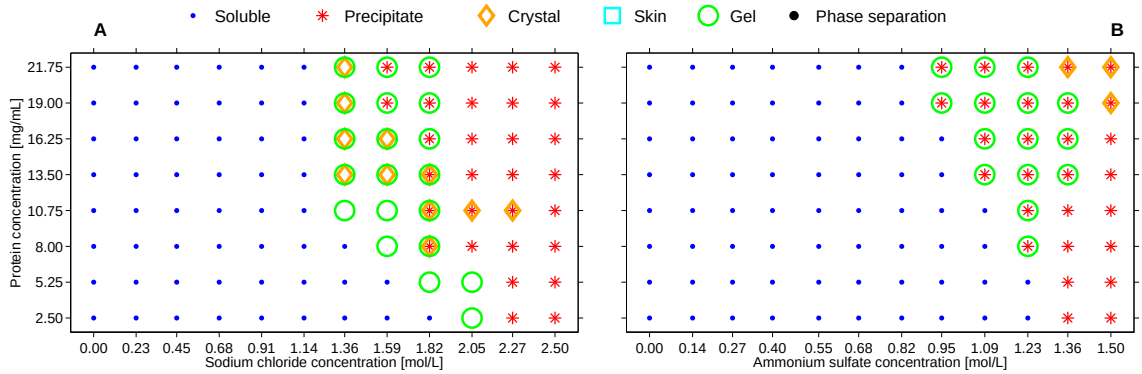

Fig. 3. Phase diagrams of human lysozyme at pH 3 using sodium chloride (A) and ammonium sulfate (B) as precipitant.

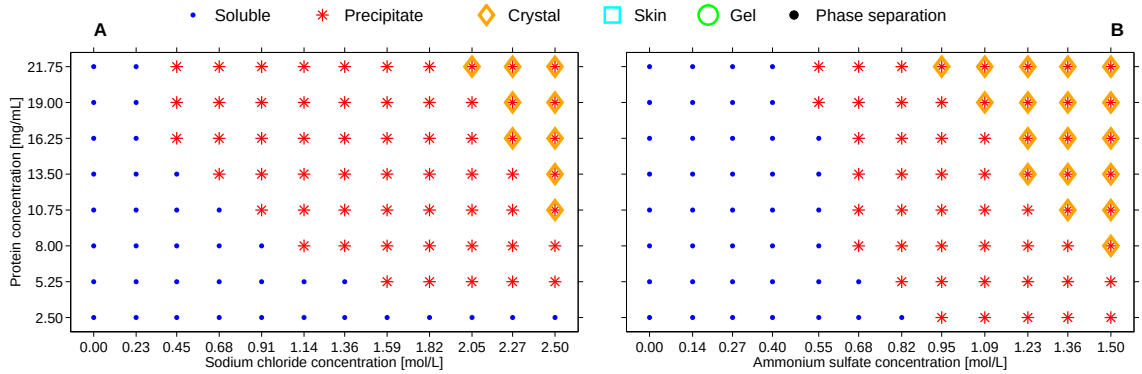

Fig. 4. Phase diagrams of human lysozyme at pH 5 using sodium chloride (A) and ammonium sulfate (B) as precipitant.

3.3.4. Glucose oxidase

For glucose oxidase only phase transitions to precipitation were observed. It was distinguished between spontaneous precipitation and precipitation of slower kinetics. The former occurred immediately after pipetting, the latter evolved over time. No other phase transitions were found at any of the examined conditions for glucose oxidase.

3.3.4.1. Sodium chloride as precipitant

For glucose oxidase and sodium chloride as precipitant precipitation with different kinetics could be observed for pH 3 and pH 5.

At pH 3 glucose oxidase was stable without salt over the whole protein concentration range and with 0.23 M sodium chloride for 2.5 – 5.25 mg/mL glucose oxidase. From and above 0.45 M sodium chloride spontaneous precipitation was induced for the highest protein concentration of 21.75 mg/mL. With increasing sodium chloride concentration the protein concentration needed for spontaneous precipitation decreased. In the transition area from soluble to spontaneous precipitation, an area with precipitation formation of slower kinetics existed.
At pH 5 slowly evolving precipitation formation was observed starting from 0.45 M sodium chloride for protein concentrations above 8.13 mg/mL. No spontaneous precipitation was found. For pH 7 and pH 9 no phase transitions were observed for sodium chloride as precipitant.

3.3.4.2. Ammonium sulfate as precipitant

For glucose oxidase and ammonium sulfate as precipitant at pH 3 (Fig. 5(A)) precipitation occurred starting at 0.14 M. At salt concentrations between 0.14 and 0.4 M this precipitate evolved over time. Ammonium sulfate induced spontaneous precipitation at salt concentrations from 0.4 M ammonium sulfate for a protein concentration of 21.75 mg/mL. At higher salt concentrations lower protein concentrations sufficed for spontaneous precipitation. At 1.5 M ammonium sulfate 2.5 mg/mL glucose oxidase sufficed to induce spontaneous precipitation. At pH 5 slowly evolving precipitate formation was monitored starting at a salt concentration of 0.27 M ammonium sulfate and protein concentrations of 8.13 mg/mL and above. Hereby it has to be mentioned that the maximum glucose oxidase concentration in the phase diagram for pH 5 was 10.88 mg/mL. At higher salt concentrations above 0.27 M ammonium sulfate lower protein concentrations sufficed for precipitate formation of slower kinetics, with 4 mg/mL glucose oxidase at 1.5 M ammonium sulfate being the lowest. At 1.5 M ammonium sulfate spontaneous precipitation occurred for protein concentrations for 9.5 and 10.88 mg/mL.
For pH 7 no phase transition was monitored.
At pH 9 (Fig. 5(B)) precipitation was found for 0.4 M ammonium sulfate and above starting at 16.25 mg/mL glucose oxidase. For higher ammonium sulfate concentrations lower protein concentrations were needed for precipitate formation. This slow evolving precipitation was followed by spontaneous precipitation at salt concentrations from 0.95 M ammonium sulfate and protein concentrations starting at 16.25 mg/mL. For 1.5 M ammonium sulfate spontaneous precipitation occurred for 5.25 mg/mL glucose oxidase and above.

3.3.4.3. PEG 300 as precipitant

Glucose oxidase showed no phase transitions for PEG 300 in the investigated concentration and pH range.

3.3.4.4. PEG 1000 as precipitant

For PEG 1000 no spontaneous precipitation formation of glucose oxidase occurred. At pH 5, close

to the protein's isoelectric point, precipitate evolved over time starting at 9.55 %(w/V) PEG 1000 and the highest protein concentration of 10.88 mg/mL. Higher precipitant concentrations let to precipitation starting at lower protein concentrations. This slowly evolving precipitation occurred at 15 %(w/V) PEG 1000 for 4 mg/mL glucose oxidase and above.

At pH 3, 7 and 9 all investigated conditions showed no phase transitions for PEG 1000.

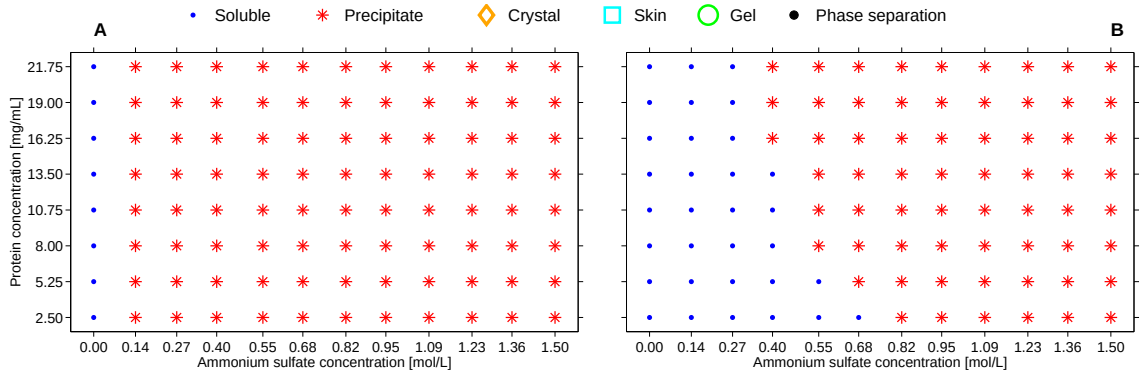

Fig. 5. Phase diagrams of glucose oxidase using ammonium sulfate as precipitant at pH 3 (A) and pH 9 (B).

3.3.5. Glucose isomerase

For glucose isomerase phase transitions to crystallization, precipitation, liquid-liquid phase separation (LLPS) and skin formation were observed. The precipitate formation was divided into spontaneous and slow evolving precipitation as described above for glucose oxidase.

3.3.5.1. Sodium chloride as precipitant

For glucose isomerase and the precipitant sodium chloride only precipitation and skin formation occurred at pH 5. Here slowly evolving precipitation and additional skin formation occurred independently of salt or protein concentration for all tested conditions.

At pH 7 and pH 9 no phase transitions occurred while using sodium chloride as precipitant.

3.3.5.2. Ammonium sulfate as precipitant

For glucose isomerase and ammonium sulfate as precipitant crystallization, precipitation and skin formation were observed. At pH 5 ammonium sulfate concentrations below 0.82 M let to evolving precipitation and skin formation for all protein concentrations, with 15 mg/mL being the highest determinable glucose isomerase concentration. At one condition within the slow evolving precipitation zone three-dimensional crystals occurred additionally (5.52 mg/mL glucose oxidase, 0.68 M ammonium sulfate). At ammonium sulfate concentrations of 0.82 M and above spontaneous protein precipitation with coexisting skin formation occurred. At 1.5 M ammonium sulfate 5.52 mg/mL sufficed to induce spontaneous precipitation.

At pH 7 and pH 9 crystallization of glucose isomerase occurred for ammonium sulfate concentrations of 1.36 M and above (Fig. 6). For pH 7 crystallization of glucose isomerase started

from 16.25 mg/mL. For pH 9 the following conditions yielded crystals: 1.36 M ammonium sulfate at 19 mg/mL glucose isomerase and 1.5 M ammonium sulfate at 16.25 – 21.75 mg/mL glucose isomerase. For pH 9 the soluble area was slightly wider compared to pH 7. All crystals found were three-dimensional.

3.3.5.3. PEG 300 as precipitant

For glucose isomerase and precipitant PEG 300 (Fig. 7(A)) only phase transitions were found at pH 5. Here, the conditions without precipitant showed slowly evolving precipitation with coexisting skin formation. For PEG 300 concentrations starting at 1.36 %(w/V) skin formation was suppressed at all investigated protein concentrations. At PEG 300 concentrations of 15 %(w/V) LLPS was observed for conditions starting from 11.21 mg/mL protein.
At pH 7 and pH 9 no phase transitions were monitored for PEG 300 as precipitant.

3.3.5.4. PEG 1000 as precipitant

For glucose isomerase and PEG 1000 phase transitions occurred only at pH 5 (Fig. 7(B)). 10.91 %(w/V) PEG 1000 let to LLPS for protein concentrations starting from 9.31 mg/mL. At PEG 1000 concentrations starting from 12.27 %(w/V) spontaneous precipitation occurred for 9.31 mg/mL glucose isomerase and above. For 15 %(w/V) PEG 1000 7.41 mg/mL glucose isomerase sufficed to induce spontaneous precipitation. For all other conditions slowly evolving precipitate was observed. Additional skin formation was observed for PEG 1000 up to 4.09 %(w/V) for all glucose isomerase concentrations. At PEG 1000 concentrations from 5.45 %(w/V) and glucose isomerase concentrations below 5.52 mg/mL skin formation was suppressed. At higher PEG concentrations skin formation was suppressed at higher protein concentrations.
At pH 7 and pH 9 no phase transitions occurred with PEG 1000 as precipitant.

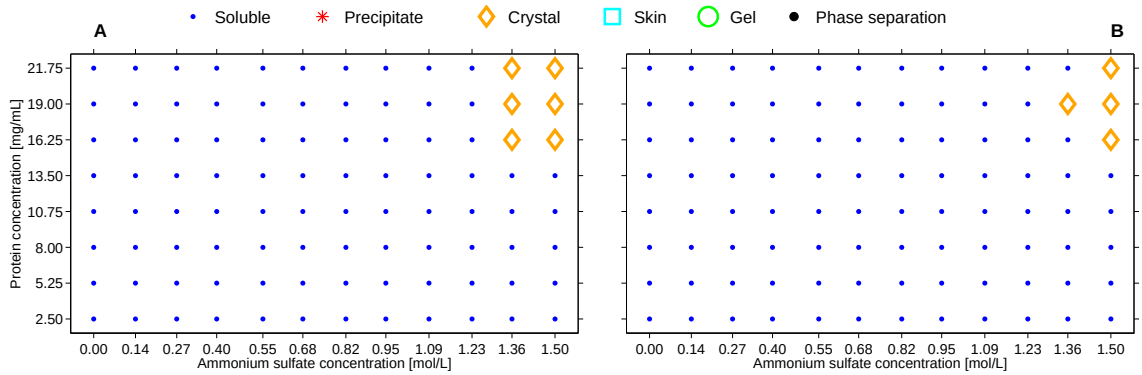

Fig. 6. Phase diagrams of glucose isomerase using ammonium sulfate as precipitant at pH 7 (A) and pH 9 (B).

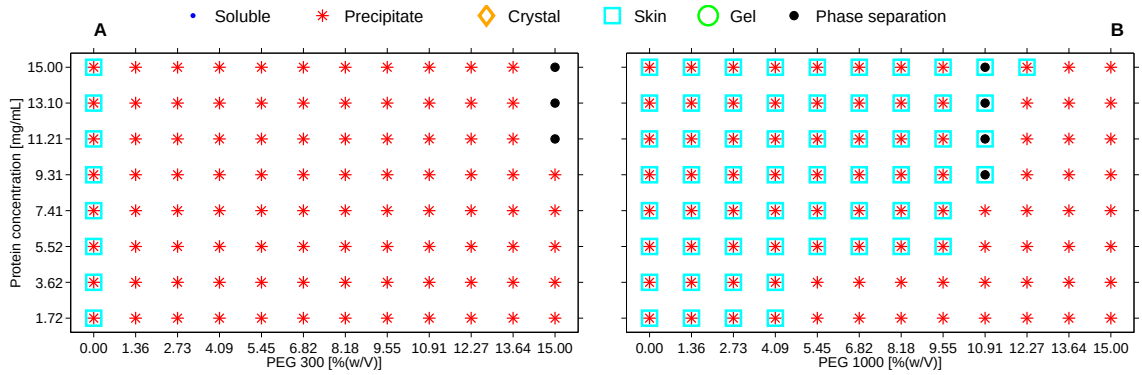

Fig. 7. Phase diagrams of glucose isomerase at pH 5 using PEG 300 (A) and PEG 1000 (B) as precipitant.

4. Discussion

4.1. pH stability in high concentrated salt solutions

When mixing a buffer without salt and a buffer with salt at the same pH the pH value of the buffer mixture can drift. This phenomenon can be described by the Debye-Hückel-Theory (Eq. (2)) (Debye and Hückel, 1923).

$$pk_a^* = pk_a - \frac{1.824 * 10^6}{(\varepsilon T)^{\frac{3}{2}}} |z^+ z^-| \sqrt{I} \qquad (2)$$

pk_a* is the corrected pk_a value, ε is the dielectric constant of the solution, T is the temperature (K), z is the charge of the ion and I is the ionic strength (M). When the ionic strength I of a buffer is increased the pk_a of the buffer substance changes to pk_a* (Eq. (2)). Thereby the pH value of the buffer solution changes. For polyprotic buffer substances the charge of the buffer ions at the second or third pk_a of the buffer substance is higher than at the first pk_a value. When the ionic strength is increased the pk_a is influenced more strongly at the second or third pk_a value compared to the first pk_a. All pk_a values were obtained from (Kröner and Hubbuch, 2013). Citrate (pk_{a1}: 3.13, pk_{a2}: 4.76, pk_{a3}: 6.40) is used at pH 3 at its first of its three pk_a values therefor the change of the ionic strength is not problematic. Acetate (pk_a: 4.76) – used at pH 5 – has only one pk_a and so there is hardly a problem when using three stock solutions with 0 M, 2.5 M and 5 M sodium chloride or 0 M, 1.5 M and 3 M ammonium sulfate. The synthetically manufactured buffer substance MOPSO (pk_a: 6.90) has only one pk_a value at pH 6.9. Hence the buffer pH stays constant in the entire salt concentration range. Bis-Tris propane has two pk_a values (pk_{a1}: 8.93, pk_{a2}: 6.59). Bis-Tris propane is a basic buffer substance, hence at pH 9 it is used at the first pk_a. The above-mentioned discussion demonstrates the importance of carefully choosing buffering components when investigating protein solutions over a wide pH and salt concentration range.

Thereby, we were able to rule out any unwanted pH shifts in the samples. Thus, it was possible to create phase diagrams at various pH values with high confidence.

4.2. Accuracy of automated liquid handling

Optimization of liquid handling is an essential step for the reproducible preparation of phase diagrams. This led to pipetting accuracies with errors below 2 % for buffers and precipitant solutions and to errors ≤ 3 % for protein solutions. Without these optimizations pipetting inaccuracies between 10 % and 30 % can be expected.

The main difference between pipetting salt buffer in air and pipetting into liquid is, that the leading air gap – that means the air gap that is dispensed following the buffer – is smaller for pipetting into liquid. The creation of too much air bubbles could be avoided through this adjustment.

With higher viscosities for PEG containing buffers compared to salt buffers the dispensing speed had to be lowered to get accurately pipetted volumes.

The main difference for protein liquid classes compared to precipitant liquid classes is, that for protein solutions no leading air gap was used. Creation of air bubbles, that could probably be critical to the protein, could be avoided.

The optimization procedures described in Section 2.2 ensured that during preparation of phase diagrams correct volumes of each liquid were pipetted and that the specified concentrations are trustworthy. As with the careful selection of buffering components, thorough optimization of liquid handling parameters is critical for generation of high quality data.

4.3. Phase diagrams

4.3.1. Performance

In this paper a method to create phase diagrams in high throughput was established. Two hours and a protein mass of less than 150 mg were needed to cover a broad protein and precipitant concentration range and to determine the phase behavior of the protein. One phase diagram consists of 96 different conditions with variation of protein and precipitant concentration at a constant pH value and at 20 °C. It was thus shown, that the developed methodology lends itself well for a standardized approach for phase diagram generation and can easily be ported to other protein and precipitant systems. However, system composition and operating conditions should be taken into consideration carefully.

4.3.2. Lysozyme from chicken egg white

4.3.2.1. Sodium chloride as precipitant

The distance from the isoelectric point and the sodium chloride concentration had an influence on lysozyme solubility, aggregation kinetics and crystal morphology.

The influence on lysozyme solubility was mainly driven by the pH value, i.e. the distance from the

isoelectric point. At pH 9 the first phase transition occurred at 0.23 M sodium chloride where lysozyme still stayed soluble at pH 3, 5 and 7. Hence, we assume that for low sodium chloride concentrations up to 0.23 M the solubility increased with increasing distance to the pI. This fits to the commonly accepted observation that for most proteins at low salt concentrations solubility decreases while approaching the pI and is lowest at the pI (Scopes, 1994). For sodium chloride concentrations above 0.23 M phase transitions occurred at lower protein concentrations at lower pH values. Hence, solubility decreased with increasing distance to the pI. In a low-electrolyte region, which was identified to reach up to 0.23 M sodium chloride, the long-range electrostatic protein interactions are significant (Velev et al., 1998). These long-range electrostatic forces are repulsive, which means that the higher the protein charge, the higher the repulsion between the equally charged protein molecules. Lysozyme solubility in the low-electrolyte region thus is highest at pH 3 because of the highest electrostatic repulsion and lowest at pH 9. Our experimental results in the low-electrolyte range can further be supported by second osmotic virial coefficient (B_{22}) measurements of Velev et al. (Velev et al., 1998). The B_{22} values up to 0.1 M sodium chloride are highest at pH 3 and decrease with increasing pH value.

The long-range electrostatic repulsion between charged proteins is reduced by the presence of salt ions. These salt ions screen the electrostatic fields of neighboring proteins and short-range attractive forces become noticeable. The importance of short-range forces in protein interactions and for protein phase behavior was described in literature before (Beretta et al., 2000; Piazza, 2000; Stradner et al., 2004; Tardieu et al., 2002). Short-range attractive forces can be van der Waals forces, hydrophobic or osmotic forces (Leckband and Israelachvili, 2001). Attractive osmotic forces due to a high electrolyte concentration have been described earlier (Curtis et al., 1998; Vlachy et al., 1993; Vlachy and Prausnitz, 1992). Hydrophobic interactions occur due to hydrophobic patches on the protein surface and are strong at high ionic strengths (Scopes, 1994).

For lysozyme from chicken egg white and sodium chloride concentrations above 0.23 M the effect of electrostatic repulsion seems to vanish and short-range forces become increasingly important. At pH values with a higher distance from the pI, lower lysozyme concentrations are needed to induce phase transitions in the high-electrolyte region, beginning with crystallization. In the high-electrolyte region the common belief that solubility is lowest at or near the pI seems not to be valid for lysozyme from chicken egg white with sodium chloride as precipitant.

As mentioned before, the sodium chloride concentration and the distance from the pI also had an influence on the lysozyme aggregation kinetics and crystal morphology. A comparison between the phase diagrams of lysozyme from chicken egg white with sodium chloride as precipitant for different pH values shows that most phase transitions occurred at pH 3. Three-dimensional crystals evolved for sodium chloride concentrations between 0.45 and 1.36 M sodium chloride

but crystal size decreased with increasing sodium chloride and protein concentration resulting in needle-shaped crystals and microcrystals. A broad area with precipitation and skin formation occurred, which is a hint for protein denaturation (Zeelen, 2009). At pH 5 fewer phase transitions occurred compared to pH 3. Aggregation kinetics also had been slower than at pH 3, resulting in bigger crystal sizes at identic sodium chloride concentrations and the absence of microcrystals and a precipitation area. At pH 7 and pH 9 only isolated phase transitions occurred. Aggregation kinetics were slow resulting in single three-dimensional, big crystals and a crystallization area that was interspersed with conditions where lysozyme stayed soluble. As phase transitions at every investigated pH value happened in a sodium chloride concentration range where long-range electrostatic forces are screened, this shows that the attractive short-range forces causing aggregation need to be pH dependent. The attractive short-range forces increased for lysozyme from chicken egg white with increasing distance to the pI resulting in smaller protein crystals. Attractive short-range forces also increased in strength with increasing sodium chloride concentration. That resulted in accelerated aggregation kinetics and thus altered crystal morphology and caused pronounced precipitation. The intermolecular short-range forces at pH 3 between lysozyme from chicken egg white molecules seem to be strong enough to induce complete or partial protein denaturation resulting in skin formation.
To the best of our knowledge the here proposed and experimentally supported pH dependency of short-range attractive forces was not described in literature earlier.

4.3.2.2. Ammonium sulfate as precipitant

For lysozyme from chicken egg white using ammonium sulfate as precipitant phase transitions only occurred at pH 3 and pH 5 at high ammonium sulfate concentrations. Except of one condition where only precipitation occurred, precipitation and skin formation always coexisted. In general, kosmotropes such as ammonium sulfate are useful for salting-out proteins, i.e. decrease protein solubility, but tend to reduce protein denaturation. Our results showed no salting-out, i.e. amorphous precipitation, up to 1.09 M ammonium sulfate at pH 3 and 1.36 M at pH 5 but induction instead of reduction of protein denaturation for higher ammonium sulfate concentrations. Riess-Kautt and Ducruix (Ries-Kautt and Ducruix, 1989) found that lysozyme from chicken egg white follows the inverse Hofmeister series for anions where solubility is highest for strong kosmotropes such as ammonium sulfate and lower for sodium chloride. Zhang and Cremer (Zhang and Cremer, 2009) suggested that in general positively charged macromolecular systems should show inverse Hofmeister behavior only at relatively low salt concentrations, but revert to a direct Hofmeister series as the salt concentration is increased. Up to 1.09 M ammonium sulfate at pH 3 and 1.36 M at pH 5 solutions stayed soluble analogous to the systems with 0 M ammonium sulfate. No statement about solubility enhancement and thus direction of the Hofmeister series in this ammonium sulfate concentration range is possible. But

solubility in ammonium sulfate solutions is higher than in sodium chloride solutions at the same concentrations. With increasing ammonium sulfate concentration almost exclusively precipitation along with skin formation, i.e. protein denaturation, occurred. This observation negates the reversion to a direct Hofmeister series for high salt concentrations described by Zhang and Cremer (Zhang and Cremer, 2009).

Attractive short-range forces causing aggregation are pH dependent as it was observed for sodium chloride as precipitant, too. The short-range forces increased with increasing distance to the pI resulting in more phase transitions. Attractive short-range forces also increased in strength with increasing ammonium sulfate concentration what resulted in pronounced precipitation and skin formation.

4.3.2.3. PEG as precipitant

No phase transitions were observed for PEG 300 and PEG 1000 up to 15 %(w/V) in the investigated protein concentration and pH range. Hence, PEG 300 and PEG 1000 have no effect on the phase behavior of lysozyme from chicken egg white in the investigated protein concentration range.

4.3.3. Human lysozyme

4.3.3.1. Sodium chloride as precipitant

For human lysozyme with sodium chloride as precipitant at pH 3, a broad gelation area was found. The observed gels were translucent. The development of gelation at pH 3 might be due to the appearance of partial denaturation. For globular proteins denaturation is often considered a prerequisite to gelation (Ziegler and Foegeding, 1990) and a pH value far away from the isoelectric point might lead to partial or complete denaturation for some proteins (Wang et al., 2010).

Gelation occurred in a sodium chloride concentration range between 1.36 and 2.05 M. For low human lysozyme concentrations gelation appeared exclusively. Increasing protein concentration led to formation of crystals or precipitate or both of them next to gelation. We assume that crystals and precipitate developed in the supernatant of the gelled phase for human lysozyme concentrations that are high enough for short-range forces to be effective. At 1.59 M sodium chloride for example crystallization co-occurred to gelation for 13.5 – 16.25 mg/mL human lysozyme, precipitation co-occurred to gelation for 19 – 21.75 mg/mL. This shows, that short-range forces either increase or gain more impact with increasing protein concentration. This concurs with the results of Stradner et al. (Stradner et al., 2004) who also described a protein concentration dependency of short-range intermolecular forces. The phase behavior of human lysozyme at pH 5, 7 and 9 showed that aggregation propensity decreased with decreasing distance to the isoelectric point. The short-range attractive forces responsible for aggregation of

the protein molecules are pH dependent as it was shown for lysozyme from chicken egg white. Aggregation propensity decreases while the short-range attractive forces decrease.

Amorphous structures emerge if short-range forces are very strong (George and Wilson, 1994) because the protein molecules do not have time to orient themselves into a crystal lattice (Curtis et al., 2002). Short-range forces have to be strong for human lysozyme at pH 5, 7 and 9 using sodium chloride as precipitant because no crystallization area evolved prior to the precipitation area.

Crystallization could only be observed at pH 5 and the crystals arose from the precipitate. It could not be determined if the crystals evolved in the supernatant of the precipitated solution or through restructuring of amorphous precipitate into a crystal lattice. It is also controversial whether the occurrence of amorphous aggregate hinders or promotes crystal growth, but the formation of amorphous aggregate prior to crystallization is a known phenomenon (Piazza, 2000). Piazza (Piazza, 2000) explains the evolution of crystals out of precipitate by restructuring of the protein molecules. Crystal growth out of the supernatant of precipitated solutions might be explained as follows. Short-range forces are too strong to enable ordered structures, especially at high protein concentrations, and precipitation occurs. In the supernatant of precipitated protein solutions the protein concentration could be low enough to form ordered structures despite the strong short-ranged attraction. This would mean that short-ranged attraction is also driven by protein concentration and decreases in its effectiveness with decreasing protein concentration. Protein concentration dependency of short-range forces was described by Stradner et al. (Stradner et al., 2004). The overall strength of the short-range attraction at pH 5 has to be higher than at pH 7 and pH 9 to explain why the reduced short-range force in the supernatant only at pH 5 is high enough to result in crystallization.

4.3.3.2. Ammonium sulfate as precipitant

At pH 3 the soluble area is followed by a gelation area. Protein denaturation is often considered a prerequisite for gel formation (Ziegler and Foegeding, 1990). As for human lysozyme in sodium chloride solutions this could be due to the pH value far away from the isoelectric point leading to partial or complete denaturation (Wang et al., 2010).

In contrast to the phase behavior of lysozyme from chicken egg white at pH 5, 7 and 9, ammonium sulfate had a similar but slightly stronger influence on human lysozyme than sodium chloride. Human lysozyme crystals coexisting to precipitate occurred at pH 5, 7 and 9 when using ammonium sulfate whereas using sodium chloride they occurred only at pH 5. As for sodium chloride, the soluble area was followed by a precipitation area. Precipitation started in a medium to high electrolyte region, which means that repulsive electrostatic forces were not significant. For high salt and protein concentrations crystals coexisted with precipitate. The propensity for crystallization decreased with increasing pH value. Short-range attractive forces,

as they were described for lysozyme from chicken egg white and sodium chloride as precipitant, again seem to be pH dependent and were stronger for pH values with a higher distance to the isoelectric point.

A comparison between the phase behavior of human lysozyme and lysozyme from chicken egg white at pH 5, 7 and 9 shows that also the short-range forces are extremely salt-specific. This was earlier described by Piazza (Piazza, 2000) who additionally mentioned that the short-range interparticle potential is closely related to the hydrophilic-hydrophobic 'patching' of the protein surface. Schwierz et al. (Schwierz et al., 2010) specified these observations and showed that direct, reversed and also partially reversed Hofmeister series are possible depending on the protein surface charge and surface polarity, i.e. hydrophobic patching. This leads to the conclusion that a reversal or transposition of the salt order in the Hofmeister series does not mainly depend on a pH below or above the isoelectric point as it was suggested by Ries-Kautt and Ducruix (Ries-Kautt and Ducruix, 1989). We go along with Piazza (Piazza, 2000) and Schwierz et al. (Schwierz et al., 2010) as we saw that the short-range forces leading to various aggregation processes are extremely salt- and protein-specific and hardly predictable.

4.3.3.3. PEG as precipitant

PEG 300 has no effect on the phase behavior of human lysozyme whereas PEG 1000 has an influence on the aggregation propensity of human lysozyme at pH 7 and pH 9. PEG 1000 causes a higher depletion attraction than PEG 300 due to its larger size. The origin of this depletion attraction is explained in detail in Section 4.3.5 for glucose isomerase.

The fact that PEG 1000 only influences aggregation propensity at pH 7 and pH 9 could be a hint for a pH dependent hydrophobic force. Lee and Lee (Lee and Lee, 1981) found that the magnitude of protein solution destabilization by PEG depends on the average hydrophobicity of proteins. Destabilization was stronger for proteins with a higher extent of hydrophilic residues. This might be explained by binding of PEG to sufficiently large hydrophobic protein surface patches as it was described by Baynes and Trout (Baynes and Trout, 2004) and by that reduction of attractive hydrophobic forces. This means that the average hydrophobicity of human lysozyme at pH 3 and pH 5 is higher than at pH 7 and pH 9 and solution destabilization, i.e. protein aggregation, at pH 3 and pH 5 is reduced by the presence of PEG. Discussion of the destabilizing effect of PEG on glucose isomerase will additionally show that destabilizing effects through depletion attraction occur when electrostatic repulsion is at its minimum, i.e. near the isoelectric point.

According to Lee and Lee (Lee and Lee, 1981) we also conclude that the average hydrophobicity of lysozyme from chicken egg white has to be higher than that of human lysozyme because neither PEG 300 nor PEG 1000 showed a destabilizing effect on lysozyme from chicken egg white.

4.3.4. Glucose oxidase

4.3.4.1. Sodium chloride as precipitant

Glucose oxidase with sodium chloride as precipitant showed only transitions from soluble to precipitated states at pH 3 and pH 5. No crystallization was observed in the transition area between soluble and precipitated states.

In the pH range of pH 3 to pH 5, close to the isoelectric point of glucose oxidase, long-ranged repulsive electrostatic forces are at the minimum. This is generally observed to decrease protein solubility in solution in a low-electrolyte region where conductivity is low (Scopes, 1994). The lower obtainable stock solution concentration for glucose oxidase at pH 5 was found to be in agreement. In higher salt concentrations the shielding of long-ranged electrostatic repulsion strengthens the influence of short-range interactions such as van der Waals, osmotic and hydrophobic interactions as described in Section 4.3.2 for lysozyme from chicken egg white. In the high-electrolyte region the extent of glucose oxidase precipitation with sodium chloride as precipitant was found to be lower at pH 5 compared to pH 3 for similar protein concentrations. This behavior of decreasing solubility with increasing distance to the isoelectric point in the pH range of pH < pI indicates a pH dependency of short-range attractive interactions as proposed for lysozyme from chicken egg white. Further it is assumed that acidic pH values can lead to partial or complete protein denaturation (Wang et al., 2010) which consequently increases hydrophobicity. This is in agreement with our experimentally observed phase behavior and might explain the observed increase of short-range forces at pH 3 leading to destabilization of the protein solution.

At pH 7 and pH 9 glucose oxidase was stable in the presence of sodium chloride as precipitant within the experimental range. In this pH range, pH > pI, short-range attractive forces were weak for glucose oxidase molecules, as no phase transitions were observed here.

4.3.4.2. Ammonium sulfate as precipitant

For glucose oxidase with ammonium sulfate as precipitant transitions from soluble to precipitated phase states were found at pH 3, 5 and 9.

At pH 3 the precipitation extend was higher compared to pH 5 at similar protein concentrations as was seen for sodium chloride as precipitant. At pH 7 no phase transition was monitored. But at further distance to the isoelectric point, at pH 9, precipitation was observed in the high-electrolyte region for ammonium sulfate.

This precipitation at increasing distance to the isoelectric point in a high-electrolyte region encourages the assumption that pH depended short-range attractive forces are of significant impact.

Further, our results for glucose oxidase showed salting-out for ammonium sulfate when compared to sodium chloride. This salt specificity of short-range forces, as also seen for human

lysozyme, was described earlier by Piazza (Piazza, 2000). The order of salts promoting salting-out of glucose oxidase is in agreement with the Hofmeister series (Hofmeister, 1888). Kosmotropes, such as ammonium sulfate, are strongly hydrated ions (water structure maker) which can lead to destabilization of the protein hydration layer. Therefore the protein reduces solvent accessible surface area by decreasing solubility. This is in contrast to the mode of action of chaotropes, such as sodium chloride. Chaotropic ions are large monovalent ions of low charge density which are weakly hydrated (water structure breaker). Those ions lead to a stabilized hydration shell around the protein maximizing its solvent accessible surface area. At high chaotrope concentrations this can lead to denaturation/unfolding of the protein molecule (Collins, 2004).

4.3.4.3. PEG as precipitant

PEG 300 had no effect on the phase behavior of glucose oxidase up to 15 %(w/V) in the investigated pH range. PEG 1000 triggered precipitation of glucose oxidase at pH 5 at PEG concentration of 9.55 %(w/V) and above. The destabilizing mechanism of PEG as precipitant can be explained with depletion effects driven by osmotic pressure as described in Section 4.3.5 for glucose isomerase.

This described PEG destabilization mechanism is in agreement with the results presented here for glucose oxidase and PEG 1000 at pH 5. The observed protein destabilization could only emerge at conditions where electrostatic repulsion between protein molecules is at its minimum, close to its isoelectric point.

4.3.5. Glucose isomerase

4.3.5.1. Sodium chloride as precipitant

For glucose isomerase and sodium chloride as precipitant only phase transitions at pH 5 were observed. At this pH, all conditions were instable and showed evolving precipitation and skin formation independent of sodium chloride concentration. In the low-electrolyte region stabilizing electrostatic forces of repulsive nature are suspected to be low due to the vicinity to the isoelectric point of glucose isomerase. The observed skin formation is an indicator for non-native aggregation (Zeelen, 2009). This observed phase behavior is supported by findings of partially folded intermediates at pH 5 with higher hydrophobicity and tendency to form aggregation (Pawar and Deshpande, 2000). This tendency to denaturation at low pH was not found for glucose oxidase, a protein with similar isoelectric point to glucose isomerase. Lysozyme from chicken egg white showed skin formation at pH 3 for both investigated salts and at pH 5 for ammonium sulfate in the high-electrolyte region whereas human lysozyme showed no skin formation but gelation.

At pH 7 and pH 9 no phase transitions occurred for glucose isomerase and sodium chloride.

Thus, both the low and high-electrolyte environment was governed by stabilizing effects. In the low-electrolyte region this effect is likely to be caused by electrostatic repulsive forces. The high-electrolyte region the solubility mediating effect of the chaotropic salt sodium chloride is likely to be the dominating effect as described in Section 4.3.4 for glucose oxidase.

4.3.5.2. Ammonium sulfate as precipitant

Ammonium sulfate as precipitant caused more phase transitions of glucose isomerase than sodium chloride in the entire investigated pH range.

At pH 5 additional to the observed evolving precipitation and skin formation without salt and in the low-electrolyte region, spontaneous precipitation was observed for ammonium sulfate in the high-electrolyte region. This leads to the assumption of stronger short-range attractive forces in high-electrolyte region for ammonium sulfate when compared with sodium chloride. This emphasizes the salting-out effect of ammonium sulfate and is in agreement with the Hofmeister series (Hofmeister, 1888).

Crystallization of glucose isomerase was found for pH 5, 7 and 9. At pH 5 one condition within a precipitation zone of evolving precipitation and skin formation occurred. At pH 7 and pH 9 crystallization evolved at high ammonium sulfate concentrations. The soluble area increased with increasing pH value. This is in agreement with findings from Chayen et al. (Chayen et al., 1988) in the pH range of 5.5 to 6.5 at 1.5 M ammonium sulfate. This observation for glucose isomerase leads to the assumption that short-range attractive interactions are weaker with increasing distance to the isoelectric point. Those attractive interactions are weak enough to form three dimensional crystals. Nonspecific aggregation would appear if attractive interactions are stronger.

Our findings concerning the pH dependency of short-range attractive forces vary for the investigated alkaline proteins, such as lysozyme from chicken egg white and human lysozyme, and acidic proteins glucose oxidase and glucose isomerase. For the alkaline proteins the short-range forces increase with increasing distance to the isoelectric point whereas for the acidic proteins lower short-range attractive forces are found in greater distance of the proteins isoelectric point.

4.3.5.3. PEG as precipitant

Polyethylene glycol (PEG) resulted in phase transitions towards precipitation and liquid-liquid phase separation (LLPS) for glucose isomerase at pH 5.

Thus, an influence of PEG was only seen at conditions close to the isoelectric point of the protein where electrostatic repulsive forces are at their minimum. Therefore attractive protein-protein interactions encouraged by PEG are at a maximum where electrostatic repulsion is at its minimum. Phase transitions were found for glucose isomerase and glucose oxidase at pH 5 and

human lysozyme at pH 7 and pH 9. No phase transition was found for lysozyme from chicken egg white.

PEG molecules are non-adsorbing and non-polar polymers. PEG influence has previously been described using the depletion attraction mechanism which was first proposed in 1958 by Asakura and Oosawa (Asakura and Oosawa, 1958) for colloidal particles in a suspension of macromolecules. The depletion attraction effect appears when the distance between protein molecules becomes smaller than the diameter of the polymer molecules. A concentration gradient of polymer between the inter-protein area and the bulk solution is the result. As a consequence of this "polymer depletion", water molecules are drawn out of the gap between the protein molecules resulting in an osmotic pressure change and encourages protein-protein interactions (Asakura and Oosawa, 1958). The extent of the depletion effect is dependent on PEG molecular weight and concentration. PEG at higher concentrations and with higher molecular weight induces a stronger depletion effect (Vivarès et al., 2002). Small PEG molecules with molecular size below the distance of two protein molecules can enter into the protein-protein interspace. This may help to shield possible attractive protein-protein interactions, thus lead to protein solution stabilization.

The consequence of PEG destabilization is protein aggregation in form of crystallization (McPherson, 1976), precipitation (Atha and Inghamg, 1981), or liquid-liquid phase separation (Annunziata et al., 2002). The latter is described to be the consequence of PEG destabilization and short-range attractive protein interactions (Hagen and Frenkel, 1994). This metastable phase separation state parts the solution into a dilute protein state and rich "dense" protein droplets (Dumetz et al., 2008).

In agreement with the described theories we found that PEG stabilizing and destabilizing interaction mechanisms are dependent on PEG molecular weight and concentration. The PEG induced attractive protein-protein interactions were only large enough to emerge in solution conditions where long-range electrostatic repulsive forces are at its minimum, close to the isoelectric point, and for the higher molecular weight PEG molecule, PEG 1000. This led to precipitation at high PEG concentration for glucose oxidase at pH 5 and human lysozyme at pH 7 and pH 9.

For glucose isomerase the PEG 1000 effect at pH 5 is more complex. The destabilizing effects are more pronounced: LLPS and spontaneous precipitation occurred at lower precipitant concentration when compared to PEG 300. LLPS occurred in the transition zone just before spontaneous precipitation. This is in agreement with literature describing this phase state as metastable (Dumetz et al., 2008). Skin formation was suppressed at high PEG 1000 concentrations. However, this was likely not a stabilizing effect of the PEG, but rather the result of PEG promoting protein-protein interactions leading to rapid precipitation in the native state rather than denaturing and skin formation. For the lower molecular weight PEG 300 a balance of

stabilizing and destabilizing effects was observed for glucose isomerase at pH 5. PEG 300 stabilized the protein solution by suppressing skin formation starting at the lowest examined PEG concentration. Destabilization was observed at the highest PEG concentration of 15 %(w/V) where LLPS formation occurred. This balance of stabilizing and destabilizing effects is in agreement with the proposed PEG mechanism at low molecular weight. PEG 300 is small enough to penetrate the inter-protein area. While separating the protein molecules it is proposed to stabilize the native protein structure leading to stabilization whilst suppressing skin formation. At higher PEG 300 concentration this stabilization is obscured by the destabilizing effect due to depletion attraction leading to formation of LLPS.

5. Conclusion

A method to generate protein phase diagrams in high throughput on an automated liquid handling station in microbatch scale was successfully developed. It was shown that thorough optimization of liquid handling parameters is critical for generation of high quality data. Buffering components need to be selected carefully, especially when high salt concentrations are used, in order to prevent unwanted pH shifts in the samples.

The method allowed for realization of an extensive and systematic study of protein phase behavior in a reasonable amount of time. Thereby the general knowledge on phase behavior of proteins, the influence of salts and the understanding of effects of extreme pH values and effects of PEG on protein structure and solution stability could be increased.

For the investigated proteins it was possible to induce all known phase transitions. The investigated salts thereby showed various effects on the protein phase behavior. It was shown that extremely acidic pH values induce specific phase transitions such as skin formation and gelation. Gelation and skin formation are assumed to be related to partial or complete protein denaturation.

To the best of our knowledge the here proposed and experimentally supported pH dependency of short-range attractive forces was not described in literature earlier.

It can be supported that the phase behavior with PEG precipitants is dependent on PEG molecular size, concentration and protein charge. For the lower molecular weight PEG 300 stabilizing effects could be found for glucose isomerase. At higher molecular weight, PEG 1000, destabilization was triggered at conditions close to the isoelectric point of glucose oxidase, glucose isomerase and human lysozyme. Here electrostatic repulsion is at its minimum and destabilizing effects due to depletion attraction can resume.

Acknowledgements

We gratefully acknowledge the expertise and support by Sigrid K. Hansen and Florian Dismer. We also gratefully acknowledge the financial support by the Federal Ministry of Education and Research (BMBF)(0315342B).

Supplementary material

A total of 64 protein phase diagrams were determined as experimental fundament for this work. Selected phase diagrams of each protein were described in detail in Section 3. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpharm.2014.12.027.

References

Ahamed, T., Esteban, B.N.A., Ottens, M., van Dedem, G.W.K., van der Wielen, L.A.M., Bisschops, M.A.T., Lee, A., Pham, C., Thömmes, J., 2007. Phase behavior of an intact monoclonal antibody. Biophysical journal 93, 610-619.

Anandakrishnan, R., Aguilar, B., Onufriev, A.V., 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic acids research 40, W537-541.

Annunziata, O., Asherie, N., Lomakin, A., Pande, J., Ogun, O., Benedek, G.B., 2002. Effect of polyethylene glycol on the liquid-liquid phase transition in aqueous protein solutions. Proceedings of the National Academy of Sciences of the United States of America 99, 14165-14170.

Asakura, S., Oosawa, F., 1958. Interaction between Particles Suspended in Solutions of Macromolecules. Journal of Polymer Science 33, 183-192.

Asherie, N., 2004. Protein crystallization and phase diagrams. Methods 34, 266-272.

Atha, D.H., Inghamg, K.C., 1981. Mechanism of Precipitation of Proteins by Polyethylene Glycols. The Journal of Biological Chemistry 256, 12108-12117.

Baynes, B.M., Trout, B.L., 2004. Rational design of solution additives for the prevention of protein aggregation. Biophysical journal 87, 1631-1639.

Beretta, S., Chirico, G., Baldini, G., 2000. Short-Range Interactions of Globular Proteins at High Ionic Strengths. Macromolecules 33, 8663-8670.

Boistelle, R., Astier, J.P., 1988. Crystallization mechanisms in solution. Journal of Crystal Growth 90, 14-30.

Chayen, N., Akins, J., Campbell-Smith, S., Blow, D.M., 1988. Solubility of glucose isomerase in ammonium sulphate solutions. Journal of Crystal Growth 90, 112-116.

Chayen, N.E., Saridakis, E., 2008. Protein crystallization: from purified protein to diffraction-quality crystal. Nature Methods 5, 147-153.

Collins, K.D., 2004. Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. Methods 34, 300-311.

Curtis, R.A., Prausnitz, J.M., Blanch, H.W., 1998. Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. Biotechnology and bioengineering 57, 11-21.

Curtis, R.A., Ulrich, J., Montaser, A., Prausnitz, J.M., Blanch, H.W., 2002. Protein-protein interactions in concentrated electrolyte solutions. Biotechnology and bioengineering 79, 367-380.

D'Arcy, A., Mac Sweeney, A., Stihle, M., Haber, A., 2003. The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. Acta Crystallographica Section D Biological Crystallography 59, 396-399.

Debye, P., Hückel, E., 1923. Zur Theorie der Elektrolyte. Physikalische Zeitschrift 24, 185-206.

Dumetz, A.C., Chockla, A.M., Kaler, E.W., Lenhoff, A.M., 2008. Protein phase behavior in aqueous solutions: crystallization, liquid-liquid phase separation, gels, and aggregates. Biophysical journal 94, 570-583.

Dumetz, A.C., Chockla, A.M., Kaler, E.W., Lenhoff, A.M., 2009. Comparative Effects of Salt, Organic, and Polymer Precipitants on Protein Phase Behavior and Implications for Vapor Diffusion. Crystal Growth & Design 9, 682-691.

Faber, C., Hobley, T.J., 2006. Measurement and Prediction of Protein Phase Behaviour and Protein-Protein-Interactions. Ph.D. thesis, Technical University of Denmark.

George, A., Wilson, W.W., 1994. Predicting protein crystallization from a dilute solution property. Acta crystallographica. Section D, Biological crystallography 50, 361-365.

Hagen, M.H.J., Frenkel, D., 1994. Determination of phase diagrams for the hard-core attractive Yukawa system. The Journal of Chemical Physics 101, 4093-4097.

Hofmeister, F., 1888. Zur Lehre von der Wirkung der Salze. Archiv für experimentelle Pathologie und Pharmakologie 25, 1-30.

Kalisz, H.M., Hecht, H.-J., Schomburg, D., Schmis, R.D., 1990. Crystallization and preliminary X-ray diffraction studies of a deglycosylated glucose oxidase from Aspergillus niger. Journal of Molecular Biology 213, 207-209.

Kröner, F., Hubbuch, J., 2013. Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application. Journal of chromatography A 1285, 78-87.

Leckband, D., Israelachvili, J., 2001. Intermolecular forces in biology. Quarterly reviews of biophysics 34, 105-267.

Lee, J.C., Lee, L.L.Y., 1981. Preferential solvent interactions between proteins and polyethylene glycols. The Journal of biological chemistry 256, 625-631.

Lin, Y.-B., Zhu, D.-W., Wang, T., Song, J., Zou, Y.-S., Zhang, Y.-L., Lin, S.-X., 2008. An Extensive Study of Protein Phase Diagram Modification: Increasing Macromolecular Crystallizability by Temperature Screening. Crystal Growth & Design 8, 4277-4283.

Lu, J., Wang, X.-J., Ching, C.-B., 2003. Effect of Additives on the Crystallization of Lysozyme and Chymotrypsinogen A. Crystal Growth & Design 3, 83-87.

Luft, J.R., DeTitta, G.T., 2009. Rational Selection of Crystallization Techniques, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 11-46.

McPherson, A., 1976. Crystallization of Proteins from Polyethylene Glycol. The Journal of Biological Chemistry 251, 6300-6303.

McPherson, A., 2001. A comparison of salts for the crystallization of macromolecules. Protein Science 10, 418-422.

McPherson, A., Cudney, B., 2006. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. Journal of structural biology 156, 387-406.

Oelmeier, S.A., Dismer, F., Hubbuch, J., 2011. Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation. Biotechnology and bioengineering 108, 69-81.

Pawar, S.A., Deshpande, V.V., 2000. Characterization of acid-induced unfolding intermediates of glucose/xylose isomerase. European journal of biochemistry / FEBS 267, 6331-6338.

Piazza, R., 2000. Interactions and phase transitions in protein solutions. Current Opinion in Colloid & Interface Science 5, 38-43.

Ries-Kautt, M.M., Ducruix, A.F., 1989. Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. The Journal of biological chemistry 264, 745-748.

Schwierz, N., Horinek, D., Netz, R.R., 2010. Reversed anionic Hofmeister series: the interplay of surface charge and surface polarity. Langmuir : the ACS journal of surfaces and colloids 26, 7370-7379.

Scopes, R.K., 1994. Separation by Precipitation, Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag New York, Inc., pp. 71-101.

Sleutel, M., Willaert, R., Gillespie, C., Evrard, C., Wyns, L., Maes, D., 2009. Kinetics and Thermodynamics of Glucose Isomerase Crystallization. Crystal Growth & Design 9, 497-504.

Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W.C.K., Egelhaaf, S.U., Schurtenberger, P., 2004. Equilibrium cluster formation in concentrated protein solutions and colloids. Nature 432, 492-495.

Tardieu, A., Bonneté, F., Finet, S., Vivarès, D., 2002. Understanding salt or PEG induced attractive interactions to crystallize biological macromolecules. Acta Crystallographica Section D Biological Crystallography 58, 1549-1553.

Velev, O.D., Kaler, E.W., Lenhoff, A.M., 1998. Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen. Biophysical journal 75, 2682-2697.

Vivarès, D., Belloni, L., Tardieu, a., Bonneté, F., 2002. Catching the PEG-induced attractive interaction between proteins. The European physical journal. E, Soft matter 9, 15-25.

Vlachy, V., Blanch, H.W., Prausnitz, J.M., 1993. Liquid-liquid phase separations in aqueous solutions of globular proteins. AIChE Journal 39, 215-223.

Vlachy, V., Prausnitz, J.M., 1992. Donnan Equilibrium. Hypernetted-Chain Study of One-Component and Multicomponent Models for Aqueous Polyelectrolyte Solutions. The Journal of Physical Chemistry 96, 6465-6469.

Wang, W., Li, N., Speaker, S., 2010. External Factors Affecting Protein Aggregation, in: Wang, W., Roberts, C.J. (Eds.), Aggregation of Therapeutic Proteins, 1st ed. John Wiley & Sons, pp. 119-204.

Zeelen, J.P., 2009. Interpretation of the Crystallization Drop Results, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 175-194.

Zhang, Y., Cremer, P.S., 2009. The inverse and direct Hofmeister series for lysozyme. Proceedings of the National Academy of Sciences of the United States of America 106, 15249-15253.

Ziegler, G.R., Foegeding, E.A., 1990. The Gelation of Proteins, in: Kinsella, J.E. (Ed.), Advances in Food and Nutrition Research. Academic Press, pp. 203-298.

3.2 Manipulation of lysozyme phase behavior by additives as function of conformational stability

Lara Galm, Josefine Morgenstern, Jürgen Hubbuch*

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

** Corresponding author. Tel.: +49 721 608 47526; fax: +49 721 608 46240. E-mail address: juergen.hubbuch@kit.edu.*

submitted to: ***International Journal of Pharmaceutics.***

Abstract

Undesired protein aggregation in general and non-native protein aggregation in particular need to be inhibited during biopharmaceutical processing to ensure patient safety and to maintain product activity. In this work the potency of different additives, namely glycerol, PEG 1000, and glycine, to prevent lysozyme aggregation and selectively manipulate lysozyme phase behavior was investigated. The results revealed a strong pH dependency of the additive impact on lysozyme phase behavior, lysozyme solubility, crystal size and morphology. This work aims to link this pH dependent impact to a protein-specific parameter, the conformational stability of lysozyme. At pH 3 the addition of 10 %(w/v) glycerol, 10 %(w/v) PEG 1000 and 1 M glycine stabilized or destabilized lysozyme's native conformation resulting in a modified size of the crystallization area without influencing lysozyme solubility, crystal size and morphology. Addition of 1 M glycine even promoted non-native aggregation at pH 3 whereas addition of PEG 1000 completely inhibited non-native aggregation. At pH 5 the addition of 10 %(w/v) glycerol, 10 %(w/v) PEG 1000 and 1 M glycine did not influence lysozyme's native conformation, but strongly influenced the position of the crystallization area, lysozyme solubility, crystal size and morphology. The observed pH dependent impact of the additives could be linked to a differing lysozyme conformational stability in the binary systems without additives at pH 3 and pH 5. However, in any case lysozyme phase behavior could selectively be manipulated by addition of glycerol, PEG 1000 and glycine. Furthermore, at pH 5 crystal size and morphology could selectively be manipulated.

Keywords: *Osmolytes, Additives, Solubility line, Crystallization area, Non-native aggregation, FT-IR*

1. Introduction

The term protein aggregation describes the assembly of native or non-native protein monomers to protein multimers, i.e. aggregation characterizes both the formation of protein crystals and amorphous precipitates and includes native and non-native aggregation forms. Protein aggregation can occur through different mechanisms (Philo and Arakawa, 2009) and during different steps of a production process (Cromwell et al., 2006). However, crystallization and precipitation are also acknowledged process steps in biopharmaceutical industries either for formulation or purification purposes (Scopes, 1994). Crystalline drug formulations for example

have shown significant benefits in the delivery of protein therapeutics to achieve high-concentration, high-stability, low-viscosity and controlled-release formulations (Basu et al., 2004; Jen and Merkle, 2001). Crystalline insulin formulations are market approved (Basu et al., 2004; Brange and Vølund, 1999; Vajo et al., 2001) and crystalline antibody formulations are studied, too (Yang et al., 2003). For formulated protein therapeutics, the presence of precipitates is typically considered to be undesirable because of the concern that especially non-native precipitates may lead to immunogenic reactions (Cromwell et al., 2006). The widespread opinion exists that aggregation processes are usually associated with a conformational change, i.e. partial unfolding of the proteins (Chi et al., 2003; Fink, 1998) and aggregation processes that resulted in non-native protein conformations have been observed (Dong et al., 1995; Dzwolak et al., 2003; Feng et al., 2012; Kendrick et al., 1998; Matheus et al., 2009). Moreover, aggregation processes might influence biological activity of protein therapeutics. Thus, in either case it is essential to ensure that the target protein remains in its native conformation and that biological activity is preserved despite aggregation. Thus, in cases where non-native aggregation is likely to occur aggregation needs to be prevented completely unless there are possibilities to stabilize the native conformational state. In cases where native aggregation occurs, the selective control of phase states is considered to be beneficial as sometimes either crystalline or precipitated forms are preferred e.g. due to a better bioavailability in the respective aggregate state (Vajo et al., 2001). Particular additives are thought to stabilize the proteins' native state, for example stabilize the protein against thermal denaturation, and thus might be used to prevent non-native aggregation processes. According to Harris and Rösgen (Harries and Rösgen, 2008) particular additives influence protein solubility as well, resulting in a manipulation of the protein phase behavior, i.e. protein aggregation might be completely inhibited or protein phase states (e.g. crystallization, precipitation) might be selectively changed. Frequently used additives are polymers (polyethylene glycol, PEG) and osmolytes. Osmolytes are low molecular weight additives, that can be grouped into the major categories of free amino acids and derivates (e.g. glycine), polyols and uncharged sugars (e.g. glycerol), methylamines, and urea (Yancey, 2001). The impact of these additives on protein stability is described to be due to a preferential binding or a preferential exclusion of the additives from the proteins' local domain. In cases where they are preferentially excluded from the proteins' local domain they are known to stabilize the proteins' native state. (Arakawa and Timasheff, 1982, 1983, 1985a; Arakawa and Timasheff, 1985b; Lee and Lee, 1981; Timasheff and Arakawa, 1988; Webb et al., 2001) The impact of additives on protein solubility is according to Harris and Rösgen (Harries and Rösgen, 2008) not as easy as predicting protein stability and no general models exist. Though, the mode of action of osmolytes and PEG on protein stability is as well not as easy as it might sound since for some additives stabilizing as well as destabilizing effects have been observed. Parameters that strongly influence the stabilizing or destabilizing character of additives are the additive concentration,

the molecular weight of the additive, and the solvent pH. PEG, dimethylglycine, and betaine for example have been found to stabilize proteins up to a certain concentration and destabilize them for higher concentrations (Arakawa and Timasheff, 1985a; Santoro et al., 1992). High molecular weights of PEG have been found to be destabilizing as well (Lee and Lee, 1981), whereas high molecular weight polyols stabilize proteins better than their low molecular weight counterparts (Poddar et al., 2008). Additive impact as function of solvent pH is even harder to generalize. According to Singh et al. (Singh et al., 2011), polyols have a higher potency to stabilize the proteins' native state at low pH, whereas methylamines are described to act most stabilizing at neutral pH and destabilizing at low pH and amino acids were found to stabilize the native state of proteins almost independent of the pH (Macchi et al., 2012). This pH dependent osmolyte action has been related to the chemical nature of the osmolytes, e.g. the pK_a values (Granata et al., 2006; Natalello et al., 2009; Singh et al., 2009), but could not explain all pH dependent observations (Kaushik and Bhat, 2003). In contrast, there are indications that the pH dependent mode of action of additives might have its origin in the nature of the proteins instead (Kaushik and Bhat, 2003; Macchi et al., 2012). It is for example known that proteins at extreme pH values are conformationally unstable, i.e. prone to at least partial unfolding (Wang et al., 2010). Thus, our work aims to find a link between the pH dependent mode of action of additives and the conformational stability of a model protein in its initial state without additive. To the best of our knowledge up to now there are no investigations on the pH dependent mode of action as a function of conformational stability of proteins in their initial state without additive. Furthermore, this publication examines if the additives impact protein solubility and thus protein phase behavior. Lysozyme from chicken egg white was studied as a model protein, but the presented approaches can easily be transferred to other biopharmaceutical proteins. Lysozyme was investigated at pH 3 and pH 5. Sodium chloride was added as precipitant to induce phase transitions of lysozyme (e.g. crystallization and precipitation) in order to study the phase behavior of lysozyme. In the following the term binary describes lysozyme in aqueous sodium chloride solutions ranging from 0 M to 2.5 M sodium chloride. This was also referred to as the initial state of lysozyme above. The term ternary in the following describes lysozyme in aqueous sodium chloride solutions ranging from 0 M to 2.5 M sodium chloride and with a constant additive concentration. Lysozyme conformational stability, solubility, and phase behavior in these ternary systems (with additive) will be compared to lysozyme conformational stability, solubility, and phase behavior in the binary systems (without additive), i.e. to lysozyme's initial state. A schematic illustration of the experimental setup that was employed during this work is shown in Fig. 1. Glycerol and glycine as additives were chosen as representatives of two osmolyte classes and PEG 1000 as additive beyond the osmolyte class. Fourier transform infrared (FT-IR) spectroscopy was applied to monitor lysozyme conformation and to account for non-native conformational changes. This allows evaluating the impact of the

additives on conformational stability and their potency to stabilize or destabilize the proteins' native state. Ternary phase diagrams, consisting of lysozyme, sodium chloride and the respective additive as solution components, were generated and compared to binary ones, consisting of lysozyme and sodium chloride. The comparison of the phase diagrams reveals information about how strong the additives manipulate lysozyme phase behavior, i.e. if they completely prevent aggregation, delay it or if they can be used to selectively control phase states, e.g. transfer former precipitated to crystalline phase states. The phase diagrams additionally allow to experimentally determine lysozyme solubility in cases where crystallization occurs as the supernatant of a crystalline solution is saturated and the lysozyme concentration in the supernatant thus reflects lysozyme solubility (Asherie, 2004; Howard et al., 1988; Retailleau et al., 1997). Experimentally determined solubility data points are fitted to an empirically found equation, resulting in continuous solubility lines. Comparison between the binary and ternary systems gives the additive impact on lysozyme solubility lines.

Altogether this publication aims to elucidate the potential origin of pH dependent additive action and tries to expand the basic knowledge on additive impact on protein solubility and phase behavior.

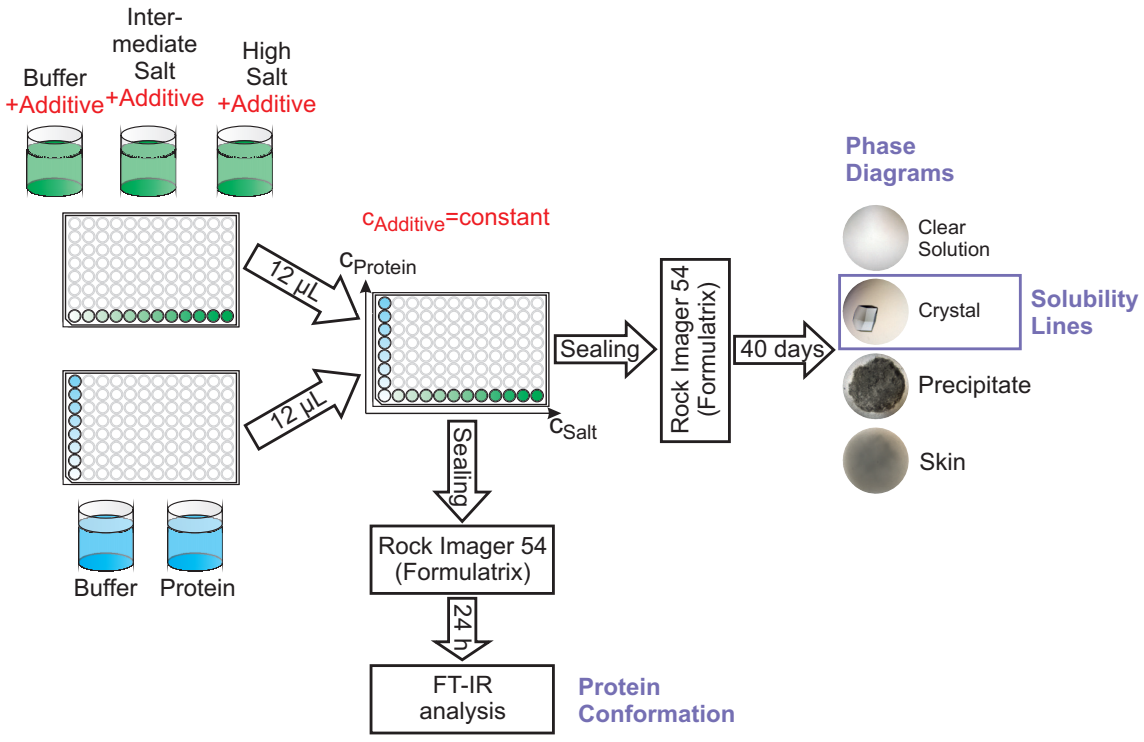

Fig. 1. Schematic illustration of the experimental setup that was employed during this work for investigation of the additive impact on conformational stability, phase behavior and solubility of lysozyme at pH 3 and pH 5.

2. Materials and methods

2.1. Materials

The used buffer substances were citric acid (Merck, Darmstadt, Germany) and sodium citrate (Sigma-Aldrich, St. Louis, MO, USA) for pH 3 and sodium acetate (Sigma-Aldrich, St. Louis, MO, USA) and acetic acid (Merck, Darmstadt, Germany) for pH 5. PEG 300 and PEG 1000 were obtained from Sigma-Aldrich (St. Louis, MO, USA). Sodium chloride as well as glycine were purchased from Merck (Darmstadt, Germany), glycerol was from Alfa Aesar (Ward Hill, MA, USA) and lactose from Carl Roth GmbH & Co. KG (Karlsruhe, Germany). Hydrochloric acid and sodium hydroxide for pH adjustment were obtained from Merck (Darmstadt, Germany). pH adjustment was performed using a five-point calibrated pH-meter (HI-3220, Hanna Instruments, Woonsocket, RI, USA). All buffers were filtered through 0.2 µm cellulose acetate filters (Sartorius, Goettingen, Germany).

Lysozyme from chicken egg white was purchased from Hampton Research (Aliso Viejo, CA, USA). The lysozyme solutions were filtered through 0.2 µm syringe filters with cellulose acetate membranes (VWR, Radnor, PA, USA) previous to further desalting via size exclusion chromatography. Size exclusion chromatography was conducted using a HiTrap Desalting Column (GE Healthcare, Uppsala, Sweden) on an AEKTAprime™ plus system (GE Healthcare, Uppsala, Sweden). A subsequent protein concentration step was performed using Vivaspin® centrifugal concentrators (Sartorius, Goettingen, Germany) with PES membranes and molecular weight cutoffs of 3 kDa.

Binary and ternary protein phase diagrams, consisting of lysozyme and sodium chloride or lysozyme, sodium chloride and additive as solution components, were prepared on MRC Under Oil 96 Well Crystallization Plates (Swissci, Neuheim, Switzerland) in 24 µL scale microbatch experiments with a Freedom EVO® 100 (Tecan, Maennedorf, Switzerland) automated liquid handling station. The MRC Under Oil 96 Crystallization Plates were covered with HDclear sealing tape (ShurTech Brands, Avon, OH, USA) to prevent evaporation. The sealed plates were stored at 20 °C and for 40 days in a Rock Imager 182/54 (Formulatrix, Waltham, MA, USA) automated imaging system.

Lysozyme concentration in the supernatant of crystalline phase states was determined using a NanoDrop 2000c UV-Vis spectroscopic device (Thermo Fisher Scientific, Waltham, MA, USA). An extinction coefficient of $\varepsilon_{280nm}^{1\%} = 22.00$ was used.

Fourier transform infrared (FT-IR) spectra of selected binary and ternary lysozyme systems were measured with a Bruker Tensor 27 (Bruker Corporation, Billerica, MA, USA) Fourier transform infrared (FT-IR) spectrophotometer with CONFOCHECK configuration at 20 °C. Infrared analysis was conducted using the attenuated total reflection technique in a Bio-ATR II cell (Bruker Corporation, Billerica, MA, USA).

2.2 Methods

2.2.1. Preparation of stock solutions

To set up the buffers all substances were weighed in and dissolved in ultrapure water of approximately 90 % of the final buffer volume. Buffer capacity was 100 mM for all buffers. pH was adjusted with the appropriate titrant with an accuracy of ±0.05 pH units. After pH adjustment the buffers were brought to their final volume using ultrapure water. All buffers were filtered through 0.2 µm cellulose acetate filters. Buffers were used at the earliest one day after preparation and after repeated pH verification. Buffer capacity was 100 mM for all buffers.

To set up the protein stock solutions, protein was weighed in and dissolved in the appropriate buffer. The protein solution then was filtered through a 0.2 µm syringe filter, and further desalted using size exclusion chromatography. Protein concentration was adjusted to 65.25 ± 1 mg/mL via centrifugal concentrators. A volume of 1 mL of protein stock solution was required for generation of one ternary phase diagram.

2.2.2. Generation of phase diagrams

Binary protein phase diagrams for lysozyme from chicken egg white were generated earlier (Baumgartner et al., 2015). Binary phase diagrams of lysozyme from chicken egg white at pH 3 and pH 5 using sodium chloride as precipitant showed multiple phase transitions. Therefore, lysozyme in aqueous sodium chloride solutions at pH 3 and pH 5 was selected for further investigation in ternary phase diagrams using 1 M glycine, 10 %(m/v) glycerol and 10 %(m/v) PEG 1000 as additives. In total six ternary phase diagrams were generated. Additive concentration was constant in any of the 96 wells on the crystallization plates for the respective pH and additive type. Only lysozyme and sodium chloride concentrations were varied in exactly the same way as in the binary phase diagrams. The resulting protein and sodium chloride concentrations on the crystallization plates thus ranged between 2.5 and 21.75 mg/mL lysozyme and between 0 and 2.5 M sodium chloride. For generation of 96 different ternary system points on the crystallization plates 12 µL of solutions with varying sodium chloride concentration and a fixed concentration of additive were mixed with 12 µL of solutions with varying lysozyme concentrations in the respective buffer without sodium chloride and additive. After mixing crystallization plates were treated as described by Baumgartner et al. (Baumgartner et al., 2015). The images after 40 days were visually examined to distinguish between the following phase states: soluble, crystalline, precipitated, skin formation, gelation, liquid-liquid phase separation.

2.2.3. Determination of solubility lines

After 40 days and after scoring of the phase states the plates were removed from the Rock Imager. In cases where crystalline phase states occurred, lysozyme concentration in the supernatant of these crystalline solutions was determined. The supernatant of a crystalline

solution is saturated and the lysozyme concentration in the supernatant thus reflects lysozyme solubility (Asherie, 2004; Howard et al., 1988; Retailleau et al., 1997). Therefore, experimental data points for lysozyme concentration in the supernatant of crystalline solutions will be referred to as solubility hereafter and were used to fit solubility lines. For determination of lysozyme solubility 3 µL of the supernatant of crystalline solutions were transferred to a Nanodrop 2000c UV-Vis spectroscopic device to determine the lysozyme concentration in the supernatant. This was conducted two- or threefold depending on the available supernatant volume. Values were averaged and solubility lines were fitted using Eq. (1) resulting from a Box Lucas 1 model with S being the protein solubility, S_0 the theoretical protein solubility for a sodium chloride concentration of 0 M. A and R_0<0 are adaptable parameters and c_{NaCl} the sodium chloride concentration.

$$\boldsymbol{S = S_0 + A * e^{(R_0 * c_{NaCl})}} \quad \textbf{(1)}$$

2.2.4. FT-IR analysis

Lysozyme solutions were scanned 24 h after preparation in absorbance mode with 700 scans at a spectral resolution of 2 cm^{-1}. Spectra were recorded from 4000 to 1000 cm^{-1}. Background spectra were recorded for the respective solution the protein was dissolved in also with 700 scans at a spectral resolution of 2 cm^{-1}. Sample volume was 20 µL. All measurements were conducted in duplicate. The OPUS 6.5 spectroscopy software was used for recording and data processing of the recorded FT-IR spectra. Data pre-processing was performed as follows: atmospheric compensation and vector normalization (Euclidean norm) were carried out to delete the influence of water vapor bands on the spectra and to delete protein concentration dependent effects. After data pre-processing second derivative spectra were calculated. Calculation of second derivative spectra was conducted in a wavenumber region from 1700 to 1600 cm^{-1} and with 25 smoothing points. Changes in the second derivate spectra correlate to conformational changes of lysozyme.

3. Results

3.1. FT-IR spectra of binary lysozyme systems at pH 3 and pH 5

The upper left section of Fig. 2 shows results from FT-IR spectroscopic measurements of lysozyme at pH 3 in different binary systems (without additive). FT-IR second derivative spectra of soluble, crystalline, crystalline and coexisting precipitated phase states of lysozyme and for phase states where skin formation occurred are shown. The different phase states were induced by addition of different concentrations of sodium chloride following the conditions from the phase diagrams. For the binary systems at pH 3 the spectrum of crystalline lysozyme was very similar to the native soluble one. The sample with lysozyme in coexisting crystalline and

precipitated states deviated slightly from the native soluble spectrum but the most evident deviation occurred for the sample with skin formation. Here, a significant deviation occurred in a wavenumber region between 1640 and 1620 cm^{-1}, which corresponds to the β–sheet region. The upper left section of Fig. 3 shows results from FT-IR spectroscopic measurements of lysozyme at pH 5 in different binary systems (without additive). FT-IR second derivative spectra of soluble, crystalline, precipitated phase states of lysozyme are shown. At pH 5 the second derivative spectra of the investigated phase states in the binary systems were very similar, even the precipitated state showed only small deviations from the native soluble state.

3.2. Impact of investigated additives on lysozyme conformation as function of pH

3.2.1. Impact on lysozyme conformation at pH 3

The upper right and the lower sections of Fig. 2 show the second derivative spectra of the FT-IR spectroscopic measurements of lysozyme at pH 3 in different ternary systems (with additives). Addition of 10 %(m/v) glycerol (Fig. 2, upper right) smoothed the deviation between the second derivative spectra of different phase states in comparison to the spectra of different phase states in binary systems (Fig. 2, upper left). The spectrum for coexisting crystallization and precipitation was much more similar to the native soluble one than in the binary case. Even the spectrum for the sample with skin formation was more similar to the soluble, native spectrum than in the binary case. The same accounted for the addition of 10 %(m/v) PEG 1000 (Fig. 2, lower right). Deviation between the second derivative spectra of different phase states was smoothed by addition of PEG 1000. In contrast, addition of 1 M glycine (Fig. 2, lower left) increased deviation between the second derivative spectra of the different phase states. The spectrum of the sample with coexisting crystalline and precipitated lysozyme deviated stronger from soluble, native spectrum than in the binary case, especially in a wavenumber region between 1660 and 1640 cm^{-1}. A decrease of the peak depth in this region corresponds to a dissipation of α–helix structures. The spectrum of the sample with skin formation again deviated strongest in the wavenumber region between 1640 and 1620 cm^{-1}.

The FT-IR analysis results for pH 3 showed that addition of glycerol and PEG 1000 reduced deviations between the second derivative FT-IR spectra of different phase states. Glycine in contrast increased deviations between the second derivative FT-IR spectra compared to the samples without additive.

3.2.2. Impact on lysozyme conformation at pH 5

The upper right and the lower sections of Fig. 3 show the second derivative spectra of the FT-IR spectroscopic measurements of lysozyme at pH 5 in different ternary systems (with additives). The addition of 10 %(m/v) glycerol, 10 %(m/v) PEG 1000 and 1 M glycine hardly influenced the

second derivative FT-IR spectra of lysozyme in different phase states. Even for addition of glycine the spectra hardly deviated.

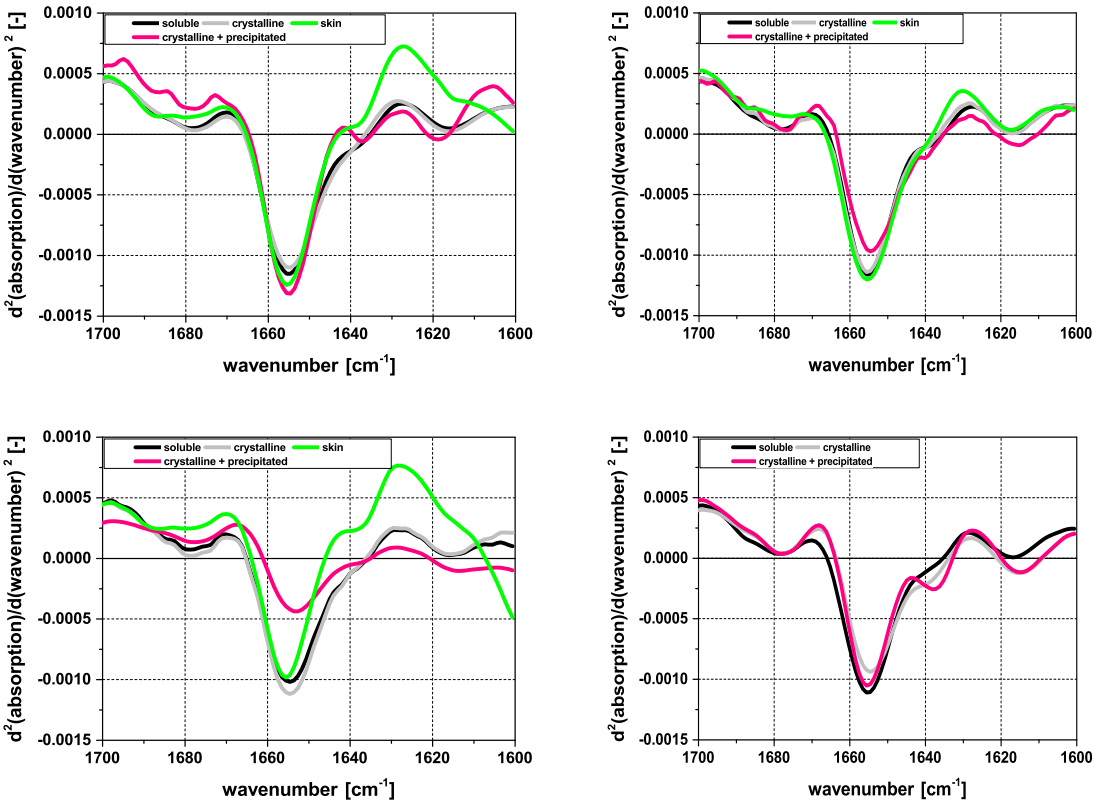

Fig. 2. Second derivative spectra from FT-IR measurements of lysozyme at pH 3 with different sodium chloride concentrations and without additive (upper left), with 10 %(m/v) glycerol (upper right), with 1 M glycine (lower left), with 10 %(m/v) PEG 1000 (lower right).

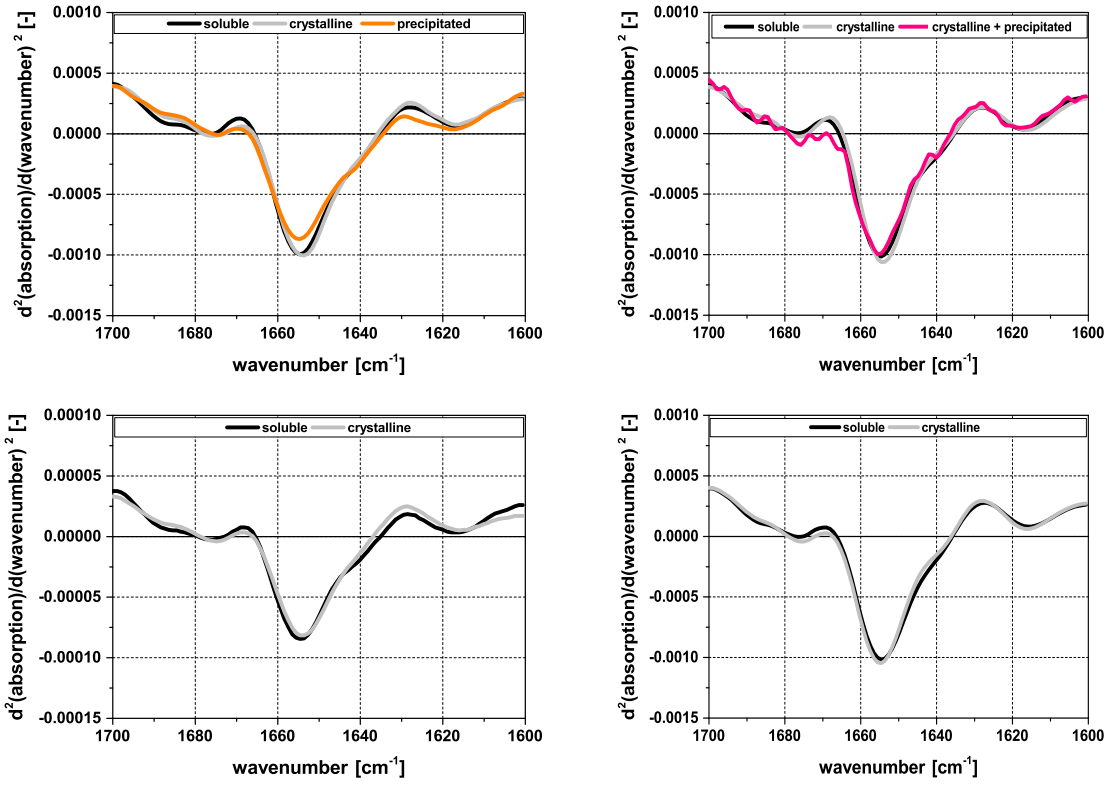

Fig. 3. Second derivative spectra from FT-IR measurements of lysozyme at pH 5 with different sodium chloride concentrations and without additive (upper left), with 10 %(m/v) glycerol (upper right), with 1 M glycine (lower left), with 10 %(m/v) PEG 1000 (lower right).

3.3. Impact of investigated additives on lysozyme phase behavior as function of pH

For examination of the impact of selected additives on lysozyme phase behavior ternary phase diagrams with lysozyme, sodium chloride and additive as solution components were compared to binary phase diagrams, consisting of lysozyme and sodium chloride, at pH 3 and pH 5. In particular, the sodium chloride concentration for initiation of crystallization, the size of the crystallization area and the crystal size and morphology were compared.

3.3.1. Impact on lysozyme phase behavior at pH 3

Fig. 4 shows the binary (upper left) and ternary (upper right, lower right, and lower left) phase diagrams of lysozyme at pH 3. The sodium chloride concentration for initiation of lysozyme crystallization did not change by addition of glycerol, PEG 1000 or glycine. In any case lysozyme crystallization started at 0.45 M sodium chloride. This said, the lysozyme concentration for crystallization was influenced and was higher than in the binary case for addition of 10 %(m/v) glycerol (upper right) and lower than in the binary case for addition of 10 %(m/v) PEG 1000 and 1 M glycine (lower right and left). Adding 10 %(m/v) glycerol to the solutions resulted in a wider crystallization area, a smaller precipitation area and reduced skin formation. Glycerol reduced

the number of conditions where skin formation occurred to 4 in contrast to 20 in the binary phase diagram. Addition of 10 %(m/v) PEG 1000 enlarged the size of the crystallization area and delayed precipitation even further than glycerol. The most remarkable difference was that the addition of PEG 1000 completely inhibited skin formation. The addition of 1 M glycine decreased the size of the crystallization area dramatically and enlarged the precipitation area as well as the area where skin formation occurred. The number of conditions where skin formation occurred was increased to 28.

In summary, at pH 3 the additives did not strongly affect the sodium chloride concentration needed for initiation of aggregation (crystallization) but they did influence the size of the crystallization and precipitation area, and the size of the area where skin formation occurred. Thus, the phase states or rather the aggregate states could be successfully selectively manipulated, for example from former precipitated states to crystalline ones. No significant influence on crystal size and morphology was observed for addition of glycerol, PEG 1000 or glycine as can be seen in Fig. 6(a-d).

3.3.2. Impact on lysozyme phase behavior at pH 5

Fig. 5 shows the binary (upper left) and ternary (upper right, lower right, and lower left) phase diagrams of lysozyme at pH 5. In contrast to pH 3 the sodium chloride concentration for initiation of lysozyme crystallization was influenced by addition of glycerol and glycine. The addition of 10 %(m/v) glycerol at pH 5 caused a shift of the crystallization area to higher sodium chloride concentrations. The addition of 1 M glycine shifted the crystallization area slightly towards lower sodium chloride concentrations. PEG 1000 did not shift the sodium chloride concentration for initiation of lysozyme crystallization but shifted the crystallization area to higher lysozyme concentrations. Moreover, addition of PEG 1000 caused a crystallization area that was interspersed with soluble conditions. These shifts of the crystallization areas caused a decrease of the size of the crystallization areas by addition of 10 %(m/v) glycerol and 10 %(m/v) PEG 1000 and an increase by addition of 1 M glycine. Addition of 10 %(m/v) glycerol and 10 %(m/v) PEG 1000 turned some former crystalline phase states to soluble phase states, i.e. inhibited aggregation. Compared to the binary phase diagram with 51 crystalline phase states and one precipitated phase state, addition of 10 %(m/v) glycerol reduced the number of crystalline phase states to 27 and one phase state with coexisting crystallization and precipitation. Addition of 10 %(m/v) PEG 1000 reduced the number of crystalline phase states to 31 and completely inhibited precipitation. Addition of 1 M glycine increased the number of crystalline phase states to 61 but also completely inhibited precipitation. The influence of the investigated additives on crystal size and morphology was strong as can be seen in Fig. 6(e-h). In the binary systems three-dimensional crystals with a size of averagely 400 µm occurred up to sodium chloride concentrations of 2.05 M (exemplarily shown in Fig. 6(e)). For higher sodium

chloride concentrations small, sea-urchin shaped two-dimensional crystals evolved instead. Addition of 10 %(m/v) glycerol and PEG 1000 (Fig. 6(f) and Fig. 6(g)) caused the formation of single, three-dimensional crystals with averagely twice the size of the crystals in the binary case over the whole sodium chloride concentration range. Glycine in contrast inhibited the formation of three-dimensional crystals from 1.36 M sodium chloride (Fig. 6(h)) and small, sea-urchin shaped two-dimensional crystals evolved instead.

In summary, at pH 5 the additives strongly affected either the sodium chloride concentration and/or the lysozyme concentration for initiation of aggregation (crystallization), i.e. caused a shift of the crystallization area. This as well influenced the size of the crystallization area. Thus, aggregation could be successfully inhibited or accelerated. Furthermore, a striking influence on lysozyme crystal size and morphology was observed.

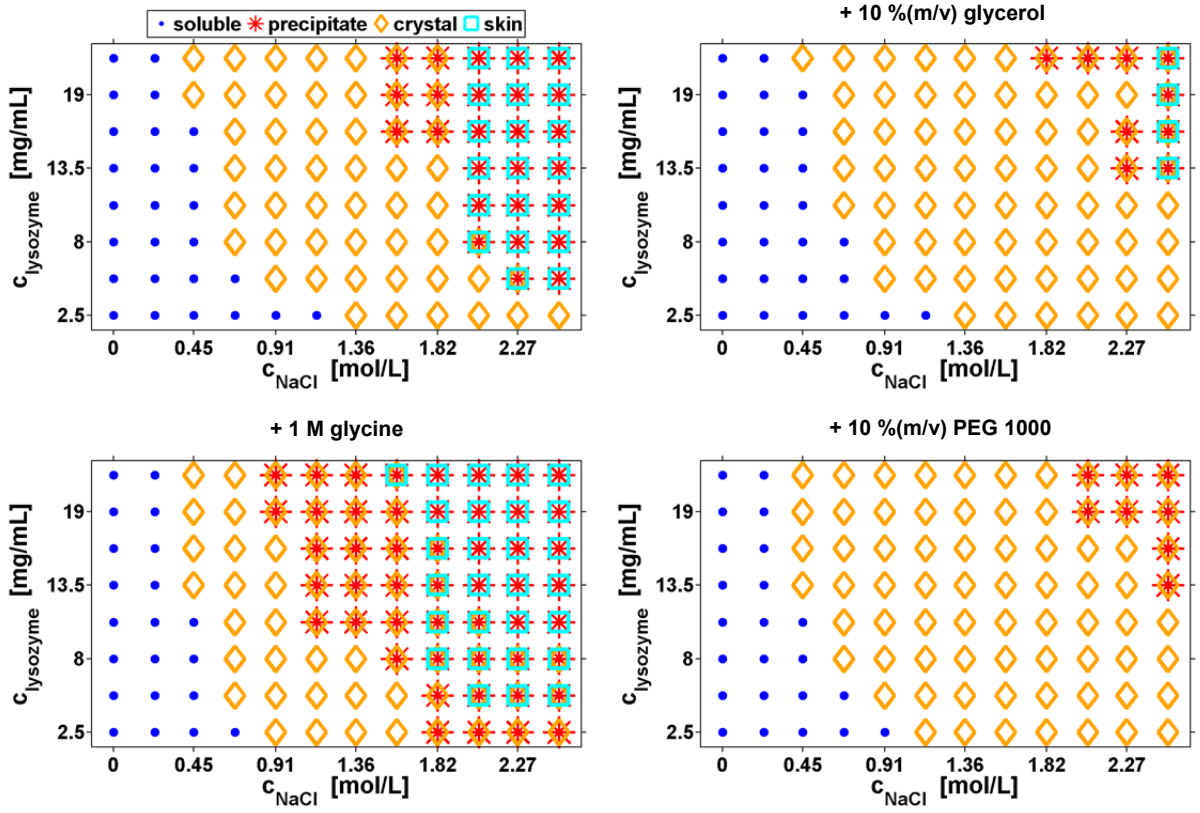

Fig. 4. Phase diagrams of lysozyme at pH 3 depending on sodium chloride concentration without additive (upper left), with 10 %(m/v) glycerol (upper right), with 1 M glycine (lower left), and with 10 %(m/v) PEG 1000 (lower right).

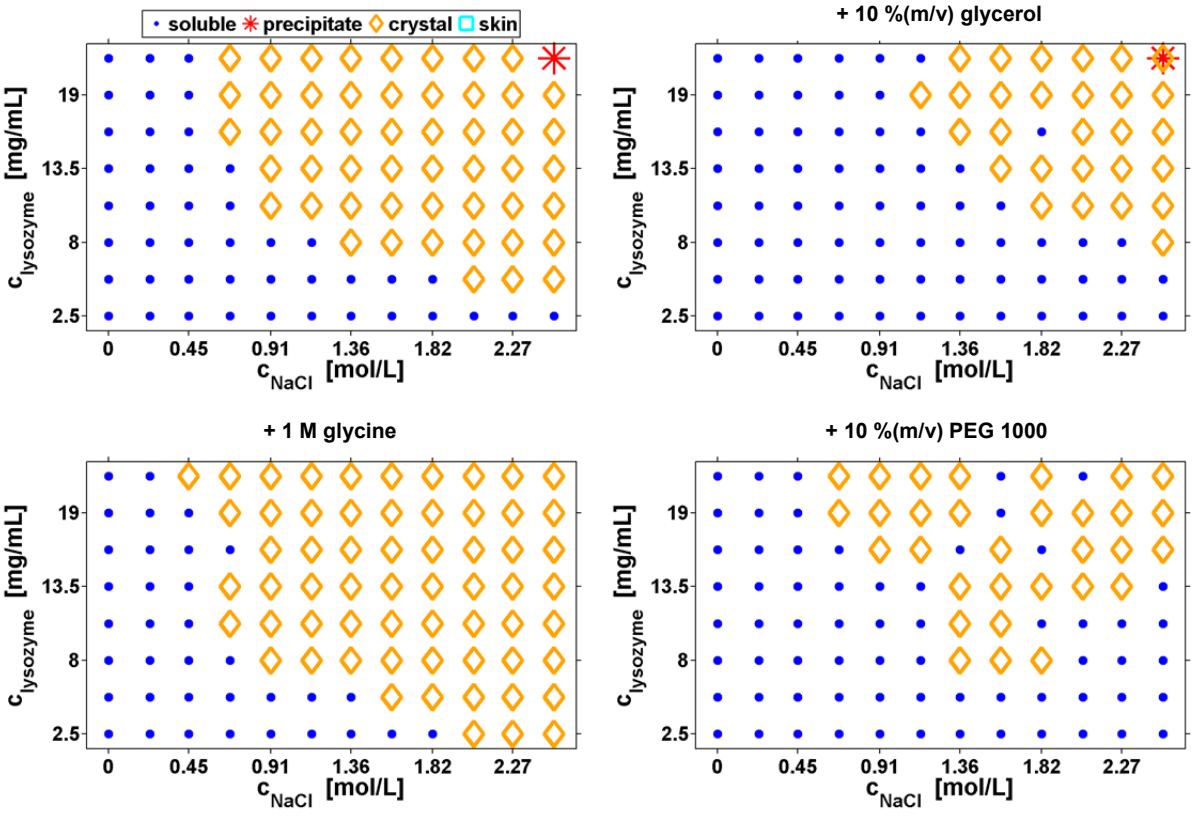

Fig. 5. Phase diagrams of lysozyme at pH 5 depending on sodium chloride concentration without additive (upper left), with 10 %(m/v) glycerol (upper right), with 1 M glycine (lower left), and with 10 %(m/v) PEG 1000 (lower right).

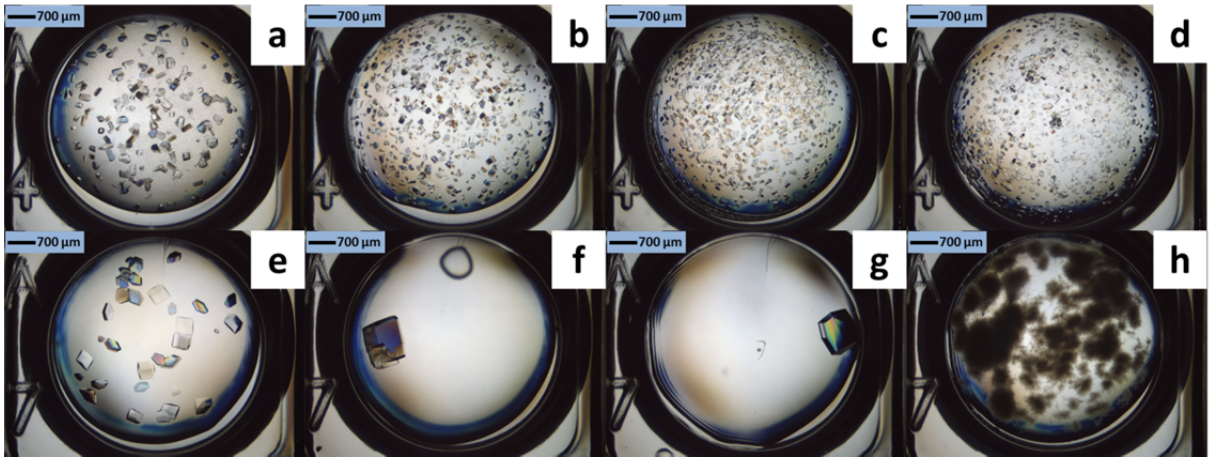

Fig. 6. Phase states of 21.75 mg/mL lysozyme after 40 days at pH 3 with 0.68 M sodium chloride (upper row) and pH 5 with 1.36 M sodium chloride (lower row). Additive concentrations: without additive (a), 10 %(m/v) glycerol (b), 10 %(m/v) PEG 1000 (c), 1 M glycine (d), without additive (e), 10 %(m/v) glycerol (f), 10 %(m/v) PEG 1000 (g), 1 M glycine (h).

3.4. Impact of investigated additives on lysozyme solubility

3.4.1. Determination of lysozyme solubility and fit of solubility lines

Lysozyme concentration in the supernatant of crystalline solutions at pH 3 could be determined with a maximum relative deviation of 6.4 % for the binary system and 13 % for the ternary systems. Lysozyme concentration in the supernatant of crystalline solutions at pH 5 could be determined with a maximum relative deviation of 16.9 % for the binary system and 24.9 % for the ternary systems. Inaccuracies could have been due to development of incomplete equilibria in the ternary systems as described by Asherie (Asherie, 2004). There, the presented approach to determine protein solubility starting with a supersaturated solution is described to be more difficult and less accurate than to dissolve a protein crystal in an undersaturated protein solution. This is however not possible with the experimental approach described in this study. Nevertheless, solubility lines could be fitted to these experimental solubility data points with a high accuracy using Eq. (1) which arose from a Box Lucas 1 fit model. This fit model was found to give the best approximation to the solubility data points, but exponential correlations between protein solubility and salt concentration as described in Eq. (1) were empirically found earlier, too, employed for example in the Cohn equation (Cohn, 1925). Eq. (1) though could even fit the ternary systems with a high accuracy as can be seen in Tab. 1, i.e. the addition of additives did not influence the validity of the used exponential correlation. The only exception from this was the solubility line for the ternary system consisting of lysozyme, sodium chloride and 10 %(m/v) glycerol at pH 5. In this case no correlation between the experimental solubility data points and Eq. (1) could be found. In the other cases the coefficient of correlation was above 0.98. The fit parameters as well as the coefficients of correlation for the individual binary and ternary systems are listed in Tab. 1. The calculated lysozyme solubility lines as well as the experimental data points are shown in Fig. 7.

3.4.2. Impact on lysozyme solubility at pH 3

At pH 3 (Fig. 7, left) the addition of glycerol, PEG 1000 and glycine did not significantly influence lysozyme solubility in comparison to the binary system. The addition of 10 %(m/v) glycerol slightly enhanced lysozyme solubility in a sodium chloride concentration range between 0.68 and 1.82 M. The addition of 1 M glycine slightly decreased the lysozyme solubility below 0.68 M sodium chloride. The experimental data points as well as the fitted solubility line for the ternary system with PEG 1000 are almost identical to the binary system.

3.4.3. Impact on lysozyme solubility at pH 5

At pH 5 (Fig. 7, right) the addition of glycerol, PEG 1000 and glycine significantly influenced the lysozyme solubility. The addition of 10 %(m/v) glycerol caused an enhancement of the lysozyme solubility for sodium chloride concentrations from 1.14 M to 2.5 M. The addition of 10 %(m/v)

PEG 1000 enhanced lysozyme solubility in the whole investigated sodium chloride concentration range: The addition of glycine increased lysozyme solubility up to 0.68 M sodium chloride and decreased lysozyme solubility for sodium chloride concentrations from 0.91 M to 2.5 M. Thus, the effect of glycine on lysozyme solubility at pH 5 depended on the sodium chloride concentration.

Tab. 1. Fit parameters and coefficients of correlation for fit of the experimental data to Eq. (1).

	pH	S_0	A	R_0	R^2
without additive	3	0.956	115.707	-4.771	0.988
+10 %(m/v) glycerol	3	1.272	48.950	-2.941	0.991
+1 M glycine	3	0.835	61.116	-3.982	0.993
+10 %(m/v) PEG 1000	3	1.428	199.846	-5.643	0.994
without additive	5	3.296	42.615	-3.400	0.989
+10 %(m/v) glycerol	5	-	-	-	-
+1 M glycine	5	2.080	125.102	-4.413	0.995
+10 %(m/v) PEG 1000	5	5.891	88.651	-3.711	0.988

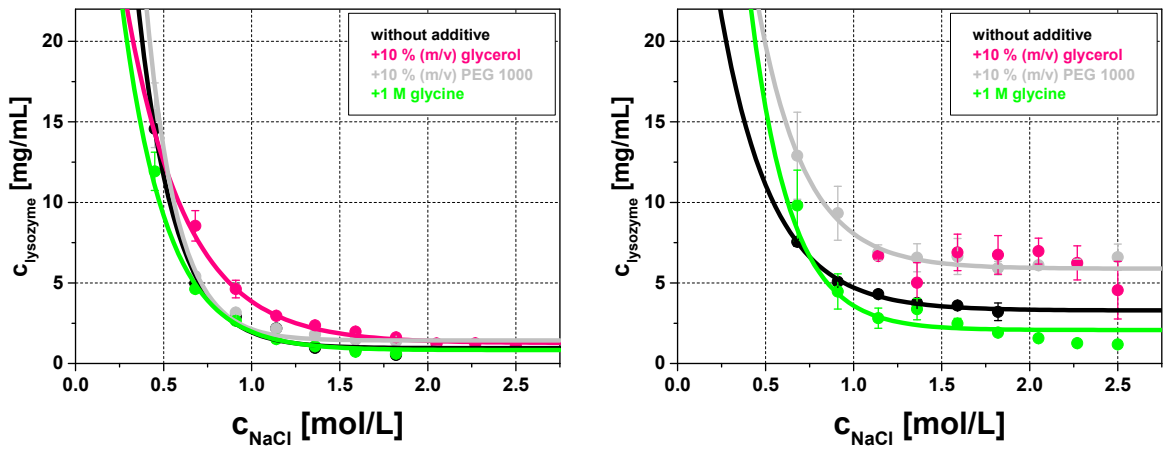

Fig. 7. Solubility lines of lysozyme at pH 3 (left) and pH 5 (right) depending on sodium chloride concentration without additive and with 10 %(m/v), 1 M glycine and 10 %(m/v) PEG 1000. The solid lines depict the solubility lines fitted using equation (1) with fit parameters as provided in Tab. .

4. Discussion

4.1. Impact of the additives on lysozyme phase behavior, conformation and solubility at pH 3

Deviation between the FT-IR spectra of different phase states of lysozyme in the binary systems at pH 3 is high, i.e. conformational instability of lysozyme at pH 3 in different phase states is high. Even non-native changes in lysozyme conformation during aggregation could be observed in the binary system as the spectrum of the sample with skin formation showed a significant deviation and according to Zeelen (Zeelen, 2009) this skin is a layer of denatured protein. Addition of

glycerol and PEG 1000 reduced the deviation between the FT-IR spectra of different phase states of lysozyme in the ternary systems, i.e. glycerol and PEG 1000 increase conformational stability of lysozyme. In contrast, addition of glycine increased the deviation between the FT-IR spectra, i.e. glycine further increased conformational instability. The differing impact of glycerol, PEG 1000 and glycine on conformational stability was also reflected in a differing impact on lysozyme phase behavior in the ternary phase diagrams compared to the binary ones. Addition of glycerol and PEG 1000 enlarged the size of the crystallization area, whereas addition of glycine reduced the size of the crystallization area. Under the assumption that crystallization is a oligomerization of native protein molecules (McPherson, 2004; Philo and Arakawa, 2009) the above further supports the results from the FT-IR experiments and indicates that glycerol as well as PEG 1000 stabilized the native state of lysozyme whereas glycine destabilized the native state of lysozyme. Additionally, skin formation could be reduced or completely inhibited by addition of glycerol and PEG 1000 at pH 3 and was enhanced by addition of glycine. As this skin represents a non-native phase state (Zeelen, 2009), reduction or inhibition of skin formation shows the potency of glycerol and PEG 1000 to delay or completely inhibit non-native aggregation. In contrast, addition of glycine promoted non-native aggregation. This again further supports the results from the FT-IR experiments. The impact of glycerol, PEG 1000 and glycine on conformational stability was not reflected in a significant influence on lysozyme solubility lines at pH 3. Lysozyme solubility lines were hardly influenced by the addition of glycerol, PEG 1000 and glycine. As protein solubility in many cases is described as directly related to protein-protein interactions (Curtis et al., 1998; Guo et al., 1999; Haas et al., 1999; Ruppert et al., 2001), no influence on the solubility lines implies no influence on protein-protein interactions by addition of glycerol, PEG 1000 and glycine. Fig. 6(a-d) shows that glycerol, PEG 1000 and glycine did not significantly influence lysozyme crystal size and morphology. This indicates that they as well did not influence aggregation kinetics (Durbin and Feher, 1986; Heidner, 1978; Kam et al., 1978). This is further supported by the fact that both, lysozyme solubility lines and the initiation of crystallization, were not significantly influenced by addition of glycerol, PEG 1000 and glycine. However, lysozyme phase states could be selectively controlled by addition of glycerol, PEG 1000 and glycine. For example, precipitation and skin formation could be transformed to crystallization or the other way around.

Altogether, additive action at pH 3 was identified to be simply driven by structure stabilizing or destabilizing tasks. This is in agreement with the preferential interaction theory. Addition of glycerol and PEG 1000 stabilized the native lysozyme conformation and thus seem to cause a preferential hydration of lysozyme, i.e. glycerol and PEG 1000 are preferentially excluded from lysozyme's local domain. Addition of glycine destabilized the native state, i.e. it is preferentially bound to or enriched at the lysozyme surface. The destabilizing properties of glycine were surprising, though. Glycine was earlier described as stabilizing osmolyte even for concentrations

up to 5 M (Bruzdziak et al., 2013; Santoro et al., 1992). However, it has to be mentioned that according to Singh et al. (Singh et al., 2011) osmolyte action is extremely pH dependent and the above named investigations with lysozyme and glycine whose results are in opposition to ours were conducted at pH 5.8-6.5.

In summary, at pH 3, glycerol and glycine exhibited typical osmolyte character that could be explained by the preferential interaction theory. PEG 1000 action could also be explained using the preferential interaction theory, but its stabilizing impact was stronger compared to glycerol as it could even completely inhibit non-native aggregation. Protein-protein interactions and aggregation kinetics were not only or only weakly influenced by addition of glycerol, glycine, and PEG 1000 as far as could be evaluated here.

4.2. Impact of the additives on lysozyme phase behavior, conformation and solubility at pH 5

Deviation between the FT-IR spectra of different phase states of lysozyme in the binary systems at pH 5 is low. Thus, conformational stability of lysozyme at pH 5 is high, even during precipitation. The additives investigated did have no apparent influence on the FT-IR spectra, i.e. conformational stability of lysozyme seems to be untouched by addition of glycerol, PEG 1000 and glycine. In contrast to the impact of the additives on lysozyme phase behavior at pH 3, where only the size of the crystallization area was changed, at pH 5 the investigated additives caused a shift of the crystallization area to higher or lower sodium chloride and/or lysozyme concentrations. The addition of 10 %(m/v) glycerol significantly shifted the crystallization area to higher sodium chloride and lysozyme concentrations. The addition of PEG 1000 caused a shift of the crystallization area to higher lysozyme concentrations and the ternary crystallization area is interspersed with soluble conditions. Therefore, glycerol and PEG 1000 successfully completely inhibited aggregation in some cases compared to the binary system. Glycine shifted the crystallization area to lower sodium chloride concentrations and thus promoted aggregation in comparison to the binary system. This shift of the crystallization area caused by the additives at pH 5 is not due to an impact of glycerol, PEG 1000 and glycine on lysozyme conformational stability. Furthermore, in contrast to pH 3 glycine at pH 5 did not induce or promote non-native aggregation. However, at pH 5 lysozyme solubility lines were significantly changed by addition of glycerol, PEG 1000 and glycine. Addition of glycerol and PEG 1000 enhanced lysozyme solubility at pH 5, whereas glycine enhanced lysozyme solubility up to 0.68 M sodium chloride and decreased solubility for higher sodium chloride concentrations. Thus, instead of affecting lysozyme conformational stability glycerol, PEG 1000 and glycine seem to strongly influence protein-protein interactions. Addition of glycerol and PEG 1000 probably reduced attractive protein-protein interactions in comparison to the binary system over the whole sodium chloride concentration range. Attractive protein-protein interactions were either reduced or enhanced by addition of glycine dependent on sodium chloride concentration, yielding in increased or

decreased solubility. The impact of the additives on lysozyme solubility and intermolecular interactions was also reflected in crystal size and morphology, i.e. in aggregation or crystallization kinetics. Much bigger lysozyme crystals than in the binary system at pH 5 evolved in presence of 10 %(m/v) glycerol and PEG 1000, i.e. glycerol and PEG 1000 slowed aggregation or crystallization kinetics. The diverse impact of glycine on lysozyme solubility and intermolecular interactions is also reflected in the crystal size and morphology for the ternary systems as can be seen in Fig. 8. Up to 1.14 M sodium chloride crystal size was increased, i.e. aggregation or crystallization kinetics slowed in comparison to the binary system, whereas for higher sodium chloride concentrations small, two-dimensional, sea-urchin shaped crystals occurred, i.e. aggregation or crystallization kinetics were accelerated.

In summary, the presented results for pH 5 indicate that additives that do not influence lysozyme conformational stability mainly act on lysozyme solubility, i.e. protein-protein interactions, and on crystal size and morphology, i.e. aggregation or crystallization kinetics. The mode of action of the additives at pH 5 could not be correlated to the preferential interaction theory as no influence on lysozyme conformation could be detected and lysozyme's native conformational state was very stable.

4.3. Impact of investigated additives as function of pH

A pH dependency of osmolyte action was generally described earlier (Macchi et al., 2012; Singh et al., 2011). So far, pH dependent osmolyte action was connected to the chemical nature of the osmolyte, e.g. the pk_a value (Granata et al., 2006; Natalello et al., 2009; Singh et al., 2009). This could be excluded in our case as no pk_a is exceeded between pH 3 and pH 5 for the investigated additives. More likely the differences between pH 3 and pH 5 are associated with the conformational stability of lysozyme in the binary systems.

At pH 3 FT-IR investigations of the binary system showed that lysozyme is conformationally unstable. Here, glycerol, PEG 1000 and glycine exhibited no impact on lysozyme solubility lines, crystal size and crystal morphology. At pH 3 the additives however stabilized (glycerol and PEG 1000) or destabilized (glycine) the native conformation of lysozyme and reduced (glycerol), completely inhibited (PEG 1000) or promoted (glycine) non-native aggregation. Stabilization or destabilization of lysozyme's native conformation resulted in modified sizes of the crystallization area, the precipitation area and the area where skin formation occurred. Thus, a selective control of lysozyme phase states is possible.

At pH 5 FT-IR investigations of the binary system showed that lysozyme is conformationally stable. Here, glycerol, PEG 1000 and glycine exhibited a significant impact on lysozyme solubility lines, the initiation of crystallization, crystal size and morphology, but no impact on lysozyme conformation could be identified. Aggregation or crystallization in particular could be selectively inhibited or promoted and crystal size and morphology could be selectively manipulated.

This in summary indicates that in cases where protein conformational stability is threatened additives adopt stabilization tasks. Their mode of action in these cases can be described following the preferential interaction theory (Arakawa and Timasheff, 1982, 1983, 1985a; Arakawa and Timasheff, 1985b; Lee and Lee, 1981; Timasheff and Arakawa, 1988; Webb et al., 2001). In cases where protein conformational stability is not threatened the ability of the additives to stabilize proteins is not needed, but their ability to increase the solubility of native proteins is essential. This was described earlier by Auton et al. (Auton et al., 2011) and is further supported by Bolen and Baskakov (Bolen and Baskakov, 2001) who stated that stabilizing osmolytes act mainly on the denatured state leaving the native state largely unperturbed.

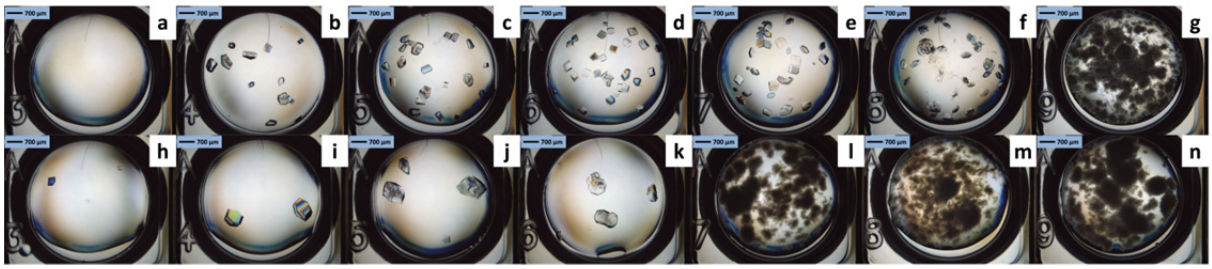

Fig. 8. Phase states of 21.75 mg/mL lysozyme after 40 days at pH 5 with varying sodium chloride concentrations. The upper row shows the phase states for binary systems without additive, the lower row the phase states for ternary systems with 1 M glycine. Sodium chloride concentrations are 0.45 M (a, h), 0.68 M (b, i), 0.91 M (c, j), 1.14 M (d, k), 1.36 M (e, l), 1.59 M (f, m), and 1.82 M (g, n).

5. Conclusion

The impact of the additives on lysozyme phase behavior and solubility lines was observed to be pH dependent. The pH dependent impact could be explained by differences in the conformational stability of lysozyme at pH 3 and pH 5 in presence of sodium chloride. pH dependent additive impact was described earlier but the presented work allowed to reduce this to lysozyme conformational stability. This publication additionally expanded the basic knowledge about additive impact on protein solubility and could present a mathematical description for a solubility line that could capture lysozyme solubility in binary and ternary systems. For conditions where the native conformational state of lysozyme was unstable in the binary system the impact caused by the investigated additives could be explained by the preferential interaction theory. Native conformation was stabilized by glycerol and PEG 1000 and destabilized by addition of glycine. PEG 1000 in particular even completely inhibited non-native aggregation, whereas glycine even promoted non-native aggregation. Stabilization or destabilization of the native conformational state resulted in a modified lysozyme phase behavior without affecting solubility lines and crystal size and morphology. Phase states could be successfully selectively controlled. For conditions where the native conformational state of

lysozyme was stable in the binary system no impact of the additives on lysozyme's native conformation could be observed. Though, the additives exhibited a significant impact on lysozyme solubility lines, the initiation of crystallization, crystal size and morphology. Glycerol and PEG 1000 increased lysozyme solubility and crystal size. Glycine increased or decreased lysozyme solubility and crystal size depending on sodium chloride concentration in the ternary systems. All of the three additives showed characteristics that could be of interest for biopharmaceutical industries: aggregation processes could be successfully manipulated with or without effects on solution characteristics like solubility or aggregate characteristics like crystal size and morphology.

Acknowledgement

We gratefully acknowledge the financial support by the Federal Ministry of Education and Research (BMBF)(0315342B).

References

Arakawa, T., Timasheff, S.N., 1982. Stabilization of Protein Structure by Sugars. Biochemistry 21, 6536-6544.

Arakawa, T., Timasheff, S.N., 1983. Preferential Interactions of Proteins with Solvent Components in Aqueous Amino Acid Solutions. Archives of Biochemistry and Biophysics 224, 169-177.

Arakawa, T., Timasheff, S.N., 1985a. Mechanism of Poly(ethylene glycol) Interaction with Proteins. Biochemistry 24, 6756-6762.

Arakawa, T., Timasheff, S.N., 1985b. The Stabilization of Proteins by Osmolytes. Biophysical Journal 47, 411-414.

Asherie, N., 2004. Protein crystallization and phase diagrams. Methods 34, 266-272.

Auton, M., Rösgen, J., Sinev, M., Holthauzen, L.M.F., Bolen, D.W., 2011. Osmolyte effects on protein stability and solubility: a balancing act between backbone and side-chains. Biophysical chemistry 159, 90-99.

Basu, S.K., Govardhan, C.P., Jung, C.W., Margolin, A.L., 2004. Protein crystals for the delivery of biopharmaceuticals. Expert Opinion on Biological Therapy 4, 301-317.

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

Bolen, D.W., Baskakov, I.V., 2001. The osmophobic effect: natural selection of a thermodynamic force in protein folding. Journal of molecular biology 310, 955-963.

Brange, J., Vølund, A., 1999. Insulin analogs with improved pharmacokinetic profiles. Advanced Drug Delivery Reviews 35, 307-335.

Bruzdziak, P., Panuszko, A., Stangret, J., 2013. Influence of Osmolytes on Protein and Water Structure: A Step To Understanding the Mechanism of Protein Stabilization. The journal of physical chemistry. B 117, 11502-11508.

Chi, E.Y., Krishnan, S., Randolph, T.W., Carpenter, J.F., 2003. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. Pharmaceutical research 20, 1325-1336.

Cohn, E.J., 1925. The physical chemistry of proteins. Physiological Reviews 5, 349-437.

Cromwell, M.E.M., Hilario, E., Jacobson, F., 2006. Protein aggregation and bioprocessing. The AAPS journal 8, E572-579.

Curtis, R.A., Prausnitz, J.M., Blanch, H.W., 1998. Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. Biotechnology and bioengineering 57, 11-21.

Dong, A., Prestrelski, S.J., Allison, S.D., Carpenter, J.F., 1995. Infrared Spectroscopic Studies of Lyophilization- and Temperature-Induced Protein Aggregation. Journal of Pharmaceutical Sciences 84, 415-424.

Durbin, S.D., Feher, G., 1986. Crystal growth studies of lysozyme as a model for protein crystallization. Journal of Crystal Growth 76, 583-592.

Dzwolak, W., Ravindra, R., Lendermann, J., Winter, R., 2003. Aggregation of bovine insulin probed by DSC/PPC calorimetry and FTIR spectroscopy. Biochemistry 42, 11347-11355.

Feng, Y.W., Ooishi, A., Honda, S., 2012. Aggregation factor analysis for protein formulation by a systematic approach using FTIR, SEC and design of experiments techniques. Journal of pharmaceutical and biomedical analysis 57, 143-152.

Fink, A.L., 1998. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding & design 3, R9-23.

Granata, V., Palladino, P., Tizzano, B., Negro, A., Berisio, R., Zagari, A., 2006. The Effect of the Osmolyte Trimethylamine N-Oxide on the Stability of the Prion Protein at Low pH. Biopolymers 82, 234-240.

Guo, B., Kao, S., McDonald, H., Asanov, A., Combs, L.L., William Wilson, W., 1999. Correlation of second virial coefficients and solubilities useful in protein crystal growth. Journal of Crystal Growth 196, 424-433.

Haas, C., Drenth, J., Wilson, W.W., 1999. Relation between the Solubility of Proteins in Aqueous Solutions and the Second Virial Coefficient of the Solution. The Journal of Physical Chemistry B 103, 2808-2811.

Harries, D., Rösgen, J., 2008. A practical guide on how osmolytes modulate macromolecular properties. Methods in Cell Biology 84, 679-735.

Heidner, E., 1978. The functional dependence of the nucleation rate on the protein concentration and the solubility. Journal of Crystal Growth 44, 139-144.

Howard, S.B., Twigg, P.J., Baird, J.K., Meehan, E.J., 1988. The solubility of hen egg-white lysozyme. Journal of Crystal Growth 90, 94-104.

Jen, A., Merkle, H.P., 2001. Diamonds in the rough: protein crystals from a formulation perspective. Pharmaceutical research 18, 1483-1488.

Kam, Z., Shore, H.B., Feher, G., 1978. On the crystallization of proteins. Journal of Molecular Biology 123, 539-555.

Kaushik, J.K., Bhat, R., 2003. Why Is Trehalose an Exceptional Protein Stabilizer? The Journal of Biological Chemistry 278, 26458-26465.

Kendrick, B.S., Cleland, J.L., Lam, X., Nguyen, T., Randolph, T.W., Manning, M.C., Carpenter, J.F., 1998. Aggregation of Recombinant Human Interferon Gamma : Kinetics and Structural Transitions. Journal of Pharmaceutical Sciences 87, 1069-1076.

Lee, J.C., Lee, L.L.Y., 1981. Preferential solvent interactions between proteins and polyethylene glycols. The Journal of biological chemistry 256, 625-631.

Macchi, F., Eisenkolb, M., Kiefer, H., Otzen, D.E., 2012. The effect of osmolytes on protein fibrillation. International Journal of Molecular Sciences 13, 3801-3819.

Matheus, S., Friess, W., Schwartz, D., Mahler, H.-C., 2009. Liquid High Concentration IgG1 Antibody Formulations by Precipitation. Journal of Pharmaceutical Sciences 98, 3043-3057.

McPherson, A., 2004. Introduction to protein crystallization. Methods (San Diego, Calif.) 34, 254-265.

Natalello, A., Liu, J., Ami, D., Doglia, S.M., de Marco, A., 2009. The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. Proteins 75, 509-517.

Philo, J.S., Arakawa, T., 2009. Mechanisms of Protein Aggregation. Current Pharmaceutical Biotechnology 10, 348-351.

Poddar, N.K., Ansari, Z.A., Singh, R.K.B., Moosavi-Movahedi, A.A., Ahmad, F., 2008. Effect of monomeric and oligomeric sugar osmolytes on ΔG_D, the Gibbs energy of stabilization of the protein at different pH values: Is the sum effect of monosaccharide individually additive in a mixture? Biophysical Chemistry 138, 120-129.

Retailleau, P., Ries-Kautt, M., Ducruix, A., 1997. No Salting-In of Lysozyme Chloride Observed at Low Ionic Strength over a Large Range of pH. Biophysical Journal 73, 2156-2163.

Ruppert, S., Sandler, S.I., Lenhoff, A.M., 2001. Correlation between the osmotic second virial coefficient and the solubility of proteins. Biotechnology progress 17, 182-187.

Santoro, M.M., Liu, Y., Khan, S.M.A., Hou, L.-X., Bolen, D.W., 1992. Increased Thermal Stability of Proteins in the Presence of Naturally Occurring Osmolytest. Biochemistry 31, 5278-5283.

Scopes, R.K., 1994. Separation by Precipitation, Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag New York, Inc., pp. 71-101.

Singh, L.R., Dar, T.A., Rahman, S., Jamal, S., Ahmad, F., 2009. Glycine betaine may have opposite effects on protein stability at high and low pH values. Biochimica et biophysica acta 1794, 929-935.

Singh, L.R., Poddar, N.K., Dar, T.A., Kumar, R., Ahmad, F., 2011. Protein and DNA destabilization by osmolytes: the other side of the coin. Life sciences 88, 117-125.

Timasheff, S.N., Arakawa, T., 1988. Mechanism of Protein Precipitation and Stabilization by Co-Solvents. Journal of Crystal Growth 90, 39-46.

Vajo, Z., Fawcett, J., Duckworth, W.C., 2001. Recombinant DNA Technology in the Treatment of Diabetes: Insulin Analogs. Endocrine Reviews 22, 706-717.

Wang, W., Li, N., Speaker, S., 2010. External Factors Affecting Protein Aggregation, in: Wang, W., Roberts, C.J. (Eds.), Aggregation of Therapeutic Proteins, 1st ed. John Wiley & Sons, pp. 119-204.

Webb, J.N., Webb, S.D., Cleland, J.L., Carpenter, J.F., Randolph, T.W., 2001. Partial molar volume, surface area, and hydration changes for equilibrium unfolding and formation of aggregation transition state: high-pressure and cosolute studies on recombinant human IFN-γ. Proceedings of the National Academy of Sciences of the United States of America 98, 7259-7264.

Yancey, P.H., 2001. Water Stress, Osmolytes and Proteins. American Zoologist 41, 699-709.

Yang, M.X., Shenoy, B., Disttler, M., Patel, R., McGrath, M., Pechenov, S., Margolin, A.L., 2003. Crystalline monoclonal antibodies for subcutaneous delivery. Proceedings of the National Academy of Sciences of the United States of America 100, 6934-6939.

Zeelen, J.P., 2009. Interpretation of the Crystallization Drop Results, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 175-194.

3.3 Non-invasive high throughput approach for protein hydrophobicity determination based on surface tension

Sven Amrhein[1], Katharina Christin Bauer[1], Lara Galm[1], Jürgen Hubbuch*

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

** Corresponding author. Tel.: +49 721 608 42557; fax: +49 721 608 46240. E-mail address: juergen.hubbuch@kit.edu.*

[1] *Contributed equally.*

accepted by: ***Biotechnology and bioengineering***. doi: 10.1002/bit.25677

Abstract

The surface hydrophobicity of a protein is an important factor for its interactions in solution and thus the outcome of its production process. Yet most of the methods are not able to evaluate the influence of these hydrophobic interactions under natural conditions. In the present work we have established a high resolution stalagmometric method for surface tension determination on a liquid handling station, which can cope with accuracy as well as high throughput requirements. Surface tensions could be derived with a low sample consumption (800 µL) and a high reproducibility (<0.1 ‰ for water) within a reasonable time (3.5 min per sample). This method was used as a non-invasive HTP compatible approach to determine surface tensions of protein solutions dependent on protein content. The protein influence on the solutions' surface tension was correlated to the hydrophobicity of lysozyme, human lysozyme, BSA, and α-lactalbumin. Differences in proteins' hydrophobic character depending on pH and species could be resolved. Within this work we have developed a pH dependent hydrophobicity ranking, which was found to be in good agreement with literature. For the studied pH range of 3 to 9 lysozyme from chicken egg white was identified to be the most hydrophilic. α-lactalbumin at pH 3 exhibited the most pronounced hydrophobic character. The stalagmometric method occurred to outclass the widely used spectrophotometric method with bromophenol blue sodium salt as it gave reasonable results without restrictions on pH and protein species.

Keywords: *Stalagmometer, Bromophenol blue, Hydrophilicity, Protein solution, Protein-solvent interaction*

1. Introduction

Hydrophobic interactions play a key role in the outcome of the production process of biopharmaceutical therapeutics passing through fermentation, the purification process, formulation and storage. During protein expression in fermentation hydrophobic forces regulate the formation of the globular protein molecule (Dill, 1990; Tanford, 1962). The resulting surface characteristics and the protein concentration govern its solubility for all following production steps. Already small changes in the hydrophobic character can provoque changes in solubility and in the aggregation tendency of the molecule (Brems et al., 1988; Nieba et al., 1997). Undesired aggregation during the process or storage can cause denaturation and thus product

loss. However, for the purification process this changes regarding solubility can also be turned into advantage in terms of protein crystallization or precipitation as purification steps. The hydrophobic character of a protein can additionally be exploited to separate complex protein mixtures by using aqueous two-phase systems (ATPS) (Andrews et al., 2005; Asenjo and Andrews, 2011; Diederich et al., 2013), reversed phase (RP) chromatography or hydrophobic interaction chromatography (HIC) (Janson, 2012). The knowledge of protein surface hydrophobicity therefore helps to predict, control and manipulate the influence of hydrophobic interactions during processing and storage.

Within the last decades a huge research effort has been spend on the development of experimental and *in silico* approaches for assessing protein surface hydrophobicity by the identification of highly hydrophobic single amino acids, that are ranked in different hydrophobicity scales (Black and Mould, 1991; Eisenberg, 1984; Janin, 1979; Kyte and Doolittle, 1982; Rose and Wolfenden, 1993). This amino acid hydrophobicity has been measured mainly in terms of their solubility in organic and denaturant solutions (Dooley and Castellino, 1972; Nozaki and Tanford, 1963, 1965, 1970; Nozaki and Tanford, 1971; Whitney et al., 1962) or their partition between an aqueous and an organic phase (Fendler et al., 1975; Radzicka and Wolfenden, 1988).

Efforts have been made on hydrophobicity rankings of whole proteins based on partitioning in aqueous two-phase systems (Shanbhag and Axelsson, 1975), retention factors in HIC (Keshavarz and Nakai, 1979) or using hydrophobic dyes (Bertsch et al., 2003; Cardamone and Puri, 1992; Hawe et al., 2008; Hendriks et al., 2002; Kato and Nakai, 1980). Yet most experimental methods are invasive and not adequate to consider the influence of solution characteristics on the protein. In most cases organic liquids are inevitable which may influence the tertiary structure of the protein or even denature it.

In silico methods apply experimentally determined hydrophobicity scales to quantify protein hydrophobicity. These methods calculate the protein hydrophobicity either based on the amino acid sequence (Salgado et al., 2005) or based on the three-dimensional structure (Chennamsetty et al., 2009; Lijnzaad et al., 1996; Miller et al., 1987). Salgado et al. (Salgado et al., 2006a, b) found that methods based on the three-dimensional structure show a better predictive performance for protein adsorption mechanisms in HIC than the ones based on the amino acid sequence. Sophisticated approaches based on quantitative structure property relationship (QSAR) or molecular dynamics (MD) simulations (Amrhein et al., 2014a; Reißer et al., 2014) give the most detailed description of the protein surface. However, these sophisticated *in silico* approaches require either an enormous computational effort or use theoretical hydrophobicity scales as described above and thus cannot account for the influence of the solvent. Generally, approaches using theoretically derived hydrophobicity scales are highly influenced by the selected hydrophobicity scale (Biswas et al., 2003; Trinquier and Sanejouand, 1998).

One promising way of considering the environmental dependency is to measure hydrophobicity via surface tension. This experimental method is able of capturing the environmental dependency because it is non-invasive. The existence of a correlation between surface or interfacial tension and hydrophobicity of single amino acids or macromolecules is known for a long time (Absolom et al., 1987; Bull and Breese, 1974; Kato and Nakai, 1980; Keshavarz and Nakai, 1979). Bull and Breese (Bull and Breese, 1974) pointed out the potency of sorting amino acids by hydrophobicity according to their surface tension increment. According to this, hydrophobic amino acids reduce the surface tension with increasing concentration. Conversely, hydrophilic amino acids increase the surface tension.

Keshavarz and Nakai (Keshavarz and Nakai, 1979) could show a significant negative correlation between the effective hydrophobicities of bovine serum albumin, ovalbumin, lysozyme, γ-globulin, myoglobin, β-lactoglobulin, trypsin, conalbumin, and α–chymotrypsin and interfacial tension. The more hydrophobic the protein, the greater the depression in the interfacial tension. Recently Genest et al. (Genest et al., 2013) found a correlation between the depression of the surface tension and the polymer hydrophobicity. There are quite a number of methods for the determination of surface tensions such as the Capillary rise method, the Wilhelmy plate method, pendant and sessile drop methods, and the stalagmometric method. This method combines low sample consumption and compatibility to high throughput liquid-handling devices. Additionally, we are confident that this gravimetric approach is superior to imaging approaches in terms of precision and robustness.

In the present work we chose the stalagmometric method and developed a high resolution experimental setup which can be integrated into high throughput (HTP) work flow by a liquid handling station. This method was used as a non-invasive high throughput compatible approach to determine protein hydrophobicity on base of the proteins' surface tension increments. Lysozyme, human lysozyme, BSA, and α-lactalbumin were characterized regarding hydrophobicity dependent on pH value. This set of proteins covers a wide range of molecular weight, isoelectric points, and includes two similar lysozyme species differing in amino acid composition. The hydrophobicity values for BSA derived from this approach were found to be in good agreement to values obtained by the widely used method of absorption difference spectroscopy of bromophenol blue sodium salt (BPB Na). Instead of this pH and protein species dependent absorption method the developed stalagmometric method was able to cover the full pH range in a completely automated way by considering the environmental protein complexity.

2. Materials and methods

2.1. Stalagmometric method for determination of surface tension

The stalagmometric method was chosen for the determination of the surface tension due to its

ability to be transformed into a fully automated procedure. In this method the specific fluid is purged very slowly through a needle in a vertical direction in order to form drops. The drop grows up to a specific maximum volume and falls onto a high precision mass balance. The Tate's law (Tate, 1864) expresses the correlation between the drop's weight (W_{drop}) and its surface tension (γ) at the moment the drop falls from the needle with an outer radius r:

$$\boldsymbol{W_{drop} = 2\pi r\gamma} \quad \textbf{(1)}$$

However, the instrumental setup has an influence on the drop size and the registered weight will be W'_{drop}. W_{drop} and W'_{drop} are correlated by an instrumental setup correction factor f_{inst}. Considering this correction factor the surface tension γ can be expressed as

$$\boldsymbol{\gamma = \frac{m_{drop} g}{2\pi r f_{inst}}} \quad \textbf{(2)}$$

where m_{drop} represents the mass of the drop, g the acceleration of gravity, r the radius of the needle, and f_{inst} the correction factor. By using a liquid with a known surface tension γ_{ref}, resulting in a specific drop mass $m_{drop,ref}$, the correction factor can be evaluated and the surface tension γ_{sample} of the liquid of interest can be determined by applying Eq. (3) to its specific drop mass $m_{drop,sample}$.

$$\boldsymbol{\frac{\gamma_{ref}}{m_{drop,ref}} = \frac{\gamma_{sample}}{m_{drop,sample}}} \quad \textbf{(3)}$$

2.2. Automation of stalagmometric method for the purpose of HTP

In the following the automation of this stalagmometric method, which principle can be found in subsection 2.1, is described. This method was optimized towards analysis speed, accuracy and precision. The validation was conducted using liquids with a wide range of surface tension values.

2.2.1. Automation using a liquid handling station

This method was established on a fully automated robotic liquid handling station, namely Freedom EVO® 100 purchased from Tecan (Crailsheim, Germany), equipped with stainless steel fixed tips and 1 mL dilutors. The liquid handling station was controlled using Evoware 2.4 SP3. The complete experimental setup, as is illustrated schematically by a partially section view in Fig. 1, consisted of two major subunits, namely a *Tip2World* interface mounted on a docking station and the stalagmometric setup, interconnected via a capillary tubing (PEEK, ID: 0.25 mm).

The *Tip2World* interface enables to supply liquid samples with a robotic liquid handling arm via connected standard capillary tubings and is described in detail elsewhere (Amrhein et al., 2014b). The stalagmometric system setup consisted of the distribution block, connecting the capillary tubing coming from the *Tip2World* interface with the drop generating needle, and a 250 mL container to collect the drops (Azlon®Specimen, SciLabware Limited, Stoke-on-Trent, United Kingdom). This container was placed on an Excellence WXTS205DU high performance balance unit (Mettler Toledo, Greifensee, Switzerland). The distribution block possesses a standard 10-32 coned input port and a vertically aligned standard 10-32 coned output port, where the drop generating stainless steel needle was fixed. This needle had an inner diameter of 0.5 mm and an outer diameter of 1.6 mm. Both ports were connected by an internal capillary (ID: 0.5 mm). The container held a drop trap to collect the falling drops in a gentle way. This was realized by a customized design shown in Fig. 1 and enhanced the balance signal stability. To prevent the drop from evaporation a cylindrical evaporation trap was developed. This consisted of a cylinder, which was placed on a bottom ring with a water trough and a lid. All components were custom designed and 3D printed. Apart from bottom and lid (Sculpteo, Issy-les-Moulineaux, France) they were manufactured by a high resolution 3D printer (Stratasys, Eden Prairie, MN, USA). The construction work was carried out with the 2D/3D CAD software SolidEdge (Siemens PLM Software, Plano, TX, USA). For the surface tension determination all samples were stored at room temperature in 1.3 mL 96-well Deep Well plates (Nalgene Nunc, Rochester, NY, USA) and sealed with a pre-slit well cap (Nalgene Nunc, Rochester, NY, USA), in order to minimize evaporation. The sample to be analyzed was pumped with a flow rate of 5 µL/s using the *Tip2World* interface to the stalagmometer through a PEEK capillary tubing with a diameter of 0.25 mm. The falling drop masses were recorded continuously via serial communication realized by automated routines written in Matlab®R2013b (The Mathworks, Natick, ME, USA). By evaluating the step profile of the balance signal, the distinguished masses of each drop were calculated. In order to increase the accuracy, instabilities of the weight signal were sorted out. Processing of the experimental results of the stalagmometric method was performed by means of fully automated routines written in Matlab®R2013b (The Mathworks, Natick, ME, USA).

2.2.2. Validation of surface tension determination

The stalagmometric approach for surface tension determination was tested with respect to reproducibility by measuring 70 samples of ultrapure water with a volume of 800 µL each. For the validation of the procedure, surface tension measurements were performed with aqueous mixtures of sodium chloride (NaCl) and ethylene glycol solutions in order to cover a wide range of surface tensions. In particular, NaCl solutions were used with a molar fraction of 0 % to about 10 %. Ethylene glycol solutions were used within a molar fraction range of 0 % to 26 %. NaCl,

analysis grade, was purchased from Merck KGaA (Darmstadt, Germany). Ethylene glycol, analysis grade, was purchased from VWR International (Radnor, PA, USA). Ethylene glycol and NaCl solutions were prepared by mixing the respective masses of ethylene glycol or NaCl with ultrapure water in order to reach the desired mass fractions. Ultrapure water was used as reference liquid. Cross contamination of samples was minimized by purging the tubing with air. All liquids were analyzed 8-fold with a volume of 800 µL each.

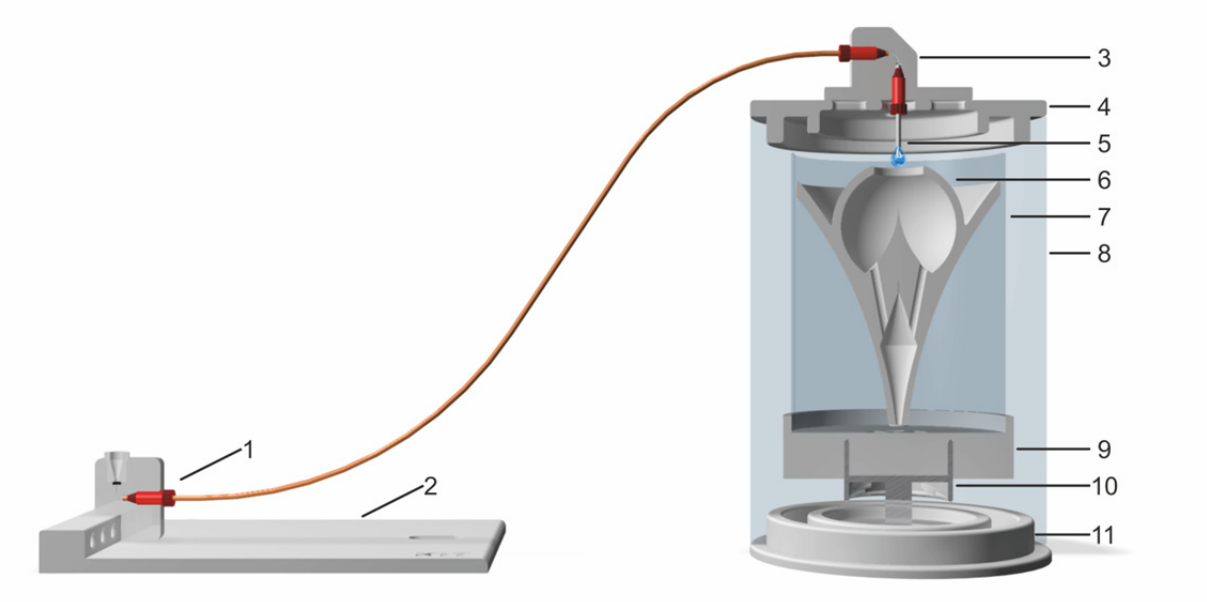

Fig. 1. Illustration in partially section view of the setup of the stalagmometer, consisting of (1) *Tip2World* interface, (2) docking station, (3) the distribution block with a standard input 10-32 port, (4) a lid for fixation and centering of the distribution block, (5) a stainless steel needle in vertical orientation, (6) a drop trap, (7) a trough to collect and weight the drops, (8) a cylinder for evaporation protection, (9) a carrier, (10) balance unit, and (11) bottom part with water trough.

2.3. Correlation of surface tension and hydrophobicity

In order to test the applicability of using the surface tension increment of an analyte to deduce its hydrophobicity, polyethylene glycol (PEG) with varying molecular weights were used. PEG is known to expose increased hydrophobicity with increasing molecular weight (Harris, 1992). The polymer sample preparation is described in detail in Section 2.3.1. The method then was applied to protein solutions. These solutions varied in protein species, pH and buffer composition as explained in Section 2.3.2. The derived hydrophobicity measures were compared to hydrophobicity values derived from an orthogonal colorimetric method using BPB Na as described in Section 2.3.2.

2.3.1. Stalagmometric determination of polymer hydrophobicity

The surface tension increments of PEG species with a molecular weight of 200, 300, 400, 600 and 1000 Da were determined by measuring aqueous PEG solutions with molar fractions

varying from 0.02 to 0.11 %. These concentrations correlate to mass fractions from 0.2 to 1.2 % (w/w) for PEG200, PEG300 and PEG400; from 0.4 to 2 % (w/w) for PEG600 and from 0.8 to 4 % (w/w) for PEG1000. Samples were prepared by mixing the respective masses of PEG with ultrapure water. 800 µL of each solution were measured 8-fold. All PEG species were of analysis grade and purchased from Merck KGaA (Darmstadt, Germany).

2.3.2. Determination of protein hydrophobicity

2.3.2.1. Sample preparation

The used buffer substances were citric acid (Merck, Darmstadt, Germany) and sodium citrate (Sigma-Aldrich, St. Louis, MO, USA) for pH 3, sodium acetate (Sigma-Aldrich, St. Louis, MO, USA) and acetic acid (Merck, Darmstadt, Germany) for pH 5, MOPSO (AppliChem, Darmstadt, Germany) for pH 7 and Bis-Tris propane (Molekula, Dorset, United Kingdom) for pH 9. Buffer capacity was 100 mM for all buffers. Hydrochloric acid and sodium hydroxide for pH adjustment were obtained from Merck (Darmstadt, Germany). pH adjustment was performed using a five-point calibrated pH-meter (HI-3220, Hanna Instruments, Woonsocket, RI, USA) equipped with a SenTix®62 pH electrode (Xylem Inc., White Plains, NY, USA). pH was adjusted with the appropriate titrant with an accuracy of ±0.05 pH units. All buffers were filtered through 0.2 µm cellulose acetate filters (Sartorius, Göttingen, Germany). Buffers were used at the earliest one day after preparation and repeated pH verification. Lysozyme from chicken egg white was purchased from Hampton Research (Aliso Viejo, CA, USA), human lysozyme, bovine serum albumin (BSA), and calcium depleted α-lactalbumin were purchased from Sigma-Aldrich. To set up the protein stock solutions, protein was weighed in and dissolved in the appropriate buffer yielding the desired concentration. All protein solutions were filtered through 0.2 µm syringe filters with cellulose acetate membrane (VWR, Radnor, PA, USA) and further desalted via size exclusion chromatography using a HiTrap Desalting Column (GE Healthcare, Uppsala, Sweden) on an AEKTAprime™ plus system (GE Healthcare, Uppsala, Sweden). The desired concentration was achieved by using Vivaspin centrifugal concentrators (Sartorius, Goettingen, Germany) with PES membranes. Protein concentration determination of the collected fractions was conducted using a NanoDrop2000c UV-Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). (Extinction coefficients were $E^{1\%}(280\ nm)_{lysozyme}$ = 22.00, $E^{1\%}(280\ nm)_{human\ lysozyme}$ = 16.00, $E^{1\%}(280\ nm)_{BSA}$= 6.70, $E^{1\%}(280\ nm)_{\alpha\text{-lactalbumin}}$ = 16.81)

2.3.2.2. Stalagmometric determination of protein hydrophobicity

Buffers and protein stock solutions of lysozyme, human lysozyme, BSA, and α-lactalbumin were prepared as described above. Protein solutions were prepared by dilution of the protein stock solutions with the respective buffer. The following solutions were laid eightfold in a 1.3 mL 96-well Deep Well plate: for reference ultrapure water and the respective buffer, protein solutions

varying in protein molar fractions from $5.6 \cdot 10^{-6}$ to $1 \cdot 10^{-2}$ % for lysozyme and from $1.4 \cdot 10^{-6}$ to $3.9 \cdot 10^{-5}$ % for human lysozyme, BSA, and α-lactalbumin. 800 µL of each of the 96 samples were pumped to the stalagmometer with a flow rate of 5 µL/s. The surface tension of all samples was analyzed according to Section 2.2.1 in 8-fold replicates.

2.3.2.3. Spectrophotometric determination of protein hydrophobicity

The determination of protein hydrophobicity using bromophenol blue sodium salt (BPB Na) as a hydrophobicity sensitive dye was conducted for lysozyme, human lysozyme, BSA, and α-lactalbumin. Buffers and protein stock solutions were prepared as described earlier. BPB Na, purchased from Sigma-Aldrich (St. Louis, MO, USA), was dissolved in the respective buffer to reach a dye concentration of 0.02 mg/mL for the BPB Na stock solution. The spectrophotometric method was conducted according to Bertsch et al. (Bertsch et al., 2003). No measurements were conducted at pH 3, because BPB Na is a well-known pH indicator dye within the pH range of 3.0 to 4.6. Protein stock solutions were diluted using the respective buffers and mixed with a fixed volume of BPB Na stock solution, yielding protein solutions with a BPB Na concentration of 7.99 µM and protein molar fractions up to $3.3 \cdot 10^{-5}$ % for BSA and up to $4.9 \cdot 10^{-3}$ % for lysozyme, human lysozyme, and α-lactalbumin. The absorption spectra of a BPB Na solution without protein and of the BPB Na solutions with added protein were measured between 550 and 650 nm in 1 nm steps using an Infinite®M200 microplate reader (Tecan, Crailsheim, Germany). The measurements of each solution were conducted in triplicate. In contrast to Bertsch et al. we examined the shift of the absorption maximum ΔA_{max} in dependency of the protein concentration instead of the absorption difference at 620 nm ΔA_{620}. The value of the absorption maximum shift was calculated as mean value from the triplicate spectrophotometric measurements. The shift of the absorption maximum ΔA_{max} (in nm) was fitted using a Box Lucas model

$$\Delta \boldsymbol{A_{max}} = \boldsymbol{a} \cdot (\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{b}\tilde{\boldsymbol{x}}}) \qquad \textbf{(4)}$$

where $\tilde{x}$ reflects the molar protein fraction, a and b are adjustable constants. Parameter a describes the limit of absorption maximum shift, a·b can be interpreted as relative surface hydrophobicity. It reflects the slope of the fitted curve for a protein molar fraction $\tilde{x}$ approaching zero. The higher a·b the higher the protein surface hydrophobicity.

3. Results

3.1. Development of HTP compatible stalagmometric method for surface tension determination

We were able to create HTP compatibility of the stalagmometric method by connecting the stalagmometer to the prior developed *Tip2World* interface (Amrhein et al., 2014b). As mentioned in Section 2.2.2 the HTP compatible stalagmometric method was tested for

reproducibility using 70 samples of ultrapure water. These measurements showed a standard deviation of less than 0.1 ‰ which correlates to a 99.7 % confidence interval (±3σ), which equals a deviation of less than ±0.2 mN/m for water. The validation, described in Section 2.2.2, was conducted by comparison of derived surface tensions with published data (Jańczuk et al., 1989; Melinder, 2007; Tsierkezos and Molinou, 1998). The experimentally obtained surface tensions were in good agreement with reported data as illustrated in Fig. 2. Ethylene glycol exposed a decreasing impact on the surface tension, NaCl increased the surface tension following a linear trend within the studied concentration interval. Thus the developed HTP compatible stalagmometric method showed high accuracy and covered a wide range of surface tensions from 57 to 82 mN/m. The developed method required about 3.5 minutes per sample resulting in less than 5.5 hours for analyzing a set of 96 samples without requiring sophisticated instrumentation and man power.

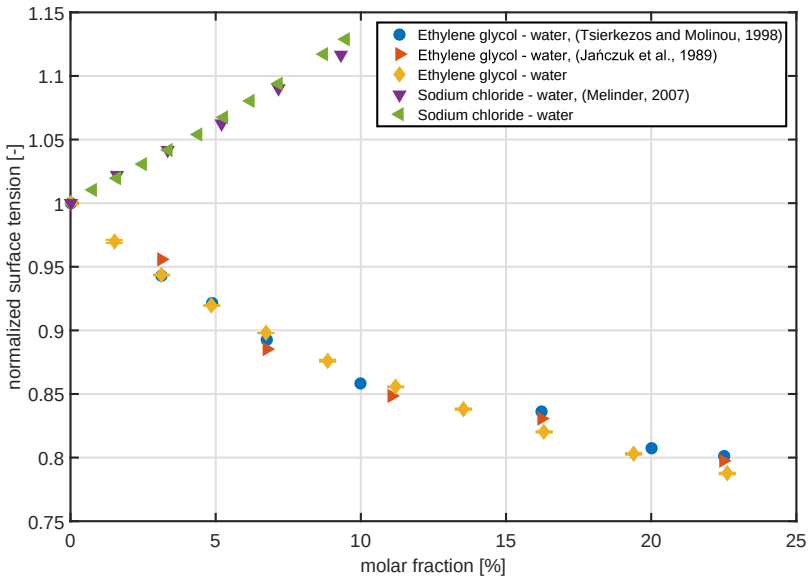

Fig. 2. Comparison of experimentally determined surface tensions with literature values (Jańczuk et al., 1989; Melinder, 2007; Tsierkezos and Molinou, 1998). The error bars refer to the 95 % confidence interval (±2σ).

3.2. Stalagmometric determination of hydrophobicity

3.2.1. Hydrophobicity of polymers

The influence of polyethylene glycol (PEG) on the surface tension was studied to assess hydrophobicity. PEG species of different molecular weights starting from 200 to 1000 Da were studied. All surface tensions were normalized on ultrapure water. As illustrated in Fig. 3 all surface tensions could be derived precisely with relative standard deviations less than 1.5 ‰ throughout all PEG samples. It is apparent from Fig. 3 that all studied PEG species followed a decreasing trend with increasing concentration. The higher the PEG molecular weight, the higher the decrease in surface tension.

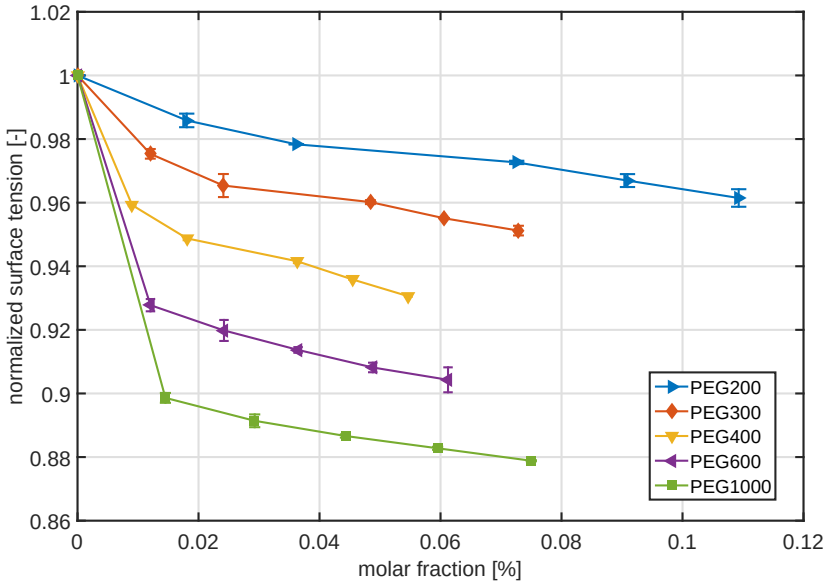

Fig. 3. Normalized surface tension profiles of PEG varying in molecular weight. Surface tensions were normalized to ultrapure water. The error bars correlate to the 95 % confidence interval (±2σ).

3.2.2. Hydrophobicity of proteins

Surface tensions of lysozyme, human lysozyme, BSA, and α-lactalbumin were analyzed in dependency of protein concentration and pH in order to characterize their hydrophobicity within the respective buffer system. Fig. 4 illustrates the normalized profiles of the surface tensions of these proteins. The surface tensions were normalized on the respective pure buffer. Human lysozyme, BSA, as well as α-lactalbumin decreased the surface tension even at very low molar fractions of less than $4 \cdot 10^{-5}$ %. It is apparent from Fig. 4 that the decrease of surface tension is highly influenced by the pH value.

In particular, for BSA pH 3 had the strongest influence on surface tension followed by pH 5, pH 7, and pH 9. The normalized surface tension profiles of pH 3 and pH 5 and the ones of pH 7 and pH 9 were close to each other. Regarding α-lactalbumin, the most distinctive decrease of the surface tension was observed at pH 3, followed by pH 5. The slopes of the normalized surface tensions at pH 7 and pH 9 were similar and slightly negative. In contrast to pH 3 and pH 5, the normalized surface tension of pH 7 and pH 9 followed a rather linear trend within the studied concentration range. In comparison to BSA and α-lactalbumin the influence of pH on the surface tension of human lysozyme was clearly lower, whereas the strongest decrease was observed at pH 3 again.

Lysozyme in contrast to BSA, α-lactalbumin and human lysozyme showed less impact on the surface tension. For lysozyme a significant reduction in surface tension could only be observed at molar fractions 300 times higher. The profiles were similar for all pH values with the strongest

reduction of surface tension to 0.92 at pH 3 for a molar fraction of 10^{-2} % within the studied concentration range.

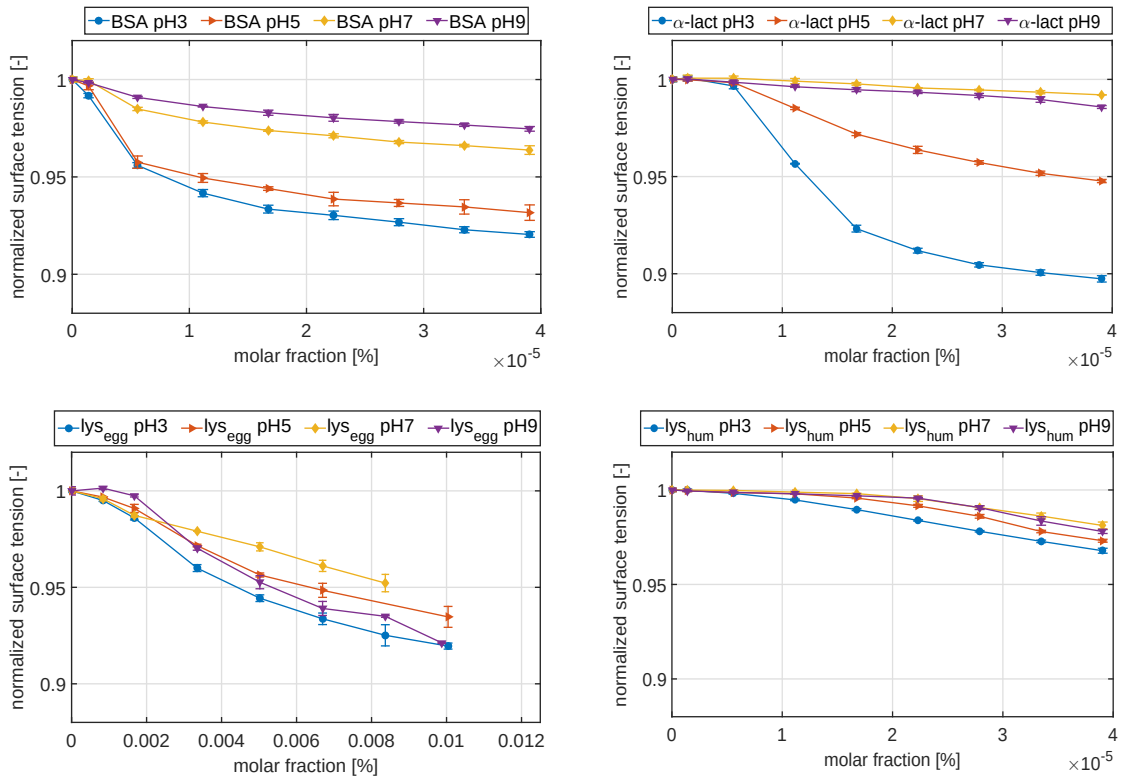

Fig. 4. Surface tension profiles of bovine serum albumin (BSA) (upper left), α-lactalbumin (α-lact) (upper right), lysozyme (lys_{egg}) (lower left), and human lysozyme (lys_{hum}) (lower right). The error bars refer to the 95 % confidence interval (±2σ).

In order to derive a hydrophobicity scale from the surface tension profiles they were fitted to Eq. (5)

$$\gamma_{norm} = 1 - \frac{d \cdot e \cdot (\tilde{x} + c)}{1 + e \cdot (\tilde{x} + c)} \qquad (5)$$

where γ_{norm} stands for the normalized surface tension, $\tilde{x}$ represents the molar protein fraction. The fitting parameter d and e could be fitted to all profiles within this study with a coefficient of correlation larger than 0.98. Exclusively in case of α-lactalbumin the normalized surface tension of pure buffer and the lowest concentrated sample were excluded from the calculation, which was considered by setting the parameter c properly. For all other proteins, parameter c was set to zero. Eq. (5) describes a saturation function of the normalized surface tension regression, where d·e can be interpreted as the surface tension activity of the protein and e as the maximal regression of the normalized surface tension. By means of the surface tension activity, the studied proteins can be ranked according to their hydrophobic character as illustrated in Fig. 5.

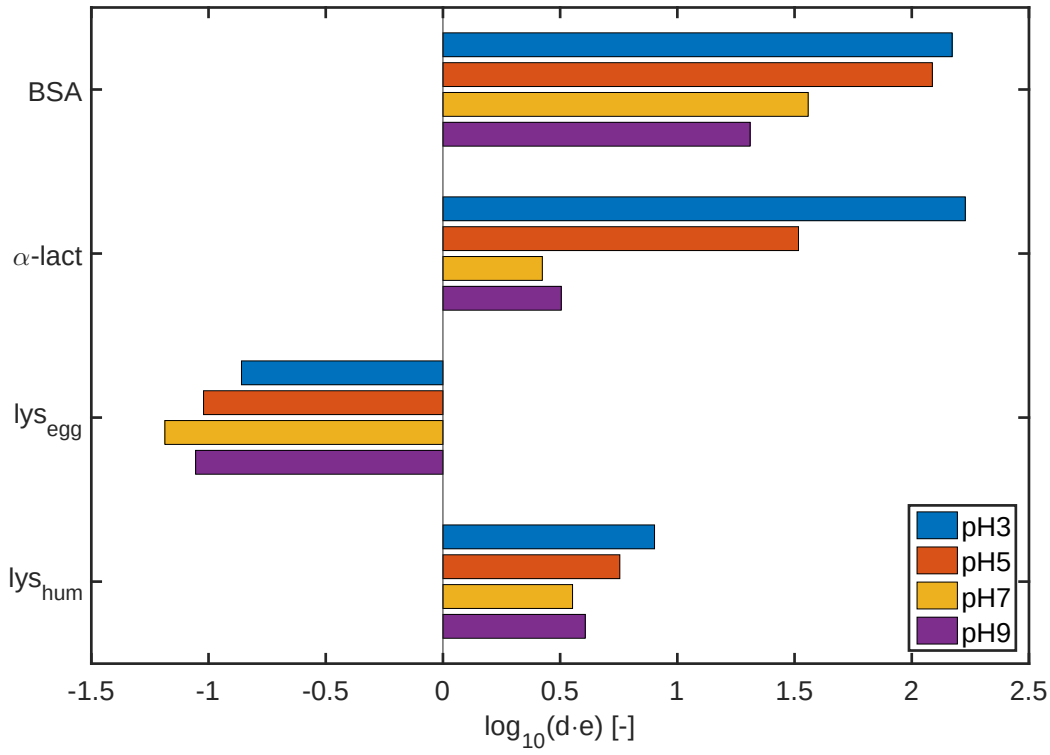

Fig. 5. Hydrophobicity ranking of bovine serum albumin (BSA), α-lactalbumin (α-lact), lysozyme (lys_{egg}), and human lysozyme (lys_{hum}) by means of the surface activity.

3.3. Spectrophotometric determination of hydrophobicity

The shift of the absorption maximum ΔA_{max} of BPB Na in presence of BSA was measured at pH 5, pH 7, and pH 9 and is illustrated in Fig. 6. There is a clear pH dependency of ΔA_{max} as function of the BSA molar fraction. The change of ΔA_{max} in the range between 0 and $1 \cdot 10^{-5}$ % was strongest for pH 5, followed by pH 7 and 9. The same order applied for the upper limit of ΔA_{max}. The data points could be fitted using Eq. (4) with a coefficient of correlation better than 0.97. This dye method was also used to compare different proteins. In addition to BSA, human lysozyme, α-lactalbumin, and lysozyme were investigated at pH 5, 7, and 9. For these proteins the molar fractions had to be increased 150 fold compared to BSA in order to identify significant shifts of the absorption maximum. For all of the investigated proteins a pH dependency of the absorption maximum shift was detected. A comparison between the four proteins at pH 7 is exemplarily shown in Fig. 7. BSA showed a significantly steeper slope and higher upper limit of the fitted curve compared to human lysozyme, α-lactalbumin, and lysozyme. Apart from BSA the highest values of ΔA_{max} were reached by human lysozyme and α-lactalbumin. However, the slope of the fitted curve of lysozyme was steepest.

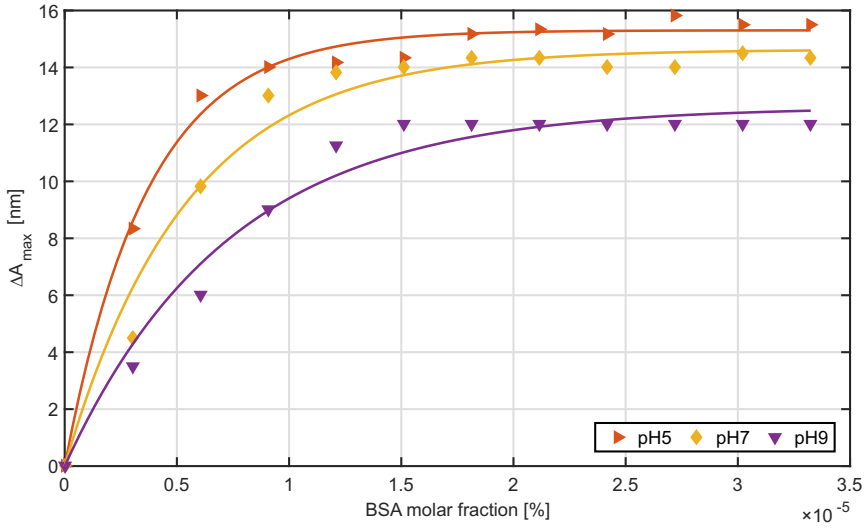

Fig. 6. Comparison of the absorption maximum shift (ΔA_{max}) of BPB Na in presence of bovine serum albumin (BSA) at pH 5, 7, and 9 and its corresponding fits.

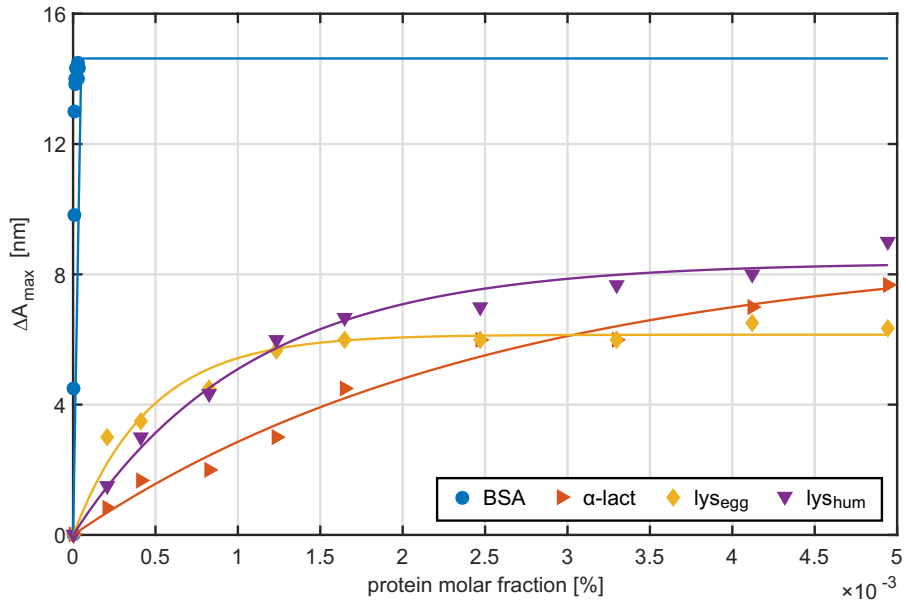

Fig. 7. Comparison of absorption maximum shift (ΔA_{max}) of BPB Na in presence of bovine serum albumin (BSA), α-lactalbumin (α-lact), lysozyme (lys_{egg}), and human lysozyme (lys_{hum}) at pH 7 and its corresponding fits.

4. Discussion

4.1. Development of HTP compatible stalagmometric method for surface tension determination

In the presented work a highly accurate, robust and HTP compatible stalagmometric method was established onto a liquid handling station using a high precision mass balance and two custom made subunits, the *Tip2World* interface and the stalagmometric setup in particular. Both subunits can be produced using 3D printing technology with a high precision and to a reasonable price. The method was validated with solutions containing NaCl or ethylene glycol. According to

literature NaCl increases the surface tension of water (Pegram and Record, 2007). Due to its lower polar character compared to water ethylene glycol exposes attractive interactions with the air-water interface and thus reduces its surface tension. With these two additives we were able to show the validity of the stalagmometric method for a wide range of surface tension values from 57 to 82 mN/m. The sample throughput of the used design was about 18 samples per hour. As mentioned before this results in an operation time of 5.5 hours for analyzing a set of 96 samples. It is important to note, that in this work we only used one tip of the liquid handling arm at a time and the setup allows the integration of multiple stalagmometric devices. Thus, using the full span of 8 tips of the liquid handling arm in combination with eight stalagmometric devices in parallel would speed up the measurement 8 fold resulting in a throughput of 144 samples per hour.

4.2. Stalagmometric determination of hydrophobicity

4.2.1. Polymers' hydrophobicity

To examine the correlation between surface tension and hydrophobicity PEG of different molecular weight was used. PEG species of higher molecular weight are known to be more hydrophobic (Harris, 1992). From our stalagmometric measurements we observed a stronger decrease of surface tension with increasing molecular weight and thus conclude that this is caused by the increase of hydrophobicity. This is in agreement with the findings of Genest et al. (Genest et al., 2013) on hydrophobically modified polyelectrolytes that also exhibited lower surface tension values the more hydrophobic they were. Thus, the presented stalagmometric method is capable of deducing an analyte's hydrophobicity from its influence on the surface tension. Distinct differences in surface tensions and thus in hydrophobicity could even be identified for small changes in molecular weight by 100 Da, as could be seen for PEG200, PEG300 and PEG400 in Section 3.2.1. This highlights the sensitivity of this stalagmometric approach. Moreover, the change of surface tension dependent on PEG molar fraction demonstrates the high resolution and precision of the developed HTP stalagmometric method. As it could be proved that the here designed HTP stalagmometric setup is able to capture polymer hydrophobicity with a high sensitivity, resolution and precision and was applied to protein solutions.

4.2.2. Proteins' hydrophobicity

In order to estimate the influence of protein species and pH we studied the surface tension increments of lysozyme, human lysozyme, BSA and α-lactalbumin at pH 3, 5, 7, and 9.

For each protein we observed an influence of pH on the surface tension increment. This is reasonable as the pH influences the protonation of ionizable amino acids and thus the charge distribution on the protein surface what causes an impact on hydrophobic surface patches. Like

for PEG an increase in hydrophobic surface character favors the interaction of the protein molecule with the less polar air-water interface. This results in a decrease of the surface tension. The derived ranking for the investigated proteins and pH values starts with α-lactalbumin at pH 3 as the sample with the most pronounced hydrophobic character. This pronounced hydrophobic character of α-lactalbumin at pH 3 can be explained by transformation of α-lactalbumin into a so called molten globule state where the protein is partially unfolded and inner hydrophobic patches are exposed on the protein surface (Permyakov and Berliner, 2000). The ranking continues with decreasing hydrophobic character as can be seen in Fig. 5. For each of the investigated proteins the hydrophobic character is highest at pH 3. This might be due to a modification of the charge distribution on the protein surface resulting from the protonation of amino acids. For example at pH 3 the amino acids glutamic acid and aspartic acid are protonated in contrast to pH 5. The protonation of the respective amino acids results in a loss of a negative charge and leads to a more pronounced hydrophobic character of the surface patches and a less favored hydration. In addition, the protonation of these amino acids influences the H-bond-network and thus could lead to minor effects on the formation of secondary structures such as α-helix and β-sheet. These minor structural changes might result in slightly different interactions with the solvent, which we were able to resolve with the presented approach. Major structural changes such as partial unfolding due to the acidic pH which would potentially result in a much stronger influence on hydrophobicity and thus surface tension could be excluded by FT-IR analysis (data not shown). Highest hydrophobic character was not always exhibited nearest to the isoelectric points of the investigated proteins as was shown for lysozyme and human lysozyme. Both exhibited the highest hydrophobic character far away from their isoelectric points. Thus, protein surface charge also influences surface tension but is not the critical factor. It is important to point out that the differences between the surface tension profiles and thus the pH dependent hydrophobic character of lysozyme were much smaller compared to BSA, α-lactalbumin and human lysozyme. Additionally, significant changes for the normalized surface tension could be found only at molar fractions 300 times higher than for the other proteins. Hence, lysozyme occurred to be much less hydrophobic compared to BSA, α-lactalbumin and human lysozyme. This significant different hydrophobic character of lysozyme is in agreement with Bigelow (Bigelow, 1967) who found BSA and α-lactalbumin to be very hydrophobic and lysozyme to be only weakly hydrophobic. This discrepancy between the structural similar lysozyme species (root mean square deviation of backbone atoms less than 1 Å) underlines the strong impact of the amino acid composition on the protein surface (sequence identity less than 65 %).

In summary, the presented HTP stalagmometric setup emerged as promising approach for deducing the hydrophobic character of whole protein molecules in their three-dimensional

conformation and in aqueous solution. Differences in the hydrophobic character depending on pH and protein species could be resolved.

4.3. Spectrophotometric determination of hydrophobicity

The dye measurements for BSA resulted in the highest hydrophobicity at pH 5, followed by pH 7 and 9. This hydrophobicity order is in agreement with our findings using the stalagmometric method. For lysozyme, human lysozyme and α-lactalbumin we had to increase the molar fraction 150 fold to see a distinct and reproducible shift of the absorption maximum. Despite the high molar fractions ΔA_{max} of BPB Na caused by lysozyme, human lysozyme and α-lactalbumin reached only half the value of ΔA_{max} for BSA. Additionally, the course of ΔA_{max} depending on lysozyme, human lysozyme and α-lactalbumin molar fraction were very similar.

Though BPB and BPB Na are well-known polarity sensitive dyes (Bertsch et al., 2003) they were so far mainly used for investigation of BSA and HSA (Bertsch et al., 2003; Bjerrum, 1968; Cao et al., 2003; Kragh-Hansen et al., 1974; Murakami et al., 1981; Tayyab and Qasim, 1990; Waldmann-Meyer and Schilling, 1956; Wei et al., 1996). It has been shown before, that the affinity for BPB and BPB Na is highly dependent on protein species (Cao et al., 2003; Flores, 1978; Waldmann-Meyer and Schilling, 1956). Cao et al. observed a great difference between the magnitude of the redshift of the BPB absorption maximum by BSA and Chitosan, a biopolymer 6 times smaller than BSA. However, protein size is not the only factor influencing the BPB-protein interaction. Investigations of Flores and Waldmann-Meyer and Schilling showed that even big proteins like ovalbumin and γ-globulin exhibited a significantly lower affinity for BPB and thus a significantly lower redshift in consequence. Likewise, small proteins like cytochrome c and myoglobin (Flores, 1978; Mayburd et al., 2000) were shown to cause pronounced redshifts. Consequently, the BPB Na method seems to be feasible for protein molecules with a high affinity to BPB Na like BSA. In this case a pH dependent hydrophobic character could be resolved. For lysozyme, human lysozyme and α-lactalbumin the dye method turned out to be inappropriate. For these proteins the dye method yielded in very similar results independent of the protein species and a low hydrophobic character compared to BSA, which is in disagreement with literature and the results of our stalagmometric method. Thus, spectrophotometric determination of protein hydrophobicity is highly dependent on protein-dye interactions and not an universally applicable method (Alizadeh-Pasdar and Li-Chan, 2000). In contrast, the stalagmometric method is able to characterize small hydrophobic and hydrophilic proteins and no limitations regarding solution composition need to be considered.

5. Conclusions

In the present work we have developed a high throughput stalagmometric method which is able to measure surface tensions in a highly accurate way and can be operated by liquid handling

stations and thus be integrated into high throughput work flow. This method was used to develop an innovative non-invasive approach for characterization of protein hydrophobicity on base of its impact on surface tension. The correlation between surface activity and hydrophobicity was validated using PEG of different molecular weights and applied to four different proteins, namely lysozyme, human lysozyme, BSA, and α-lactalbumin at four pH values. It was possible to rank protein hydrophobicity in dependency of the pH. Lysozyme was found to be hydrophilic, whereas α-lactalbumin turned out to be the most hydrophobic at pH 3. The derived ranking was in good agreement with literature and theoretical considerations regarding pH depending charge distributions.

The stalagmometric method was compared to an orthogonal and established spectrophotometric method for estimating protein hydrophobicity. Results of the spectrophotometric method regarding pH dependency of BSA were in agreement with the stalagmometric method. Spectrophotometric results for lysozyme, human lysozyme, and α-lactalbumin could not be used to derive protein hydrophobicity as only BSA caused reasonable shifts of the absorption maximum of BPB Na. Dye based methods are often restricted by pH and protein size and are highly influenced by the aromatic and aliphatic nature of the dye molecule. Contrary, the stalagmometric method is non-invasive and not restricted by pH and protein size. Differences in the hydrophobic character of proteins depending on protein species and pH could be resolved. Thus, we are convinced that the presented approach is an appropriate way to determine protein hydrophobicity in a non-invasive and highly accurate way.

Acknowledgement

We thank Matthias Franzreb and Jonas Wohlgemuth for performing the 3D printing. This research work is part of the projects ‘Molecular Interaction Engineering: From Nature’s Toolbox to Hybrid Technical Systems’ and ‘Proteinaggregation bei der Herstellung moderner Biopharmazeutika’, which are funded by the German Federal Ministry of Education and Research (BMBF)(031A095B and 0315342B).

References

Absolom, D.R., Zingg, W., Neumann, A.W., 1987. Protein adsorption to polymer particles: role of surface properties. J. Biomed. Mater. Res. 21, 161-171.

Alizadeh-Pasdar, N., Li-Chan, E.C.Y., 2000. Comparison of Protein Surface Hydrophobicity Measured at Various pH Values Using Three Different Fluorescent Probes. J. Agric. Food Chem. 48, 328-334.

Amrhein, S., Oelmeier, S.A., Dismer, F., Hubbuch, J., 2014a. Molecular dynamics simulations approach for the characterization of peptides with respect to hydrophobicity. J. Phys. Chem. B 118, 1707-1714.

Amrhein, S., Schwab, M.-L., Hoffmann, M., Hubbuch, J., 2014b. Characterization of aqueous two phase systems by combining lab-on-a-chip technology with robotic liquid handling stations. J. Chromatogr. A 1367, 68-77.

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., 2011. Aqueous two-phase systems for protein separation: A perspective. J. Chromatogr. A 1218, 8826-8835.

Bertsch, M., Mayburd, A.L., Kassner, R.J., 2003. The identification of hydrophobic sites on the surface of proteins using absorption difference spectroscopy of bromophenol blue. Anal. Biochem. 313, 187-195.

Bigelow, C.C., 1967. On the average hydrophobicity of proteins and the relation between it and protein structure. J. Theor. Biol. 16, 187-211.

Biswas, K.M., DeVido, D.R., Dorsey, J.G., 2003. Evaluation of methods for measuring amino acid hydrophobicities and interactions. J. Chromatogr. A 1000, 637-655.

Bjerrum, O.J., 1968. Interaction of Bromphenol Blue and Bilirubin with Bovine and Human Serum Albumin Determined by Gel Filtration. Scand. J. Clin. Lab. Investig. 22, 41-48.

Black, S.D., Mould, D.R., 1991. Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. Anal. Biochem. 193, 72-82.

Brems, D.N., Plaisted, S.M., Havel, H.A., Tomich, C.S., 1988. Stabilization of an associated folding intermediate of bovine growth hormone by site-directed mutagenesis. Proc. Natl. Acad. Sci. U. S. A. 85, 3367-3371.

Bull, H.B., Breese, K., 1974. Surface tension of amino acid solutions: A hydrophobicity scale of the amino acid residues. Arch. Biochem. Biophys. 161, 665-670.

Cao, W.G., Jiao, Q.C., Fu, Y., Chen, L., Liu, Q., 2003. Mechanism of the Interaction Between Bromophenol Blue and Bovine Serum Albumin. Spectrosc. Lett. 36, 197-209.

Cardamone, M., Puri, N.K., 1992. Spectrofluorimetric assessment of the surface hydrophobicity of proteins. Biochem. J. 282, 589-593.

Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., Trout, B.L., 2009. Design of therapeutic proteins with enhanced stability. Proc. Natl. Acad. Sci. U. S. A. 106, 11937-11942.

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.

Dill, K.A., 1990. Dominant forces in protein folding. Biochemistry 29, 7133-7155.

Dooley, K.H., Castellino, F.J., 1972. Solubility of amino acids in aqueous guanidinium thiocyanate solutions. Biochemistry 11, 1870-1874.

Eisenberg, D., 1984. Three-dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. Biochem. 53, 595-623.

Fendler, J.H., Nome, F., Nagyvary, J., 1975. Compartmentalization of Amino Acids in Surfactant Aggregates. J. Mol. Evol. 6, 215-232.

Flores, R., 1978. A Rapid and Reproducible Assay for Quantitative Estimation of Proteins using Bromophenol Blue. Anal. Biochem. 88, 605-611.

Genest, S., Schwarz, S., Petzold-Welcke, K., Heinze, T., Voit, B., 2013. Characterization of highly substituted, cationic amphiphilic starch derivatives: Dynamic surface tension and intrinsic viscosity. Starch/Stärke 65, 999-1010.

Harris, J.M., 1992. Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications. Springer.

Hawe, A., Sutter, M., Jiskoot, W., 2008. Extrinsic fluorescent dyes as tools for protein characterization. Pharm. Res. 25, 1487-1499.

Hendriks, J., Gensch, T., Hviid, L., van der Horst, M.A., Hellingwerf, K.J., van Thor, J.J., 2002. Transient exposure of hydrophobic surface in the photoactive yellow protein monitored with Nile Red. Biophys. J. 82, 1632-1643.

Jańczuk, B., Białopiotrowicz, T., Wójcik, W., 1989. The components of surface tension of liquids and their usefulness in determinations of surface free energy of solids. J. Colloid Interface Sci. 127, 59-66.

Janin, J., 1979. Surface and inside volumes in globular proteins. Nature 277, 491-492.

Janson, J.-C., 2012. Protein purification: principles, high resolution methods, and applications. John Wiley & Sons.

Kato, A., Nakai, S., 1980. Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. Biochim. Biophys. Acta 624, 13-20.

Keshavarz, E., Nakai, S., 1979. The relationship between hydrophobicity and interfacial tension of proteins. Biochim. Biophys. Acta 576, 269-279.

Kragh-Hansen, U., Møller, J.V., Lind, K.E., 1974. Relation between binding of phenolsulfophtalein dyes and other ligands with a high affinity for human serum albumin. Biochim. Biophys. Acta 365, 360-371.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Lijnzaad, P., Berendsen, H.J.C., Argos, P., 1996. A method for detecting hydrophobic patches on protein surfaces. Proteins 26, 192-203.

Mayburd, A.L., Tan, Y., Kassner, R.J., 2000. Complex formation between Chromatium vinosum ferric cytochrome c' and bromophenol blue. Arch. Biochem. Biophys. 378, 40-44.

Melinder, Å., 2007. Thermophysical properties of aqueous solutions used as secondary working fluids. Ph.D. thesis, School of Industrial Engineering and Management, Royal Institute of Technology, KTH Stockholm.

Miller, S., Janin, J., Lesk, A.M., Chothia, C., 1987. Interior and surface of monomeric proteins. J. Mol. Biol. 196, 641-656.

Murakami, K., Sano, T., Yasunaga, T., 1981. Kinetic Studies of the Interaction of Bromophenol Blue with Bovine Serum Albumin by Pressure-jump Method. Bull. Chem. Soc. Jpn. 54, 862-868.

Nieba, L., Honegger, A., Krebber, C., Pluckthun, A., 1997. Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. Protein Eng. 10, 435-444.

Nozaki, Y., Tanford, C., 1963. The solubility of amino acids and related Compounds in aqueous urea solutions. J. Biol. Chem. 238, 4074-4081.

Nozaki, Y., Tanford, C., 1965. The solubility of amino acids and related compounds in equeous ethylene glycol solutions. J. Biol. Chem. 240, 3568-3573.

Nozaki, Y., Tanford, C., 1970. The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. J. Biol. Chem. 245, 1648-1652.

Nozaki, Y., Tanford, C., 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. J. Biol. Chem. 246, 2211-2217.

Pegram, L.M., Record, M.T., 2007. Hofmeister Salt Effects on Surface Tension Arise from Partitioning of Anions and Cations between Bulk Water and the Air−Water Interface. J. Phys. Chem. B 111, 5411-5417.

Permyakov, E.A., Berliner, L.J., 2000. α-Lactalbumin: structure and function. FEBS Lett. 473, 269-274.

Radzicka, A., Wolfenden, R., 1988. Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry 27, 1664-1670.

Reißer, S., Strandberg, E., Steinbrecher, T., Ulrich, A.S., 2014. 3D Hydrophobic Moment Vectors as a Tool to Characterize the Surface Polarity of Amphiphilic Peptides. Biophys. J. 106, 2385-2394.

Rose, G.D., Wolfenden, R., 1993. Hydrogen bonding, hydrophobicity, packing, and protein folding. Annu. Rev. Biophys. Biomol. Struct. 22, 381-415.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2005. Is it possible to predict the average surface hydrophobicity of a protein using only its amino acid composition? Journal of Chromatography A 1075, 133-143.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2006a. Predicting the behaviour of proteins in hydrophobic interaction chromatography 1: Using the hydrophobic imbalance (HI) to describe their surface amino acid distribution. Journal of Chromatography A 1107, 110-119.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2006b. Predicting the behaviour of proteins in hydrophobic interaction chromatography 2. Using a statistical description of their surface amino acid distribution. Journal of Chromatography A 1107, 120-129.

Shanbhag, V.P., Axelsson, C.-G., 1975. Hydrophobic interaction determined by partition in aqueous two-phase systems. Eur. J. Biochem. 60, 17-22.

Tanford, C., 1962. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. J. Am. Chem. Soc. 84, 4240-4247.

Tate, T., 1864. XXX. On the magnitude of a drop of liquid formed under different circumstances. London, Edinburgh, Dublin Philos. Mag. J. Sci. 27, 176-180.

Tayyab, S., Qasim, M.A., 1990. Binding of bromophenol blue to bovine serum albumin and its succinylated forms. Int. J. Biol. Macromol. 12, 55-58.

Trinquier, G., Sanejouand, Y.-H., 1998. Which effective property of amino acids is best preserved by the genetic code? Protein Eng. 11, 153-169.

Tsierkezos, N.G., Molinou, I.E., 1998. Thermodynamic Properties of Water + Ethylene Glycol at 283.15, 293.15, 303.15, and 313.15 K. J. Chem. Eng. Data 43, 989-993.

Waldmann-Meyer, H., Schilling, K., 1956. The Interaction of Bromophenol Blue with Serum Albumin and γ-Globulin in Acid Medium. Arch. Biochem. Biophys. 64, 291-301.

Wei, Y.-J., Li, K.-A., Tong, S.-Y., 1996. The interaction of Bromophenol Blue with proteins in acidic solution. Talanta 43, 1-10.

Whitney, P.L., Tanford, C., W, P.L., Tanford, C., 1962. Solubility of amino acids in aqueous urea solutions and its implications for the denaturation of proteins by urea. J. Biol. Chem. 237, 1735-1737.

3.4 Predictive power of protein surface characteristics and conformational flexibility for protein aggregation propensity

Lara Galm[1], Sven Amrhein[1], Jürgen Hubbuch*

Institute of Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 3, Karlsruhe 76131, Germany

** Corresponding author. Tel.: +49 721 608 47526; fax: +49 721 608 46240. E-mail address: juergen.hubbuch@kit.edu.*

[1] *Contributed equally.*

in preparation

Abstract

Protein aggregation is a highly complex objective and challenges the production and development of biopharmaceuticals at every stage. Thus, gaining understanding of protein aggregation is crucial to ensure quality and safety of biopharmaceuticals. Protein aggregation has been described to be mainly influenced by surface hydrophobicity and electrostatics, though different publications disagree about electrostatics or hydrophobicity being the key driver for protein aggregation. Furthermore, protein conformational flexibility has been described as the main mechanistic determinant of protein aggregation propensity. In the present manuscript the predictive power of protein surface characteristics and conformational flexibility for the aggregation behavior of α-lactalbumin, human lysozyme, and lysozyme from chicken egg white at pH 3, 5, 7, and 9 without precipitants and with precipitants, namely sodium chloride and ammonium sulfate, was tested. Therefore, the correlation between protein zeta potential, protein surface hydrophobicity, protein conformational flexibility and aggregation propensity was probed. Protein zeta potential and surface hydrophobicity were determined experimentally, whereas conformational flexibility was determined using molecular dynamics (MD) simulations. Protein aggregation propensity was extracted from protein phase diagrams. For the protein solutions without precipitants the hydrophobic character of the protein surface and conformational flexibility could be identified to be main determinants for aggregation propensity. For representation of pH dependent differences in the aggregation propensity within one protein species electrostatic surface properties, i.e. the zeta potential, needed to be taken into account additionally. Investigation of the correlation between surface hydrophobicity, conformational flexibility, and aggregation propensity with sodium chloride and ammonium sulfate as precipitants revealed a strong dependency from the salt species and the isoelectric point (pI) of the proteins. For sodium chloride as precipitant and pH<pI protein flexibility was best suited for estimation of aggregation propensity at extreme pH values (pH 3 and pH 9), whereas surface hydrophobicity was best suited for moderate to physiological pH values (pH 5 and pH 7). For ammonium sulfate as precipitant and pH<pI surface hydrophobicity was best suited for estimation of aggregation propensity of different proteins at all pH values investigated. Though, neither for sodium chloride nor for ammonium sulfate a clear correlation between surface hydrophobicity or protein flexibility and aggregation propensity as function of pH within one protein species could be

found. Overall, the hydrophobic character of the protein surface and protein conformational flexibility could be identified to be suitable parameters for prediction of aggregation propensity, albeit with some limitations.

Keywords: *Protein surface hydrophobicity, Zeta potential, Hydrophilicity, Protein phase diagrams, Molecular dynamics simulations*

1. Introduction

Protein aggregation describes the assembly of protein monomers to protein multimers. Thus, protein aggregation includes the emergence of crystalline and amorphous precipitated protein phase states and also incorporates cross-linked structures like protein gels or non-native structures like protein skins. Occurrence of protein aggregation is a challenge in biopharmaceutical industries as it might be a desirable or an undesirable event. On the one hand protein aggregation is applied as acknowledged process step in biopharmaceutical industries either for formulation or purification purposes (Basu et al., 2004; Brange and Vølund, 1999; Jen and Merkle, 2001; Scopes, 1994; Vajo et al., 2001; Yang et al., 2003). On the other hand aggregation processes might influence biological activity and bioavailability of protein therapeutics and might cause immunogenic reactions after application (Brange and Vølund, 1999; Jen and Merkle, 2001; Vajo et al., 2001) which is why the level of protein aggregates has to be thoroughly monitored as critical quality attribute (Paul et al., 2012). Protein aggregation propensity in aqueous solution depends on the conformational stability or flexibility and on the colloidal stability of proteins in solution (Chi et al., 2003; Mahler et al., 2009; Valerio et al., 2005). Conformational stability hereby describes profound conformational changes such as partial unfolding or complete denaturation, whereas conformational flexibility describes minor structural fluctuations. Proteins are not rigid and static, but highly flexible and in constant motion between different conformational states with similar energies (Carlson and McCammon, 2000; Petsko and Ringe, 2004b). This conformational flexibility of proteins is amongst others essential for protein function, ligand binding, and metabolism (Carlson and McCammon, 2000; Petsko and Ringe, 2004b; Teague, 2003). Protein conformational flexibility is as well connected to protein phase behavior, aggregation propensity and thermostability (Sousa, 1995; Valerio et al., 2005; Vihinen, 1987). Twenty years ago, Sousa found that increasing conformational flexibility in a protein will increasingly favor amorphous precipitation over crystallization. Ten years ago, Valerio et al. postulated that conformational flexibility is the main mechanistic determinant of protein aggregation propensity (the higher the conformational flexibility the higher the aggregation propensity). As described above, in addition to conformational stability or flexibility, the colloidal stability of proteins in solution is a main influencing factor for protein

aggregation. Colloidal stability hereby depends on protein-protein interactions arising due to protein surface characteristics. Protein surface hydrophobicity and electrostatics are thought to be of major importance during protein aggregation, though different publications disagree about electrostatics (Price et al., 2011; Yadav et al., 2012; Yadav et al., 2011) or hydrophobicity (Chennamsetty et al., 2009; Kumar et al., 2011; Zhang and Topp, 2012) being the main determinant. Lauer et al. (Lauer et al., 2012) found that the higher the protein surface hydrophobicity the higher the aggregation propensity, whereas the influence of the protein net charge needed to be evaluated individually.

However, despite this general knowledge about main influencing factors for protein aggregation, aggregation conditions are mainly found by trial and error, whereas undesired protein aggregation is often discovered not until process set-up. General approaches for prediction of protein aggregation propensity are consequently rare but would be highly valuable due to an immense reduction of financial, material and time effort in comparison to classical empirical approaches. The prediction of protein aggregation propensity is highly challenging due to the complexity of protein molecules and the diversity of protein-solute and protein-protein interactions. This complexity requires investigation on protein structural level as possible via *in silico* methodologies such as quantitative structure–activity relationship (QSAR) models or molecular dynamics (MD) simulations. For examples of the few *in silico* approaches for prediction of protein aggregation propensity see for instance (Brunsteiner et al., 2013; Lauer et al., 2012; Valerio et al., 2005). Thus, the present manuscript aims to develop an approach for prediction of protein aggregation propensity based on conformational flexibility and protein surface characteristics. With reference to Valerio et al. (Valerio et al., 2005) a combination of conformational flexibility and of electrostatic and hydrophobic protein surface properties is supposed to be best suited for development of a predictive approach for protein aggregation propensity and was applied in this manuscript. In the present study the correlation between conformational flexibility, protein surface characteristics, and aggregation propensity was examined for three different model proteins, namely α-lactalbumin (α -lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}), at four different pH values (pH 3, 5, 7, and 9). Conformational flexibility was determined by an *in silico* approach using molecular dynamics (MD) simulations, whereas protein surface characteristics were determined experimentally. Conformational flexibility was assessed by analysis of the root-mean-square-fluctuation (RMSF) of the C_α, C, and N backbone atoms of the proteins. To assess protein surface characteristics, the zeta potential and the hydrophobic character were determined as function of pH. The zeta potential describes the overall net charge of the protein surface and was determined by measuring the electrophoretic mobility of proteins in the respective solution. The hydrophobic character was determined according to Amrhein et al. (Amrhein et al., 2015). To correlate conformational flexibility and surface characteristics to protein aggregation propensity, protein

phase states were determined in a protein concentration range between 2.5 and 21.75 mg/mL as function of pH. Additionally, the influence of sodium chloride and ammonium sulfate on protein aggregation propensity was investigated and the predictive power of conformational flexibility and surface characteristics for the aggregation propensity of solutions with precipitants was tested.

2. Material and methods

2.1. Selection of proteins

α-lactalbumin (α-lact), human lysozyme (lys_{hum}) and lysozyme from chicken egg white (lysozyme, lys_{egg}) were investigated at pH 3, 5, 7, and 9. α-lactalbumin from bovine milk in its calcium depleted form (L6010) and human lysozyme (L1667) were purchased from Sigma-Aldrich (St. Louis, MO, USA), lysozyme from chicken egg white (HR7-110) was purchased from Hampton Research (Aliso Viejo, CA, USA). These proteins were chosen for investigation due to their similar size and the high similarity in secondary and tertiary structure. α-lact has a molecular weight of 14.175 kDa (Smith et al., 1990), lys_{hum} has a molecular weight between 14 and 15.6 kDa (Lüllig et al., 1983) and lys_{egg} has a molecular weight of 14.3 kDa (Smith et al., 1990). The similarity of the secondary and tertiary structures between the three proteins is displayed by superposition of their crystal structures in Fig. 1. This selection of proteins thus enables to exclude size and secondary and tertiary structure depending effects while studying pH depending surface properties and charge distributions respectively. Additionally, the chosen proteins are small enough to conduct molecular dynamics (MD) simulations with a reasonable computational effort. Though, despite this similarity in size and secondary and tertiary structure, α-lact, lys_{hum}, and lys_{egg} differ in the location of their isoelectric points and in their primary structure and are supposed to exhibit differing phase behavior and aggregation propensity. The isoelectric point of α-lact is located between pH 4.25 and 4.5 (Robbins et al., 1965), between 10.0 and 11.0 for lys_{hum} (Lundblad et al., 1972; Righetti and Caravaggio, 1976) and between 10.0 and 11.0 for lys_{egg} as well (Lüllig et al., 1983; Righetti et al., 1981). The sequence identities of the primary sequences are low as is shown in Tab. 1.

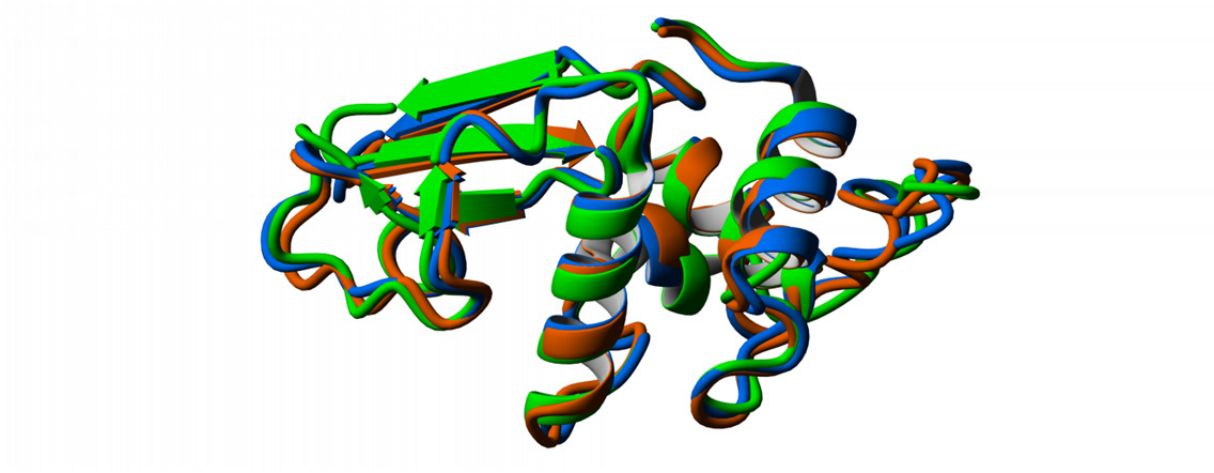

Fig. 1. Superposition of the crystal structures of α-lactalbumin (pdb entry 1F6R) in green, human lysozyme (pdb entry 2NWD) in blue and lysozyme from chicken egg white (pdb entry 2VB1) in orange, all depicted in ribbon style. Protein alignment was conducted applying the MUSTANG method (Konagurthu et al., 2006). Molecular graphics were created using Yasara (www.yasara.org) and POVRay (www.povray.org).

Tab. 1. Percental sequence identity between α-lactalbumin (pdb entry 1F6R), human lysozyme (pdb entry 2NWD) and lysozyme from chicken egg white (pdb entry 2VB1), listed pairwise.

Sequence identity [%]	1F6R	2NWD	2VB1
1F6R		38.94	40.37
2NWD	38.94		60.94
2VB1	40.37	60.94	

2.2. Preparation of buffers and protein solutions

Used buffer substances were citric acid (Merck, Darmstadt, Germany) and sodium citrate (Sigma–Aldrich, St. Louis, MO, USA) for pH 3, sodium acetate (Sigma–Aldrich, St. Louis, MO, USA) and acetic acid (Merck, Darmstadt, Germany) for pH 5, MOPSO (AppliChem, Darmstadt, Germany) for pH 7 and Bis-Tris propane (Molekula, Dorset, UK) for pH 9. Buffers were prepared as described in Baumgartner et al. (Baumgartner et al., 2015). Buffer capacity was 100 mM. The pH was adjusted with the appropriate titrant with an accuracy of ±0.05 pH units. Proteins were weighed in, dissolved in the respective buffer, filtered through 0.2 µm syringe filters with PTFE membranes (VWR, Radnor, PA, USA) to remove particulates and desalted using a HiTrap Desalting Column (GE Healthcare, Uppsala, Sweden) on an AEKTAprime™ plus system (GE Healthcare, Uppsala, Sweden). A subsequent protein concentration step, if needed, was performed using Vivaspin centrifugal concentrators (Sartorius, Goettingen, Germany) with PES membranes and a molecular weight cutoff of 3 kDa.

2.3. Protein surface characterization

The protein net charge and the hydrophobic character of the molecular surfaces of α-lactalbumin

(α-lact), human lysozyme (lys_{hum}) and lysozyme from chicken egg white (lys_{egg}) were determined at different pH values, as charge and hydrophobicity of the proteins are thought to be of major importance for aggregation propensity. Measurements for determination of net charge and hydrophobicity of the protein surface were conducted without any additional salts apart from buffer components in solution.

2.3.1. Determination of electrostatic properties (zeta potential)

Protein net charge was assessed using the zeta potential of the protein surface. Determination of the zeta potential of α-lact, lys_{hum}, and lys_{egg} at pH 3, 5, 7, and 9 was performed with a Zetasizer Nano ZSP (Malvern Instruments Ltd, Malvern, United Kingdom). The measurements were conducted using the so-called diffusion barrier technique. Therefore, folded capillary cells (DTS1070, Malvern Instruments Ltd, Malvern, United Kingdom) were filled completely with the respective buffer and hereafter 30 µL of the protein solution were injected to the measurement zone of the buffer filled cell using a Corning® Costar gel-loading pipet tip (4853, Corning Incorporated, Corning, NY, USA). To ensure conformational as well as colloidal integrity of the proteins during zeta potential determination, a measurement procedure was set up with a dynamic light scattering measurement for determination of the protein size first, followed by the zeta potential determination and a dynamic light scattering measurement again. During the zeta potential determination the light scattered by the protein undergoing electrophoresis is combined with a reference beam. The rate of change of phase of the combined beam is measured and used to calculate the electrophoretic mobility and thus the zeta potential by using the Smoluchowski equation (Hunter, 1981). A significant Joule heating was avoided by setting the voltage to limit the current to a maximum of ±5 mA and by limiting the number of runs, i.e. number of phase changes, to maximum 15. The measurements were conducted at least threefold.

2.3.2 Determination of hydrophobic properties

Determination of the hydrophobic character of α-lact, lys_{hum}, and lys_{egg} at pH 3, 5, 7, and 9 was conducted as described by Amrhein et al. (Amrhein et al., 2015). There, it could be shown that the surface tension increment as a function of protein concentration is related to the pH depending hydrophobic character of proteins. Surface tension increments were determined using a novel, non-invasive stalagmometric approach. Surface tension increments were used to rank the proteins according to their hydrophobic character.

2.4 In silico *protein characterization by means of molecular dynamics (MD) simulations*

Molecular dynamics (MD) simulations of the proteins of interest in solutions varying in pH were performed in order to estimate the proteins' flexibility and the root-mean-square-fluctuation (RMSF) of atomic positions in particular as a measure of conformational flexibility. The complete MD simulation workflow starting from protein structure preparation over the calculation of the

initial spatial solvent configuration to the final MD simulation and evaluation is explained in detail in the following. MD simulations of the proteins of interest in solutions varying not solely in pH, but also in salt concentration and salt species to estimate protein conformational flexibility as function of salt concentration and salt species are currently running.

2.4.1. Starting structures

Selection of protein structure data driven by sequence completeness and resolution is listed in Tab. 2. The starting structures were initially prepared with the MD software package Yasara Structure 14.6.23 (Krieger et al., 2002) by applying force field parameters of AMBER03 (Duan et al., 2003).

Default pK_a values were assigned to derive a rough protonation, followed by an optimization of the hydrogen bonding network and a simulated annealing energy minimization until the energy deviated by less than 10 J/mol per atom during 200 steps at 293.15 K. These protein structures then were optimized in terms of protonation and structure to match the buffer pH by applying the H++ algorithm (Anandakrishnan et al., 2012). For neutralization purposes the force field parameters of the ions sodium and chloride are included in the AMBER03 force field.

Tab. 2. Selection of protein structure data for molecular dynamics simulation.

Protein Name (Abbreviation)	PDB ID	Reference
α-lactalbumin apo (α-lact)	1F6R	(Chrysina et al., 2000)
lysozyme human (lys_{hum})	2NWD	(Durek et al., 2007)
lysozyme from chicken egg white (lys_{egg})	2VB1	(Wang et al., 2007)

2.4.2. Initial solvation

The dimension of the cubic MD simulation cell was chosen to be 30 Å larger than the protein's largest spatial dimension according to the van der Waals radii defined in the AMBER03 force field. The solvent volume and mass were calculated by means of the protein's molecular volume and the total box volume in combination with the respective density at 293.15 K to fit the desired protein concentration. The TIP3P model was used for water molecules.

Fully randomized solvation assembles were derived in 10-fold for each protein and pH value by applying the PACKMOL algorithm (Martínez and Martínez, 2003; Martínez et al., 2009).

2.4.3. Molecular dynamics (MD) simulations

Starting with the thus derived packed assembles the final explicit MD simulations were performed via Yasara on a high performance distributed memory parallel computer (InstitutsCluster II) at the Steinbuch Centre for Computing (SCC) of the Karlsruhe Institute of Technology (KIT).

The AMBER03 force field was applied in combination with calculating electrostatic Coulomb forces without cutoff using the particle mesh Ewald (PME) algorithm (Essmann et al., 1995). The van der Waals forces were calculated with a cutoff of 7.86 Å, as Yasara is optimized to cutoffs multiple of 2.62 Å. The temperature was set to 293.15 K and was controlled by a reassignment of atom velocities and thus the resulting kinetic energy.

After an annealing energy minimization the whole system was slowly heated up from 0 K to 293.15 K in 1 K steps. Each step was hold while the maximal speed within the system was larger than 1000 m/s. This procedure inhibited steric clashes. After the desired temperature of 293.15 K was reached, the final MD simulations were performed over 5 ns.

2.4.4. Assessment of conformational flexibility

Molecular dynamics of each protein were analyzed for root-mean-square-fluctuations (RMSF) of the backbone atoms (C_α, C, and N) with respect to the starting structure by applying a Yasara implemented algorithm. The RMSF of atom *i* was determined by Eq. (1), where *j* runs over the three cartesian components *x*, *y* and *z* of the atom position vector *P*, and *k* runs over the *N* coordinate sets.

$$\boldsymbol{RMSF_i = \sqrt{\sum_{j=1}^{3}\left(\frac{1}{N}\sum_{k=1}^{N} P_{ikj}^2 - \bar{P}_{ij}^2\right)}} \qquad \textbf{(1)}$$

The evaluation was conducted by fully automated routines written in Matlab® R2015a (MathWorks®, Natick, MA, USA) and employing Yasara.

2.5. Generation of protein phase diagrams

Protein phase diagrams are essential for investigation of protein phase states and thus for determination of protein aggregation propensity as function of precipitant species and precipitant concentration. Protein phase diagrams were generated for pH 3, 5, 7 and 9 using sodium chloride and ammonium sulfate as precipitants. Sodium chloride was obtained from Merck (Darmstadt, Germany) and ammonium sulfate was from VWR (Radnor, PA, USA). The precipitant stock solutions were prepared as described in detail in Baumgartner et al. (Baumgartner et al., 2015). Phase diagrams for human lysozyme (lys_{hum}) and lysozyme from chicken egg white (lys_{egg}) were generated earlier (Baumgartner et al., 2015). α-lactalbumin (α-lact) concentration in the protein stock solutions at pH 3, 7, and 9 was adjusted to 43.5±1 mg/mL. At pH 5 α-lact could only be concentrated to 7±1 mg/mL due to solubility limitations. The protein phase diagrams were generated by adding 12 µL of diluted protein solution to 12 µL of diluted precipitant solution on MRC under Oil 96 Well Crystallization Plates (Swissci, Neuheim, Switzerland) by using a Freedom EVO® 100 (Tecan, Maennedorf,

Switzerland) automated liquid handling station. Resulting α-lact concentration on the crystallization plates ranged between 2.5 and 21.75 mg/mL for pH 3, 7, and 9 and between 0.4 and 3.5 mg/mL for pH 5, sodium chloride concentration ranged between 0 and 2.5 M, ammonium sulfate concentration ranged between 0 and 1.5 M. Crystallization plates were then centrifuged for 1 min at 1000 rpm to remove air bubbles and covered using optically clear and UV compatible sealing tape (HDclear™ sealing tape (ShurTech Brands, Avon, OH, USA)). The sealed plates were stored in a Rock Imager 54 automated imaging device (Formulatrix, Waltham, MA, USA) for 40 days at 20 °C that collected microscopic images of the individual wells of the crystallization plates. The images were examined after 40 days. Six possible phase states were classified: clear solution, crystallization, precipitation, skin formation, gelation and phase separation. For further details on generation of protein phase diagrams in microbatch experiments see Baumgartner et al. (Baumgartner et al., 2015).

3. Results

3.1. Protein surface characteristics

3.1.1. Electrostatic properties (zeta potential)

Fig. 2 shows the values of the zeta potential of α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}) at pH 3, 5, 7, and 9. Zeta potentials could be determined with a maximum standard deviation of ±2.6 mV. α-lact exhibited a positive zeta potential for pH 3 and a negative zeta potential for pH 5, 7, and 9. The course of the zeta potential as function of pH is very similar for lys_{hum} and lys_{egg}. Lys_{hum} and lys_{egg} exhibited a positive zeta potential at pH 3 and pH 5 and a negative zeta potential for pH 7 and pH 9, though the isoelectric points are expected to be located between pH 10.0 and pH 11.0 (Lüllig et al., 1983; Lundblad et al., 1972; Righetti and Caravaggio, 1976; Righetti et al., 1981). However, values for the zeta potential of lys_{hum} and lys_{egg} at pH 7 and pH 9 are only slightly negative. α-lact obtained the highest absolute value of the zeta potential at pH 9 where it reached 18.9 mV. Lys_{hum} and lys_{egg} obtained the highest absolute value of the zeta potential at pH 3 with 7.4 mV for lys_{hum} and 10.7 mV for lys_{egg}.

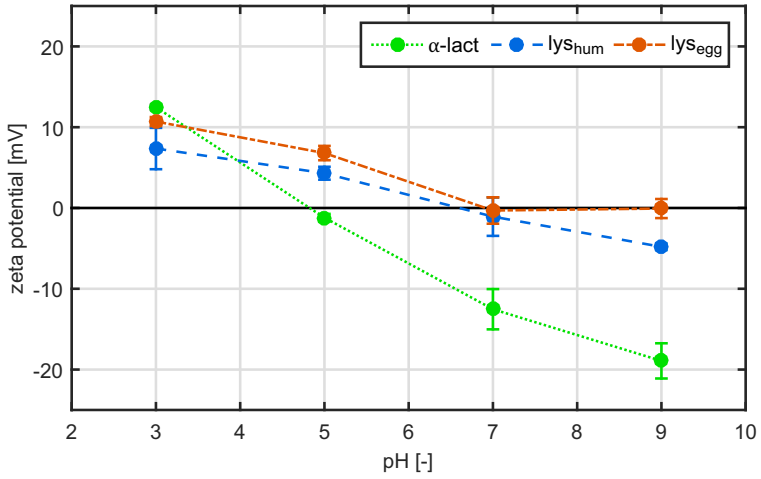

Fig. 2. Values for the zeta potential of α–lact, lys$_{hum}$, and lys$_{egg}$ at pH 3, 5, 7, and 9 without salt. Lines should guide the viewer's eyes.

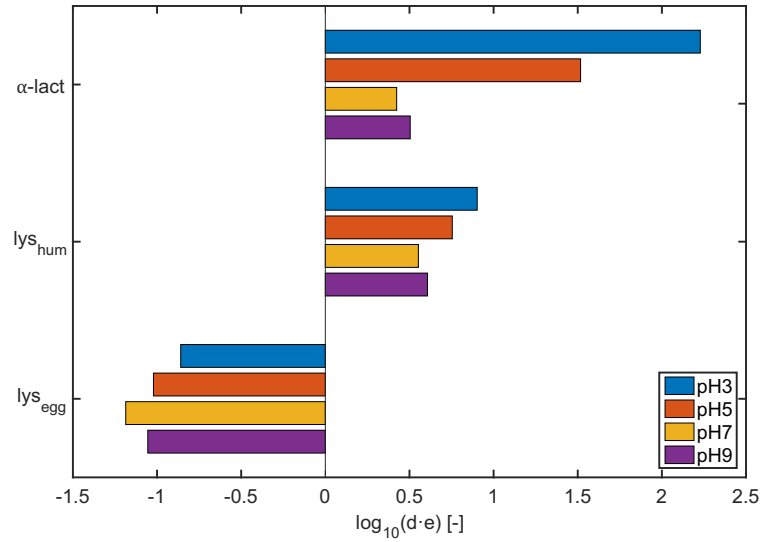

Fig. 3. Hydrophobic character of α–lact, lys$_{hum}$, and lys$_{egg}$ at pH 3, 5, 7, and 9 without salt. Data was taken from Amrhein et al. (Amrhein et al., 2015).

3.1.2. Hydrophobic properties

Fig. 3 shows the ranking of the hydrophobic character of α-lact, lys$_{hum}$, and lys$_{egg}$ at pH 3, 5, 7, and 9 as it was determined by Amrhein et al. (Amrhein et al., 2015). The strongest hydrophobic character (quantified by $\log_{10}$(d·e)) among the investigated proteins was obtained by α-lact at pH 3 and pH 5, followed by lys$_{hum}$ at pH 3 and pH 5. Though, the hydrophobic character of α-lact at pH 3 and pH 5 is significantly higher than the hydrophobic character of lys$_{hum}$ at pH 3 and pH 5. α-lact and lys$_{hum}$ both obtained their lowest hydrophobic character among the investigated pH values at pH 7. Negative values for the hydrophobic character, as obtained by lys$_{egg}$, indicated lys$_{egg}$ to be rather hydrophilic than hydrophobic. Here, the most pronounced hydrophilic

character of lys_{egg} was obtained at pH 7 and the least pronounced hydrophilic character was obtained at pH 3. The hydrophobic character as function of pH in descending order is: α-lact (pH 3) > α-lact (pH 5) >> lys_{hum} (pH 3) > lys_{hum} (pH 5) > lys_{hum} (pH 9) > lys_{hum} (pH 7) > α-lact (pH 9) > α-lact (pH 7) >> lys_{egg} (pH 3) > lys_{egg} (pH 5) > lys_{egg} (pH 9) > lys_{egg} (pH 7).

3.2. Protein conformational flexibility without precipitants

Fig. 4 shows the root-mean-square-fluctuations (RMSF) of the C_α, C, and N backbone atoms of α-lact (1F6R, top), lys_{hum} (2NWD, middle), and lys_{egg} (2VB1, bottom) at pH 3, 5, 7, and 9 plotted against the primary sequence in 1-letter code. The RMSF as function of the primary sequence was calculated as 95 % confidence trimmed mean value from the 10-fold simulations. The highest RMSF value for α-lact (1F6R) is slightly above 2 Å, whereas maximum fluctuations of lys_{hum} (2NWD) and lys_{egg} (2VB1) backbone atoms are around 1.8 Å. High RMSF values in this context correspond to high protein flexibility. It is apparent from Fig. 4 that there are conserved sections within the primary sequence of α-lact (1F6R), lys_{hum} (2NWD), and lys_{egg} (2VB1) and sections that exposed enhanced flexibility in dependency of the pH. In order to visualize the flexibility of the three dimensional structure, the flexibility of each amino acid as the mean of the respective backbone atom RMSF is projected onto the protein structures in a color code in Fig. 5. As illustrated in Fig. 5 the highly conserved regions can be allocated to highly organized structures and α-helix- and β-sheet- structures in particular, stabilized by regular hydrogen-bond networks. The highest protein flexibility occurred in random coil structures and loops, i.e. secondary structure elements without regular hydrogen-bonding. For further investigation the proteins were ranked pH dependent according to their mean flexibility. The pH dependent flexibility ranking is shown in Fig. 6. Conformational flexibility of α-lact (1F6R) was highest amongst the investigated protein species for all pH values, followed by lys_{egg} (2VB1). The highest mean RMSF value for α-lact (1F6R) was 0.796 Å at pH 3. Lys_{egg} (2VB1) exposed the highest flexibility at pH 3 as well, with a mean RMSF of 0.67 Å. The mean RMSF values of lys_{egg} at pH 5, 7, and 9 are very similar and around 0.64 Å. Lys_{hum} (2NWD) exhibited the lowest conformational flexibility amongst the investigated proteins. The highest mean RMSF value and thus the highest flexibility of lys_{hum} was observed at pH 7 (0.59 Å), followed by pH 9 (0.58 Å) and pH 3 (0.56 Å). The lowest flexibility of lys_{hum} was observed at pH 5 (0.55 Å). Generally, it has to be mentioned that mean RMSF values within one protein species as function of pH only vary slightly whereas mean RMSF values of different protein species vary significantly.

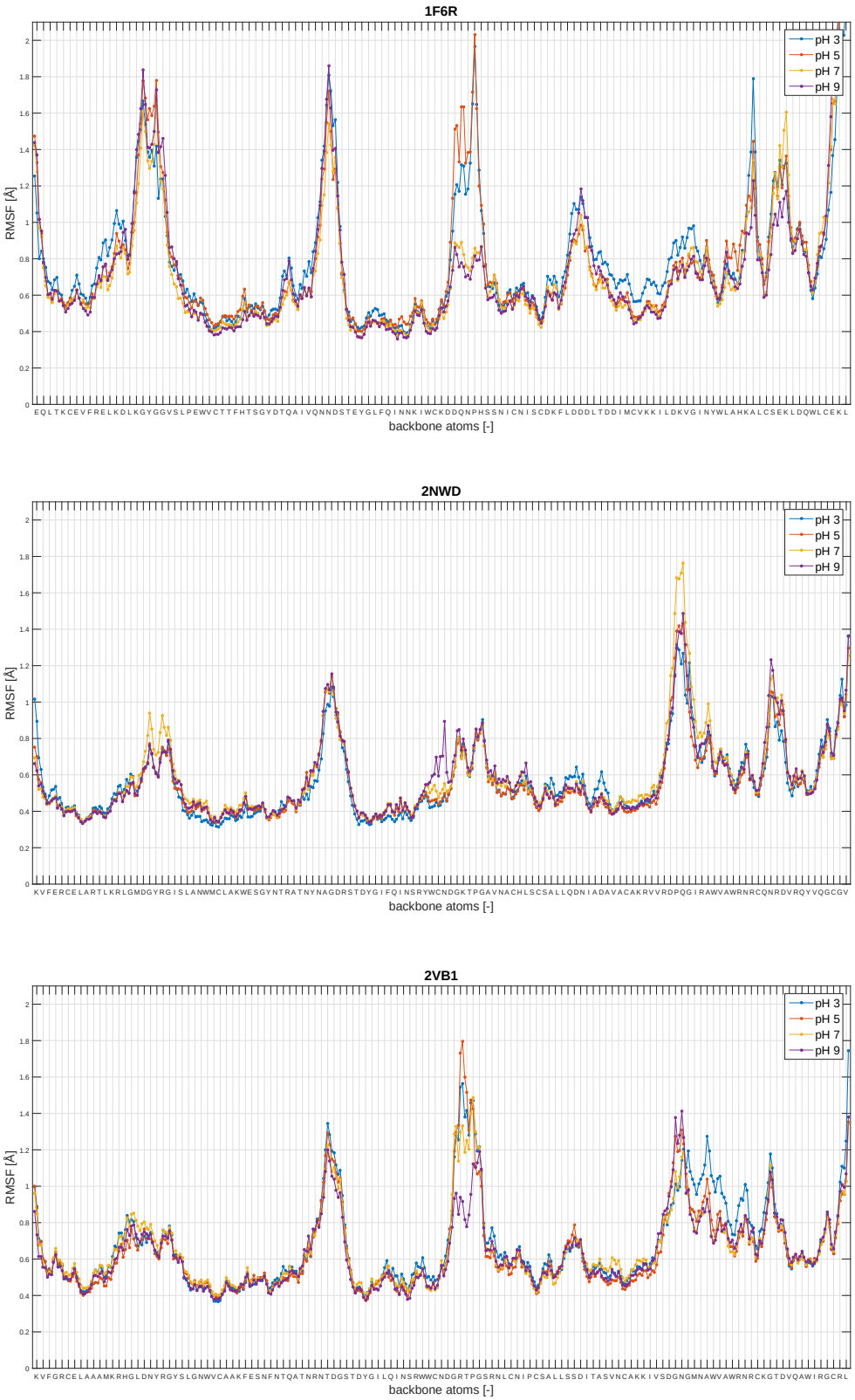

Fig. 4. Root-mean-square-fluctuation (RMSF) of the C_α, C, and N backbone atoms of α-lact (1F6R, top), lys_{hum} (2NWD, middle), and lys_{egg} (2VB1, bottom) at pH 3, 5, 7, and 9 plotted against the primary sequence in 1-letter code.

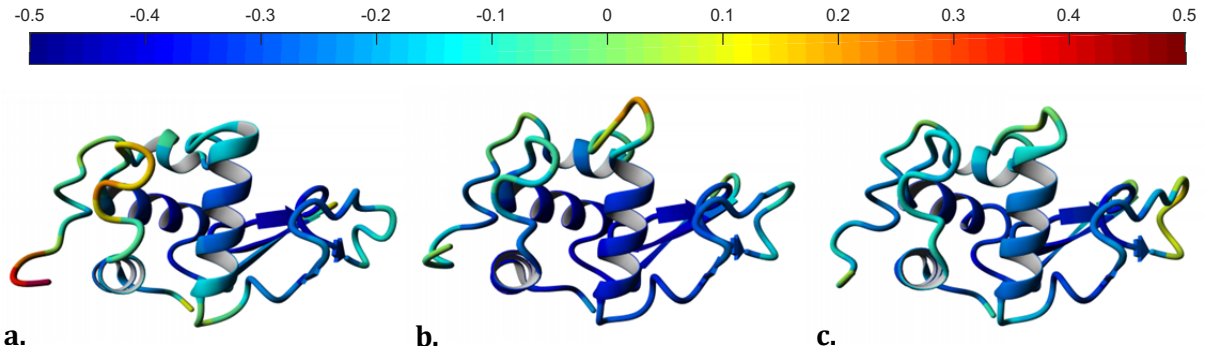

Fig. 5. Common logarithm of RMSF values (log_{10}(RMSF)) projected onto the ribbon depictedd protein structures of (a.) α-lact (1F6R), (b.) lys_{hum} (2NWD), and (c.) lys_{egg} (2VB1) at pH 7. The color bar on top represents the color code for the logarithmized RMSF values. Molecular graphics were created using Yasara (www.yasara.org) and POVRay (www.povray.org).

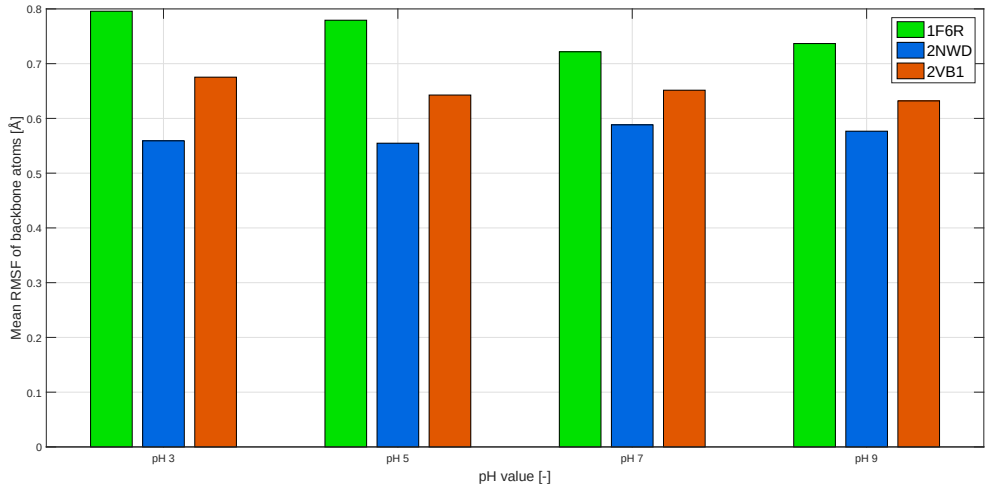

Fig. 6. Mean root-mean-square-fluctuation (RMSF) of the C_α, C, and N backbone atoms of α-lact (1F6R), lys_{hum} (2NWD), and lys_{egg} (2VB1) as function of pH. Mean RMSF is interpreted as mean protein conformational flexibility.

3.3. Phase behavior without precipitants

Fig. 7 shows the phase states of α-lact, lys_{hum} and lys_{egg} at pH 3, 5, 7, and 9 without precipitants. Phase transitions in the investigated protein concentration range between 2.5 and 21.75 mg/mL were observed for α-lact at pH 3 and pH 5. At pH 3 α-lact solutions gelate for α-lact concentrations of 19 and 21.75 mg/mL. At pH 5 α-lact precipitates for concentrations from 2.17 to 3.5 mg/mL. As described before, at pH 5 the α-lact stock solution could only be concentrated up to 7±1 mg/mL. Lys_{hum} and lys_{egg} stayed soluble in the investigated pH and protein concentration range.

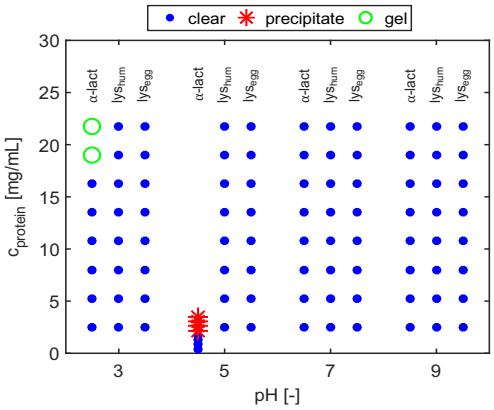

Fig. 7. Phase states of α–lact, lys_{hum}, and lys_{egg} at pH 3, 5, 7, and 9 without salt. Data for lys_{hum} and lys_{egg} was taken from Baumgartner et al. (Baumgartner et al., 2015).

3.4. Protein phase behavior using sodium chloride and ammonium sulfate as precipitants

Fig. 8 and Fig. 9 show the phase diagrams of α-lact, lys_{hum}, and lys_{egg} at pH 3 (first row), pH 5 (second row), pH 7 (third row), and pH 9 (fourth row) with sodium chloride and ammonium sulfate as precipitants.

Phase behavior of α-lact is identical for pH 3 and pH 5 regardless of using sodium chloride or ammonium sulfate as precipitant. At pH 9 phase transitions of α-lact could only be observed for ammonium sulfate as precipitant. At pH 3 the first phase transition of α-lact occurred for 0.23 M sodium chloride and for 0.14 M ammonium sulfate. At pH 5 addition of sodium chloride and ammonium sulfate stabilized α-lact as the former precipitated states could be transformed into soluble phase states. At pH 9 addition of 1.23 M ammonium sulfate induced the first phase transition of α-lact. No influence of both precipitants on α-lact phase behavior could be observed at pH 7 and no influence of sodium chloride could be observed at pH 9.

For lys_{hum} phase transitions induced by sodium chloride and ammonium sulfate were observed at pH 3, 5, 7, and 9. Using sodium chloride as precipitant the first phase transition of lys_{hum} was observed for 1.36 M sodium chloride at pH 3, for 0.45 M sodium chloride at pH 5 and pH 7, and for 1.59 M sodium chloride at pH 9. Using ammonium sulfate as precipitant the first phase transition of lys_{hum} was observed for 0.95 M ammonium sulfate at pH 3, for 0.55 M ammonium sulfate at pH 5, and for 0.40 M ammonium sulfate at pH 7 and pH 9.

For lys_{egg} phase transitions were observed at pH 3, 5, 7, and 9 using sodium chloride as precipitant and at pH 3 and pH 5 using ammonium sulfate as precipitant. The first phase transition of lys_{egg} induced by sodium chloride was observed for 0.45 M sodium chloride at pH 3, for 0.68 M sodium chloride at pH 5 and pH 7, and for 0.23 M sodium chloride at pH 9. The first phase transition of lys_{egg} induced by ammonium sulfate was observed for 1.23 M ammonium

sulfate at pH 3 and for 1.5 M ammonium sulfate at pH 5.

For α-lact the predominant phase state beside the soluble state was precipitate independent of the precipitant used. The same accounted for lys_{hum}, whereas here for very high sodium chloride or ammonium sulfate concentrations in some cases crystals and precipitate coexisted. Something special to mention is that at pH 3 for α-lact and lys_{hum} gelation was observed. The predominant phase state beside the soluble one for lys_{egg} depends on the precipitant used. For sodium chloride as precipitant lys_{egg} readily crystallized whereas for ammonium sulfate as precipitant skin formation and precipitation were observed.

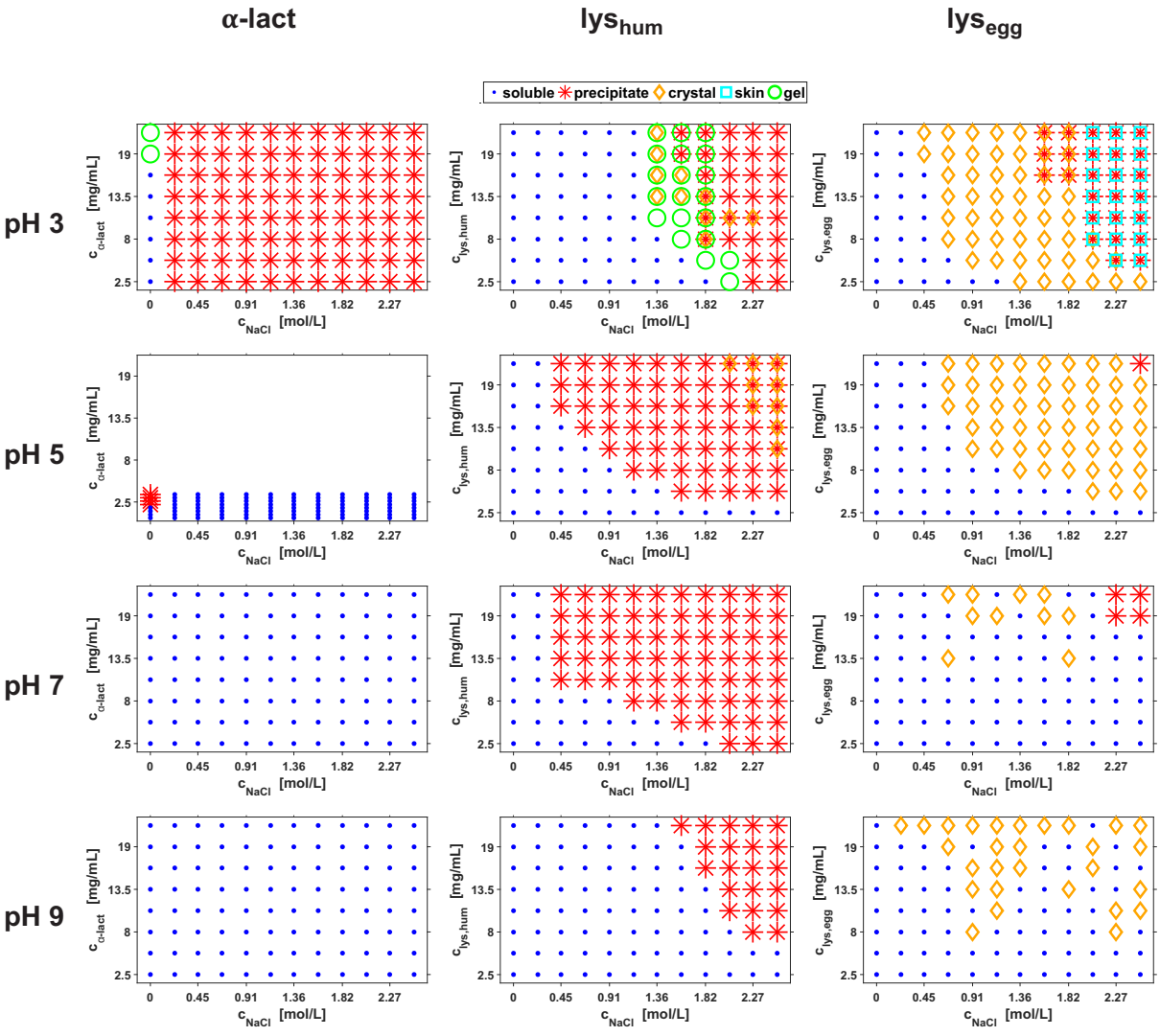

Fig. 8. Phase diagrams of α–lact, lys_{hum}, and lys_{egg} at pH 3 (first row), pH 5 (second row), pH 7 (third row), and pH 9 (fourth row) using sodium chloride as precipitant. Data for the phase diagrams of lys_{hum} and lys_{egg} was taken from Baumgartner et al. (Baumgartner et al., 2015).

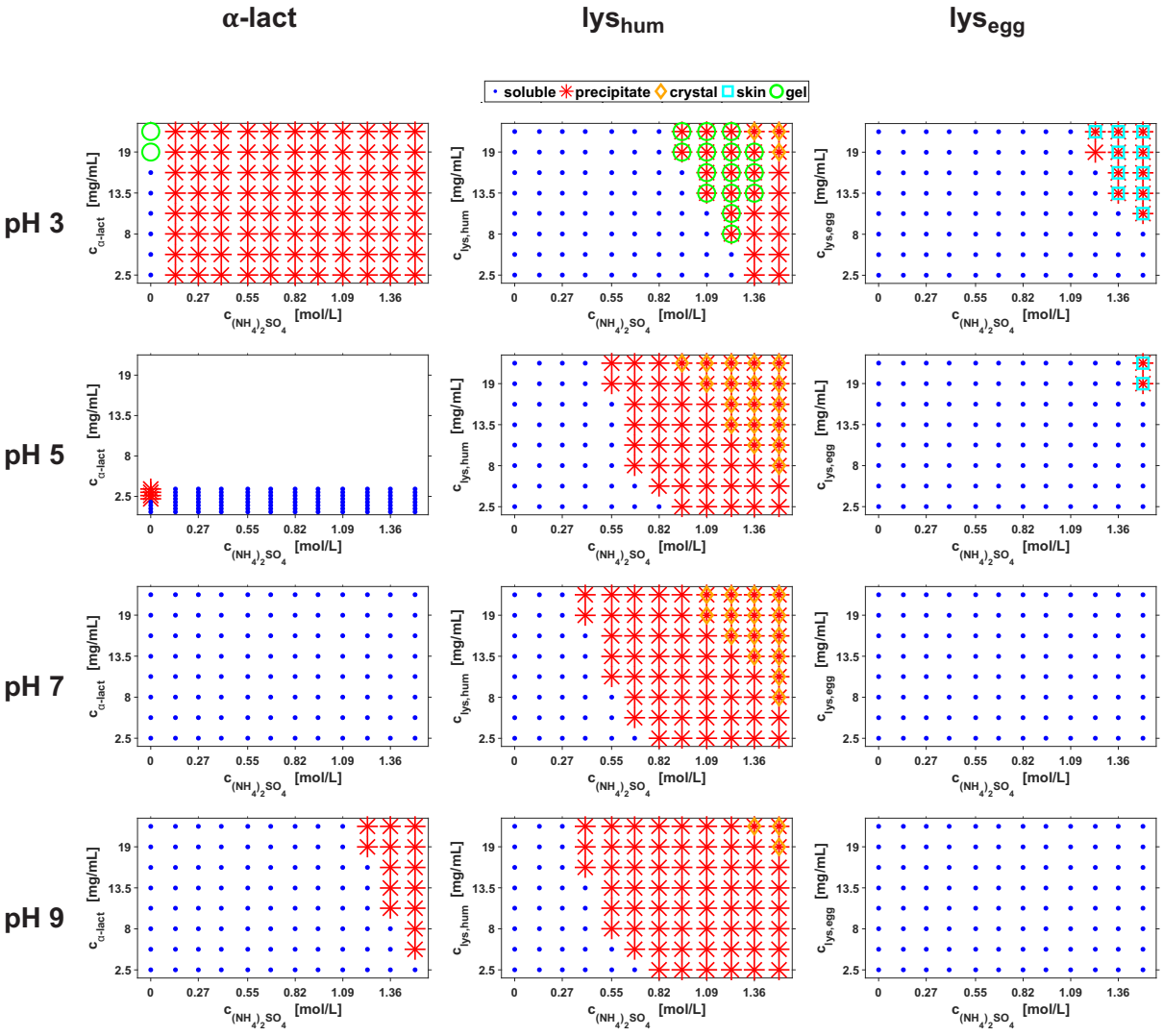

Fig. 9. Phase diagrams of α–lact, lys_{hum}, and lys_{egg} at pH 3 (first row), pH 5 (second row), pH 7 (third row), and pH 9 (fourth row) using ammonium sulfate as precipitant. Data for the phase diagrams of lys_{hum} and lys_{egg} was taken from Baumgartner et al. (Baumgartner et al., 2015).

3.5 Determination of aggregation lines

Using the information from Fig. 8 and Fig. 9 aggregation lines could be determined, representing the occurrence of the first phase transition, i.e. the protein concentration of the first phase transition, as function of precipitant concentration. At fixed precipitant concentration the theoretical protein concentration between the last soluble and the first aggregated phase state was determined. These theoretical protein concentration values as function of the sodium chloride or ammonium sulfate concentration were exponentially fitted using Eq. (2).

$$c_{protein} = A * e^{(-B*c_{precipitant})} \quad (2)$$

In Eq. (2) *A* and *B* are adaptable parameters. Parameter *A* for example represents the protein concentration $c_{protein}$ for the precipitant concentration $c_{precipitant}$ approaching zero, i.e. *A* gives the

theoretical protein concentration of the first phase transition without precipitant. Tab. 3 exemplarily lists the values of *A* and *B* for lys_{hum}.

Tab. 3. Adaptable parameters *A* and *B* and coefficient of correlation R^2 resulting from the fit of aggregation lines to experimental data of lys_{hum} using Eq. (2).

	Precipitant	***A***	***B***	**R^2**
lys_{hum} at pH 3	NaCl	110,11	1,80	0,99
lys_{hum} at pH 3	$(NH_4)_2SO_4$	385,69	3,23	0,98
lys_{hum} at pH 5	NaCl	21,62	0,89	0,95
lys_{hum} at pH 5	$(NH_4)_2SO_4$	659,07	6,60	0,99
lys_{hum} at pH 7	NaCl	13,86	0,63	0,84
lys_{hum} at pH 7	$(NH_4)_2SO_4$	122,27	4,82	0,99
lys_{hum} at pH 9	NaCl	215,94	1,49	0,98
lys_{hum} at pH 9	$(NH_4)_2SO_4$	112,91	5,08	0,98

4. Discussion

4.1. Aggregation propensity of α-lactalbumin, human lysozyme, and lysozyme from chicken egg white without precipitants and its correlation to the zeta potential, the hydrophobic character and conformational flexibility

Aggregation propensity was estimated in the following by the minimal protein concentration at which a phase transition, i.e. aggregation, of α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}) without precipitants was observed. In the investigated protein concentration range between 2.5 and 21.75 mg/mL phase transitions could only be observed for α-lact at pH 3 and pH 5, which is why aggregation propensity without precipitants is highest for α-lact amongst the investigated proteins. Here, α-lact was identified to have the strongest hydrophobic character of all protein species and pH values investigated. Thus, aggregation propensity correlates with the hydrophobic character of the proteins, i.e. the higher the hydrophobic character, the higher the aggregation propensity without precipitants. However, α-lact aggregation propensity was clearly higher at pH 5 than at pH 3 though the hydrophobic character of α-lact at pH 3 is stronger than at pH 5. This discrepancy between the hydrophobic character and aggregation propensity might be explained (regarding colloidal stability) by the differing zeta potential of α-lact at pH 3 and pH 5. At pH 3 the absolute value of the zeta potential of α-lact is 12.5 mV whereas it is 1.2 mV and thus close to zero at pH 5. Thus, at pH 3 electrostatic repulsion seems to be strong enough to delay aggregation. Though, it has to be mentioned that electrostatic repulsion between the α-lact molecules is not high enough to prevent aggregation. This is in accordance with the classical DLVO theory for colloids, where zeta potentials below ±30 mV are considered to be insufficient for stabilization of colloidal suspensions (Greenwood,

2003). Therefore, the zeta potential in any of the investigated solutions is not high enough to stabilize the proteins in solution simply by electrostatic repulsion. Protein flexibility as well correlates with aggregation propensity as conformational flexibility of α-lact at pH 3 and pH 5 is significantly higher than for all other investigated protein species and pH values. But, just like for hydrophobicity, conformational flexibility of α-lact at pH 3 was determined to be higher than conformational flexibility of α-lact at pH 5, assuming aggregation propensity of α-lact at pH 5 to be higher than at pH 3. This discrepancy between protein flexibility and aggregation propensity might be explained by the occurrence of α-lact gelation at pH 3 as gelation represents a very special aggregate state. Denaturation is considered to be a prerequisite to gelation (Ziegler and Foegeding, 1990) and as mentioned above, denaturation is a profound conformational change that can neither be reflected in molecular dynamics (MD) simulations over the studied time range nor in protein flexibility. No statement about the predictive power of protein surface characteristics and conformational flexibility for aggregation propensity of lys_{hum} and lys_{egg} was possible as both proteins stayed soluble in the investigated protein concentration and pH range. In summary, differences regarding aggregation propensities of α-lact, lys_{hum}, and lys_{egg} in aqueous solutions without precipitants at pH 3, 5, 7, and 9 could be related to the hydrophobic character and to conformational flexibility of these proteins. Differences regarding the aggregation propensity of α-lact as function of pH could as well be identified to be rather driven by hydrophobic properties and conformational flexibility than by electrostatic properties, but herein pH dependent differences in the aggregation propensity could only be explained by electrostatic properties of α-lact. Thus, hydrophobic surface characteristics and protein flexibility have the power to predict differing aggregation propensities for different protein species in aqueous solutions without precipitants. pH dependent effects on aggregation within one protein species might not be predicted solely using hydrophobic surface characteristics and protein flexibility, but by additionally taking electrostatic surface properties into account.

4.2. Aggregation propensity of α-lactalbumin, human lysozyme, and lysozyme from chicken egg white using sodium chloride and ammonium sulfate as precipitants and its correlation to the zeta potential, the hydrophobic character and conformational flexibility

As described in Section 3 protein surface characteristics and protein flexibility of α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}) were determined in solutions without precipitants. In the previous section it could be shown, that these electrostatic and hydrophobic surface characteristics in combination with protein flexibility could be directly correlated to the aggregation propensity without precipitants. The predictive power of protein surface characteristics and protein flexibility for aggregation propensity was further studied for addition of precipitants, namely sodium chloride and ammonium sulfate. Long-range electrostatic repulsion between charged proteins is reduced by the presence of salt

ions. These salt ions shield the electrostatic fields of proteins and short-range attractive forces become noticeable (Baumgartner et al., 2015). As could be shown in Baumgartner et al. (Baumgartner et al., 2015) for lysozyme from chicken egg white, for sodium chloride concentrations above 0.23 M the effect of electrostatic repulsion seemed to vanish. Additionally, the zeta potential of the proteins investigated here is not high enough to significantly influence the aggregation propensity of the proteins in solution and is moreover further reduced by addition of sodium chloride and ammonium sulfate. This is why electrostatic repulsion due to the zeta potential of proteins was considered negligible in the following.

4.2.1. Influence of salt species on aggregation propensity

Tab. 3 shows the results of the aggregation line fit. Parameter *A* gives the theoretical protein concentration for extrapolation of the aggregation lines to protein solutions without precipitants. Thus, if salt species was without influence to protein aggregation propensity, parameter *A* should be identical for the aggregation line fits at fixed pH value regardless of the salt species. It is apparent from Tab. 3 that parameter *A* differed strongly and that aggregation propensity is highly influenced by the salt species used. This is in accordance with the Hofmeister series (Hofmeister, 1888), where different salt species are ranked according to their potency to increase or decrease protein stability and thus aggregation propensity. This salt specific effect on protein stability can be connected to hydrophobic interactions. Kosmotropic salt ions (e.g. PO_4^{3-}, SO_4^{2-}, and NH_4^{+}) promote hydrophobic interactions and protein precipitation whereas chaotropic salt ions (e.g. SCN^{-}, Ba^{2+}) decrease the strength of hydrophobic interactions and thus increase proteins' conformational and colloidal stability (Wu and Karger, 1996). Therefore, in the following the correlation between protein hydrophobicity, protein flexibility, and aggregation propensity induced by sodium chloride and ammonium sulfate will be discussed separately.

4.2.2. Influence of salt species on aggregation propensity as function of the isoelectric point

In addition to the salt species itself, the location of the isoelectric point (pI) of a protein with respect to the pH value of the aqueous solution the protein is dissolved in is of high importance. Salts were observed to expose contrary effects on protein stability for pH values above and below the pI (Boström et al., 2005). The pI of α-lact is between pH 4.25 and 4.5 (Robbins et al., 1965). This might explain, why sodium chloride as well as ammonium sulfate induce precipitation of α-lact at pH 3 and stabilize α-lact against precipitation at pH 5. As the isoelectric points of lys_{hum} and lys_{egg} are between pH 10.0 and 11.0 (Lüllig et al., 1983; Lundblad et al., 1972; Righetti and Caravaggio, 1976; Righetti et al., 1981), all pH values investigated in this manuscript are below their isoelectric point. In order to exclude effects related to the location of the pI with respect to the pH value, α-lact was for the following examination only considered at pH 3, as then

protein aggregation propensity is compared for pH values below the isoelectric points of the three proteins.

4.2.3. Predictive power of hydrophobic surface character and conformational flexibility for estimation of aggregation propensity using sodium chloride as precipitant

Aggregation propensity was estimated in the following by the minimal sodium chloride concentration at which a phase transition, i.e. aggregation, of α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}) was observed. In the following aggregation propensity of these three proteins will be compared at fixed pH value, first, and hereafter pH dependent aggregation propensity within one protein species will be discussed. At pH 3 α-lact had the highest aggregation propensity among the three proteins, followed by lys_{egg} and lys_{hum}. At pH 5 and pH 7, the aggregation propensity of lys_{hum} is higher than the aggregation propensity of lys_{egg} in contrast to pH 9, where lys_{egg} is more prone to aggregation than lys_{hum}. Thus, for pH 3 and pH 9 aggregation propensity as function of protein species (for pH<pI) directly correlates with protein conformational flexibility and for pH 5 and pH 7 it directly correlates with protein hydrophobicity. Furthermore, the results indicate that the classification of a protein as hydrophobic or hydrophilic strongly influences protein phase behavior in terms of aggregate species under addition of salt. For α-lact and lys_{hum} the predominant phase transition induced by addition of sodium chloride was from soluble to precipitate. For lys_{egg} instead the first phase transition induced by sodium chloride was from soluble to crystal. Lys_{egg} thus seems to adopt a special behavior as it was the only hydrophilic protein investigated in this work and the only one that readily crystallized under addition of sodium chloride. The two proteins with hydrophobic character instead readily precipitated and crystals could only be found in coexistence with precipitate for high precipitant concentrations. For very high hydrophobicities, as observed for α-lact and lys_{hum} at pH 3, those proteins readily gelled. As for the aggregation propensity without precipitants, no correlation between protein surface hydrophobicity or protein flexibility within one protein species as function of the pH value could be found.

In summary, the results lead to the conclusion that protein flexibility and surface hydrophobicity are well-suited parameters for estimation of aggregation propensity of different proteins under addition of sodium chloride. Though, aggregation propensity could not be reduced to the same parameter for all conditions. Protein flexibility seems to be best suited for estimation of aggregation propensity at extreme pH values (pH 3 and pH 9), whereas surface hydrophobicity seems to be best suited for moderate to physiological pH values (pH 5 and pH 7). Protein surface hydrophobicity furthermore might indicate the aggregate state to evolve as hydrophilicity could be related to crystallization and high hydrophobicity to gelation under addition of sodium chloride.

4.2.4. Predictive power of hydrophobic surface character and conformational flexibility for estimation of aggregation propensity using ammonium sulfate as precipitant

In the following aggregation propensity was estimated by the minimal ammonium sulfate concentration at which a phase transition, i.e. aggregation, of α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}) was observed. At pH 3 α-lact exhibited the highest aggregation propensity among the three proteins, followed by lys_{hum} and lys_{egg}. At pH 5 the aggregation propensity of lys_{hum} is higher than the aggregation propensity of lys_{egg}. At pH 7 and pH 9, the aggregation propensity of lys_{hum} is clearly higher than the aggregation propensity of lys_{egg}, as lys_{hum} readily precipitated, whereas lys_{egg} stayed soluble. Thus, for all pH values aggregation propensity as function of protein species (for pH<pI) directly correlates with protein hydrophobicity. This is in agreement with the general observation that ammonium sulfate as kosmotropic salt strengthens hydrophobic interactions (Wu and Karger, 1996), resulting in increased aggregation propensity for the proteins with hydrophobic character (for pH<pI) compared to a precipitant with chaotropic properties. The predominant phase transition for α-lact and lys_{hum} induced by addition of ammonium sulfate was from soluble to precipitate and crystals could only be found in coexistence with precipitate for high precipitant concentrations. For the protein with hydrophilic character the kosmotropic salt induced skin formation which is according to Zeelen (Zeelen, 2009) synonymous with protein denaturation. Thus, proteins with hydrophobic character easily precipitated by addition of kosmotropic and chaotropic salts, whereas proteins with a hydrophilic character seemed to easily crystallize by addition of chaotropic salts but denaturate by addition of kosmotropic salts. In summary, the results lead to the conclusion that protein surface hydrophobicity is best suited for estimation of aggregation propensity of different proteins under addition of ammonium sulfate. Though, again no clear correlation between surface hydrophobicity or protein flexibility and aggregation propensity as function of pH within one protein species could be found.

5. Conclusion and Outlook

α-lactalbumin (α-lact), human lysozyme (lys_{hum}), and lysozyme from chicken egg white (lys_{egg}), three proteins very similar in size and secondary and tertiary structure, could be shown to exhibit different surface characteristics regarding electrostatics and hydrophobicity and different values for conformational flexibility. In protein solutions without precipitants it could be shown that differences in the aggregation propensity among the different protein species directly correlate to protein surface hydrophobicity and protein flexibility. Differences in the aggregation propensity as function of pH within one protein species could be related to protein net charge, protein surface hydrophobicity and protein flexibility. Thus, protein net charge, protein surface hydrophobicity and protein flexibility could be identified to be suitable parameters for

prediction of aggregation propensity of α-lact, lys_{hum}, and lys_{egg} in aqueous solution without precipitants. However, protein surface hydrophobicity and protein flexibility were identified to be of major importance for protein aggregation.

Prediction of aggregation propensity in solutions with precipitants, based on the values of protein surface hydrophobicity and protein flexibility determined in solutions without precipitants, could only be realized with some limitations. For both precipitants no clear correlation between surface hydrophobicity or protein flexibility and aggregation propensity as function of pH within one protein species could be found. Aggregation propensity could only be correlated to surface hydrophobicity or flexibility at fixed pH values by comparison of different protein species. For sodium chloride as precipitant, correlation between surface hydrophobicity or protein flexibility and aggregation propensity revealed a pH dependency. Furthermore, investigation of aggregation propensity with precipitants was restricted to pH values below the isoelectric points of α-lact, lys_{hum}, and lys_{egg}. In upcoming molecular dynamic (MD) simulations sodium chloride and ammonium sulfate are considered and thus protein flexibility will be obtained as function of precipitant concentration. We are confident that these comprehensive MD simulations will circumvent these limitations and will improve the explanatory power of protein flexibility and protein aggregation propensity and thus increase the predictive power of protein flexibility for protein aggregation propensity.

Acknowledgement

We gratefully acknowledge the financial support by the Federal Ministry of Education and Research (BMBF)(0315342B, 031A095B).

References

Amrhein, S., Bauer, K.C., Galm, L., Hubbuch, J., 2015. Non-invasive high throughput approach for protein hydrophobicity determination based on surface tension. Biotechnology and bioengineering, n/a-n/a.

Anandakrishnan, R., Aguilar, B., Onufriev, A.V., 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic acids research 40, W537-541.

Basu, S.K., Govardhan, C.P., Jung, C.W., Margolin, A.L., 2004. Protein crystals for the delivery of biopharmaceuticals. Expert Opinion on Biological Therapy 4, 301-317.

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

Boström, M., Tavares, F.W., Finet, S., Skouri-Panet, F., Tardieu, A., Ninham, B.W., 2005. Why forces between proteins follow different Hofmeister series for pH above and below pI. Biophysical Chemistry 117, 217-224.

Brange, J., Vølund, A., 1999. Insulin analogs with improved pharmacokinetic profiles. Advanced Drug Delivery Reviews 35, 307-335.

Brunsteiner, M., Flock, M., Nidetzky, B., 2013. Structure based descriptors for the estimation of colloidal interactions and protein aggregation propensities. PloS one 8, e59797.

Carlson, H.A., McCammon, J.A., 2000. Accommodating protein flexibility in computational drug design. Molecular pharmacology 57, 213-218.

Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., Trout, B.L., 2009. Design of therapeutic proteins with enhanced stability. Proc. Natl. Acad. Sci. U. S. A. 106, 11937-11942.

Chi, E.Y., Krishnan, S., Randolph, T.W., Carpenter, J.F., 2003. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. Pharmaceutical research 20, 1325-1336.

Chrysina, E.D., Brew, K., Acharya, K.R., 2000. Crystal structures of Apo- and holo-bovine α-lactalbumin at 2.2-Å resolution reveal an effect of calcium on inter-lobe interactions. Journal of Biological Chemistry 275, 37021-37029.

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., Kollman, P., 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. Journal of Computational Chemistry 24, 1999-2012.

Durek, T., Torbeev, V.Y., Kent, S.B.H., 2007. Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme. Proc Natl Acad Sci USA 104, 4846-4851.

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., 1995. A smooth particle mesh Ewald method. J Chem Phys 103, 8577-8593.

Greenwood, R., 2003. Review of the measurement of zeta potentials in concentrated aqueous suspensions using electroacoustics. Advances in Colloid and Interface Science 106, 55-81.

Hofmeister, F., 1888. Zur Lehre von der Wirkung der Salze. Archiv für experimentelle Pathologie und Pharmakologie 25, 1-30.

Hunter, R.J., 1981. The Calculation of Zeta Potential, Zeta potential in colloid science. ACADEMIC PRESS LIMITED, pp. 59-124.

Jen, A., Merkle, H.P., 2001. Diamonds in the rough: protein crystals from a formulation perspective. Pharmaceutical research 18, 1483-1488.

Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., Lesk, A.M., 2006. MUSTANG: A multiple structural alignment algorithm. Proteins: Structure, Function, and Bioinformatics 64, 559-574.

Krieger, E., Koraimann, G., Vriend, G., 2002. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. Proteins: Structure, Function, and Bioinformatics 47, 393-402.

Kumar, V., Dixit, N., Zhou, L., Fraunhofer, W., 2011. Impact of short range hydrophobic interactions and long range electrostatic forces on the aggregation kinetics of a monoclonal antibody and a dual-variable domain immunoglobulin at low and high concentrations. International Journal of Pharmaceutics 421, 82-93.

Lauer, T.M., Agrawal, N.J., Chennamsetty, N., Egodage, K., Helk, B., Trout, B.L., 2012. Developability index: A rapid in silico tool for the screening of antibody aggregation propensity. Journal of Pharmaceutical Sciences 101, 102-115.

Lüllig, H., Frank, T., Fleck, H., Ebert, W., Vogt-Moykopf, J., 1983. Isolation and characterization of human pneumocytes. Praxis und Klinik der Pneumologie 37 Suppl 1, 842-844.

Lundblad, G., Vesterberg, O., Zimmerman, R., Lind, J., 1972. Studies on lysozyme from human leucemic urine by isoelectric focusing. Acta Chem. Scand. 26, 1711-1713.

Mahler, H.-C., Friess, W., Grauschopf, U., Kiese, S., 2009. Protein aggregation: pathways, induction factors and analysis. Journal of pharmaceutical sciences 98, 2909-2934.

Martínez, J.M., Martínez, L., 2003. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. Journal of Computational Chemistry 24, 819-825.

Martínez, L., Andrade, R., Birgin, E.G., Martínez, J.M., 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. Journal of computational chemistry 30, 2157-2164.

Paul, R., Graff-Meyer, A., Stahlberg, H., Lauer, M.E., Rufer, A.C., Beck, H., Briguet, A., Schnaible, V., Buckel, T., Boeckle, S., 2012. Structure and function of purified monoclonal antibody dimers induced by different stress conditions. Pharmaceutical Research 29, 2047-2059.

Petsko, G.A., Ringe, D., 2004b. Protein Flexibility, Protein Structure and Function. New Science Press Ltd, pp. 46-48.

Price, W.N., Handelman, S.K., Everett, J.K., Tong, S.N., Bracic, A., Luff, J.D., Naumov, V., Acton, T., Manor, P., Xiao, R., Rost, B., Montelione, G.T., Hunt, J.F., 2011. Large-scale experimental studies show unexpected amino acid effects on protein expression and solubility in vivo in E. coli. Microbial Informatics and Experimentation 1, 1-20.

Righetti, P.G., Caravaggio, T., 1976. Isoelectric point and molecular weights of proteins. A table. Journal of chromatography 127, 1-28.

Righetti, P.G., Tudor, G., Ek, K., 1981. Isoelectric points and molecular weights of proteins. A new table. Journal of chromatography 220, 115-194.

Robbins, F.M., Kronman, M.J., Andreotti, R.E., 1965. Inter- and intramolecular interactions of α-lactalbumin. V. The effect of amidination on association and aggregation. Biochimica et biophysica acta 109, 223-233.

Scopes, R.K., 1994. Separation by Precipitation, Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag New York, Inc., pp. 71-101.

Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J., Udseth, H.R., 1990. New developments in biochemical mass spectrometry: electrospray ionization. Analytical chemistry 62, 882-899.

Sousa, R., 1995. Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization. Acta Crystallographica - Section D Biological Crystallography 51, 271-277.

Teague, S.J., 2003. Implications of protein flexibility for drug discovery. Nature reviews. Drug discovery 2, 527-541.

Vajo, Z., Fawcett, J., Duckworth, W.C., 2001. Recombinant DNA Technology in the Treatment of Diabetes: Insulin Analogs. Endocrine Reviews 22, 706-717.

Valerio, M., Colosimo, A., Conti, F., Giuliani, A., Grottesi, A., Manetti, C., Zbilut, J.P., 2005. Early events in protein aggregation: Molecular flexibility and hydrophobicity/charge interaction in amyloid peptides as studied by molecular dynamics simulations. Proteins: Structure, Function and Genetics 58, 110-118.

Vihinen, M., 1987. Relationship of protein flexibility to thermostability. Protein Engineering 1, 477-480.

Wang, J., Dauter, M., Alkire, R., Joachimiak, A., Dauter, Z., 2007. Triclinic lysozyme at 0.65 Å resolution. Acta Crystallographica Section D: Biological Crystallography 63, 1254-1268.

Wu, S.L., Karger, B.L., 1996. Hydrophobic interaction chromatography of proteins. Methods in enzymology 270, 27-47.

Yadav, S., Laue, T.M., Kalonia, D.S., Singh, S.N., Shire, S.J., 2012. The influence of charge distribution on self-association and viscosity behavior of monoclonal antibody solutions. Molecular Pharmaceutics 9, 791-802.

Yadav, S., Shire, S.J., Kalonia, D.S., 2011. Viscosity analysis of high concentration bovine serum albumin aqueous solutions. Pharmaceutical Research 28, 1973-1983.

Yang, M.X., Shenoy, B., Disttler, M., Patel, R., McGrath, M., Pechenov, S., Margolin, A.L., 2003. Crystalline monoclonal antibodies for subcutaneous delivery. Proceedings of the National Academy of Sciences of the United States of America 100, 6934-6939.

Zeelen, J.P., 2009. Interpretation of the Crystallization Drop Results, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 175-194.

Zhang, J., Topp, E.M., 2012. Protein G, protein A and protein A-derived peptides inhibit the agitation induced aggregation of IgG. Molecular Pharmaceutics 9, 622-628.

Ziegler, G.R., Foegeding, E.A., 1990. The Gelation of Proteins, in: Kinsella, J.E. (Ed.), Advances in Food and Nutrition Research. Academic Press, pp. 203-298.

4. Conclusion and Outlook

The extensive and systematic study of protein phase behavior increased the basic understanding of protein aggregation and allowed to identify main influencing factors. For the investigated proteins it was possible to induce all known phase transitions, which are crystallization, precipitation, skin formation, gelation and liquid-liquid phase separation. The investigated salts thereby showed various effects on the protein phase behavior that could not be generalized regarding salt type. High salt concentrations though strongly increased aggregation propensity. Extremely acidic pH values induced specific phase transitions such as skin formation and gelation. The depletion attraction of PEG on the protein phase behavior could be shown to be most pronounced where long-range electrostatic repulsion is at its minum, i.e. close to the isoelectric point of the investigated proteins. Using Fourier transform infrared (FT-IR) spectroscopy lysozyme conformation during aggregation could be monitored and phase transitions accompanied by a conformational change of the protein could be identified. However, lysozyme aggregation in most cases was native. In cases where it was non-native, glycerol and PEG 1000 could be shown to be very effective in preventing non-native aggregation. FT-IR spectroscopy furthermore aided to increase the basic understanding of the mode of action of typical solution additives as the pH dependent impact of these additives on lysozyme phase behavior, lysozyme solubility, and crystal size and morphology could be correlated to the conformational stability of the protein. A novel non-invasive approach to determine protein surface hydrophobicity based on surface tension was presented. This approach aided to verify, that the hydrophobic character of proteins is pH dependent, which was assumed up to know, but could not be proved experimentally. Furthermore, combining the results from this protein surface hydrophobicity studies and from calculation of the conformational flexibility from an *in silico* appraoch by means of molecular dynamics simulations, and correlating it to protein phase behavior, protein surface hydrophobicity and conformational flexibility could be shown to be the main influencing factors for protein aggregation propensity. Thus, the experimental approach for determination of protein surface hydrophobicity and the *in silico* approach for determination of conformational flexibilty give a powerful tool to predict protein aggregation propensity. In order to use it as a universal predictive tool it needs to be verified by screening a wider range of protein solutions in the future.

In any of the presented approaches sample volume was minimized to save product consumption and high throughput methods, whenever possible, were used to reduce time consumption. The presented approaches were developed and validated using selected model proteins (lysozyme from chicken egg white, human lysozyme, glucose oxidase, glucose isomerase, bovine serum albumin, and α-lactalbumin), but can easily be transferred to any protein of interest, such as monoclonal antibodies, and will provide valuable information for optimization of existing

biopharmaceutical purification and formulation processes as well as for the design of novel purification and formulation processes.

Abbreviations and Symbols

ANS	1-anilino-8-naphthalene sulfonate
AS	ammonium sulfate
ATPS	aqueous two-phase system
ATR	attenuated total reflection
BMBF	Federal Ministry of Education and Research
BPB	bromophenol blue
BPB Na	bromophenol blue sodium salt
BSA	bovine serum albumin
FT-IR	Fourier transform infrared
HIC	hydrophobic interaction chromatography
HTP	high throughput
IEX	ion-exchange chromatography
IgG	immunoglobulin G
LLPS	liquid-liquid phase separation
MD	molecular dynamics
NaCl	sodium chloride
NMR	nuclear magnetic resonance
PDB	protein data bank
PEEK	polyether ether ketone
PEG	polyethylene glycol
PES	polyether sulfone
pI	isoelectric point
PME	particle mesh Ewald
PTFE	polytetrafluoroethylene
QSAR	quantitative structure-activity relationship
RMSF	root-mean-square-fluctuation
RP	reversed phase

References

Absolom, D.R., Zingg, W., Neumann, A.W., 1987. Protein adsorption to polymer particles: role of surface properties. J. Biomed. Mater. Res. 21, 161-171.

Ahamed, T., Esteban, B.N.A., Ottens, M., van Dedem, G.W.K., van der Wielen, L.A.M., Bisschops, M.A.T., Lee, A., Pham, C., Thömmes, J., 2007. Phase behavior of an intact monoclonal antibody. Biophysical journal 93, 610-619.

Alizadeh-Pasdar, N., Li-Chan, E.C.Y., 2000. Comparison of Protein Surface Hydrophobicity Measured at Various pH Values Using Three Different Fluorescent Probes. J. Agric. Food Chem. 48, 328-334.

Amrhein, S., Bauer, K.C., Galm, L., Hubbuch, J., 2015. Non-invasive high throughput approach for protein hydrophobicity determination based on surface tension. Biotechnology and bioengineering, n/a-n/a.

Amrhein, S., Oelmeier, S.A., Dismer, F., Hubbuch, J., 2014a. Molecular dynamics simulations approach for the characterization of peptides with respect to hydrophobicity. J. Phys. Chem. B 118, 1707-1714.

Amrhein, S., Schwab, M.-L., Hoffmann, M., Hubbuch, J., 2014b. Characterization of aqueous two phase systems by combining lab-on-a-chip technology with robotic liquid handling stations. J. Chromatogr. A 1367, 68-77.

Anandakrishnan, R., Aguilar, B., Onufriev, A.V., 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic acids research 40, W537-541.

Anandakrishnan, R., Aguilar, B., Onufriev, A.V., 2013. H++ v3.1.

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.

Annunziata, O., Asherie, N., Lomakin, A., Pande, J., Ogun, O., Benedek, G.B., 2002. Effect of polyethylene glycol on the liquid-liquid phase transition in aqueous protein solutions. Proceedings of the National Academy of Sciences of the United States of America 99, 14165-14170.

Arakawa, T., Timasheff, S.N., 1982. Stabilization of Protein Structure by Sugars. Biochemistry 21, 6536-6544.

Arakawa, T., Timasheff, S.N., 1983. Preferential Interactions of Proteins with Solvent Components in Aqueous Amino Acid Solutions. Archives of Biochemistry and Biophysics 224, 169-177.

Arakawa, T., Timasheff, S.N., 1985a. Mechanism of Poly(ethylene glycol) Interaction with Proteins. Biochemistry 24, 6756-6762.

Arakawa, T., Timasheff, S.N., 1985b. The Stabilization of Proteins by Osmolytes. Biophysical Journal 47, 411-414.

Asakura, S., Oosawa, F., 1958. Interaction between Particles Suspended in Solutions of Macromolecules. Journal of Polymer Science 33, 183-192.

Asenjo, J.A., Andrews, B.A., 2011. Aqueous two-phase systems for protein separation: A perspective. J. Chromatogr. A 1218, 8826-8835.

Asherie, N., 2004. Protein crystallization and phase diagrams. Methods 34, 266-272.

Asherie, N., Lomakin, A., Benedek, G., 1996. Phase Diagram of Colloidal Solutions. Physical Review Letters 77, 4832-4835.

Atha, D.H., Inghamg, K.C., 1981. Mechanism of Precipitation of Proteins by Polyethylene Glycols. The Journal of Biological Chemistry 256, 12108-12117.

Auton, M., Rösgen, J., Sinev, M., Holthauzen, L.M.F., Bolen, D.W., 2011. Osmolyte effects on protein stability and solubility: a balancing act between backbone and side-chains. Biophysical chemistry 159, 90-99.

Barth, A., 2007. Infrared spectroscopy of proteins. Biochimica et biophysica acta 1767, 1073-1101.

Basu, S.K., Govardhan, C.P., Jung, C.W., Margolin, A.L., 2004. Protein crystals for the delivery of biopharmaceuticals. Expert Opinion on Biological Therapy 4, 301-317.

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

Baynes, B.M., Trout, B.L., 2004. Rational design of solution additives for the prevention of protein aggregation. Biophysical journal 87, 1631-1639.

Beretta, S., Chirico, G., Baldini, G., 2000. Short-Range Interactions of Globular Proteins at High Ionic Strengths. Macromolecules 33, 8663-8670.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic acids research 28, 235-242.

Bertsch, M., Mayburd, A.L., Kassner, R.J., 2003. The identification of hydrophobic sites on the surface of proteins using absorption difference spectroscopy of bromophenol blue. Anal. Biochem. 313, 187-195.

Bigelow, C.C., 1967. On the average hydrophobicity of proteins and the relation between it and protein structure. J. Theor. Biol. 16, 187-211.

Biswas, K.M., DeVido, D.R., Dorsey, J.G., 2003. Evaluation of methods for measuring amino acid hydrophobicities and interactions. J. Chromatogr. A 1000, 637-655.

Bjerrum, O.J., 1968. Interaction of Bromphenol Blue and Bilirubin with Bovine and Human Serum Albumin Determined by Gel Filtration. Scand. J. Clin. Lab. Investig. 22, 41-48.

Black, S.D., Mould, D.R., 1991. Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. Anal. Biochem. 193, 72-82.

Boistelle, R., Astier, J.P., 1988. Crystallization mechanisms in solution. Journal of Crystal Growth 90, 14-30.

Bolen, D.W., Baskakov, I.V., 2001. The osmophobic effect: natural selection of a thermodynamic force in protein folding. Journal of molecular biology 310, 955-963.

Boström, M., Tavares, F.W., Finet, S., Skouri-Panet, F., Tardieu, A., Ninham, B.W., 2005. Why forces between proteins follow different Hofmeister series for pH above and below pI. Biophysical Chemistry 117, 217-224.

Brange, J., Vølund, A., 1999. Insulin analogs with improved pharmacokinetic profiles. Advanced Drug Delivery Reviews 35, 307-335.

Brems, D.N., Plaisted, S.M., Havel, H.A., Tomich, C.S., 1988. Stabilization of an associated folding intermediate of bovine growth hormone by site-directed mutagenesis. Proc. Natl. Acad. Sci. U. S. A. 85, 3367-3371.

Brunsteiner, M., Flock, M., Nidetzky, B., 2013. Structure based descriptors for the estimation of colloidal interactions and protein aggregation propensities. PloS one 8, e59797.

Bruzdziak, P., Panuszko, A., Stangret, J., 2013. Influence of Osmolytes on Protein and Water Structure: A Step To Understanding the Mechanism of Protein Stabilization. The journal of physical chemistry. B 117, 11502-11508.

Bull, H.B., Breese, K., 1974. Surface tension of amino acid solutions: A hydrophobicity scale of the amino acid residues. Arch. Biochem. Biophys. 161, 665-670.

Byler, D.M., Susi, H., 1986. Examination of the Secondary Structure of Proteins by Deconvolved FTIR Spectra. Biopolymers 25, 469-487.

Cao, W.G., Jiao, Q.C., Fu, Y., Chen, L., Liu, Q., 2003. Mechanism of the Interaction Between Bromophenol Blue and Bovine Serum Albumin. Spectrosc. Lett. 36, 197-209.

Cardamone, M., Puri, N.K., 1992. Spectrofluorimetric assessment of the surface hydrophobicity of proteins. Biochem. J. 282, 589-593.

Carlson, H.A., McCammon, J.A., 2000. Accommodating protein flexibility in computational drug design. Molecular pharmacology 57, 213-218.

Chayen, N., Akins, J., Campbell-Smith, S., Blow, D.M., 1988. Solubility of glucose isomerase in ammonium sulphate solutions. Journal of Crystal Growth 90, 112-116.

Chayen, N.E., 2004. Turning protein crystallisation from an art into a science. Current Opinion in Structural Biology 14, 577-583.

Chayen, N.E., Saridakis, E., 2008. Protein crystallization: from purified protein to diffraction-quality crystal. Nature Methods 5, 147-153.

Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., Trout, B.L., 2009. Design of therapeutic proteins with enhanced stability. Proc. Natl. Acad. Sci. U. S. A. 106, 11937-11942.

Chi, E.Y., Krishnan, S., Randolph, T.W., Carpenter, J.F., 2003. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. Pharmaceutical research 20, 1325-1336.

Chrysina, E.D., Brew, K., Acharya, K.R., 2000. Crystal structures of Apo- and holo-bovine α-lactalbumin at 2.2-Å resolution reveal an effect of calcium on inter-lobe interactions. Journal of Biological Chemistry 275, 37021-37029.

Cleland, J., Powell, M., Shire, S., 1993. The development of stable protein formulations: a close look at protein aggregation, deamidation, and oxidation. Crit Rev Ther Drug Carrier Syst 10, 307-377.

Cohn, E.J., 1925. The physical chemistry of proteins. Physiological Reviews 5, 349-437.

Collins, K.D., 2004. Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. Methods 34, 300-311.

Cromwell, M.E.M., Hilario, E., Jacobson, F., 2006. Protein aggregation and bioprocessing. The AAPS journal 8, E572-579.

Curtis, R.A., Prausnitz, J.M., Blanch, H.W., 1998. Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. Biotechnology and bioengineering 57, 11-21.

Curtis, R.A., Ulrich, J., Montaser, A., Prausnitz, J.M., Blanch, H.W., 2002. Protein-protein interactions in concentrated electrolyte solutions. Biotechnology and bioengineering 79, 367-380.

D'Arcy, A., Mac Sweeney, A., Stihle, M., Haber, A., 2003. The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. Acta Crystallographica Section D Biological Crystallography 59, 396-399.

Daugherty, A.L., Mrsny, R.J., 2006. Formulation and delivery issues for monoclonal antibody therapeutics. Advanced drug delivery reviews 58, 686-706.

Debye, P., Hückel, E., 1923. Zur Theorie der Elektrolyte. Physikalische Zeitschrift 24, 185-206.

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.

Dill, K.A., 1990. Dominant forces in protein folding. Biochemistry 29, 7133-7155.

Dong, A., Huang, P., Caughey, W.S., 1990. Protein secondary structures in water from second-derivative amide I infrared spectra. Biochemistry 29, 3303-3308.

Dong, A., Prestrelski, S.J., Allison, S.D., Carpenter, J.F., 1995. Infrared Spectroscopic Studies of Lyophilization- and Temperature-Induced Protein Aggregation. Journal of Pharmaceutical Sciences 84, 415-424.

Dooley, K.H., Castellino, F.J., 1972. Solubility of amino acids in aqueous guanidinium thiocyanate solutions. Biochemistry 11, 1870-1874.

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., Kollman, P., 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. Journal of Computational Chemistry 24, 1999-2012.

Dumetz, A.C., Chockla, A.M., Kaler, E.W., Lenhoff, A.M., 2008. Protein phase behavior in aqueous solutions: crystallization, liquid-liquid phase separation, gels, and aggregates. Biophysical journal 94, 570-583.

Dumetz, A.C., Chockla, A.M., Kaler, E.W., Lenhoff, A.M., 2009. Comparative Effects of Salt, Organic, and Polymer Precipitants on Protein Phase Behavior and Implications for Vapor Diffusion. Crystal Growth & Design 9, 682-691.

Durbin, S.D., Feher, G., 1986. Crystal growth studies of lysozyme as a model for protein crystallization. Journal of Crystal Growth 76, 583-592.

Durek, T., Torbeev, V.Y., Kent, S.B.H., 2007. Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme. Proc Natl Acad Sci USA 104, 4846-4851.

Dzwolak, W., Ravindra, R., Lendermann, J., Winter, R., 2003. Aggregation of bovine insulin probed by DSC/PPC calorimetry and FTIR spectroscopy. Biochemistry 42, 11347-11355.

Eisenberg, D., 1984. Three-dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. Biochem. 53, 595-623.

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., 1995. A smooth particle mesh Ewald method. J Chem Phys 103, 8577-8593.

Faber, C., Hobley, T.J., 2006. Measurement and Prediction of Protein Phase Behaviour and Protein-Protein-Interactions. Ph.D. thesis, Technical University of Denmark.

Feher, G., Kam, Z., 1985. Nucleation and growth of protein crystals: general principles and assays. Methods in Enzymology 114, 77-115.

Fendler, J.H., Nome, F., Nagyvary, J., 1975. Compartmentalization of Amino Acids in Surfactant Aggregates. J. Mol. Evol. 6, 215-232.

Feng, Y.W., Ooishi, A., Honda, S., 2012. Aggregation factor analysis for protein formulation by a systematic approach using FTIR, SEC and design of experiments techniques. Journal of pharmaceutical and biomedical analysis 57, 143-152.

Fink, A.L., 1998. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding & design 3, R9-23.

Flores, R., 1978. A Rapid and Reproducible Assay for Quantitative Estimation of Proteins using Bromophenol Blue. Anal. Biochem. 88, 605-611.

Frauenfelder, H., Parak, F., Young, R.D., 1988. Conformational substates in proteins. Annual review of biophysics and biophysical chemistry 17, 451-479.

García-Ruiz, J.M., 2003. Nucleation of protein crystals. Journal of Structural Biology 142, 22-31.

Genest, S., Schwarz, S., Petzold-Welcke, K., Heinze, T., Voit, B., 2013. Characterization of highly substituted, cationic amphiphilic starch derivatives: Dynamic surface tension and intrinsic viscosity. Starch/Stärke 65, 999-1010.

George, A., Wilson, W.W., 1994. Predicting protein crystallization from a dilute solution property. Acta crystallographica. Section D, Biological crystallography 50, 361-365.

Granata, V., Palladino, P., Tizzano, B., Negro, A., Berisio, R., Zagari, A., 2006. The Effect of the Osmolyte Trimethylamine N-Oxide on the Stability of the Prion Protein at Low pH. Biopolymers 82, 234-240.

Greenwood, R., 2003. Review of the measurement of zeta potentials in concentrated aqueous suspensions using electroacoustics. Advances in Colloid and Interface Science 106, 55-81.

Guo, B., Kao, S., McDonald, H., Asanov, A., Combs, L.L., William Wilson, W., 1999. Correlation of second virial coefficients and solubilities useful in protein crystal growth. Journal of Crystal Growth 196, 424-433.

Haas, C., Drenth, J., 1998. The Protein−Water Phase Diagram and the Growth of Protein Crystals from Aqueous Solution. The Journal of Physical Chemistry B 102, 4226-4232.

Haas, C., Drenth, J., Wilson, W.W., 1999. Relation between the Solubility of Proteins in Aqueous Solutions and the Second Virial Coefficient of the Solution. The Journal of Physical Chemistry B 103, 2808-2811.

Hagen, M.H.J., Frenkel, D., 1994. Determination of phase diagrams for the hard-core attractive Yukawa system. The Journal of Chemical Physics 101, 4093-4097.

Harries, D., Rösgen, J., 2008. A practical guide on how osmolytes modulate macromolecular properties. Methods in Cell Biology 84, 679-735.

Harris, J.M., 1992. Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications. Springer.

Hawe, A., Sutter, M., Jiskoot, W., 2008. Extrinsic fluorescent dyes as tools for protein characterization. Pharm. Res. 25, 1487-1499.

Heidner, E., 1978. The functional dependence of the nucleation rate on the protein concentration and the solubility. Journal of Crystal Growth 44, 139-144.

Hendriks, J., Gensch, T., Hviid, L., van der Horst, M.A., Hellingwerf, K.J., van Thor, J.J., 2002. Transient exposure of hydrophobic surface in the photoactive yellow protein monitored with Nile Red. Biophys. J. 82, 1632-1643.

Hofmeister, F., 1888. Zur Lehre von der Wirkung der Salze. Archiv für experimentelle Pathologie und Pharmakologie 25, 1-30.

Howard, S.B., Twigg, P.J., Baird, J.K., Meehan, E.J., 1988. The solubility of hen egg-white lysozyme. Journal of Crystal Growth 90, 94-104.

Hunter, R.J., 1981. The Calculation of Zeta Potential, Zeta potential in colloid science. ACADEMIC PRESS LIMITED, pp. 59-124.

Jackson, M., Mantsch, H.H., 1995. The use and isuse of FTIR spectroscopy in the determination of protein structure. Critical Reviews in Biochemistry and Molecular Biology 30, 95-120.

Jańczuk, B., Białopiotrowicz, T., Wójcik, W., 1989. The components of surface tension of liquids and their usefulness in determinations of surface free energy of solids. J. Colloid Interface Sci. 127, 59-66.

Janin, J., 1979. Surface and inside volumes in globular proteins. Nature 277, 491-492.

Janson, J.-C., 2012. Protein purification: principles, high resolution methods, and applications. John Wiley & Sons.

Jen, A., Merkle, H.P., 2001. Diamonds in the rough: protein crystals from a formulation perspective. Pharmaceutical research 18, 1483-1488.

Johnson-Léger, C., Power, C.A., Shomade, G., Shaw, J.P., Proudfoot, A.E.I., 2006. Protein therapeutics-lessons learned and a view of the future. Expert opinion on biological therapy 6, 1-7.

Josuran, R., 2015. Prot pi - Protein Tool v1.2.0.27.

Judge, R.A., Forsythe, E.L., Pusey, M.L., 1998. The effect of protein impurities on lysozyme crystal growth. Biotechnology and Bioengineering 59, 776-785.

Kalisz, H.M., Hecht, H.-J., Schomburg, D., Schmis, R.D., 1990. Crystallization and preliminary X-ray diffraction studies of a deglycosylated glucose oxidase from Aspergillus niger. Journal of Molecular Biology 213, 207-209.

Kam, Z., Shore, H.B., Feher, G., 1978. On the crystallization of proteins. Journal of Molecular Biology 123, 539-555.

Kato, A., Nakai, S., 1980. Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. Biochim. Biophys. Acta 624, 13-20.

Kaushik, J.K., Bhat, R., 2003. Why Is Trehalose an Exceptional Protein Stabilizer? The Journal of Biological Chemistry 278, 26458-26465.

Kendrick, B.S., Cleland, J.L., Lam, X., Nguyen, T., Randolph, T.W., Manning, M.C., Carpenter, J.F., 1998. Aggregation of Recombinant Human Interferon Gamma : Kinetics and Structural Transitions. Journal of Pharmaceutical Sciences 87, 1069-1076.

Keshavarz, E., Nakai, S., 1979. The relationship between hydrophobicity and interfacial tension of proteins. Biochim. Biophys. Acta 576, 269-279.

Khurana, R., Fink, A.L., 2000. Do Parallel β-Helix Proteins Have a Unique Fourier Transform Infrared Spectrum ? Biophysical Journal 78, 994-1000.

Kiess, M., Hecht, H.-J., Kalisz, H.M., 1998. Glucose oxidase from Penicillium amagasakiense. European Journal of Biochemistry 252, 90-99.

Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., Lesk, A.M., 2006. MUSTANG: A multiple structural alignment algorithm. Proteins: Structure, Function, and Bioinformatics 64, 559-574.

Kong, J., Yu, S., 2007. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. Acta Biochimica et Biophysica Sinica 39, 549-559.

Kragh-Hansen, U., Møller, J.V., Lind, K.E., 1974. Relation between binding of phenolsulfophtalein dyes and other ligands with a high affinity for human serum albumin. Biochim. Biophys. Acta 365, 360-371.

Krieger, E., Koraimann, G., Vriend, G., 2002. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. Proteins: Structure, Function, and Bioinformatics 47, 393-402.

Kröner, F., Hubbuch, J., 2013. Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application. Journal of chromatography A 1285, 78-87.

Kumar, V., Dixit, N., Zhou, L., Fraunhofer, W., 2011. Impact of short range hydrophobic interactions and long range electrostatic forces on the aggregation kinetics of a monoclonal antibody and a dual-variable domain immunoglobulin at low and high concentrations. International Journal of Pharmaceutics 421, 82-93.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Lauer, T.M., Agrawal, N.J., Chennamsetty, N., Egodage, K., Helk, B., Trout, B.L., 2012. Developability index: A rapid in silico tool for the screening of antibody aggregation propensity. Journal of Pharmaceutical Sciences 101, 102-115.

Leckband, D., Israelachvili, J., 2001. Intermolecular forces in biology. Quarterly reviews of biophysics 34, 105-267.

Lee, J.C., Lee, L.L.Y., 1981. Preferential solvent interactions between proteins and polyethylene glycols. The Journal of biological chemistry 256, 625-631.

Lehermayr, C., Mahler, H.-C., Mäder, K., Fischer, S., 2010. Assessment of net charge and protein-protein interactions of different monoclonal antibodies. Journal of pharmaceutical sciences 100, 2551-2562.

Lijnzaad, P., 2007. Hydrophobic patches on protein surfaces. PhD Thesis U. Utrecht 343, 333-343.

Lijnzaad, P., Berendsen, H.J.C., Argos, P., 1996. A method for detecting hydrophobic patches on protein surfaces. Proteins 26, 192-203.

Lin, Y.-B., Zhu, D.-W., Wang, T., Song, J., Zou, Y.-S., Zhang, Y.-L., Lin, S.-X., 2008. An Extensive Study of Protein Phase Diagram Modification: Increasing Macromolecular Crystallizability by Temperature Screening. Crystal Growth & Design 8, 4277-4283.

Lu, J., Wang, X.-J., Ching, C.-B., 2003. Effect of Additives on the Crystallization of Lysozyme and Chymotrypsinogen A. Crystal Growth & Design 3, 83-87.

Luft, J.R., DeTitta, G.T., 2009. Rational Selection of Crystallization Techniques, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 11-46.

Lüllig, H., Frank, T., Fleck, H., Ebert, W., Vogt-Moykopf, J., 1983. Isolation and characterization of human pneumocytes. Praxis und Klinik der Pneumologie 37 Suppl 1, 842-844.

Lundblad, G., Vesterberg, O., Zimmerman, R., Lind, J., 1972. Studies on lysozyme from human leucemic urine by isoelectric focusing. Acta Chem. Scand. 26, 1711-1713.

Macchi, F., Eisenkolb, M., Kiefer, H., Otzen, D.E., 2012. The effect of osmolytes on protein fibrillation. International Journal of Molecular Sciences 13, 3801-3819.

Mahler, H.-C., Friess, W., Grauschopf, U., Kiese, S., 2009. Protein aggregation: pathways, induction factors and analysis. Journal of pharmaceutical sciences 98, 2909-2934.

Martínez, J.M., Martínez, L., 2003. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. Journal of Computational Chemistry 24, 819-825.

Martínez, L., Andrade, R., Birgin, E.G., Martínez, J.M., 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. Journal of computational chemistry 30, 2157-2164.

Matheus, S., Friess, W., Schwartz, D., Mahler, H.-C., 2009. Liquid High Concentration IgG1 Antibody Formulations by Precipitation. Journal of Pharmaceutical Sciences 98, 3043-3057.

Mayburd, A.L., Tan, Y., Kassner, R.J., 2000. Complex formation between Chromatium vinosum ferric cytochrome c' and bromophenol blue. Arch. Biochem. Biophys. 378, 40-44.

McPherson, A., 1976. Crystallization of Proteins from Polyethylene Glycol. The Journal of Biological Chemistry 251, 6300-6303.

McPherson, A., 2001. A comparison of salts for the crystallization of macromolecules. Protein Science 10, 418-422.

McPherson, A., 2004. Introduction to protein crystallization. Methods (San Diego, Calif.) 34, 254-265.

McPherson, A., Cudney, B., 2006. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. Journal of structural biology 156, 387-406.

Melinder, Å., 2007. Thermophysical properties of aqueous solutions used as secondary working fluids. Ph.D. thesis, School of Industrial Engineering and Management, Royal Institute of Technology, KTH Stockholm.

Miller, S., Janin, J., Lesk, A.M., Chothia, C., 1987. Interior and surface of monomeric proteins. J. Mol. Biol. 196, 641-656.

Murakami, K., Sano, T., Yasunaga, T., 1981. Kinetic Studies of the Interaction of Bromophenol Blue with Bovine Serum Albumin by Pressure-jump Method. Bull. Chem. Soc. Jpn. 54, 862-868.

Muschol, M., Rosenberger, F., 1997. Liquid-Liquid Phase Separation in Supersaturated Lysozyme Solutions and Associated Precipitate Formation/Crystallization. The Journal of Chemical Physics 107, 1953-1962.

Natalello, A., Liu, J., Ami, D., Doglia, S.M., de Marco, A., 2009. The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. Proteins 75, 509-517.

Nieba, L., Honegger, A., Krebber, C., Pluckthun, A., 1997. Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. Protein Eng. 10, 435-444.

Nozaki, Y., Tanford, C., 1963. The solubility of amino acids and related Compounds in aqueous urea solutions. J. Biol. Chem. 238, 4074-4081.

Nozaki, Y., Tanford, C., 1965. The solubility of amino acids and related compounds in equeous ethylene glycol solutions. J. Biol. Chem. 240, 3568-3573.

Nozaki, Y., Tanford, C., 1970. The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. J. Biol. Chem. 245, 1648-1652.

Nozaki, Y., Tanford, C., 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. J. Biol. Chem. 246, 2211-2217.

Oelmeier, S.A., Dismer, F., Hubbuch, J., 2011. Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation. Biotechnology and bioengineering 108, 69-81.

Paul, R., Graff-Meyer, A., Stahlberg, H., Lauer, M.E., Rufer, A.C., Beck, H., Briguet, A., Schnaible, V., Buckel, T., Boeckle, S., 2012. Structure and function of purified monoclonal antibody dimers induced by different stress conditions. Pharmaceutical Research 29, 2047-2059.

Pawar, S.A., Deshpande, V.V., 2000. Characterization of acid-induced unfolding intermediates of glucose/xylose isomerase. European journal of biochemistry / FEBS 267, 6331-6338.

Pegram, L.M., Record, M.T., 2007. Hofmeister Salt Effects on Surface Tension Arise from Partitioning of Anions and Cations between Bulk Water and the Air−Water Interface. J. Phys. Chem. B 111, 5411-5417.

Permyakov, E.A., Berliner, L.J., 2000. α-Lactalbumin: structure and function. FEBS Lett. 473, 269-274.

Petsko, G.A., Ringe, D., 2004a. From Sequence to Structure, Protein Structure and Function. New Science Press Ltd, pp. 1-47.

Petsko, G.A., Ringe, D., 2004b. Protein Flexibility, Protein Structure and Function. New Science Press Ltd, pp. 46-48.

Philo, J.S., Arakawa, T., 2009. Mechanisms of Protein Aggregation. Current Pharmaceutical Biotechnology 10, 348-351.

Piazza, R., 2000. Interactions and phase transitions in protein solutions. Current Opinion in Colloid & Interface Science 5, 38-43.

Poddar, N.K., Ansari, Z.A., Singh, R.K.B., Moosavi-Movahedi, A.A., Ahmad, F., 2008. Effect of monomeric and oligomeric sugar osmolytes on ΔG_D, the Gibbs energy of stabilization of the protein at different pH values: Is the sum effect of monosaccharide individually additive in a mixture? Biophysical Chemistry 138, 120-129.

Price, W.N., Handelman, S.K., Everett, J.K., Tong, S.N., Bracic, A., Luff, J.D., Naumov, V., Acton, T., Manor, P., Xiao, R., Rost, B., Montelione, G.T., Hunt, J.F., 2011. Large-scale experimental studies show unexpected amino acid effects on protein expression and solubility in vivo in E. coli. Microbial Informatics and Experimentation 1, 1-20.

Putnam, C., 2013. Protein Calculator v3.4.

Radzicka, A., Wolfenden, R., 1988. Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry 27, 1664-1670.

Reißer, S., Strandberg, E., Steinbrecher, T., Ulrich, A.S., 2014. 3D Hydrophobic Moment Vectors as a Tool to Characterize the Surface Polarity of Amphiphilic Peptides. Biophys. J. 106, 2385-2394.

Retailleau, P., Ries-Kautt, M., Ducruix, A., 1997. No Salting-In of Lysozyme Chloride Observed at Low Ionic Strength over a Large Range of pH. Biophysical Journal 73, 2156-2163.

Ries-Kautt, M.M., Ducruix, A.F., 1989. Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. The Journal of biological chemistry 264, 745-748.

Righetti, P.G., Caravaggio, T., 1976. Isoelectric point and molecular weights of proteins. A table. Journal of chromatography 127, 1-28.

Righetti, P.G., Tudor, G., Ek, K., 1981. Isoelectric points and molecular weights of proteins. A new table. Journal of chromatography 220, 115-194.

Robbins, F.M., Kronman, M.J., Andreotti, R.E., 1965. Inter- and intramolecular interactions of α-lactalbumin. V. The effect of amidination on association and aggregation. Biochimica et biophysica acta 109, 223-233.

Rose, G.D., Wolfenden, R., 1993. Hydrogen bonding, hydrophobicity, packing, and protein folding. Annu. Rev. Biophys. Biomol. Struct. 22, 381-415.

Ruppert, S., Sandler, S.I., Lenhoff, A.M., 2001. Correlation between the osmotic second virial coefficient and the solubility of proteins. Biotechnology progress 17, 182-187.

Sackett, D.L., Wolff, J., 1987. Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein surfaces. Analytical biochemistry 167, 228-234.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2005. Is it possible to predict the average surface hydrophobicity of a protein using only its amino acid composition? Journal of Chromatography A 1075, 133-143.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2006a. Predicting the behaviour of proteins in hydrophobic interaction chromatography 1: Using the hydrophobic imbalance (HI) to describe their surface amino acid distribution. Journal of Chromatography A 1107, 110-119.

Salgado, J.C., Rapaport, I., Asenjo, J.A., 2006b. Predicting the behaviour of proteins in hydrophobic interaction chromatography 2. Using a statistical description of their surface amino acid distribution. Journal of Chromatography A 1107, 120-129.

Santoro, M.M., Liu, Y., Khan, S.M.A., Hou, L.-X., Bolen, D.W., 1992. Increased Thermal Stability of Proteins in the Presence of Naturally Occurring Osmolytest. Biochemistry 31, 5278-5283.

Schwierz, N., Horinek, D., Netz, R.R., 2010. Reversed anionic Hofmeister series: the interplay of surface charge and surface polarity. Langmuir : the ACS journal of surfaces and colloids 26, 7370-7379.

Scopes, R.K., 1994. Separation by Precipitation, Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag New York, Inc., pp. 71-101.

Shanbhag, V.P., Axelsson, C.-G., 1975. Hydrophobic interaction determined by partition in aqueous two-phase systems. Eur. J. Biochem. 60, 17-22.

Singh, L.R., Dar, T.A., Rahman, S., Jamal, S., Ahmad, F., 2009. Glycine betaine may have opposite effects on protein stability at high and low pH values. Biochimica et biophysica acta 1794, 929-935.

Singh, L.R., Poddar, N.K., Dar, T.A., Kumar, R., Ahmad, F., 2011. Protein and DNA destabilization by osmolytes: the other side of the coin. Life sciences 88, 117-125.

Sleutel, M., Willaert, R., Gillespie, C., Evrard, C., Wyns, L., Maes, D., 2009. Kinetics and Thermodynamics of Glucose Isomerase Crystallization. Crystal Growth & Design 9, 497-504.

Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J., Udseth, H.R., 1990. New developments in biochemical mass spectrometry: electrospray ionization. Analytical chemistry 62, 882-899.

Sousa, R., 1995. Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization. Acta Crystallographica - Section D Biological Crystallography 51, 271-277.

Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W.C.K., Egelhaaf, S.U., Schurtenberger, P., 2004. Equilibrium cluster formation in concentrated protein solutions and colloids. Nature 432, 492-495.

Susi, H., Byler, D.M., 1983. Protein Structure by Fourier Transform Infrared Spectroscopy: Second Derivative Spectra. Biochemical and Biophysical Research Communications 115, 391-397.

Susi, H., Byler, D.M., 1987. Fourier Transform Infrared Study of Proteins with Parallel β-Chains. Archives of biochemistry and biophysics 258, 465-469.

Talreja, S., Perry, S.L., Guha, S., Bhamidi, V., Zukoski, C.F., Kenis, P.J.A., 2010. Determination of the phase diagram for soluble and membrane proteins. The journal of physical chemistry. B 114, 4432-4441.

Tanford, C., 1962. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. J. Am. Chem. Soc. 84, 4240-4247.

Tardieu, A., Bonneté, F., Finet, S., Vivarès, D., 2002. Understanding salt or PEG induced attractive interactions to crystallize biological macromolecules. Acta Crystallographica Section D Biological Crystallography 58, 1549-1553.

Tate, T., 1864. XXX. On the magnitude of a drop of liquid formed under different circumstances. London, Edinburgh, Dublin Philos. Mag. J. Sci. 27, 176-180.

Tayyab, S., Qasim, M.A., 1990. Binding of bromophenol blue to bovine serum albumin and its succinylated forms. Int. J. Biol. Macromol. 12, 55-58.

Teague, S.J., 2003. Implications of protein flexibility for drug discovery. Nature reviews. Drug discovery 2, 527-541.

Timasheff, S.N., Arakawa, T., 1988. Mechanism of Protein Precipitation and Stabilization by Co-Solvents. Journal of Crystal Growth 90, 39-46.

Trinquier, G., Sanejouand, Y.-H., 1998. Which effective property of amino acids is best preserved by the genetic code? Protein Eng. 11, 153-169.

Tsierkezos, N.G., Molinou, I.E., 1998. Thermodynamic Properties of Water + Ethylene Glycol at 283.15, 293.15, 303.15, and 313.15 K. J. Chem. Eng. Data 43, 989-993.

Vajo, Z., Fawcett, J., Duckworth, W.C., 2001. Recombinant DNA Technology in the Treatment of Diabetes: Insulin Analogs. Endocrine Reviews 22, 706-717.

Valerio, M., Colosimo, A., Conti, F., Giuliani, A., Grottesi, A., Manetti, C., Zbilut, J.P., 2005. Early events in protein aggregation: Molecular flexibility and hydrophobicity/charge interaction in amyloid peptides as studied by molecular dynamics simulations. Proteins: Structure, Function and Genetics 58, 110-118.

Vekilov, P.G., 2012. Phase diagrams and kinetics of phase transitions in protein solutions. Journal of Physics: Condensed Matter 24, 193101.

Vekilov, P.G., Monaco, L.a., Thomas, B.R., Stojanoff, V., Rosenberger, F., 1996. Repartitioning of NaCl and protein impurities in lysozyme crystallization. Acta Crystallographica Section D: Biological Crystallography 52, 785-798.

Velev, O.D., Kaler, E.W., Lenhoff, A.M., 1998. Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen. Biophysical journal 75, 2682-2697.

Vihinen, M., 1987. Relationship of protein flexibility to thermostability. Protein Engineering 1, 477-480.

Vivarès, D., Belloni, L., Tardieu, a., Bonneté, F., 2002. Catching the PEG-induced attractive interaction between proteins. The European physical journal. E, Soft matter 9, 15-25.

Vlachy, V., Blanch, H.W., Prausnitz, J.M., 1993. Liquid-liquid phase separations in aqueous solutions of globular proteins. AIChE Journal 39, 215-223.

Vlachy, V., Prausnitz, J.M., 1992. Donnan Equilibrium. Hypernetted-Chain Study of One-Component and Multicomponent Models for Aqueous Polyelectrolyte Solutions. The Journal of Physical Chemistry 96, 6465-6469.

Voet, D., Voet, J.G., W, P.C., 1999. Amino Acids, Fundamentals of Biochemistry, 1st ed. John Wiley & Sons, Inc., pp. 77-92.

Volmer, M., Weber, A., 1926. Keimbildung in übersättigten Gebilden. Zeitschrift für Physikalische Chemie 119, 207-301.

Waldmann-Meyer, H., Schilling, K., 1956. The Interaction of Bromophenol Blue with Serum Albumin and γ-Globulin in Acid Medium. Arch. Biochem. Biophys. 64, 291-301.

Wang, J., Dauter, M., Alkire, R., Joachimiak, A., Dauter, Z., 2007. Triclinic lysozyme at 0.65 Å resolution. Acta Crystallographica Section D: Biological Crystallography 63, 1254-1268.

Wang, W., Li, N., Speaker, S., 2010. External Factors Affecting Protein Aggregation, in: Wang, W., Roberts, C.J. (Eds.), Aggregation of Therapeutic Proteins, 1st ed. John Wiley & Sons, pp. 119-204.

Webb, J.N., Webb, S.D., Cleland, J.L., Carpenter, J.F., Randolph, T.W., 2001. Partial molar volume, surface area, and hydration changes for equilibrium unfolding and formation of aggregation transition state: high-pressure and cosolute studies on recombinant human IFN-γ. Proceedings of the National Academy of Sciences of the United States of America 98, 7259-7264.

Wei, Y.-J., Li, K.-A., Tong, S.-Y., 1996. The interaction of Bromophenol Blue with proteins in acidic solution. Talanta 43, 1-10.

Whitney, P.L., Tanford, C., W, P.L., Tanford, C., 1962. Solubility of amino acids in aqueous urea solutions and its implications for the denaturation of proteins by urea. J. Biol. Chem. 237, 1735-1737.

Winzor, D.J., 2004. Determination of the net charge (valence) of a protein: A fundamental but elusive parameter. Analytical Biochemistry 325, 1-20.

Wu, S.L., Karger, B.L., 1996. Hydrophobic interaction chromatography of proteins. Methods in enzymology 270, 27-47.

Yadav, S., Laue, T.M., Kalonia, D.S., Singh, S.N., Shire, S.J., 2012. The influence of charge distribution on self-association and viscosity behavior of monoclonal antibody solutions. Molecular Pharmaceutics 9, 791-802.

Yadav, S., Shire, S.J., Kalonia, D.S., 2011. Viscosity analysis of high concentration bovine serum albumin aqueous solutions. Pharmaceutical Research 28, 1973-1983.

Yancey, P.H., 2001. Water Stress, Osmolytes and Proteins. American Zoologist 41, 699-709.

Yang, M.X., Shenoy, B., Disttler, M., Patel, R., McGrath, M., Pechenov, S., Margolin, A.L., 2003. Crystalline monoclonal antibodies for subcutaneous delivery. Proceedings of the National Academy of Sciences of the United States of America 100, 6934-6939.

Zeelen, J.P., 2009. Interpretation of the Crystallization Drop Results, in: Bergfors, T.M. (Ed.), Protein Crystallization, 2nd ed. Internat'l University Line, pp. 175-194.

Zhang, J., Topp, E.M., 2012. Protein G, protein A and protein A-derived peptides inhibit the agitation induced aggregation of IgG. Molecular Pharmaceutics 9, 622-628.

Zhang, Y., Cremer, P.S., 2009. The inverse and direct Hofmeister series for lysozyme. Proceedings of the National Academy of Sciences of the United States of America 106, 15249-15253.

Ziegler, G.R., Foegeding, E.A., 1990. The Gelation of Proteins, in: Kinsella, J.E. (Ed.), Advances in Food and Nutrition Research. Academic Press, pp. 203-298.

Curriculum Vitae

Lara Galm lara.galm@kit.edu

Education

since 12/2011	**PhD candidate** at the Karlsruhe Institute of Technology, Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering *"Characterization, manipulation, and prediction of protein aggregation in model systems"* supervisor: Prof. Dr. Jürgen Hubbuch
10/2006 – 07/2011	**Studies of Bioengineering** at the Karlsruhe Institute of Technology
01/2011 – 07/2011	**Diploma thesis** at the Karlsruhe Institute of Technology, Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering *"Charakterisierung von Proteinwechselwirkungen für die Proteinkristallisation mittels statischer Lichtstreuung"* supervisors: Prof. Dr. Jürgen Hubbuch, Dr.-Ing. Natalie Rakel
11/2009 – 09/2010	**Undergraduate thesis** at the Karlsruhe Institute of Technology, Institute of Thermal Process Engineering, Thin Film Technology "*Untersuchung der Trocknung von Sensorfilmen zur Blutzuckermessung*" supervisors: Prof. Dr.-Ing. Wilhelm Schabel, Dr.-Ing. Philip Scharfer, Dipl.-Ing. Sibylle Kachel

Work experience

since 12/2011	**Research assistant** at the Karlsruhe Institute of Technology, Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering
10/2007 – 03/2011	**Student assistant** at the Karlsruhe Institute of Technology, Institute for Algebra and Geometry
03/2006 – 07/2006	**Internship** at Unisensor Sensorsyteme GmbH, Department for Research and Development for Soft Drink Industry

Scientific publications

Talks

L. Galm, J. Morgenstern, S. A. Oelmeier, J. Hubbuch

Investigating the influence of glycerol, PEG 1000, and glycine on the phase behavior of lysozyme and their impact on the stability of the native conformational state

Biological and Complex Fluids II: Novel Trends in Characterizing Interactions, Microstructure and Rheology (ECI Conference Series), Durham, North Carolina, USA, August 10-14, 2014

S. Amrhein, K. C. Bauer, L. Galm, M.-T. Schermeyer, J. Hubbuch

Predictive Approaches for Protein Phase Behavior

Recovery of Biological Products Conference XVI, Rostock, Germany, July 27-31, 2014

Papers

N. Rakel, L. Galm, K. C. Bauer, J. Hubbuch

Influence of macromolecular precipitants on phase behavior of monoclonal antibodies

Biotechnology Progress, 31 (2015) 145-153, doi: 10.1002/btpr.2027

N. Rakel, K.C. Bauer, L. Galm, J. Hubbuch

From osmotic second virial coefficient (B_{22}) to phase behavior of a monoclonal antibody

Biotechnology Progress, 31 (2015) 438-451, doi: 10.1002/btpr.2065

K. Baumgartner, L. Galm, J. Nötzold, H. Sigloch, J. Morgenstern, K. Schleining, S. Suhm, S. A. Oelmeier, J. Hubbuch

Determination of protein phase diagrams by microbatch experiments: Exploring the influence of precipitants and pH

International Journal of Pharmaceutics, 479 (2015) 28-40, doi: 10.1016/j.ijpharm.2014.12.027

S. Amrhein, K.C. Bauer, L. Galm, J. Hubbuch

Non-invasive high throughput approach for protein hydrophobicity determination based on surface tension

Biotechnology and Bioengineering, doi: 10.1002/bit.25677

L. Galm, J. Morgenstern, J. Hubbuch

Manipulation of lysozyme phase behavior by additives as function of conformational stability

submitted to International Journal of Pharmaceutics