Towards in silico process development and digitized manufacturing for antibody-drug conjugates (ADCs): Mechanistic kinetic modeling and its applications

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (DR.-ING.)

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

genehmigte

DISSERTATION

von

Jan Tobias Weggen, M.Sc.

aus Ratingen

Tag der mündlichen Prüfung: 08.11.2024

Erstgutachter: Prof. Dr. Jürgen Hubbuch

Zweitgutachter: Prof. Dr. Michel Eppink (TU Delft)

Danksagung

Eine längere Reise mit vielen Hochs und Tiefs geht nun zu Ende und ich möchte mich hiermit bei allen bedanken, die mich auf diesem Weg begleitet und unterstützt haben:

An erster Stelle möchte ich mich bei Prof. Jürgen Hubbuch für die Ermöglichung dieser Doktorarbeit im Rahmen eines so spannenden Forschungsprojekts bedanken. Den großzügigen Freiraum, den ich für die Gestaltung dieses Projektes erfahren durfte, weiß ich sehr zu schätzen. Vielen Dank für das entgegengebrachte Vertrauen, die wissenschaftlichen Ratschläge und die zahlreichen, konstruktiven Diskussionen.

Zusätzlich möchte ich Prof. Michel Eppink für die Übernahme des zweiten Gutachtens und das Interesse an meiner Arbeit danken.

Ermöglicht wurde diese Arbeit auch durch eine langjährige, wissenschaftliche Kollaboration mit AstraZeneca (früher MedImmune, Gaithersburg, USA). Aus diesem Grund danke ich AstraZeneca, sowohl für die finanzielle Unterstützung als auch die Bereitstellung der verwendeten Antikörper, sowie dem Projektteam, Dr. Michaela Wendeler, Dr. Kimberly Hui und Ryan Bean, für die großartige Zusammenarbeit in den letzten Jahren. *Thank you for the countless project meetings, the fruitful discussions and all the effort you put in reviewing our manuscripts. Special thanks to Kimberly and Ryan for conducting so many experiments, which created a solid foundation for this work.*

Besonders danken möchte ich außerdem Robin Schiemer, für die erfolgreiche Zusammenarbeit, sowie die ausgedehnten Diskussionen, was schließlich zu einer gemeinsamen Publikation geführt hat. Danke auch für die unzähligen Denkanstöße für andere Projekte, die Motivation für mathematische Methoden und das jederzeit offenes Ohr bei Fragen oder Problemen.

Ein großer Dank gebührt außerdem meinen Studenten Janik Seidel, Katrin Schmitt, Michael Liu, David Fischbach, Niels Holfelder, Pedro González und Johanna Ossmann für die tatkräftige Unterstützung, im Rahmen von Abschlussarbeiten und als Hiwi im Labor. Ich bin sehr dankbar für eure engagierte und harte (Labor-)Arbeit, eure frischen Ideen und neuen Perspektiven, von denen ich sehr profitieren konnte.

Ein riesiger Dank geht außerdem an alle aktuellen und ehemaligen Kolleginnen und Kollegen des MAB-Instituts für die harmonische Arbeitsatmosphäre und das kollegiale Miteinander. Danke für die vielen lustigen Mittags- und

Kaffeepausen, sowie für spaßige Momente bei unseren Seminarfahrten, Praktika, Konferenzreisen und Partys.

Außerdem möchte ich mich bei meinen Bürokollegen für die entspannte Atmosphäre und die regen wissenschaftlichen Gespräche bedanken. Hierbei möchte ich mich besonders bei Tina für unsere Aktiv-Pausen und diverse Aktivitäten außerhalb des Institutsalltags bedanken. Annabelle bin ich sehr dankbar für ihre etlichen Ratschläge zu Manuskripten, Präsentationen oder Vorgehensweisen im Labor.

Zum Schluss möchte ich mich bei meiner Familie für die stetige Unterstützung bedanken, die mir es ermöglicht hat, mein Studium und schlussendlich den Doktor zu absolvieren. Außerdem geht ein großer Dank an meine Freundinnen und Freunde für viele lustige Momente, die mich den Arbeitsalltag vergessen ließen. Last but not least, möchte ich von Herzen bei meiner Freundin Léanne für den emotionalen Rückhalt, ihr offenes Ohr bei Problemen, aufmunternde Worte, und viele unvergessliche Augenblicke & Abendteuer bedanken. Danke für Alles!

<div align="right">
Jan Tobias Weggen

Heidelberg, 12.12.2024
</div>

Abstract

Antibody-drug conjugates (ADCs) are one of the most promising classes of biopharmaceuticals and are used for the treatment of various diseases, primarily for cancer indications. Inspired by Paul Ehrlich's "magic bullet" concept, ADCs are hybrid molecules that transport a potent cytotoxic payload ("drug") specifically to target cells, thus minimizing potential off-target effects. Since the approval of the first ADC, *Mylotarg*, in 2000, the U.S. Food and Drug Administration (FDA) has approved fourteen ADCs and hundreds more are currently proceeding in clinical trials.

Each ADC consists of a monoclonal antibody (mAb) and a cytotoxic drug covalently connected via a chemical linker. The antigen specificity of the mAb ensures that the ADC primarily accumulates inside the target cells, where it releases its payload often triggered through a chemical release mechanism. Crucial for the therapeutic efficacy is the Drug-to-Antibody Ratio (DAR), which typically ranges from 2 to 8. The DAR can be controlled through various conjugation methods, which differ in the binding sites for payload attachment. These methods can be divided into stochastic and site-directed conjugation methods. Regarding the former, the native binding sites on the mAb are utilized, such as surface lysine residues or reduced interchain disulfide bonds. This requires one or a few biochemical reactions to generate reactive bindings sites. The resulting conjugated ADC product usually exhibits a wide distribution of different DAR species. Recently, next-generation ADCs effectively reduced this heterogeneity by using site-specific conjugation technologies. Additionally, novel payloads are expanding the array of options for ADCs design aiming to improve the overall therapeutic properties of the final drug product. This evolving landscape and accelerated approval paths, like Breakthrough Therapy designation, presents unique challenges for pharmaceutical companies and requires an intensified process development.

Manufacturing of ADCs involves several biochemical reaction steps, namely mAb functionalization and conjugation, followed by additional purification steps. Such a bioprocess requires a well-defined Chemistry, Manufacturing and Control (CMC) strategy to consistently generate a high-quality product despite the variability introduced by numerous process parameters and raw materials variations. In this context, Quality by Design (QbD) is a systematic approach proposed by the FDA to develop robust manufacturing processes and is based on process understanding, risk management, and continuous improvement. A key aspect of QbD is gaining insights into the relationship between critical process parameters (CPPs) and critical quality attributes (CQAs), which may

also include scaling factors. This understanding is essentially needed to define the design space, in which process parameters can be changed aware of the risk without the need for re-approval.

In this regard, process modeling is a powerful tool for understanding and designing pharmaceutical processes and is also pushed within the current Pharma 4.0 paradigm. Process models are based on mathematical equations trying to represent the real process using either fundamental physical laws, called mechanistic modeling, or empirical correlations, called statistical modeling. While the latter has recently gained popularity through major advances in Machine Learning, it requires large amounts of data, which are challenging to generate due to the complex, cost-intensive nature of biotechnological experiments. In contrast, mechanistic models need smaller datasets for calibration, allow for extrapolation and can integrate with other modeling approaches. Herein, the main challenge is to obtain the necessary process understanding which allows for the postulation of suitable mathematical expressions. This also involves a suitable workflow for model selection and validation. For ADCs, the necessary complexity for model building increases due to the biological origin and macromolecular structure of the mAb. The primary goal for process development is to describe the influence of process parameters on the CQAs, such as the DAR, the Drug Load Distribution (DLD), the level of aggregates and of unconjugated drug. To resolve conjugated species and positional isomers, different analytical methods are available. Most commonly, methods based on High-Performance Liquid Chromatography (HPLC) and mass spectrometry are utilized.

Another major CMC task is establishing an effective monitoring and control strategy for maintaining process robustness. In agreement with QbD, the FDA introduced the Process Analytical Technology (PAT) concept, which promotes advanced sensor technologies in combination with chemometric modeling allowing for real-time, on-line monitoring of CQAs during the process, rather than traditional off-line and end-product testing. This ultimately aims to pave the way for real-time release and, in combination with process models, the establishment of a digital twin.

In the field of ADCs, the fulfillment of the mentioned QbD demands is still hindered due to a significant gap in model and PAT tools utilization. Only a few model-based approaches for individual cases have been recently established. Shortcomings remain in reliable mechanistic models that parametrize reaction kinetics and describe the concentration of species over time, also covering the entire conjugation process. In a prior PhD thesis, the focus has been on the conjugation process for site-directed ADCs including the subsequent purification step, employing a conjugation kinetic model and a PAT sensor

among other things. These studies showed promising results and benefits for process development in the case of site-directed ADCs. To increase the acceptance of model-based methods in ADC process development, open challenges and tasks can be derived: i) extension of kinetic models towards stochastic conjugation reactions including the previous reaction steps, ii) description of the effect of different payloads, iii) investigation and model-based prediction of the influence of mixing, scale, and vessel geometry for the conjugation reaction, and iv) an algorithm to recursively update the conjugation kinetic model with the newly established sensor predictions.

The objective of this thesis was to further digitalize ADC process development by expanding process models to a wider range of ADC reactions and modalities, as well as to demonstrate their benefit within model-based applications. The first thesis part focuses on the development of kinetic models for the primary reaction steps—reduction and conjugation—in the conjugation process of cysteine-based ADCs. In the second part, two case studies deal with the implementation of kinetic models specifically for the site-directed conjugation reaction: 1) studying the prediction of process scale-up and 2) establishing a soft-sensor for advanced real-time monitoring. Thus, this work aimed to gain mechanistic insights through the establishment and integration of models in order to achieve robust control of the final DAR and DLD throughout different scales. Novel ADC modalities and real payloads, which are currently in the clinical phase, were included in this work, underscoring the relevance of this work.

The conjugation reaction is a key step in the ADC process, as the payload is conjugated to the functionalized mAb forming the final therapeutic product. Here, an industry-applicable modeling approach, that can be easily transferred to similar ADCs and payloads to increase the acceptance of process models, is currently missing. In the former work, a novel kinetic model for the homogeneous conjugation reaction of a surrogate payload to the engineered cysteines of a site-directed ADC has been developed. However, this model cannot be transferred to stochastic conjugation reactions, as, e.g., the conjugation to reduced interchain disulfides produces a more heterogeneous mixture of conjugated species. To widen the spectrum of models in ADC conjugation reactions, in the first study (Chapter 3) structurally similar kinetic models were established for the conjugation of two cysteine-based ADCs, namely site-directed conjugation for DAR 2 and interchain disulfides conjugation for DAR 8. The experimental data of three different payloads and three different mAbs were included. Reversed-phase (RP) HPLC as analytics provided accurate and comparable reference data across the two related modalities. Particularly for DAR 8, the fast conjugation trajectories could be resolved using slow, controlled drug feeding. The presence of over- and under-

conjugated species was incorporated in the models via an initial cysteine distribution. The kinetic models showed good accuracy with regards to the reference data for both modalities, whereas they also highlighted that primarily the reactive, initial cysteine distribution controls the maximum achievable DAR value. Comparing surrogate and real payloads enabled to quantify variations in conjugation rates and revealed large differences in payload stability in conjugation buffer. Finally, the kinetic model's ability to aid in process design and optimize reaction conditions was showcased through an exemplary *in silico* screening.

The newly gained kinetic insights suggested that the previous reduction strongly affects the achievable DAR for interchain disulfide conjugates. However, this reaction lacks a mechanistic description of the relationship between the numerous reaction parameters and the final CQAs. Therefore, in the second study (Chapter 4), the reduction reaction of the eight mAb interchain disulfide bonds under varying reducing agent concentrations, mAb concentrations, and reaction temperatures was investigated. The reduction kinetics were herein determined in terms of the reduced mAb fragments as quantified by Capillary Gel Electrophoresis (CGE). First, various reduction kinetic models containing different reaction networks and parameter sets were formulated and the overall best model was selected using error metrics and statistical criteria. A key finding was that the reduced disulfides can undergo re-oxidation after being reduced, while this phenomenon appeared to occur exclusively between the disulfides connecting heavy and light chains. Second, multiple linear regression (MLR) models were established which used the concentrations of the reduction species at each reduction time point to predict the final conjugation outcome, namely the DAR and DLD as determined by RP-HPLC. Together, the reduction kinetic and regression models formed an integrated kinetic model which created a novel tool for *in silico* analysis of the reduction reaction conditions while directly estimating the influence on the final conjugation results. A case study emphasized the capability of the model in selecting reaction parameters to finely adjust the final DLD even for constant DAR value.

Upon entering later clinical phases or approval, scale-up or comparability studies are performed to define the large-scale process parameters in order to establish robust ADC processes. These are mainly based on side-by-side analysis of the final CQAs between the process established at the reaction screening and the large-scale process. However, these studies are resource-intensive, do not account for the reaction kinetics or provide the QbD-demanded understanding of the parameter influence at different scales. To evaluate the capability of model-based approaches for assessing conjugation reaction scalability, in the third study (Chapter 5) the integration of a kinetic model into computational fluid dynamics (CFD) was investigated. This led to a 3D-reactor model which

was generated for three commonly used reaction vessels of varying scales in ADC manufacturing. Subsequently, the effect of the reactor scale and large-scale mixing parameters on the conjugation dynamic could be analyzed *in silico*. Due to the consecutive nature and the site-selectivity of this reaction, the final DAR value and homogeneity were not remarkably influenced by the studied parameter variations or reactor geometry. As the importance of mixing for the DAR value in small-scale devices, which are used during initial process development, is also unknown, experimental kinetic studies were utilized to demonstrate the importance of proper initial homogenization and mixing devices for DAR consistency. Overall, both experimental small-scale studies and large-scale simulations demonstrated that adequate mixing is rather difficult to achieve at small scales, whereas at larger scales this is usually unproblematic due to the ratio of feeding and mixing times. A time-scale analysis was proposed as an alternative to estimate the potential impact of mixing times on the reaction kinetics.

In the conjugation reaction, the conjugated end product is usually sampled and analyzed offline, which does not provide process information allowing to adjust the process, if necessary. This is mainly due to the lack of available sensors able to monitor the conjugating species. Although the formerly developed Ultraviolet/Visible (UV/Vis)-based PAT sensor provides real-time measurements of the DAR, it cannot be used to update the single reacting species in the existing kinetic models. Additionally, both sensor and model predictions are subject to uncertainty and should be therefore combined to provide more accurate estimations of the current reaction than the individual models. In the fourth study (Chapter 6) the creation of an uncertainty-aware soft-sensor for the conjugation reaction is described, which fuses spectroscopic sensor predictions with a kinetic model using an Extended Kalman Filter (EKF). The combination of both sensor and kinetic model by the EKF algorithm enabled the update of the single reacting ADC species recursively, despite the sensor only measuring the DAR value. A newly applied Gaussian Process Regression (GPR) model was used to provide the necessary uncertainty estimates along with the sensor predictions. It was shown that the EKF must be carefully tuned so that it accounts for the time-varying model and sensor uncertainty while reducing the error compared to the uncorrected kinetic model. Finally, the developed framework proved to maintain robustness against sensor noise as well as faulty model initializations and was successfully transferred from batch to fed-batch mode.

In conclusion, this work advances the computational toolbox for model-based ADC process development and process digitalization through diverse reaction models and innovative model-based approaches. The established kinetic models particularly contributed to improving the process understanding spanning over

multiple reaction steps as well as in showcasing their potential for *in silico* optimization of reaction conditions. Through the coupling of a kinetic model and CFD, novel insights into the dominating factors for scale-up could be derived. Further, a novel soft-sensor framework showed its capabilities for advanced and robust real-time monitoring of the conjugation reaction. Overall, the achievements of this thesis paved the way for more knowledge-driven and efficient ADC process development as well as more robust manufacturing processes.

Zusammenfassung

Antikörper-Wirkstoff-Konjugate (ADCs) gehören zu den vielversprechendsten Klassen der Biopharmazeutika und werden zur Behandlung verschiedener Krankheiten, hauptsächlich Krebserkrankungen, eingesetzt. Inspiriert von Paul Ehrlichs Zauberkugel-Konzept sind ADCs, hybride Moleküle, die ein potentes Zytostatikum (den sog. „Wirkstoff') spezifisch in die Zielzellen transportieren und somit mögliche *Off-Target*-Effekte minimieren. Seit der Zulassung des ersten ADCs, Mylotarg, im Jahr 2000, hat die Food and Drug Administration der Vereinigten Staaten (FDA) vierzehn weitere ADCs zugelassen und derzeit befinden sich Hunderte weitere ADCs in klinischen Studien.

Jedes ADC besteht aus einem monoklonalen Antikörper (mAb) und einem zytotoxischen Wirkstoff, welche über einen Linker kovalent verbunden sind. Hierbei gewährleistet die Antigenspezität des mAbs, dass ADCs hauptsächlich innerhalb der Zielzellen akkumulieren und dort, meist durch einen chemischen Freisetzungsmechanismus, den Wirkstoff freisetzen. Ein entscheidender Faktor für die therapeutische Wirksamkeit ist das Wirkstoff-zu-Antikörper Verhältnis (DAR), welches typischerweise im Bereich von 2 und 8 liegt. Der DAR-Wert kann durch verschiedene Konjugationstechnologien eingestellt werden, die sich in den Bindungsstellen für die Wirkstoffkonjugation unterscheiden. Dabei unterscheidet man zwischen chemischen (auch stochastisch genannt) und spezifischen Methoden. Bei den stochastischen Methoden werden nativen Bindestellen des mAbs verwendet. Dazu gehören Oberflächen-Lysine oder reduzierte Zwischenketten-Disulfidbrücken, welche ein oder mehrere biochemische Reaktionen erfordern, um die reaktiven Bindestellen für die Konjugation zu erzeugen. Dies führt allerdings zu einem heterogenen ADC-Produkt mit einer breiten Verteilung an DAR-Spezies. Bei den ADCs der neuen Generation wird diese Heterogenität durch sog. spezifische Konjugationstechnologien reduziert, wobei die gewünschten Bindestellen über Antikörper-Engineering integriert werden. Weiterhin erweitern die Entwicklungen neuartiger Wirkstoffe die Möglichkeiten für das ADC-Design, um die therapeutischen Eigenschaften des finalen Arzneimittelprodukts zu verbessern. Diese zunehmende Palette an unterschiedlichen ADCs und beschleunigten Zulassungspfaden, wie z.B. der *Breakthrough-Therapie*, stellen einzigartige Herausforderungen für Pharmahersteller dar und erfordern eine intensivierte Prozessentwicklung.

Die Herstellungsprozesse von ADCs umfassen mehrere biochemische Reaktionsschritte, wobei zuerst die Bindestellen des mAbs aktiviert werden und danach der Wirkstoff konjugiert werden. Darauf folgen ggf. zusätzliche

Aufreinigungsschritte. Ein solcher Bioprozess erfordert eine gut definierte *Chemistry, Manufacturing and Control* (CMC)-Strategie, um trotz der Variabilität, welche durch zahlreiche Prozessparameter und Rohstoffvariationen beeinflusst wird, immer ein qualitativ hochwertiges Produkt zu erzeugen. In diesem Zusammenhang wurde das Konzept *Quality by Design* (QbD) von der FDA etabliert, um robuste Herstellungsprozesse basiert auf Prozessverständnis, Risikomanagement und kontinuierlicher Verbesserung zu entwickeln. Ein Kernaspekt hierbei ist das grundlegende Verständnis des Einflusses von kritischen Prozessparametern (CPPs) auf kritische Qualitätsmerkmale (CQAs), welches ebenfalls um Skalierungsfaktoren ergänzt werden kann. Dieses Verständnis ist entscheidend, um den sog. *design space* festzulegen, in dem Prozessparameter unter Berücksichtigung des Risikos geändert werden dürfen, ohne, dass eine erneute Zulassung erforderlich ist. Relevante CQAs für ADCs sind u.a. der DAR-Wert, die Wirkstoffbeladungsverteilung (DLD) und der Anteil an Aggregaten sowie an freiem Wirkstoff.

In dieser Hinsicht ist die Prozessmodellierung ein effektives Werkzeug zum Verstehen und Gestalten von pharmazeutischen Prozessen und wird auch im aktuellen „Pharma 4.0" Paradigma gefordert. Prozessmodelle basieren auf mathematischen Gleichungen, die den realen Prozess entweder mithilfe fundamentaler physikalischer Gesetze (mechanistische Modellierung) oder empirischer Korrelationen (statistische Modellierung) darstellen. Letzteres hat in den vergangenen Jahren durch Fortschritte im Bereich *Machine Learning* an Popularität gewonnen. Diese Methode erfordert jedoch große Datenmengen, die aufgrund der Komplexität und Kostspieligkeit biotechnologischer Experimente schwer zu generieren sind. Im Gegensatz dazu benötigen mechanistische Modelle kleinere Datensätze, ermöglichen Extrapolationen und können mit anderen Modellierungsansätzen verknüpft werden. Die Herausforderung bei der Modelbildung besteht darin, dass notwendige Prozessverständnis zu erlangen, das die Herleitung geeigneter mathematischer Gleichungen ermöglicht. Dies umfasst außerdem einen geeigneten Workflow für die Modellauswahl und dessen Validierung. In Bezug auf ADCs steigt die notwendige Modellkomplexität aufgrund des biologischen Ursprungs und der makromolekularen Struktur des mAb enorm. Das übergeordnete Ziel der Prozessentwicklung ist es, den Einfluss von Prozessparameter auf CQAs wie den DAR-Wert, die Wirkstoffbeladungsverteilung (DLD), den Anteil an Aggregation und die Konzentration des nicht konjugierten Wirkstoffs modelhaft zu beschreiben. Zur Quantifizierung der konjugierten Spezies und Positionsisomere stehen verschiedene analytische Methoden zur Verfügung, wobei am häufigsten Methoden basierend auf Hochleistungsflüssigkeitschromatographie (HPLC) und Massenspektrometrie eingesetzt werden.

Ein weiteres Ziel von CMC ist die Entwicklung einer effektiven Überwachungs- und Kontrollstrategie zur Aufrechterhaltung der Prozessrobustheit. Im Einklang mit QbD hat die FDA das *Process Analytical Technology* (PAT) Konzept etabliert, welches mithilfe neuartiger Sensoren und chemometrischer Modelle die Echtzeit- und Online-Überwachung der CQAs während des Prozesses anstelle von Offline- oder Endprodukt-Messung ermöglicht. Dadurch sollen schlussendlich Echtzeit-Freigaben und, in Kombination mit Prozessmodellen, die Etablierung eines digitalen Zwillings ermöglicht werden.

Speziell in der Prozessentwicklung von ADCs mangelt es an mechanistischen Modellen und PAT-Lösungen, die zur Erfüllung der erwähnten QbD-Anforderungen notwendig sind. Bis dato wurden nur wenige und für besondere Fälle angepasste Prozessmodelle etabliert. Dabei mangelt es vor allem an mechanistischen Modellen, die Reaktionskinetiken parametrisieren und die Konzentration der Spezies über die Zeit beschreiben und dabei möglichst den gesamten Konjugationsprozess abdecken. In einer früheren Doktorarbeit lag der Schwerpunkt auf dem Konjugationsprozess für einen bindestellenspezifischen ADC einschließlich des anschließenden Aufreinigungsschritts. Dabei wurden u.a. Kinetikmodelle und PAT-Sensoren etabliert. Diese Arbeit konnte vielversprechende Ergebnisse und Vorteile für die modellhafte Prozessentwicklung im Falle von bindestellenspezifischen ADCs liefern. Um die Akzeptanz modellbasierter Methoden in der ADC Prozessentwicklung insgesamt zu erhöhen, können offene Herausforderungen und notwendige Folgestudien abgeleitet werden: i) Ausweitung der Kinetikmodelle auf stochastische Konjugationsreaktionen und der vorherigen Reaktionsschritte, ii) Beschreibung des Einflusses verschiedener Wirkstoffe, iii) Beschreibung eines *scale-down models* und Anwendung der Modell zur Vorhersage von *scale-up* für die Konjugationsreaktion und iv) ein Algorithmus zur rekursiven Aktualisierung des bestehenden Kinetikmodells mit den online Messungen des neu entwickelten PAT-Sensors für die Konjugationsreaktion.

Das Ziel dieser Doktorarbeit ist es, die ADC-Prozessentwicklung weiter zu digitalisieren, indem Prozessmodelle auf zusätzliche ADC-Reaktionen und -Modalitäten erweitert werden und ihr Nutzen innerhalb modellbasierter Anwendungen gezeigt wird. Der erste Teil der Dissertation konzentriert sich auf die Entwicklung kinetischer Modelle für die primären Reaktionsschritte— Reduktion und Konjugation—im Konjugationsprozess von Cystein-basierten ADCs. Im zweiten Teil untersuchen zwei Studien die Implementierung kinetischer Modelle speziell am Beispiel einer bindestellenspezifischen Konjugationsreaktion: 1) Untersuchung der Vorhersage des Prozess Scale-up und 2) Etablierung eines Soft-Sensors für die erweiterte Echtzeitüberwachung. Diese Arbeit soll daher, neue mechanistische Einblicke durch die Etablierung und Integration von Modellen liefern, um eine robuste Kontrolle des finalen

DAR und DLD über verschiedene Skalen hinweg zu erreichen. Neuartige ADC-Modalitäten und reale Wirkstoffe, die sich derzeit in der klinischen Phase befinden, wurden in diese Arbeit einbezogen, was die Relevanz dieser Arbeit unterstreicht.

Im Herstellungsprozess für ADCs ist die Konjugationsreaktion ein entscheidender Prozessschritt, in dem der Wirkstoff an den funktionalisierten mAb konjugiert wird und somit das endgültige therapeutische Produkt entsteht. Hierbei fehlt ein Industrie-relevanter Modellierungsansatz, der leicht auf ähnliche ADCs oder andere Wirkstoffe übertragen werden kann, um so die Anwendbarkeit von Prozessmodellen zu erhöhen. In der vorherigen Doktorarbeit wurde ein neuartiges Kinetikmodell für die homogene Konjugationsreaktion eines Pseudo-Wirkstoffs an die zwei Cysteine eines bindestellenspezifischen ADCs entwickelt. Dieses Model kann jedoch nicht für die stochastische Konjugationreaktion angewendet werden, da z.B. bei der Konjugation an reduzierte Zwischenketten-Cysteine eine weitaus heterogenere Mischung an konjugierter Spezies entsteht. Um das Spektrum der Modelle auf verschiedene ADCs zu erweitern, wurden in der ersten Studie (Kapitel 3) strukturell ähnliche Kinetikmodelle für die Konjugationreaktion von zwei Cysteine-basierender ADCs etabliert: die bindestellenspezifische Konjugation für DAR 2 und die Zwischenketten-Disulfid-Konjugation für DAR 8. Dazu wurden experimentelle Konjugationskinetiken von drei verschiedenen Wirkstoffen und drei verschiedenen mAbs aufgenommen. Umkehrphasen (RP) HPLC wurde als Analytik verwendet, um präzise und vergleichbare Referenzdaten für die beiden verwandten Modalitäten zu erstellen. Die schnelle Konjugationskinetik im Falle von DAR 8 konnte durch langsame, kontrollierte Wirkstoffzufuhr aufgelöst werden. Die Existenz von über- und unterkonjugierten Spezies wurde in den Modellen durch eine anfängliche Cysteine-Verteilung integriert. Die erstellten kinetischen Modelle konnten eine hohe Genauigkeit bezüglich der Referenzdaten für beide Modalitäten erreichen. Dabei konnten sie nachweisen, dass die anfängliche Cysteine-Verteilung maßgeblich den maximal erreichbaren DAR-Wert bestimmt. Der Vergleich von Pseudo- und realen Wirkstoffen ermöglichte den quantitativen Vergleich bezüglich der Konjugationsraten und konnte entscheidende Unterschiede bei der Wirkstoff-Stabilität in Konjugationspuffer nachweisen. Schließlich wurde die Anwendung der Kinetikmodelle für das Prozess-Design und die Optimierung der Reaktionsbedingungen durch ein beispielhaftes *in silico*-Screening gezeigt.

Die neu gewonnenen Erkenntnisse deuteten darauf hin, dass die vorherige Reduktion den erreichbaren DAR für Zwischen-Ketten Konjugate entscheidend bestimmt. Hierbei fehlt jedoch ebenfalls eine mechanistische Beschreibung der Beziehung zwischen den zahlreichen Reaktionsparametern und den finalen CQAs. Daher wurde in der zweiten Studie (Kapitel 4) die partielle

Reduktionsreaktion der acht mAb Zwischenketten-Disulfidbrücken unter variierenden Reduktionsmittel- und mAb-Konzentrationen sowie Reaktionstemperaturen untersucht. Die Reduktionskinetik wurde hier in Form der reduzierten mAb-Fragmente bestimmt, die mittels Kapillargelelektrophorese (CGE) analytisch quantifiziert wurden. Im ersten Schritt wurden verschiedene Kinetikmodelle mit unterschiedlichen Reaktionsnetzwerken und Parametersätzen erstellt und danach das beste Modell anhand von Fehlermetriken und statistischen Kriterien ausgewählt. Ein zentrales Ergebnis war, dass die reduzierten Disulfide nach der Reduktion wieder oxidieren können, wobei dieses Phänomen ausschließlich zwischen den Disulfiden der schweren und leichten Ketten auftritt. Im zweiten Schritt wurden multiple lineare Regressionsmodelle (MLR) erstellt, die anhand der Konzentration an jedem Zeitpunkt der Reduktion das finale Konjugationsergebnis vorhersagen. Dieses beinhaltet den DAR und die DLD, welche durch RP-HPLC bestimmt wurden. Zusammen bildeten die Reduktionskinetik- und Regressionsmodelle ein integriertes Kinetikmodell, welches ein neuartiges Werkzeug darstellt, um den Einfluss der Reaktionsbedingungen in der Reduktion auf das finale Konjugationsergebnisse *in silico* zu analysieren. Eine Fallstudie konnte die Fähigkeit dieses Ansatzes zur Auswahl geeigneter Reaktionsparameter, um die DLD bei konstantem DAR-Wert fein zu justieren, aufzeigen.

Beim Eintritt in spätere klinische Phasen oder bei der Zulassung werden Scale-up- oder Vergleichbarkeitsstudien durchgeführt, um die großtechnischen Prozessparameter zu definieren und robuste ADC-Prozesse zu etablieren. Dies basiert hauptsächlich auf einer vergleichenden Analyse der endgültigen CQAs zwischen dem im Reaktionsscreening etablierten Prozess und dem großtechnischen Prozess. Diese Studien sind jedoch ressourcenintensiv, berücksichtigen nicht die Reaktionskinetik und liefern nicht das von QbD geforderte Verständnis des Parameter-Einflusses in verschiedenen Skalen. Um die Möglichkeit modellbasierter Ansätze zur Abschätzung der Skalierbarkeit der Konjugationsreaktion zu untersuchen, wurde in der dritten Studie (Kapitel 5) ein bestehendes Kinetikmodell mit numerischer Strömungssimulation (CFD) gekoppelt. Dies bildete ein sog. 3D-Reaktormodell, welche für drei häufig verwendete Reaktoren unterschiedlicher Volumina in der ADC-Herstellung erstellt wurde. Anschließend konnte der Effekt der Reaktorgröße und großtechnischer Mischparameter auf die Konjugationsdynamik *in silico* analysiert werden. Aufgrund der konsekutiven Natur und der Bindestellenspezifität dieser Reaktion wurden der endgültige DAR-Wert und die Homogenität nicht bemerkenswert durch die untersuchten Parameter-Variationen oder die Reaktorgeometrie beeinflusst. Da der Einfluss des Mischens auf den DAR-Wert in kleineren Reaktionsgefäßen, welche für die

anfängliche Prozessentwicklung verwendet werden, ebenfalls unbekannt ist, wurden experimentelle kinetische Studien genutzt, um die Bedeutung der anfänglichen Vermischung sowie der Mischgeräte für die Konsistenz des DAR zu demonstrieren. Insgesamt zeigten sowohl die experimentellen Studien im Kleinmaßstab als auch die Simulationen im großtechnischen Maßstab, dass eine adäquate Durchmischung besonders in kleineren Reaktionsgefäßen eher schwierig zu erreichen ist, während dies in größeren Maßstäben aufgrund des Verhältnisses von Zuführungs- und Mischzeiten normalerweise unproblematisch ist. Eine Analyse der Zeitskalen wurde zusätzlich als Alternative vorgeschlagen, um den potenziellen Einfluss von Mischzeiten auf die Reaktionskinetik abzuschätzen.

Bei der Konjugationsreaktion wird das konjugierte Endprodukt normalerweise nach der Reaktion durch eine Probenahme offline analysiert. Dies liefert keinerlei Prozessinformationen, die es ermöglichen, den Prozess bei Bedarf anzupassen. Ursächlich hierfür ist der Mangel an verfügbaren Sensoren, die die konjugierenden Spezies während der Reaktion überwachen. Obwohl der zuvor entwickelte PAT-Sensor Echtzeitmessungen des DARs liefert, kann dieser nicht direkt verwendet werden, um die einzelnen reagierenden Spezies im bestehenden Kinetikmodell zu aktualisieren. Darüber hinaus unterliegen sowohl Sensor- als auch Modellvorhersagen Unsicherheiten und müssten daher kombiniert werden, um eine präzisere Vorhersage der aktuellen Reaktionsfortschrittes zu liefern. In der vierten Studie (Kapitel 6), wird die Erstellung eines unsicherheitsbewussten Soft-Sensors für die Konjugationsreaktion beschrieben, welcher durch die Fusion von spektroskopischen Sensorvorhersagen mit einem Kinetikmodell mithilfe eines *Extended* Kalman-Filters (EKF) erreicht wurde. Die Kombination von Sensor und Kinetikmodell durch den EKF-Algorithmus ermöglichte die rekursive Aktualisierung der einzelnen reagierenden ADC-Spezies, obgleich der Sensor lediglich den DAR-Wert bestimmen kann. Ein Gauß-Prozess Regression (GPR) Modell wurde zusätzlich verwendet, um die notwendigen Unsicherheitsschätzungen für die Sensorvorhersagen liefern zu können. Es konnte gezeigt werden, dass der EKF sorgfältig getunt werden muss, damit dieser die zeitabhängige Modell- und Sensorunsicherheit berücksichtigt und den Fehler im Vergleich zum unkorrigierten Kinetikmodell reduziert. Schließlich erwies sich das entwickelte Framework als robust gegenüber Sensorrauschen sowie fehlerhaften Modellinitialisierungen und wurde erfolgreich vom Batch- auf den Fed-Batch-Modus übertragen.

Zusammenfassend erweitert diese Dissertation das Modell-Repertoire für die modellbasierte ADC-Prozessentwicklung und Prozessdigitalisierung mithilfe von vielfältigen Reaktionsmodellen und innovativen modellbasierten Ansätzen. Insbesondere konnten die erstellten Kinetikmodelle zum einen zur Verbesserung des notwendigen Prozessverständnisses über mehrere Reaktionsschritte hinweg

beitragen und zum anderen ihr Potential für die *in silico* Optimierung der Reaktionsbedingungen aufzeigen. Durch die Kopplung eines Modells mit CFD konnte außerdem neue Erkenntnisse in die dominierenden Faktoren hinsichtlich Scale-up gewonnen werden. Weiterhin wurde ein neuartiges Soft-Sensor Konzept etabliert, welches zur erweiterten und robusten Prozessüberwachung der Konjugation genutzt werden kann. Insgesamt ebnen diese Errungenschaften den Weg für eine wissensgetriebene und effizientere ADC-Prozessentwicklung sowie robustere Herstellungsprozesse.

Table of contents

1

Introduction

Biopharmaceuticals, or biologics, have revolutionized modern medicine, providing specialized treatments for a range of diseases, including cancer and metabolic disorders. This advancement began with the first approval of recombinant human insulin in the early 1980s [1]. Unlike traditional small molecule drugs synthesized chemically, biopharmaceuticals are produced in living cells. This origin allows them to have a complex molecular structure which is usually 100-1000 times larger compared to small-molecule drugs and capable of specifically targeting and modulating biological pathways important for disease mechanisms. The primary advantage of biopharmaceuticals over small molecule drugs is their higher specificity, which generally leads to increased efficacy and reduced nonspecific effects. Examples include biological proteins (cytokines, hormones, and clotting factors), monoclonal antibodies (mAbs), vaccines, and cell- and gene-based therapies [2]. The ongoing development and diversification of biopharmaceuticals, along with the creation of novel formats, continue to reshape the biopharmaceutical landscape.

Targeted therapy represents an advanced form of medical treatment that uses innovative, highly specific drugs for treating cancer. In this group of therapeutics, mAbs have seen vast growth within the last years [3], [4]. While unmodified mAbs can exhibit some therapeutic effectiveness, their clinical benefit often varies and they may ultimately be not curative [5]. One promising improvement strategy is their chemical modification to create novel immunoconjugates with a better therapeutic index. This concept has been used in oncology to attach chemo drugs to mAbs generating so-called antibody-drug conjugates (ADCs). The objective is to combine the specificity of an antibody with the potency of cytotoxic drugs to target and kill cancer cells while sparing healthy cells. Other types of molecules, such as radioisotopes, enzymes, or

toxins, have also been coupled, expanding this concept towards other research or therapeutic areas [6], [7].

Like other biopharmaceuticals, the main challenge in process development for ADCs consists of establishing robust manufacturing processes that consistently meet regulatory standards and therapeutic expectations. Here, complexity increases owing to multiple facets of both biologics and small-molecule properties [8]. In this context, the FDA's Quality by Design (QbD) initiative becomes particularly relevant. QbD emphasizes a proactive, science- and risk-based approach to pharmaceutical development, highlighting the need to understand and control critical quality attributes (CQAs) throughout the manufacturing process [9]. Another QbD-consistent framework is Process Analytical Technology (PAT), which aims to provide real-time data and feedback for process adjustments [10]. Process modeling and simulation have emerged as a crucial method in both QbD and PAT because it is essential for system understanding, process optimization, and control. This trend is emphasized by the increasing number of publications dealing with modeling in the field of biopharmaceuticals over the past 25 years as can be seen in Figure 1-1.

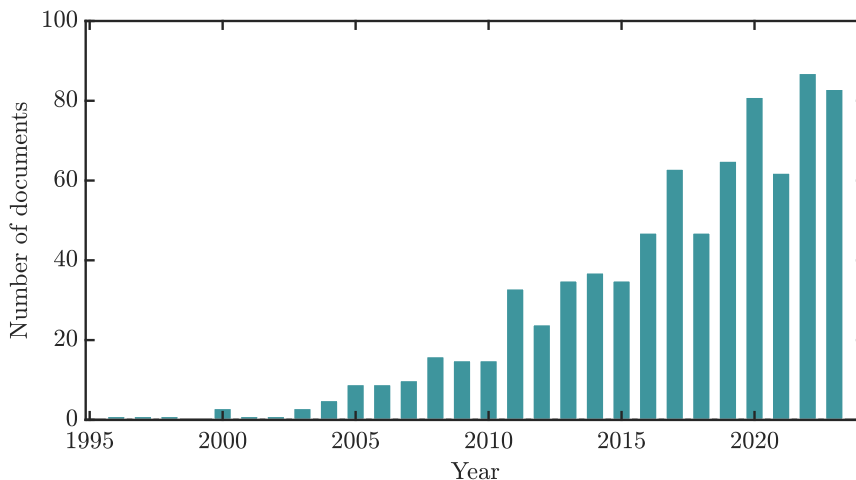Figure 1-1: Number of publications per year about modeling in the field of biopharmaceuticals. Data was extracted from Scopus database searching within article title, abstract, and keywords using the keywords "modeling" and "biopharmaceuticals".

The following sections provide the theoretical background for ADCs as well as the fundamentals for mathematical modeling, QbD, and PAT for biochemical processes.

1.1 Antibody-drug conjugates

1.1.1 Historical perspective

Targeted therapy traces its origins to the visionary work of Paul Ehrlich in the early 20th century. Ehrlich, fascinated by the selective staining of tissues by synthetic dyes, hypothesized the presence of specific receptors on cells that could be targeted by chemicals in his pioneering side-chain theory [11]. From this foundation, Ehrlich developed the "magic bullet" concept, envisioning the development of personalized, tailored therapeutics that are chemically modified in such a way that they precisely target malignant cells without damaging healthy tissue [12], [13]. The first practical applications of chemotherapy began with the accidental discovery of the anticancer properties of nitrogen mustard in patients with lymphoma by Goodman and Gilman in 1946 [14]. Antifolates were the next advancement, which started with the work of Sidney Farber in 1948, who demonstrated the potential of folic acid antagonists to induce remissions in children with acute lymphoblastic leukemia [15], [16]. These discoveries formed the basis for the development of a new class of drugs, such as antimetabolites and antimitotics in the early '40s and '50s [17].

The concept of antibody-drug conjugates stems from Ehrlich's idea to attach toxins to antibodies thereby improving their therapeutic specificity [18]. In 1958, methotrexate was conjugated to a leukemia cell-targeting antibody, constituting the first investigated ADC [19]. However, early clinical trials did not accomplish success due to suboptimal payloads, polyclonal antibodies, and unstable linkers [20], [21]. A major boost was the development of the hybridoma technique by Köhler and Milstein in 1975, which allowed for the creation of monoclonal antibodies to target specific antigens on cancer cells [22]. Initially, these mAbs were mouse-derived causing severe side effects in the human. Genetic engineering and phage display technologies led to the development of chimeric, humanized, and human mAbs which allowed for minimization of mAb immunogenicity [23], [24]. The first human clinical trial of an ADC was conducted in 1983 [25]. At the same time, multiple conventional mAbs were approved for cancer therapy, such as Trastuzumab and Rituximab, which promoted research in the field of immunotherapy [17]. In the 1990s, the first ADCs consisting of chimeric or humanized mAbs demonstrated significant potential in clinical trials but still faced challenges such as poor linker stability and insufficient delivery of the drug to cancer cells [26]. An important milestone was marked in 2000 with the approval of the first ADC, Gemtuzumab ozogamicin (*Mylotarg*), for the treatment of acute myeloid leukemia [27]. However, *Mylotarg* was later withdrawn due to safety concerns observed in the clinic caused by premature payload release leading to off-target toxicity and

was then re-approved with a lower dose in 2017 [28]. Subsequent advancements in linker technology and payload design have led to the development of next-generation ADCs that release the drug only upon reaching the targeted cancer cells thus improving therapeutic indices [29]. This resulted in the development and approval of various ADCs in the 2010s, such as Brentuximab vedotin and Trastuzumab emtansine. Recent years have seen a substantial growth of ADCs in terms of new approvals and new ADCs entering the clinic. As of 2023, 15 ADCs were approved and over a hundred were progressing in clinical trials [29]. While oncology remains the primary focus, applications of ADCs are expanding into other therapeutic areas, such as autoimmune and infectious diseases, demonstrating the versatility of the ADC platform [30].

1.1.2 Mechanism of action

ADCs are sophisticated agents which aim to selectively deliver cytotoxic drugs to targeted cells, wherein it interacts and induce a medical effect. Although the detailed mechanisms are complex and multifaceted, Figure 1-2 schematically presents the primary mechanism of action (MOA).
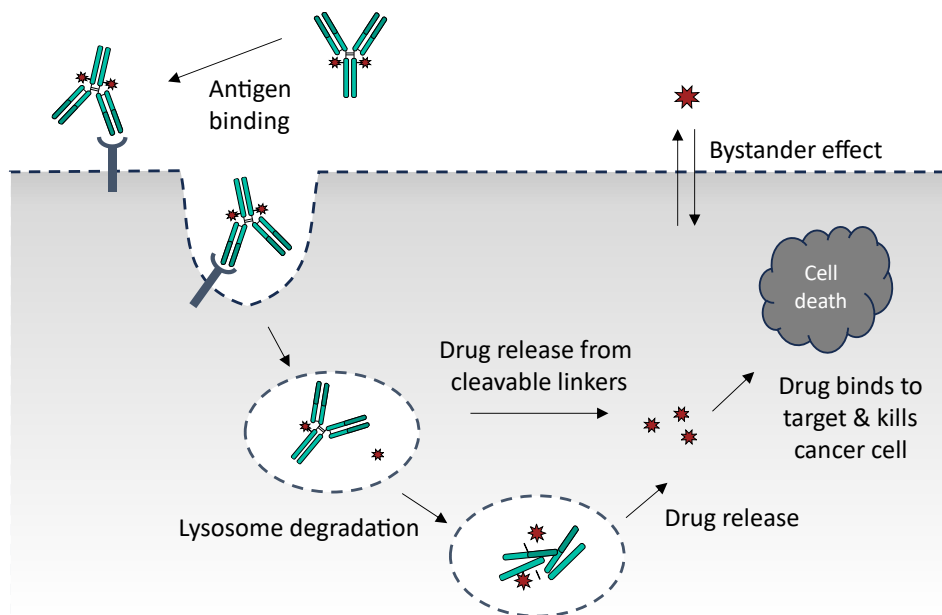
Figure 1-2: Schematic representation of the mechanism of action of an ADC inside the cell.

Upon intravenous administration in the bloodstream, the ADC binds to the target antigen on the cell surface and is then internalized through receptor-mediates endocytosis [45]. Inside the cell, the ADC is transported in endosomes, followed by lysosomes. Depending on the release mechanism, the cytotoxic payload is either released due to linker degradation in endosomes or lysosomes, or full lysosomal digestion of the ADC [49]. The free drug then induces cell death, typically by interfering with critical cellular processes such as DNA

replication or microtubule function, thereby killing the cancer cell. Additionally, a bystander effect occurs when free drug diffuses out of the cell and lead to non-selective damage of neighboring cells in the tumor microenvironment. [50]

1.1.3 Key elements

An ADC is composed of three main components which is depcited in Figure 1-3: a monoclonal antibody targeting a specific tumor-associated antigen, a linker molecule that attaches the antibody to a cytotoxic drug, and the cytotoxic drug itself, also called payload.
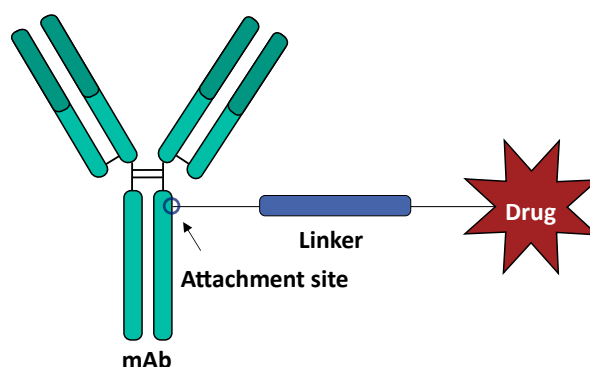
Figure 1-3: Main building blocks of an ADC.

Designing the ideal ADC for a specific indication presents several challenges due to the multifaced nature of these molecule as well as the complex biological mechanisms involved. Besides the three building blocks of an ADC, the target antigen also needs to be considered when designing an ADC for a specific purpose. These elements are discussed in more detail in the following:

Target antigen: The selection of an antigen target is a critical factor in defining the therapeutic window of an ADC. A few aspects need to be considered to ensure the targeted deliver of the payload to cancer cells. These involves a highly dense and homogeneous expression on cancerous cells with low expression on healthy cells, while being accessibility on the cell surface for the circulating mAb [31], [32]. Additionally, an efficient internalization after binding and rapid routing to lysosomes for proteolysis is regarded essential for safety and efficacy [33]. However, even non-internalized ADCs can be therapeutically effective due to the bystander effect, where the drug is released in the tumor microenvironment and affect neighboring tumor cells [18]. Antigen examples that have been utilized for ADCs include ERBB2, CD19, CD33 or CD22 [18].

Antibody moiety: The mAb must fulfill several requirements, such as high affinity to the target antigen, high target specificity, and long retention time in the blood plasma (half-life of up to three weeks [34]). To reduce immunogenic reactions of the patient, humanized or fully human mAbs are typically used

[32]. Specifically, immune globulins G (IgG) are predominantly used for ADCs, with IgG1 being the most common, because it can induce additional effector functions, such as Antibody-Dependent Cellular Cytotoxicity (ADCC) and Complement-Dependent Cytotoxicity (CDC), which contribute to the overall antitumor activity of the ADC [35]. However, other IgG subclasses, such as IgG2 or IgG3 are also used, depending on the desired ADC characteristics. Recent advancements in protein engineering allowed for the creation of bispecific mAbs, which showed to improve internalization and intracellular trafficking, ultimately resulting in higher efficacy in specific cases compared to conventional mAbs [36].

Linker: The linker is the chemical bridge between mAb and payload which needs to ensure maintaining stability over the period of blood circulation and uptake in the tumor cell. Special focus lies on controlling intracellular payload release, utilizing cleavable or non-cleavable linkers [35], [37]. With regards to the cleavable linkers, an inherent property inside the tumor cells is exploited to induce payload release [38]. Several release strategies can be applied in ADC design: The protease-sensitivity approach exploits the high tumor cell-specific concentration of proteases (e.g., cathepsin or plasmin), to cleave a peptide sequence within the linker. A popular example is the valine-citrulline (vc) dipeptide which is cleaved by the upregulated lysosomal protease cathepsin B, as pioneered by Dubowchik et al. [39]. The pH-sensitive strategy utilizes the acidic environment of endosomal (pH 5–6) and lysosomal (pH 4.8) compartments compared with the neural pH in extracellular environment to trigger hydrolysis for acid-labile groups such as hydrazones [40]. Another strategy are linkers with a disulfide bridge being reduced by glutathiones which exhibit a 1000-fold higher intracellular concentration in tumor cells than in normal cells [37], [38]. Alternatively, non-cleavable linkers (e.g. thioether and amide bands) remain stable intracellularly and release the payload only upon mAb degradation inside the lysosomes [40]. Whilst these linkers have the advantage of offering greater stability compared with cleavable linkers, they rely on payloads that evolve their potential despite being chemically modified [35]. The release strategy is selected on a case-by-case basis by considering multiple factors, such as the activity of the drug, the characteristic of the mAb, and the target disease. One particular problem is the inherent hydrophobicity of the conjugated payloads which can lead to aggregation of the ADC causing undesired biodistribution, increased immunogenicity, and higher tendency for effluxing out of the target cells [38]. To tackle this, linkers that engage hydrophilic moieties, such as polyethylene glycol (PEG), have been recently investigated, resulting in a reduction of hydrophobicity and improved pharmacokinetic profiles [41], [42].

Cytotoxic payload: The payload is the active cytotoxic warhead of an ADC and must therefore meet several criteria in order to create an effective therapeutic product. Primarily, the payload should exhibit strong cell toxicity which is crucial due to the low rate of mAb update by the tumor cells. In fact, less than 1-2% of the injected dose typically reach the targeted site of action [35]. Hence, the utilized cytotoxic agents are normally 100-1000-fold more potent compared to conventional chemo-drugs [43]. Furthermore, the payload must be modifiable to facilitate chemical coupling, remain stable from manufacturing until reaching the target site, and exhibit low immunogenic potential with a well-defined metabolic profile owing the problem that a significant portion of the drug is converted outside the tumor cells [33]. The employed drugs mainly fall into two categories: tubulin inhibitors, such as monomethyl auristatin derivatives and maytansines, and DNA-damaging drug, such as calicheamicin, pyrrolobenzodiazepines (PBD), or topoisomerase I inhibitors [44], [45]. However, a major challenge in the payload selection is overcoming multidrug resistance, which is a phenomenon where tumor cells develop resistance to various drugs and actively expel the cytotoxic agent out of the cell using glycoprotein efflux pumps [43]. Research in this field is ongoing with novel payloads constantly diversifying the payload landscape, such as ultra-potent payloads from natural products [46], topoisomerase 2 inhibitors or non-cytotoxic immunomodulatory agents [47]. Another format that is investigated are dual payload ADC, which aim to increase treatment effects by using two distinct payloads per mAb [48].

1.1.4 Conjugation chemistries

Bioconjugation refers to a chemical reaction which is used to attach synthetic groups to a desired position on proteins to alter or enhance the properties of the protein. For ADCs, this involves selecting suitable binding site on the IgG scaffold for the payload conjugation, which is referred as conjugation chemistry. Figure 1-4 gives an overview of the three most common conjugation chemistries which are discussed in more detail in the following.

1.1.4.1 Stochastic conjugation

The first generation of ADCs utilized the native antibody structure, primarily conjugation to either amine group of lysine side chains or cysteine residues which were generated by controlled reduction of the interchain disulfide bonds [52]. In lysine conjugation (Figure 1-4, left), a stable bond between the NH_2-group of a lysine residue and an amine-reactive group, mainly N-hydroxysuccinimide (NHS) esters, on the linker are formed [34]. In this case, the conjugation workflow typically involves two steps: First, a bifunctional crosslinker, usually SMCC, attaches to the mAb lysines through its NHS ester.

Next, the payload is conjugated by reacting with the maleimide of the crosslinker [53]. Since roughly 40 out of 100 lysine residues in an IgG1 are solvent accessible, this method results in different possible isomers of species with the same DAR value as well as a wide DAR distribution [54], [55]. The typical average DAR value ranges between $3.5 - 4$ [21].
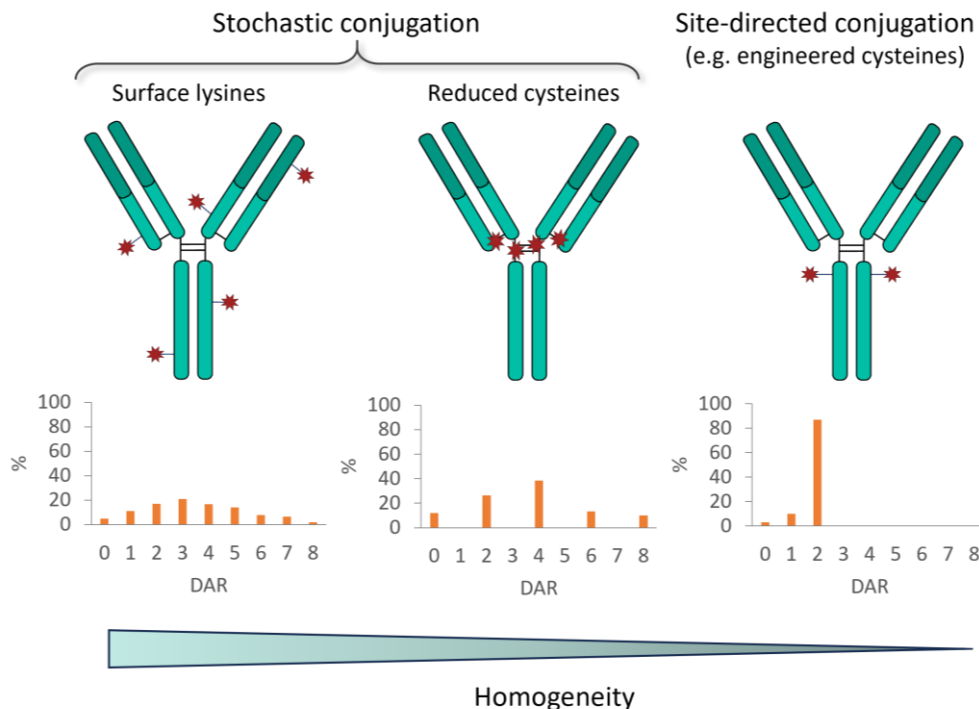
Figure 1-4: Overview of the common conjugation chemistries which schematically indicates the payload bindings positions on the IgG scaffold and typical DLDs. Adopted from [45].

In cysteine-directed conjugation (Figure 1-4, middle), the cysteines engaged in the four interchain disulfide bonds are targeted for payload conjugation. This requires a mild to full reduction of the interchain disulfide bonds using reducing agents, such as tris(2-carboxyethyl) phosphine (TCEP) or dithiothreitol (DTT) [52]. This reaction generates up to eight reactive thiol groups which are usually conjugated via maleimide-functionalized payloads resulting in a DAR distribution which is more homogeneous compared to lysine conjugation [54]. The average DAR values is mostly around 4, while ADCs with higher DAR values of 8 are also approved [56]. Overall, the ADC population heterogeneity of both so-called random conjugation methods can lead to distinct pharmacokinetic profiles for each ADC species, potentially affecting the overall *in vivo* efficacy and safety [57]. For example, it has been shown that high DAR species, such as species with a DAR 8, exhibit faster clearance rates resulting in reduced potency [58]. Additionally, the maleimide-thiols linkage is known to be affected by retro-Michael reaction in the serum causing premature payload release [56]. Despite these disadvantages, these methods has demonstrated their

effectiveness and are widely used among approved and clinically-tested ADCs [52]. Additionally, solutions to improve heterogeneity for cysteine-based coupling has been proposed, e.g. disulfide re-bridging [59], low-temperature reduction or the addition of zinc ions [56].

In clinical or commercial manufacturing, the inherent heterogeneity of stochastic conjugation approaches in particular presents challenges in ensuring batch-to-batch consistency. Therefore, significant effort is required to achieve robust and scalable processes that yield a reproducible DAR and payload distribution [60]. Due to the complexity of the DAR profile, especially for lysine conjugation, regulatory agencies demand a strategy to ensure consistency with regards to the average DAR value, as this strongly correlates with the overall DAR profile [61].

1.1.4.2 Site-directed conjugation

The second generation of ADC conjugation strategies aimed to achieve more controlled and homogeneous mixtures. One common technique is inserting additional cysteines, commonly two, into specific locations on the mAb which do not to alter the overall functionality of the mAb [50]. This site-directed approach, which is schematically shown in Figure 1-4, right, was pioneered by Junutula et al. who showed to produce well-tolerated ADCs with a highly homogeneous DAR distribution, predominantly DAR 2 (> 92 %) [62]. The resulting ADC product exhibited a similarly efficacy compared to cysteine-based ADCs with higher drug load of 3.5, highlighting its improved therapeutic window [50]. When engineering the mAb, the location of the added cysteines is an important consideration which has been shown to have a large impact on the therapeutic activity [63]. Generating site-directed ADCs involves several reaction steps: a mild reduction to break existing disulfide bonds and expose free thiol groups on the engineered cysteines, a dialysis to remove reducing agents and other impurities, and a re-oxidation to reform disulfide bonds selectively [64]. This yields unpaired, reactive cysteines to be conjugated in the following conjugation reaction. One problem that may arise is disulfide bond shuffling/scrambling, where the engineered cysteines wrongly pair with native cysteines residues [50], [65]. Despite these challenges, the success in clinical studies led to several approvals of site-directed ADCs, showcasing their potential in producing safer and better tolerated ADCs. Other noteworthy strategies include site-directed conjugation using unnatural amino acids, enzymatic-assisted ligation, and glycol-conjugation [49].

1.1.5 Conjugation process

The ADC manufacturing process belongs to downstream processing (DSP). In the multi-step conjugation workflow, the number of reaction steps varies

depending on the selected conjugation method. A schematical overview about the entire workflow is shown in Figure 1-5. The process begins with a purified mAb and contains a series of biochemical reaction steps to chemically modify the mAb creating the active bindings sites followed by the conjugation of the payload. This results in an intermediate drug product with a certain Drug-to-Antibody-Ratio (DAR) and distribution of DAR species, also called drug load distribution (DLD). Subsequently, additional purification steps are typically conducted, specifically tangential flow filtration to remove organic solvents and reaction agents, and hydrophobic interaction chromatography (HIC) to remove under- or overconjugated species, thereby ensuring tightly control over the final DAR distribution [51]. In this workflow, the conjugation reaction steps are of core interest, as they dictate the properties of the final ADC product, namely the DAR and the DLD.
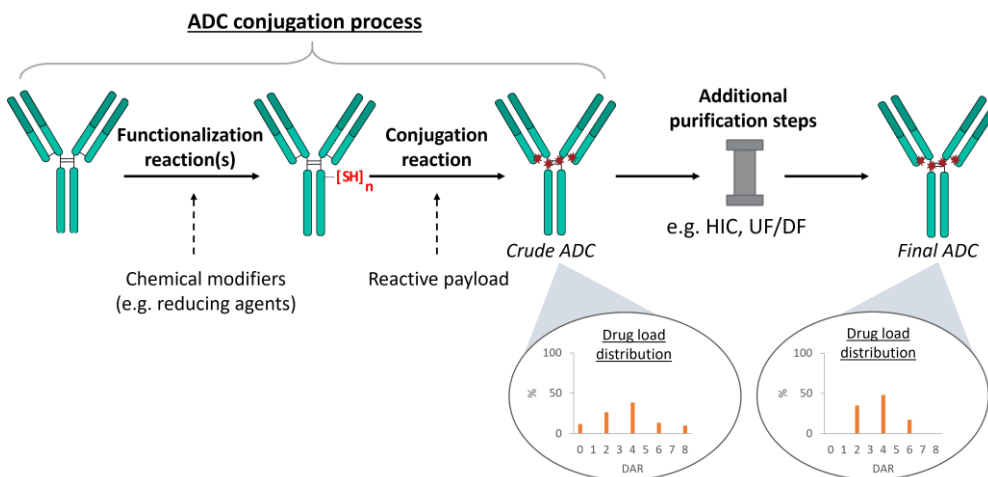
Figure 1-5 Overview of the process steps in the ADC conjugation step and the following purification steps. Exemplary drug load distributions are shown for interchain cysteine-linked ADCs.

1.1.6 Critical quality attributes and analytics

Owing to the combination of large biologics with a small-molecule drugs, various CQAs exist, of which some are ADC-specific, whereas others are related to one of the components. The most important CQAs are the DAR, which quantifies the average number of drugs conjugated to a single mAb molecule, and the DLD [66]. Both directly affect the safety, efficacy, and stability of the ADC product and must be therefore tightly controlled by the manufacturing process and post-manufacturing measures. Other important CQAs are the concentrations of unconjugated mAb and of residual payload, size variants, such as fragments or aggregates, and charge variants [67].

Various analytical methods are utilized in ADC characterization, mainly based on spectroscopy, liquid chromatography (LC) and mass spectrometry (MS),

which may vary, depending on the conjugation chemistry. The DAR can be determined using simple, non-invasive ultraviolet/visible (UV/Vis) spectroscopy, if both mAb and drug feature distinct and non-overlapping absorbance maxima within the UV/Vis range [67]. Figure 1-6 presents exemplary spectra of a pseudo drug, a mAb and the resulting ADC. The distinct absorption spectra of the individual components, mAb and the drug, are clearly emphasized. In contrast, the ADC spectrum exhibits a mixed spectrum, combining features of both the mAb and the drug components. Furthermore, a slight red shift is visible at the two peak maxima at approx. 325 nm and 340 nm, which can be observed when comparing ADC and drug spectra.
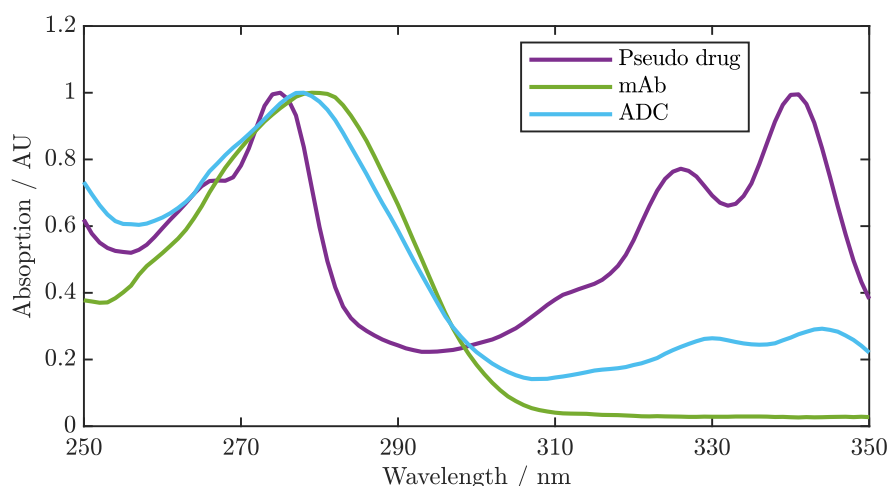
Figure 1-6: Exemplary UV/Vis spectra (normalized) of a pseudo drug, unconjugated mAb and ADC.

Commonly, HIC or revered-phase high-performance liquid chromatography (RP-HPLC) is used to determine the DAR and the DLD due to the different hydrophobicity of the ADC species with a different drug load. Whilst the separation in HIC is performed under mild conditions preserving the native ADC structure, RP requires a reducing treatment dissociating the ADC in its subunits, heavy and light chains [68]. RP has the advantage of being able to additionally quantify residual drug concentration and is applicable with subsequent MS methods [66]. Both techniques are readily applied for site-directed or cysteine-based ADCs due to their lower heterogeneity [67]. The isomeric populations for stochastic ADCs can be resolved utilizing combinatory approaches of HIC and RP or capillary electrophoresis sodium dodecyl sulfate (CE-SDS) [69], [70]. Other orthogonal, MS-based techniques, such as electrospray ionization-MS or MALDI-TOF, are more frequently used for lysine conjugates due to their increased heterogeneity which requires better-resolving techniques [67]. The mentioned analytical technologies allow for additional quantification of the unconjugated mAb. Moreover, size exclusion chromatography or CE-SDS are typically applied for the determination of size

variants, and ion exchange chromatography or isoelectric focusing is used for charge variants.

1.2 Quality by Design

The guideline Q8(R2) of the International Council for Harmonization emphasizes that pharmaceutical development should aim to design a quality product and robust manufacturing process that consistently meets the indented performance, rather than relying on end-product testing. Pharmaceutical companies should provide a comprehensive understanding of both the product and the manufacturing process to regulatory bodies, promoting flexible regulatory approaches based on the level of scientific knowledge. One important concept is QbD which involves defining a Quality Target Product Profile (QTPP) and identifying CQAs that impact the product performance. It is anticipated to apply risk assessment tools to identify the impact of material attributes and critical process parameters (CPP) on the CQA, aiming to prioritize and address critical areas in the manufacturing process. Within this context, it is promoted to use statistical tools, such as design of experiment (DoE), and/or mathematical models, to gain in-depth (mechanistic) insights leading to a higher level of process understanding. Another key aspect is the definition of the design space that describes the multidimensional combination of material attributes and process parameters to ensure consistent CQAs for a single step or spanning over multiple unit operations. Additionally, scaling factors may be incorporated to describe the design space' behavior across different manufacturing scales. [71]

1.3 Mathematical models for chemical processes

Process modeling is a discipline of process system engineering, providing a systematic framework for understanding, analyzing, and optimizing chemical processes. A process model usually encompasses a variety of interconnected sub-models describing different occurring phenomena using mass and energy balances. These, so-called mechanistic models, describe the physical, chemical, and biological phenomena, such as mass flows, thermodynamics and reaction kinetics occurring in a process, through mathematical representations. Apart from mechanistic models, also purely empirical or statistical models, also called black-box models, may be applied.

A process reaction model can be seen as a combination of a kinetic model, which describes the (bio)chemical reactions, with a reactor model, representing the physical phenomena such as flow conditions and mass transfer inside a reactor.

Since the ADC conjugation steps are reactions in the aqueous phase and typically conducted in continuously mixed/stirred reaction vessels of different sizes, the following sections focus on mechanistic modeling of homogeneous (liquid-phase) reactions in stirred vessels.

1.3.1 Kinetic modeling of homogeneous reactions

Kinetic modeling aims to elucidate the correct reaction mechanism by analyzing the time-course of a chemical reaction. Possible kinetic models are established by formulating ordinary differential equations (ODEs) based on kinetic laws and estimating kinetic rate constants using experimental concentration time-course data. For the sake of simplicity, a constant reaction volume V is assumed to derive the kinetic equations. For a simple, homogeneous reaction, e.g., $v_1 A + v_2 B \rightarrow v_3 R + v_4 S$, the rate of reaction of a species, r_i, can be written as [72]:

$$r_i = -\frac{1}{v_i V}\frac{dn_i}{dt} = -\frac{1}{v_i}\frac{dc_i}{dt} \tag{1}$$

where v_i is stochiometric coefficient, n is the number of moles and c is the molar concentration. The identification of all reactions in a composition reaction together with the stochiometric descriptions of all species yields the species balances. The rate of reaction can then be expressed according to the kinetic rate law:

$$r_i = k_i \prod_{i=1}^{n_r} c_i^{\alpha_i} \tag{2}$$

where k is the rate constant, α is the order of reaction, and n_r is the number of reactants. The reaction order can theoretically be any number, while first-, second- or thirst-order reactions are most common. The rate constant k depend on various operating conditions, such as temperature, pressure, pH or on the presence of a catalyst [73]. The temperature dependence of the rate constant k can be described by the popular Arrhenius equation:

$$k = Ae^{-\frac{E_a}{RT}} \tag{3}$$

where A is the pre-exponential factor, E_a is the activation energy, R is the universal gas constant, and T is the reaction temperature.

1.3.1.1 Parameter estimation

Once the kinetic model is established, the next step is to estimate the model parameters $\boldsymbol{\theta} = (k_i, E_{a,i}, \dots)$ based on statistical inference methods and experimental data. The inference of $\boldsymbol{\theta}$ is called an inverse problem which aims to minimize the residual difference ε between data (= observations) y and

model predictions $\hat{y} = g(t, \boldsymbol{\theta})$. In reaction kinetics, an observation usually represents the measured concentration of a species i at a discrete time point, therefore $y(t_k) = c_i(t_K)$. For each observation j, the residual is represented by:

$$\varepsilon_{ij} = y_{ij} - g(t, \boldsymbol{\theta}) \tag{4}$$

Different estimation methods, such as maximum likelihood estimation (MLE) or Bayesian estimation exist. In last squares (LS) estimation, which is a special case of MLE, the objective is to minimize the residual sum of the squared errors [74] to obtain a point estimate of $\boldsymbol{\theta}$ according to:

$$RSS(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(y_{ij} - g(t, \boldsymbol{\theta}) \right)^2 \tag{5}$$

with

$$\boldsymbol{\theta}_{LS} = \underset{\boldsymbol{\theta}}{\arg\min} \, RSS(\boldsymbol{\theta}) \, . \tag{6}$$

A variety of minimization algorithms, e.g., Newton methods, gradient descent or Nelder-Mead simplex, are available to solve the objective function [75]. Although in LS approach the inherent assumption is that the model structure is appropriately defined, in practice, the model structure might be unknown.

1.3.1.2 Assessment of model performance

If a kinetic model needs to be formulated, often different possible assumptions result in various model candidates. Here, different methodologies have been proposed to aid in selecting a suitable model among the model candidates, such as analyzing the estimation quality of the parameters, including parameter confidence, identifiability and sensitivity, and various error metrics. In the frequentist paradigm, it is common to quantify the certainty of the estimated parameters with the $100(1 - \alpha)\%$ confidence interval. This requires calculating the Jacobian matrix $\mathbf{J}(\boldsymbol{\theta})$ and estimating the covariance matrix of the estimated parameter vector σ^2 according to [76]:

$$cov(\boldsymbol{\theta}) = \sigma^2 \left(\mathbf{J}(\boldsymbol{\theta})^{\,T} \mathbf{J}(\boldsymbol{\theta}) \right)^{-1} \tag{7}$$

where σ^2 can be calculated from the LS estimation:

$$\sigma^2 = \frac{RSS(\boldsymbol{\theta})}{n - p} \tag{8}$$

where n indicates the number of samples and p is the number of model parameters. The confidence interval can be then calculated from the percentiles of the t-distribution (or normal distribution) with:

$$\theta_{1-\alpha} = \theta \pm t_{\alpha/2,n-p} \cdot \sqrt{cov(\theta)_{ii}} \tag{9}$$

where $t_{\alpha/2,n-p}$ is the critical value from the t-distribution with $n-p$ degrees of freedom. Usually, the 95% confidence interval for $\boldsymbol{\theta}$ is calculated, which signifies that this interval is 95% likely to contain the true effect size or true mean [77]. Another approach to estimate the parameter quality is to calculate parameter identifiability or the Fisher Information Matrix, which assesses the information content in the data regarding the parameters and indicates the precision of the parameter estimates. Alternatively, one can numerically estimate the parameter uncertainty by bootstrapping based on Monte Carlo sampling [78].

However, these methods only indicate the quality of the estimated parameters and using them directly for model comparison may sometimes be difficult. A common criterion for comparison of candidate models is the Akaike information criterion (AIC) which assesses the balance between model complexity and model accuracy. For small datasets, the corrected Akaike information criterion (AIC_c) is a more accurate representation, which is defined as [79]:

$$\text{AIC}_\text{c} = n \cdot \ln\left(\frac{SSE}{n}\right) + 2p + \frac{2p(p+1)}{n-p-2} + n \cdot \ln(2\pi) + n. \tag{10}$$

With regards to error metrics, the model error is commonly quantified using the root mean squared error, $RMSE$, for each species which can be calculated with:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{7}$$

where n is the number of data points. Another common metric is the coefficient of determination, R^2, which is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{8}$$

Here $\bar{y}$ denotes the mean of the measured values. To evaluate the model performance on unseen data and thus assess the generalizability of model, datasets are routinely split into training and test set, while the model is calibrated on the training data and then evaluated on the test data. Another common technique for model validation is cross-validation.

1.3.2 Computational fluid dynamics for stirred rank reactors

The kinetic models introduced in the previous section assumed the reaction volume to be perfectly mixed. However, real reactors, especially large-scale

reactors, often exhibit deviations from ideal mixing due to various factors such as incomplete mixing, flow patterns or turbulence. These become increasingly important for fast, mixing-sensitive reactions. Therefore, understanding the influence of mixing on the CQAs might be important to understand at all scales. Computational fluid dynamics (CFD) offers a powerful tool for modeling the hydrodynamics in mixed vessels. In essence, CFD simulations approximate the fluid dynamic conditions by solving a set of governing equations. This includes the continuity equation (in differential form) [86]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = S_c \tag{9}$$

where ρ stands for the fluid density, $\mathbf{v}$ denotes the velocity vector and S_c is a general source term, i.e., mass that is added (or removed) from the continuous phase, and the momentum equation, also known as Navier-Stokes equation:

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) = -\nabla p + \nabla \boldsymbol{\tau} + \rho \mathbf{g} \tag{10}$$

where p is the fluid pressure, $\boldsymbol{\tau}$ denotes the viscous stress tensor, $\mathbf{g}$ is the gravitational acceleration. To incorporate chemical reactions in the aqueous phase, the continuity equation is separately formulated for each species and homogeneous reactions can be added via the source term S_c. Various numerical methods are used to solve the system of partial differential equations for given boundary and initial conditions. This requires to divide the spatial domain into discrete elements in which the governing equations are solved, namely finite volume, finite elements, or finite difference methods. As depicted in Figure 1-7, setting up a CFD simulation for a stirred tank requires several additional steps, including the design of computer-aided design (CAD) model, the generation of an appropriate mesh, selection of solver settings, and proper post-processing to judge the goodness of the converged results.

CAD model
- Single-phase / Multiphase
- Impeller geometry

Meshing
- Grid structure
- Mesh refinement
- Mesh quality

Solving
- Steady state / transient
- Impeller modeling
- Turbulence models
- Discretization methods

Post-Processing
- Judging convergence
- Visualization
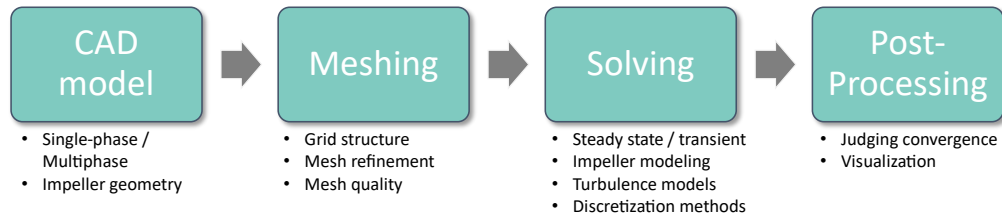
Figure 1-7: Key steps and considerations involved in for setting up a CFD simulation for a stirred tank.

Specifically designed for modeling of stirred tanks, Multiple Reference Frame (MRF) approaches are normally used to effectively simulate the impeller motion and its impact of the surrounding fluid. Since the flow in stirred tank reactors is usually turbulent or transient, turbulence models are necessary. Depending on the requirement to accurately resolve the turbulence flow, different

turbulence models are available, with Reynolds Averaged Navier Stokes (RANS) models, especially the k-ϵ-model, being widely applied due to their computational efficiency and relatively high accuracy [87], [88]. A range of open source and commercial software packages are available, such as OpenFOAM, Ansys Fluent, and Comsol, to perform CFD simulations.

In terms of biopharmaceutical applications, CFD modeling has been widely applied for modeling large-scale bioprocesses and fermentations [80]–[82], mixing in ultrafiltration/diafiltration tanks [83], low-pH virial inactivation [84] or mixing of mAb formulations [85].

1.4 Process analytical technology

In pharmaceutical production, establishing a robust process can be highly challenging as slight changes in the process might have a strong impact on the target product quality [90]. Therefore, it is desirable to implement process monitoring tools allowing to control the process, which aims to maintain process robustness. The term PAT originally stems from process analysis, a field in many sectors in industrial manufacturing, which involves real-time process analytical tools and chemometrics to monitor chemical or physical attributes [89]. This involves utilizing a wider range of analytical devices to conduct measurements throughout the process. With regards to biopharmaceutical processes, this may be challenging due to the complex structure of biologics, consisting of many sub-species, which can exhibit close structural similarity.

The PAT framework, initiated by the FDA in 2004, describes a systematic approach for "designing, analyzing and controlling manufacturing processes through timely measurements of CQAs", ensuring that the final product consistently reaches its target specification [91]. Consistent with the QbD-framework, one major goal is facilitating better process understanding as PAT enables the real-time monitoring and control of CQAs. This has the potential to improve yields, reduce waster and energy consumption and may ultimately enable real-time batch release [92]. One key pillar of PAT is to implement advanced analytical tools, so-called process analyzers, which perform measurements on-/in-/at-line providing rapid results and do not require manual handling. As opposed to conventional wet lab sample analysis, this eventually decreases the boundaries between analytics and process [89]. Furthermore, obtaining real-time feedback about the process allows to proactively adjust process parameters to reach the target product quality. Here, optical spectroscopy has emerged as a powerful tool as it provides rapid results containing valuable information about CQA in the recorded spectra [93]. Owing

to the complexity of spectroscopic data, multivariate methodologies are required for extracting the desired information.

The subsequent sections, focus on spectroscopy as PAT tool and describe the fundamentals of chemometric modeling as well as optimal state estimators.

1.4.1 Spectroscopy in general

In biopharmaceutical DSP, HPLC is still one of the most common techniques for final product quality testing, which has the main disadvantage of being time consuming due to long measurement times [94]. As mentioned above, real-time monitoring and process control require fast measurements that allow to obtain information directly, or indirectly, about the concentration or structure of certain agents as well as other relevant process variables. Particularly, monitoring of the active pharmaceutical ingredient (API) itself, and process- or product-related impurities is of interest. Spectroscopic sensors have been widely utilized to study the properties of proteins in solution and reactions of proteins, as it provides rapid and non-invasive measurements [95]. There exists a multitude of spectroscopic methods, where each has its own capabilities of probing certain protein properties. Specifically for proteins, the governing spectroscopic effects rely on the amino acid sequence affecting the protein structure (primary, secondary, tertiary, and quaternary) which can be detected by various optical sensors, such as UV/Vis, infrared and Raman spectroscopy [96].

The analytical method, as described in Section 1.1.6, do not provide the necessary measurement speed to facilitate process monitoring or comply with the PAT framework for the ADC process. Here, the current trend is rather to achieve faster measurement times for traditional analytical methods [97], then to implement novel spectroscopic tools. Only a few studies in literature focused on PAT in the field of ADCs. Andris et al. demonstrated that a UV/Vis-based sensor in combination with a chemometric model is capable to monitor the DAR value during the conjugation reaction in real-time. [98]. Similarly, polarized excitation emission matrix spectroscopy also showed to monitor both the DAR and the aggregates level during the conjugation reaction [99]. Zhang et al. used Raman spectroscopy to quantify aggregate level and particle formation in ADC formulation [100]. These examples showcased the potential for spectroscopy as PAT tool in ADC process development and manufacturing.

1.4.2 UV/Vis spectroscopy

Especially, UV/Vis spectroscopy is well-suited for protein analysis, due to its lower water interference compared to other spectroscopic techniques [101]. Various case studies could already emphasize its utility for PAT applications in

DSP unit operations [102]. In principle, UV/Vis spectroscopy is based on the widely recognized Lambert-Beer law, which correlates the light absorption by a sample to the concentration of a specific chemical substance, e.g. proteins [89]. Mainly the aromatic amino acids contribute to the overall absorption in the mid UV range between 220 nm and 300 nm, thus allowing to probe the primary structure of the protein.

1.4.3 Multivariate data analysis

To extract and analyze actual information, e.g., about the CQAs, from the recorded spectra, chemometric tools are required. For this purpose, multivariate data analysis offers several types of mathematical regression or classification algorithms, such as Principial Component Analysis (PCA) and Partial Least Squares (PLS), to decompose the valuable data from multivariate spectra [103]. Since recorded spectra often contain noise, interference from the background or uninformative effects, additional preprocessing methods might be required to ensure robust modeling results [104]. The most common methods are baseline correction, smoothing using Savitzky-Golay filtering, derivative taking or background subtraction. However, the preprocessing methods must be appropriately selected, as each combination of spectroscopy and use case presents its unique spectral effects [101]. When multiple spectra are mutually analyzed, further pretreatment including scaling, centering, or variable transformation may be necessary.

Regression modeling is used to build quantitative, chemometric models projecting the recorded spectra to reference variables. In other words, the chemometric model aims to regress multivariate predictors variables $\mathbf{X}$, containing the spectral values as columns and the samples as rows, to response variables $\mathbf{Y}$, containing the values of the reference analytics as columns and the samples as rows. In the simple case, when the variables in $\mathbf{X}$ are not intercorrelated, multiple linear regression (MLR), which is an extension of linear regression, can be applied. A MLR model is given with:

$$\mathbf{Y} = \boldsymbol{\beta}\mathbf{X} + \mathbf{E} \tag{12}$$

where $\boldsymbol{\beta}$ is the vector of regression coefficients, that needs to be determined using least-squares methods, and $\mathbf{E}$ is the error term. However, in chemometric modeling, multicollinearity commonly occurs, and the number of samples may be too small to determine the regression coefficients uniquely. PLS regression addresses these limitations by projecting the predictors and the response into a lower-dimensional space through latent variables. This is done by decomposing $\mathbf{X}$ and $\mathbf{Y}$ according to:

$$\mathbf{X} = \mathbf{TP}^{T} + E_X \tag{13}$$

and

$$\mathbf{Y} = \mathbf{U}\boldsymbol{Q}^T + \boldsymbol{E}_Y \tag{14}$$

where $\mathbf{T}$ and $\mathbf{U}$ are both the scores matrix, and $\mathbf{P}$ and $\mathbf{Q}$ are both the loadings matrix of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Within the PLS algorithm, the goal is to extract the latent variables which explain the maximum covariance between the predictors and responses. Due to the relatively good interpretability of PLS models, its robustness against noise as well as its computational efficiency, PLS is often applied for chemometric modeling. Nevertheless, one main problem lies in the assumption of a linear relationship between predictor and response. Since spectroscopic data may exhibit nonlinearities this sometimes can lead to poor approximation of PLS models [89].

Gaussian Process Regression (GPR) is a powerful Machine Learning (ML) method which is well-suited for handling nonlinearities enabling its application for chemometric modeling. In brief, a Gaussian Process (GP) is defined as a collection of random variables, any finite number of which have a joint Gaussian distribution [105], [106]. In GPR, it is assumed that the function $f(x)$, which maps inputs x (spectral data) to the outputs y (reference analytics), is a realization of a GP. The GP is characterized by a mean function $m(x)$ and a covariance function (kernel) $k(x, x')$ and is defined as:

$$f(x) \sim \mathcal{GP}\big(m(x), k(x, x')\big). \tag{15}$$

Commonly, either a Radial Basis Function (RBF) kernel, linear kernel, or polynomial kernel is employed. By choosing an appropriate kernel, GPRs can adapt to different types of data and relationships, making it highly versatile and applicable for chemometric modeling, which has been shown in various publications [107]. Additionally, GPR models provide not only single predictions but also estimates of uncertainty, which is valuable for quantifying the reliability of the predictions. Alternative ML methods which are of interest due to their ability in handling nonlinearities of spectroscopic data are, e.g., Support Vector Machine Regression (SVR) or Artificial Neural Networks (ANN) [108].

1.4.4 Nonlinear state estimation

Process monitoring plays a crucial role in the successful implementation of model-based control methods. This relies on the availability of accurate real-time information about the state of a dynamic system providing input for feedback control systems and detect process deviations. In practice, instruments for real-time measurements of process variables, especially CQAs, are typically sparse [109]. This problem can be addressed through optimal state estimators

which use the available sensors that only monitor part of the system variables to update process models and thus enable monitoring of unmeasurable species. This technique is also referred as model-based soft-sensor, as a physical sensor is combined with a mechanistic process model implemented as virtual software [110].

In the state-space formalism, the systems' dynamic is typically described by a continuous-time, nonlinear process model f defined as:

$$\frac{dx}{dt} = f[x(t), u(t), \theta] + w(t) \tag{16}$$

where $x(t)$ denotes the system state, $u(t)$ is the control inputs, $w(t)$ corresponds to the process noise. Additionally, the measurement model h represents the available sensor measurements at the discrete time points t_k given with

$$y_k = h[x(t_k)] + r(t_k) \tag{17}$$

where $r(t_k)$ denotes the measurement/sensor noise. In general, optimal state estimation is a method to determine the most accurate estimate of the "true" state (e.g. concentration of a certain species within a batch reactor) of a dynamic system from noisy and incomplete observations. This requires the use of statistical techniques or mathematical algorithms to combine observations from one or multiple sensors with a predictive model of the system. As the goal is to minimize the estimation error, the estimator must also account for the individual model and measurement uncertainties. Figure 1-8 presents the required physical/virtual components as well as the fundamental state estimation procedure consisting of two-stage: first, the model prediction provides the *a priori* state estimate and second, the state estimator corrects this estimate by taking account the measurement update providing the *a posteriori* estimate. Recursive estimation also involves using past estimates within this procedure.
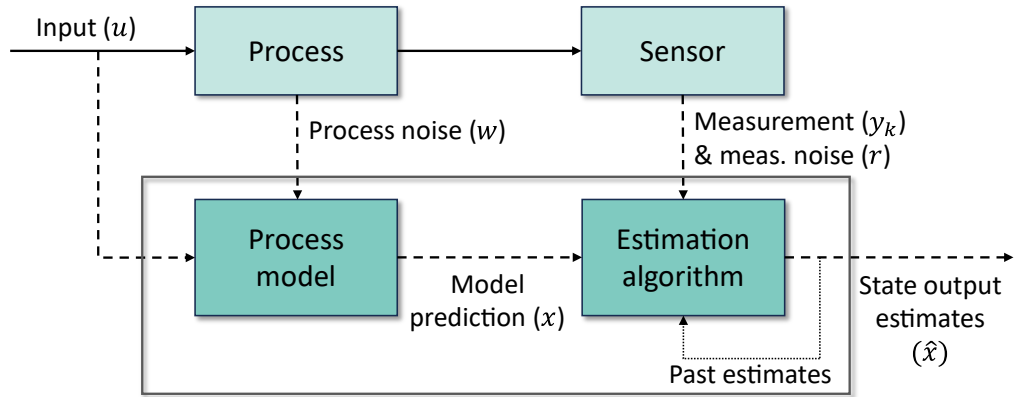
Figure 1-8: Schematic representation of optimal state estimation. The solid lines indicate systems connection, the dashed and dotted lines correspond to information signals. Adopted from [111].

Various recursive state estimation algorithms are available, specifically designed to handle nonlinear systems. Notable examples include the Extended Kalman Filter (EKF), and Unscented Kalman Filter (UKF), which are readily applied in advanced monitoring of chemical reaction [112] and bioprocesses [113], [114]. These algorithms propagate the mean and the covariance matrices stepwise, differing in how they approximate non-linear models. A critical aspect of implementing a reliable state estimator is the estimation of noise statistics, which are often unknown and can be determined e.g. using Monte Carlo-based sampling methods [115]. To accommodate additional hard constraints (e.g. non-negative concentrations in (bio)chemical reactions) other estimators such as particle filters can be used. Moreover, for correcting long-horizon forecasts, alternative optimization-based algorithms like moving horizon estimation may be suitable [111].

.

2

Thesis outline

2.1 Research proposal

Antibody–drug conjugates (ADCs) have emerged as a promising class of anticancer biopharmaceuticals by combining the target specificity of monoclonal antibodies (mAbs) with the potent cytotoxic effects of antitumor agents that are too toxic to be used alone. Especially in the treatment of cancer, ADCs offer great potential over conventional chemotherapeutics, due to higher doses of cytotoxic agent reaching the target cells and the mAb itself triggering an immune response. Over the past decades, research effort has intensified to reduce toxicity and improve efficacy, leading to fourteen approved ADCs and hundreds in clinical trials, as of 2024. In early development of ADC manufacturing processes, reaction understanding of the conjugation workflow is crucial to adjust final product properties. This is challenging due the hybrid nature of ADCs which demands the combination of conventional organic synthesis with biotechnological manufacturing. Furthermore, an increasing number of recently developed ADCs are utilizing novel linkers, payloads, and new conjugation technologies. While some of these, so-called, next-generation ADCs improved the product heterogeneity and efficacy, they consist of individual reaction sequences resulting in a growing complexity for early process development. Upon proceeding in clinical trials, the main objective is to establish robust and scalable processes which requires to systematically understand the relationship between process parameters and final ADC attributes. Despite the recent growth of ADCs, specific regulatory guidelines, best practices, or platform processes are still missing. Current industrial ADC process development heavily relies on empirical approaches, such as Design of Experiment (DoE), to gain the necessary process and product understanding.

Today's biopharmaceutical industry is confronted with increasing R&D costs, the need for faster time-to-market, and the demand for consistently delivering high-quality products. The latter is especially enforced by regulatory authorities, such as the U.S. Food and Drug Administration (FDA) and European Medicines Agency (EMA), through guidelines like Quality by Design (QbD) and Process Analytical Technology (PAT). These guidelines emphasize the importance of a profound, science-based understanding of the relationship between critical process parameters (CPP) and critical quality attributes (CQA) to develop robust manufacturing processes. To meet these requirements, modeling and model-based methods are supportive and can assist in process digitalization which ultimately aims to accelerate process development for new therapeutics. These approaches apply computational methods to transform experimental data into valuable mechanistic understanding. Additionally, PAT tools, which combine sensor data with chemometric modeling, can be employed to monitor various properties of the process in real-time, providing data to support advanced model-based concepts for process monitoring, control, or automation. Digital process models, grouped into data-driven, mechanistic or hybrid models, are regarded a transformative technology for bioprocess development. In principle, these models try to virtually represent the real process, which makes them applicable to characterize, design, and monitor processes, as well as predict scale-up or process transfer. Once established, they offer great advantages, such as reduced experimental effort, enhanced process understanding, and improved optimization of production conditions. Despite their demonstrated value in numerous use cases, process models are not widely applied in process development and manufacturing for biopharmaceuticals, because of the intrinsic complexity of bioprocesses and the high effort to decipher and quantify underlying physical or chemical processes and phenomena. For this purpose, classical Design of Experiment (DoE) approaches are still routinely applied. This statistical method is relatively versatile and easy to implement. However, the needed amount of experiments grows exponentially with the number of interacting variables, the gained insight is limited and DoEs are only valid within the design space.

In light of the above-mentioned regulatory concepts, this research aims to contribute to the growing demand for modeling approaches, facilitating process understanding, prediction, and robustness, in the process development and manufacturing of ADCs. Since the involved biochemical reactions are performed in batch mode, meaning that individual components are mixed over a fixed time, kinetic models are essential for reaction understanding as they describe the concentration of the reactants vs. time. First, the focus will be on modeling the reactions in the conjugation process, namely, reduction and conjugation reaction, representing core steps for both site-directed and stochastic ADCs. To

fully exploit the potential of these models within a QbD context, they must be integrated with other techniques. Therefore, the second part of this research will explore model-based approaches to tackle scalability issues and advance real-time monitoring, both exemplarily for the site-directed conjugation reaction.

Before this work, a kinetic model for the site-specific conjugation reaction between the two cysteines on an engineered mAb and a surrogate payload has been successfully established, demonstrating to enhance mechanistic reaction understanding. Despite these advantages, the adoption of this method to the expanding diversity of ADCs is lacking as the model was built utilizing a customized analytical method. It would be advantageous to employ a common analytical platform to enable model building which can be tailored to compare conjugation chemistries with minimal effort. This will offer a deeper and more comparative process understanding in line with the mentioned QbD principles. Another unanswered question is the comparability between surrogate and real payloads regarding their stability and conjugation rates. Consequently, the aim of the first study is to develop kinetic models for both site-directed (DAR 2) and interchain disulfide conjugation (DAR 8), as they constitute a significant portion among the current conjugation chemistries. Both approaches are closely related, as they are based on conjugation to cysteines, but differ in their complexity due to the varying number and location of the binding sites. Kinetic samples will be analyzed based on a widely applicable reducing reversed phase high-performance liquid chromatography (RP-HPLC) method. The models shall also account for the occurrence of the unwanted over- and under-conjugated species. Upon calibration, the final kinetic models may provide quantitative estimates of payload-specific conjugation rates, offering a new understanding of the governing reaction mechanisms, so that e.g., the utilized payload can be minimized.

In a broader sense, integrated process understanding is highly desirable, as process parameters in a single unit operation or reaction step being part of a series of process steps can substantially impact the final product quality. In the case of interchain disulfide ADCs, the conditions in the partial reduction step essentially determine the final conjugation outcome, namely the DAR and the DLD, after the conjugation reaction, by reducing a certain number of the disulfide bridges on the mAb thereby liberating free binding sites. However, the effect of reaction parameters on both the reduction kinetic and the conjugation outcome remains mechanistically little understood. For example, recent literature findings indicate that the species distribution can be controlled via the reduction temperature. Hence, the goal of the second story is to develop a mechanistic model for the reduction kinetic which also forecasts of the final DAR/DLD. The kinetic data will be analyzed with capillary gel electrophoresis

(CGE), as this method measures the generated mAb fragments when disulfide bonds are reduced. Given the temperature sensitivity of the reduction reaction, particular emphasis should be given to the parametrization of this phenomenon. Since the conjugation reaction was previously determined to be several timescales faster, a secondary model shall directly predict the final conjugation results from each reduction state without explicitly accounting for the conjugation kinetic. In combination with the reduction model, this approach may allow for a comprehensive analysis of reduction reaction parameters, both in terms of the reduction reaction itself and the final conjugation outcome.

Mechanistic kinetic models have already proven to be valuable in the characterization and design of ADC conjugation reactions. However, experiments to calibrate these models are typically performed using small-scale equipment, which does not allow to simulate large-scale processes. Large-scale ADC reactions are typically conducted in stirred vessels, which are difficult to probe during the reaction due to good manufacturing practice (GMP) regulations. A common approach is to design and validate a proper scale-down model based on similarity in final product quality, which does not account for the reaction dynamics. Additionally, the influence of the dynamic flow behavior on the reaction kinetic remains unknown. To overcome these limitations, computational fluid dynamics (CFD) is a powerful tool capable of describing flow behavior and mass transfer within user-defined geometries, offering the possibility to incorporate reaction kinetics. Therefore, the third study aims to gain kinetic insights into conjugation kinetics at different scales, exemplary for the site-directed conjugation reaction. Experimental conjugation kinetics in small- and mid-scales will be necessary to assess the effect of mixing parameters and scale directly on the DAR kinetics. To investigate kinetics at larger scales, an existing kinetic model will be incorporated in CFD models of different mid- and large-scale equipment, namely stirred tanks of various volumes and geometries. This will result in a 3D-reactor model that enables to predict the effect of parameter variations, such as stirrer speed or feeding times, entirely *in silico*, providing a versatile tool to gain in-depth insights into the scale-dependency of the conjugation dynamics.

Constructing a QbD-compliant digital twin of a biochemical reaction necessitates real-time monitoring of various process variables, a process model and the ability to update the model. Ultraviolet/visible (UV/Vis) spectroscopy combined with chemometric modeling has previously demonstrated its capability in monitoring the DAR during the site-directed conjugation reaction. A remaining issue is that the existing kinetic model of the reaction cannot be updated by real-time predictions to track conjugation intermediates. Optimal state estimators, such as the Extended Kalman Filter (EKF), enable the integration of on-line measurements with mechanistic models, allowing for

recursive updates and correction of model predictions. <u>Therefore, the goal of the fourth study is to establish a model-based soft sensor that fuses the chemometric model output with the kinetic reaction model</u>. An EKF will be deployed to update the reacting species included in the model that are not directly observable by the PAT sensor. The EKF should also adaptively account for the time-varying model and sensor uncertainty. Therefore, the original sensor model may be advanced to provide a quantitative uncertainty measure. Finally, the robustness of the developed soft-sensor framework and its transferability from batch to fed-batch should be closely evaluated.

2.2 Outline and author statements

The research papers written in the scope of this thesis are briefly outlined in this section. As different reaction steps and ADC modalities were addressed, a comprehensive overview of the main topics along with the respective chapter number is given in Figure 2-1. The first two papers deal with the establishment of kinetic models, whereas the latter two papers focus on case studies for model application.
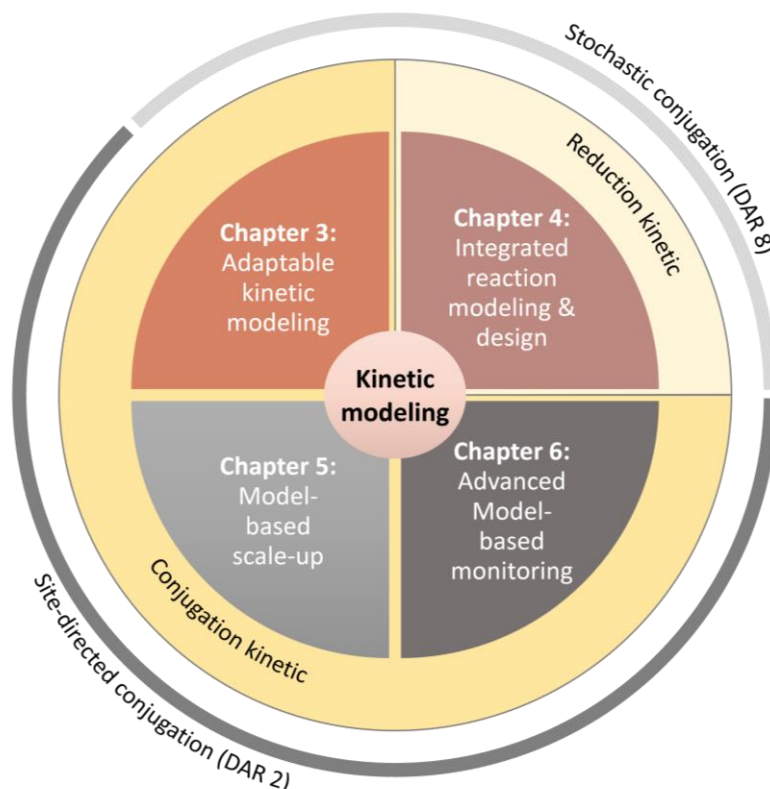

Figure 2-1: Graphical outline of the main topic of the four chapters in this thesis also depicting the reaction step and the ADC modalities of the focus of each chapter.

In Chapter 6, first authorship was shared (contributed equally) to elevate the quality of our common publication. In the following, the included research papers in this thesis are listed along with a short summary, the main findings, and the author contributions for each included paper. A detailed listing of author contributions signed by the authors can be found in the Appendix of the examination copy.

Chapter 3: Kinetic models towards an enhanced understanding of diverse ADC conjugation reactions

Jan Tobias Weggen, Ryan Bean, Kimberly Hui, Michaela Wendeler, Jürgen Hubbuch

Published in Frontiers in Bioengineering and Biotechnology (2024), Volume 12, article 1403644, https://doi.org/10.3389/fbioe.2024.1403644

In Chapter 3, a versatile modeling framework for the conjugation kinetic of different cysteine-linked ADCs, specifically site-specific (DAR 2) and interchain disulfide conjugation (DAR 8) is presented. The kinetic data, analyzed by reducing reversed-phase chromatography to measure conjugation to light and heavy chains, provided a common analytical database for both modalities. By employing various payloads, separate payload stability testing, and controlled gradual payload feeding, a detailed, comparative investigation of the reaction kinetics was possible. Several kinetic models were developed and selected based on multiple criteria, revealing a consecutive conjugation mechanism. A comparison of the payloads could quantify payload-specific conjugation rates. The calibrated model finally allowed for an *in silico* screening of the initial concentrations, offering valuable insights into the conjugation process for ADC development.

Author contributions: **Jan Tobias Weggen**: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Visualization, Writing (original draft, review and editing), **Ryan Bean**: Investigation, Data curation, Formal analysis, Writing (review and editing), **Kimberly Hui**: Investigation, Data curation, Formal analysis, Writing (review and editing), **Michaela Wendeler**: Conceptualization, Supervision, Project administration, Funding acquisition, Writing (review and editing), **Jürgen Hubbuch**: Conceptualization, Supervision, Funding acquisition, Writing (review and editing).

Chapter 4: Kinetic modeling of the antibody disulfide bond reduction reaction with integrated prediction of the drug load profile for cysteine-conjugated ADCs

Jan Tobias Weggen, Pedro González, Kimberly Hui, Ryan Bean, Michaela Wendeler, Jürgen Hubbuch

Submitted to Biotechnology and Bioengineering

The focus in this paper is on enhancing the mechanistic understanding of the reduction reaction kinetic, a critical step for adjusting the final DAR in the production of interchain cysteine-linked ADCs. CGE and RP-HPLC were used to measure the reaction kinetics of species formed during the reduction and simultaneously derive the final DAR/DLD. Initially, a reduction kinetic model was developed capable of describing the impact of reducing agent excess and reaction temperature. The kinetic findings were evaluated against a recently established reduction kinetic model. Secondly, a cross-analytical approach using multiple linear regression (MLR) models was evaluated to transform CGE data into the corresponding DAR/DLD. These two models stacked together created an integrated modeling framework to quantify both the final DAR and the DLD upon changing reduction reaction conditions. This integrated model was finally applied for an *in silico* screening to showcase the potential for gaining comprehensive, in-depth reaction understanding.

Author contributions: **Jan Tobias Weggen**: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Visualization, Writing (original draft, review and editing), **Pedro González**: Methodology, Software, Visualization, **Kimberly Hui**: Investigation, Data curation, Formal analysis, Writing (review and editing), **Ryan Bean**: Investigation, Data curation, Formal analysis, Writing (review and editing), **Michaela Wendeler**: Conceptualization, Supervision, Project administration, Funding acquisition, Writing (review and editing), **Jürgen Hubbuch**: Conceptualization, Supervision, Funding acquisition, Writing (review and editing).

Chapter 5: Kinetic studies and CFD-based reaction modeling for insights into the scalability of ADC conjugation reactions

Jan Tobias Weggen, Janik Seidel, Ryan Bean, Michaela Wendeler, Jürgen Hubbuch

Published in Frontiers in Bioengineering and Biotechnology (2023), Volume 11, article 1123842, https://doi.org/10.3389/fbioe.2023.1123842

This study investigates the conjugation kinetics in the manufacturing of site-directed antibody-drug conjugates, focusing on the influence of scale and mixing parameters. Experimental kinetic studies and CFD simulations were used to understand these effects. It was found that using internal stirrer mixing at small scales yields more robust kinetics than external mixing devices. CFD simulations coupled with a conjugation kinetic model showed that sufficient mixing can be achieved at all scales, with deviations primarily occurring during the payload addition period. The study also demonstrated that a time-scale analysis can be alternatively used to assess mixing effects during scale-up. Overall, the findings highlighted the utility of kinetic models in enhancing process understanding without expensive large-scale experiments.

Author contributions: **Jan Tobias Weggen**: Conceptualization, Methodology, Investigation, Software, Visualization, Writing (original draft, review and editing), **Janik Seidel**: Methodology, Software, Visualization, Writing (review and editing), **Ryan Bean**: Investigation, Data curation, Writing (review and editing), **Michaela Wendeler**: Conceptualization, Supervision, Project administration, Writing (review and editing), **Jürgen Hubbuch**: Conceptualization, Supervision, Funding acquisition, Writing (review and editing).

A Corrigendum of this paper is available online, which was not included in this thesis.

Chapter 6: An adaptive soft-sensor for advanced real-time monitoring of an antibody-drug conjugation reaction

Robin Schiemer[1], Jan Tobias Weggen[1], Katrin Marianne Schmitt, Jürgen Hubbuch
([1]contributed equally)

Chapter 6 focuses on improving the real-time monitoring of a site-direct conjugation reaction by fusing spectroscopic sensor data with a kinetic reaction model. An extended Kalman filter (EKF) was used to dynamically update the kinetic model forecasts with the on-line sensor prediction, while adaptively accounting for the uncertainty of the underlying models. A Gaussian process regression model was applied to account for the time-variant sensor uncertainty. The EKF framework beneficially merged sensor predictions with kinetic model forecasts enabling the correction of unmeasurable reaction species, which also

demonstrated to be robust against noisy sensor predictions and model errors. This framework was finally transferred from batch to fed-batch modes emphasizing that it holds potential for future model-based process control and other ADC conjugation reactions.

Author contributions: **Robin Schiemer**: Conceptualization, Methodology, Supervision, Investigation, Software, Dara curation, Formal analysis, Visualization, Writing (original draft, review and editing), **Jan Tobias Weggen**: Conceptualization, Methodology, Supervision, Investigation, Software, Dara curation, Formal Analysis, Visualization, Writing (original draft, review and editing), **Katrin Marianne Schmitt**: Investigation, Data Curation, Formal analysis, Software, Writing (review and editing), **Jürgen Hubbuch**: Conceptualization, Supervision, Funding acquisition, Writing (review and editing).

3

Kinetic models towards an enhanced understanding of diverse ADC conjugation reactions

Jan Tobias Weggen[1], Ryan Bean[2], Kimberly Hui[2], Michaela Wendeler[2], Jürgen Hubbuch[1]

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

[2]Purification Process Sciences, BioPharmaceuticals Development, R&D, AstraZeneca, Gaithersburg, Maryland, USA

Abstract

The conjugation reaction is the central step in the manufacturing process of antibody-drug conjugates (ADCs). This reaction generates a heterogeneous and complex mixture of differently conjugated sub-species depending on the chosen conjugation chemistry. The parametrization of the conjugation reaction through mechanistic kinetic models offers a chance to enhance valuable reaction knowledge and ensure process robustness. This study introduces a versatile modeling framework for the conjugation reaction of cysteine-conjugated ADC modalities – site-specific and interchain disulfide conjugation. Various conjugation kinetics involving different maleimide-functionalized payloads were performed, while controlled gradual payload feeding was employed to decelerate the conjugation, facilitating a more detailed investigation of the reaction

mechanism. The kinetic data were analyzed with a reducing reversed phase (RP) chromatography method, that can readily be implemented for the accurate characterization of ADCs with diverse drug-to-antibody ratios, providing the conjugation trajectories of the single chains of the monoclonal antibody (mAb). Possible kinetic models for the conjugation mechanism were then developed and selected based on multiple criteria. When calibrating the established model to kinetics involving different payloads, conjugation rates were determined to be payload specific. Further conclusions regarding the kinetic comparability across the two modalities could also be derived. One calibrated model was used for an exemplary in-silico screening of the initial concentrations offering valuable insights for profound understanding of the conjugation process in ADC development.

3.1 Introduction

Targeted anticancer therapeutics are becoming increasingly prevalent in the field of biopharmaceuticals. One important class in this toolbox are antibody-drug conjugates (ADCs), which consist of a conventional monoclonal antibody (mAb) chemically coupled with a highly potent small-molecule (so called "drug" or "payload"). The success of an ADC depends on conjugating a specific number, typically between 2-8, of cytotoxic payload onto a mAb, determining the final ADC potency and toxicity [116]. Related to this, one facet is the choice of the conjugation strategy which controls critical quality attributes (CQA) of the final product, such as the drug-to-antibody ratio (DAR) and drug load distribution (DLD) [117], [118]. To overcome problems associated with unfavorable DAR heterogeneities and to improve the overall potency, site-specific conjugation or more hydrophilic linkers have been developed [50], [119]. On the downside, these continuous advances hinder the development of a standardized platform process, and thus increase the time and effort to develop a scalable and robust manufacturing process [53], [60].

Quality by Design (QbD) is increasingly expected by regulatory agencies, aiming to ensure consistent product quality and improve process understanding [120]. One major aspect in QbD is the utilization of modeling techniques which has been proven to be an invaluable tool for gaining insights into complex systems and optimizing manufacturing operations [121], [122]. Process modeling enables parametrization of (bio)chemical effects dominating bioprocesses and hence, understanding the impact of such effects in-silico. In essence, process models serve as a basis for digital and automation technologies that can accelerate the development of efficient and robust manufacturing processes [123].

For biochemical reactions, one particularly important aspect of process modeling is kinetic modeling, which focuses on characterizing the rates and complex mechanisms of reactions involved. In contrast to statistical approaches, such as Design of Experiment (DoE) which purely rely on the statistical relationship between input and output variables, kinetic models provide a quantitative description of the underlying reaction kinetics, elucidating the impact of various factors such as temperature, pH, reactant concentrations [124]. Previous work has demonstrated the successful establishment of kinetic models in the area of bioprocessing, e.g. for biomass conversion [125], fermentation [126] or small-molecule synthesis [127], and with a special focus on protein modification, e.g. for protein PEGylation [128], antibody oxidation [129] or antibody reduction for ADCs [130]. The challenge of creating a kinetic model usually centers around finding the correct rate laws for each individual reaction step and their corresponding rate constants. This can be cumbersome due to the presence of multiple interacting species or complex reaction networks. To find the most reliable process model among possible model candidates, different techniques have been proposed in the literature to assess the quality of the estimated parameters (e.g. parameter identifiability analysis or Fisher Information Matrix) and quantifying the output uncertainty [131]–[133].

In ADC manufacturing, the conjugation reaction represents a key step as it generates the ADC molecule [134]. Comprehensive understanding of the kinetics of this reaction step is vital for process developers, as it enables the minimization of payload usage, thereby reducing the cost-of-goods. Consequently, employing less payload also facilitates the removal of free unconjugated payload, which is crucial to minimize toxicity of the final product [135]. The currently available literature guiding process development for ADCs is scarce [136]. Typically, DoE methodologies are utilized to gain knowledge about reaction parameters, such as reactant concentrations, temperature, time or pH, affecting the conjugation process [137], which can be augmented when high-throughput screenings [138] or continuous flow reactors [139] are used to automatize experimental work. However, DoEs cannot provide a deep understanding of the molecular or chemical mechanisms that drive the reaction. Andris et al. [140] developed a mechanistic kinetic model for a site-specific conjugation reaction using a pseudo payload. The study showcases the benefit of the kinetic model as a versatile in-silico decision tool for the investigation and development of the conjugation reaction. In subsequent studies, we could show that conjugation kinetic models realize their full potential by coupling a kinetic model with computational fluid dynamics (CFD) to study large-scale conjugation reactions [141], and enable real-time monitoring of non-observable ADC species during the reaction through combination with UV/Vis spectroscopy via an extended Kalman Filter [142]. However, applications of this kinetic model are still limited due to various

reasons: 1) The reliance on an analytical technique that is performed under native (non-denaturing) conditions which is usually limited to ADCs with low DAR values; 2) a relatively small concentration range up to 2.5 g/L, whereas actual operational conditions may be higher, typically around 20 g/L; 3) the unproven transferability of the kinetic model to other payloads; and 4) the unverified application to alternative conjugation modalities such as interchain cysteine conjugation.

This study focuses on mechanistic modeling of the conjugation kinetic for two prominent cysteine-based modalities, site-specific (DAR 2) and interchain disulfide (DAR 8) conjugation, which represent a main part of current conjugation chemistries. Both modalities rely on the conjugation to reactive cysteines, but differ in the number and location of the reactive cysteines on the mAb. Batch and fed-batch conjugation kinetics are generated across a broad range of initial mAb concentrations and drug excess. Time-course samples are analyzed using a widely employed and easily implemented reducing reversed phase ultrahigh performance liquid chromatography (RP-UHPLC) method that separates the conjugated mAb into its respective subunits prior to analysis, thereby producing well-resolved chromatograms that illustrate conjugation state of heavy and light chains. Consequently, the kinetic data represents the conjugation trajectories of heavy and light chains. Due to the different quantities of reacting species in the two modalities, multiple kinetic model candidates are proposed and the optimal number of kinetic rates for each model type is chosen based on multiple criteria: parameter identifiability, parameter uncertainty and prediction errors. In addition, absorbance measurements are conducted to determine the stability of each payload in conjugation buffer. Ultimately, an in-silico screening is performed accounting for the effects of reactant concentration and exemplifying the usage for process optimization.

3.2 Materials and methods

3.2.1 Experimental conjugation kinetic studies

The kinetic datasets encompass two distinct ADC modalities. The datasets 1 and 2 use an engineered IgG1 mAb – ADC1 – with two inserted cysteines in the hinge region for a site-specific "DAR 2" conjugation to engineered cysteines. The datasets 3 and 4 use two different IgG1 mAbs – ADC2 and ADC3 – for a "DAR 8" conjugation to reduced cysteines that are usually engaged in interchain disulfide bonds. Table 3.1 summarizes molecules and conjugation conditions. Datasets 1, 2 and 4 were generated at *AstraZeneca* and dataset 3 at *Karlsruhe Institute of Technology (KIT)*. Minor differences due to different raw materials in chemicals and analytical devices are expected.

Table 3.1: Summary of experimental conditions of all conjugation runs using different ADC modalities and payloads. A detailed overview for the individual experiments is given in the Appendix A.

Data-set	ADC	Type	Pay-load	c_{mAb} / (g L^{-1})	Molar drug excess	Drug addition	No. of conditions
1	ADC1	DAR2	Drug1	1.5 – 10	1x – 8x	Batch	8
2	ADC1	DAR2	NPM	1.5 – 3	3x – 5x	Batch	4
3	ADC2	DAR8	NPM	1.5 – 3	6x – 13x	Batch/Fed-Batch	10
4	ADC3	DAR8	Drug2	1.5 & 20	11x & 14x	Batch	4

3.2.1.1 Chemicals, ADCs and functionalization steps

For DAR 2 conjugation of ADC1, the antibody was initially fully reduced through treatment with tris(2-carboxyethyl) phosphine hydrochloride (TCEP, EMD Millipore), followed by a buffer exchange using Vivaspin 20 (30 kDa MWCO, Cytiva) and a re-oxidation of the interchain disulfides with (L)-dehydroascorbic acid (DHAA, Sigma-Aldrich). Conjugation was performed using a maleimide-functionalized payload. Two payloads were compared for ADC1, a cytotoxic payload ("Drug1") and a nontoxic surrogate N-(1-pyrenyl)-maleimide (NPM, Merck KGaA). For DAR 8 conjugation of ADC2 and ADC3, a full reduction of the interchain disulfides with TCEP was performed. Conjugation for ADC2 was carried out with NPM, while for ADC3 another cytotoxic payload ("Drug2") was used. For reaction quenching, N-acetyl cysteine (NAC, Merck KGaA) was used. All payloads were dissolved in DMSO (Sigma-Aldrich). All other solutions were prepared with 20 mM sodium phosphate buffer (J.T. Baker), 1 mM EDTA (EMD Millipore), pH 7.0. For sample pretreatment for DAR analysis, samples were diluted with denaturing buffer containing guanidine HCl (Thermo Fisher), Tris (Thermo Fisher), EDTA (EMD Millipore) at pH 7.6 and reduced with dithiothreitol (DTT, Thermo Fisher). For the sample analysis, a RP-UHPLC (Agilent Technologies) with acetonitrile (VWR) and HPLC water (VWR) with 0.1 % (v/v) trifluoroacetic acid (Thermo Scientific) as mobile phases were used.

3.2.1.2 Conjugation kinetics

For site-specific DAR 2 conjugation (ADC1), the antibody was initially treated with the reducing agent TCEP, which reduces both the engineered inserted cysteine residues, as well as cysteines engaged in interchain disulfide bonds. The reducing agent was subsequently removed and the buffer was exchanged, before the antibody was mildly re-oxidized with dehydroascorbic acid which allows re-

formation of interchain disulfide bonds, leaving only the two inserted cysteines available for conjugation. DAR 2 conjugation were performed by adding either Drug1 or NPM solution to the re-oxidized mAb solution in a microcentrifuge tube at 1 mL scale. Selected conditions were performed in duplicates. For conjugation to reduced interchain disulfides, the native mAb was completely reduced. In the DAR 8 conjugations of ADC2 ten out of thirteen kinetics were performed in fed-batch mode, where drug solution was continuously added to resolve the fast time-course of the individual reacting species. These reactions were performed in centrifuge tubes at a liquid volume of 4 mL and the required volume of payload solution was constantly added with a syringe pump (Nemesys S, Cetoni GmbH) over a defined period of time (10, 20 or 30 min) to the stirred reaction solution. The remaining batch conjugations were performed at 1 mL scale. For ADC3, batch conjugation and another payload (Drug2) were used. An overview of the experimental conditions for the conjugation kinetics is given in Table 3.1. A detailed overview of all experimental conditions is given in Appendix A, Table S 1. To acquire conjugation kinetics, samples were taken at discrete time points over the course of 1 h, transferred to a microcentrifuge tube and immediately quenched by adding a 12x molar excess over payload of NAC solution to terminate the conjugation reaction.

3.2.1.3   Sample treatment and reference analytics

To determine the DAR and the DLD of each sample a reducing RP method was used. A sample treatment was conducted to reduce the conjugated ADC molecule into heavy and light chains. Sample concentrations were adjusted to 1 mg/mL using ultrapure water. Samples were then mixed with denaturing buffer and DTT solution before incubating over 30 min at 37 °C on an orbital shaker (650 rpm). For the RP-UHPLC analysis, 10 μL of sample were injected onto a BioResolve RP mAb polyphenyl column (2.1*150 mm, 2.7 μm, Waters Corporation). An identical method, including flow rate, gradient combination and buffer compositions, was used at AZ and KIT, as described previously [143]. The peak areas in the resulting chromatograms were manually determined. The molar concentrations of the corresponding conjugated and unconjugated heavy or light chains were calculated by normalizing the peak area of light/heavy chain to the respective concentration in the reaction.

3.2.1.4   Payload depletion study

As the inactivation of the used surrogate NPM was reported in a previous study [140], a depletion study for each of the tested payloads was conducted in order to determine the individual stability in conjugation buffer. For that purpose, DMSO-dissolved payload solution was added to conjugation buffer reaching a payload concentration of 0.1 mM and 10% DMSO in a 1.5 mL microcentrifuge tube. The solution was allowed to mix on a tube rotator. Samples were taken

over the course of 1 h and the absorbance was measured using the UV/Vis function on the Stunner light scattering instrument (Unchained Labs). For NPM, the course of the absorbance was directly measured in a cuvette using a spectrophotometer (Spectrostar Nano, BMG Labtech). To verify the model assumption that depleted payload becomes unreactive for conjugation, the payloads NPM and Drug1 were separately dissolved in conjugation buffer (10% DMSO, 0.1 mM) and mixed for 1 h on a tube rotator instead of using it immediately for conjugation.

3.2.2 Conjugation kinetic model development

3.2.2.1 Modeling reduced conjugation kinetics and reaction schemes

Due to the sample treatment, which reduces the mAb in each sample and consequently eliminates information about the intact ADC species, the kinetic models employed in this study describe the conjugation reaction to the individual mAb chains. For the DAR 2 mAb, only conjugation to heavy chain was observed. This is expected, as the inserted cysteine residues are located in the hinge region of the heavy chain. In contrast, for DAR 8, conjugation to both heavy and light chain was observed as expected for stochastic conjugation. Given the absence of analytical data to resolve positional isomers, iso-reactivity for different binding locations on the heavy chain was assumed. These constraints narrow the possible rate equations down to consecutive reactions schemes describing the heavy chain as a "black box" with a certain number of available binding sites and exclude parallel schemes. The primary conjugation sites for the two ADC types with an overview of the assumed kinetic models are illustrated in Figure 3-1, left.

3.2.2.1.1 Initial cysteine distribution

To align with the observed DLD at the end of the reaction, i.e. the ratio of differently conjugated heavy and light chains, an ADC-specific initial reactive cysteine distribution of the starting mAb material was set prior parameter estimation. With regards to the DAR 2 conjugation, the main product in the final DLD is H1 (heavy chain + 1 drug). To model the amounts of unconjugated and over-conjugated heavy chains, H0 and H2, respectively, we assumed an initial cysteine distribution of heavy chains with zero, one and two reactive cysteines, namely $H0_{0c}$, $H0_{1c}$ and $H0_{2c}$. Over-conjugation in DAR 2 ADCs is thought to occur on cysteines as discussed in Cao et al. [143]. On the contrary, for DAR 8, this distribution contains heavy chains with up to four and light chains with up to one reactive cysteine, while $H0_{3c}$ (heavy + 3 drug) and $L0_{1c}$ (light chain + 1 drug) are the main species. In this case the distribution is primarily controlled by the conditions in the reduction reaction, which were kept constant in this study. It is important to note, that the presence of heavy

chain with four drugs in the case of DAR 8 depicts unspecific over-conjugation that is likely to be attributed to additional "mis-alkylation" of the payload to other residues such as lysine, which was demonstrated to occur for other model proteins [144]. This mis-alkylation appears to be marginal in the range of the studied drug excess, which led to the assumption that it only occurs in addition to the primary conjugation of the three cysteines. The chemical group of the fourth binding site is unknown, but is declared as being part of the initial cysteine distribution for the sake of simplicity. Furthermore, as the distribution might slightly vary for each mAb-drug combination, it was determined specifically for each ADC. Consequently, the initial cysteine distribution was determined experimentally per ADC based on the average distribution (%H0, %H1, ...) at the final steady-state of the reaction ($t_f = 60\,\text{min}$) in experiments with sufficient drug excess to reach full conjugation of all available cysteines, i.e. drug excess > 3x for DAR 2 and drug excess > 8x for DAR 8. As an example, the percentage of $H0_{1c}$ was calculated from the averaged percentage of $H1_{tf}$ among all heavy chains. The determined distribution was then used to calculate the molar concentration of the various initial species within the starting mAb. Example distributions for DAR 2 and DAR 8 are shown in Figure 3-1.
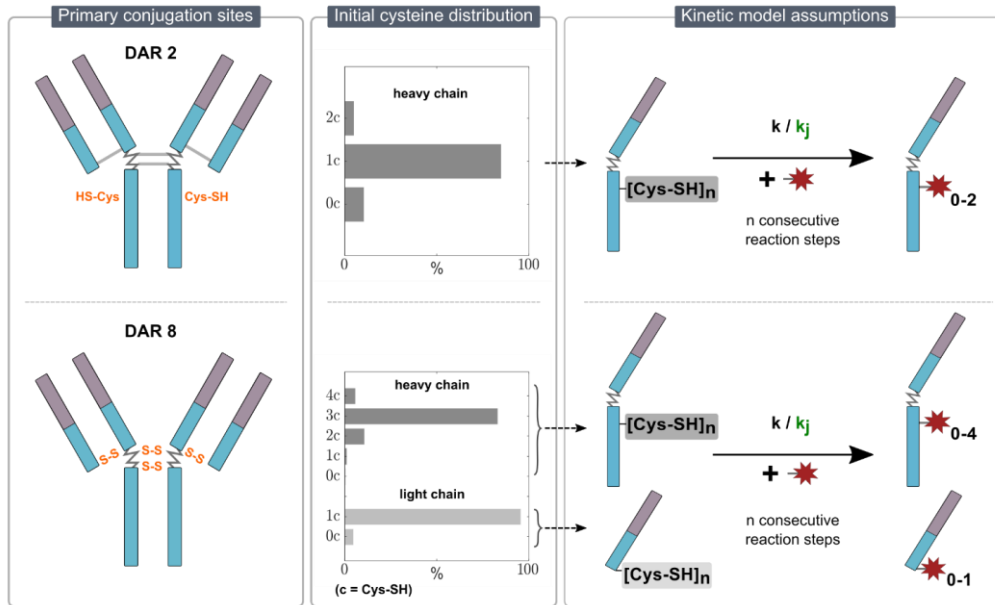
Figure 3-1: Overview of the kinetic model assumptions for DAR 2 (top row) and DAR 8 (bottom row) conjugation. The left panel illustrates the primary conjugation sites for each ADC in orange. The middle panel indicates the initial cysteine distribution for each ADC. The right panel presents the proposed stepwise conjugation reaction schemes of the payload to n reactive thiols on either light/heavy chain with either one kinetic rate (simple model) or j kinetic rates (detailed model) for each ADC.

3.2.2.1.2   Rate equations

For the conjugation reaction in a perfectly mixed system with continuous drug addition over a fixed period, the dilution rate due to added drug solution has to be considered. Therefore, the rate equation for a reaction species i can be expressed as:

$$\frac{dc_i}{dt} = r_{conj,i} - \frac{c_i(t)}{V_l(t)} q_{in} ,$$
(1)

where c_i represents the molar concentration of the ith species, $r_{\mathrm{conj},i}$ is the conjugation reaction rate of the ith species, V_l represents the reaction volume and q_{in} is the feed flow rate. In case of the free drug, the second term is changed to $+\frac{c_{\mathrm{drug,in}}}{V_l(t)} q_{\mathrm{in}}$, where $c_{\mathrm{drug,in}}$ is the drug concentration in the feed. The change in volume can be described with:

$$\frac{dV_l}{dt} = q_{in} .$$
(2)

To account for the sampling at discrete time points, the equation is integrated until each sampling time point and the volume is subtracted by the sample volume. In case of a batch conjugation reaction, Eq. (1) simplifies to

$$\frac{dc_i}{dt} = r_{conj,i} .$$
(3)

To describe the reaction rates, some assumptions were made: conjugation refers to the second-order reaction of the maleimide of a functionalized payload with the sulfhydryl residue (SH-group) of a reactive cysteine. The reaction was assumed to be irreversible as no de-conjugation was observed and temperature effects on the conjugation kinetic were neglected as the temperature did not show noticeable effects on the kinetics. Moreover, a payload-specific depletion rate k_{drug} caused by, for example, unspecific adsorption of the hydrophobic molecule to vessels wall, was considered. Based on these assumptions, a system of ordinary differential equations (ODE) describing the rate equations was formulated. Two model candidates, for both DAR 2 and DAR 8 conjugation reaction, varying in their degree of complexity were proposed: Either "simple" conjugation rates assuming one kinetic rate for all conjugation steps, or "detailed" conjugation rates assuming individual kinetic rates for each sequential reaction step. The basic reaction schemes for both DAR modalities can be found in the Appendix A, S3.2. As an example, the seven ODEs for the detailed model for the DAR 2 reaction are:

$$\frac{dc_{H0_{2c}}}{dt} = -k_1 c_{H0_{2c}} c_{drug} \tag{4}$$

$$\frac{dc_{H0_{1c}}}{dt} = -k_1 c_{H0_{1c}} c_{drug} \tag{5}$$

$$\frac{dc_{H0_{0c}}}{dt} = 0 \tag{6}$$

$$\frac{dc_{H1_{2c}}}{dt} = -k_2 c_{H1_{2c}} c_{drug} + k_1 c_{H0_{2c}} c_{drug} \tag{7}$$

$$\frac{dc_{H1_{1c}}}{dt} = k_1 c_{H0_{1c}} c_{drug} \tag{8}$$

$$\frac{dc_{H2_{2c}}}{dt} = k_2 c_{H1_{2c}} c_{drug} \tag{9}$$

$$\frac{dc_{drug}}{dt} = -k_1 c_{H0_{2c}} c_{drug} - k_1 c_{H0_{1c}} c_{drug} - k_2 c_{H1_{2c}} c_{drug} - k_{drug} c_{drug}. \tag{10}$$

Here, c_H denotes the molar concentration of the heavy chain with the number indicating the number of conjugated drugs and the index indicates the number of initial available cysteines. Furthermore, c_{drug} represents the concentration of the payload, k_j the reaction rate for the jth conjugation step and k_{drug} the depletion rate of the specific payload. The system of ODEs for the DAR 8 kinetic models can be found in the Appendix A, S3.3.

3.2.2.1.3 Payload depletion rate

Initially, the estimation of the model parameter k_{drug} led to large confidence intervals and correlation coefficients to the other conjugation rates. This could be attributed to the absence of direct kinetic data for c_{drug} in combination with the interconnected system of ODEs with regards to c_{drug} being present in all ODEs (see Eq. 4-10). Consequently, the depletion rate was considered non-identifiable given the available kinetic data and was separately determined by UV/Vis spectroscopy (cf. chapter 3.2.1.4). The UV/Vis-determined depletion rate for each payload was then set constant for the subsequent estimation of the remaining conjugation rates.

3.2.2.2 Data handling

All runs for one mAb and one payload were summarized as a kinetic dataset resulting in four distinct datasets as outlined in Table 3.1. All replicate kinetics were averaged. The dataset involving Drug1 was used as the basis for DAR 2 model candidate selection. A manual split in six training and two test runs was done. The dataset with NPM was used for external model re-calibration to compare the conjugation kinetic rates of the payloads.

In the case of DAR 8, two individual datasets were available: The first one comprised fed-batch and batch kinetics of ADC2 with the surrogate drug NPM and was utilized for DAR 8 model candidate selection. Similarly, a manual split in ten training and three test runs was performed. The second external dataset,

containing batch kinetics of ADC3 and Drug2, was later used for the in-silico screening. A detailed overview of all runs and the train/test split is provided in the Appendix A, Table S 1.

As explained above, a sufficient drug excess, such as 3x for DAR 2 and 11x for DAR 8, resulted in a saturation of both the DAR and the DLD at the steady state of the conjugation reaction. Lowering the drug excess led to lower DAR value and changing DLD. The saturated DAR and DLD is assumed as being unique to every mAb and payload. With regards to the modeling assumptions, the average DLD of kinetics with sufficiently large drug excesses was considered to compute the initial cysteine distribution for each individual kinetic subset.

3.2.2.3 Parameter estimation

All simulations were performed in Matlab R2023a. Maximum likelihood estimation of the kinetic rates in each model was conducted by minimizing the squared error between model predicted and experimental concentrations using the in-built *lsqnonlin* function. To account evenly for the entire concentration range, each run was normalized using a scaling factor corresponding to the maximum concentration in the respective run. Model predictions were performed based on the initial concentrations of the starting mAb, the initial cysteine distribution and payload concentration. The ODE system is numerically solved using the *ode15s* solver.

3.2.2.4 Model candidate selection criteria

The selection of the most appropriate kinetic model candidate was conducted based on various metrics, namely the quality of the estimated parameters, parameter identifiability ranking and model errors regarding the cross-validation of the training data and the test set.

3.2.2.4.1 Uncertainty of the estimated parameters

The uncertainty of the estimated parameters was evaluated based on their statistical uncertainty. First, the parameter covariance matrix $\mathbf{cov}(\hat{\boldsymbol{\theta}})$ was calculated using the Jacobian matrix $\mathbf{J}$ assuming independent measurement errors with Gaussian white noise

$$cov(\hat{\theta}) = s^2 (J' \cdot J)^{-1} \,, \tag{11}$$

with

$$s^2 = \frac{\sum_{i=1}^{N} \hat{\epsilon}^2}{n - p} \,, \tag{12}$$

where s^2 denotes the variance of the error, $\hat{\epsilon}$ is the error between predicted and measured concentrations, $\mathbf{n}$ indicates the number of samples and $\mathbf{p}$ is the

number of parameters. The parameter standard deviation $\boldsymbol{\sigma}$ is calculated from the diagonal elements of the parameter covariance matrix:

$$\sigma = \sqrt{diag\left(cov(\hat{\theta})\right)} \,. \tag{13}$$

Typically, parameter estimates can be considered reasonable, when the parameter standard deviations values is below 25% relative to the parameter estimate [75]. The confidence intervals of the parameters were estimated using *nlparci* in Matlab.

3.2.2.4.2 Parameter identifiability analysis

Local sensitivity analysis employing the one factor at a time (OAT) method was conducted in accordance to Sin et al. [131]. In this method, each model parameter, i.e. the kinetic rates, is systematically varied holding others constant, and the resulting impact on the model's output is observed. In practice, sensitivities for each model parameter are calculated using the first-order derivative of the model output with respect to the parameter. For each parameter, the derivatives are determined numerically by perturbing the parameter of interest by 10% of its nominal value. The sensitivities are then averaged over time and relative sensitivities $sr_{j,i}$ with respect to each species i and each model parameter j are calculated by accounting for the parameter nominal value $\hat{\theta}_j$:

$$sr_{j,i} = \frac{\partial y_i(t)}{\partial \theta_j} \theta_0. \tag{14}$$

By computing the square root of the mean of all sensitivities over all runs, the parameter significance values were calculated according to:

$$\delta^{msqr} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(sr_{j,i}\right)}. \tag{15}$$

Summing up all significance values returns the $\sum \delta^{msqr}$ value, allowing for a parameter importance ranking according to their impact on the model's output. Thus, this approach aids in identifying which kinetic rates are identifiable based on their impact. The parameter identifiability analysis was performed for the two detailed conjugation models using the previously calibrated kinetic rates.

3.2.2.4.3 Model error

The estimated kinetic rates were then used to predict the kinetic and the resulting model error regarding the species i was evaluated using the RMSE_i:

$$RMSE_i = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(c_i(t) - \tilde{c}_i(t)\right)^2}, \tag{16}$$

where $c_i(t)$ denotes the predicted species concentration, $\tilde{c}_i(t)$ is the measured species concentration and N indicates the number samples. Afterwards, the $RMSE_i$ values of each species were averaged to one RMSE value for each subset. To assess the capability of the model candidates to generalize and extrapolate, both the prediction error based on cross-validation of the training data and on the independent test set were conducted. Cross-validation was performed based on *leave-one-run-out* scheme using the training runs. This leads to two error metrics RMSECV and RMSEP.

3.2.2.5 In-silico screening for influence of initial mAb and payload concentrations

An in-silico screening was conducted for the calibrated DAR 8 kinetic model (ADC3 + Drug2) in order to observe the model output at variable initial concentrations of mAb and payload. The model outputs were the final DAR and the concentration of unreacted payload after a reaction time of 30 min as well as the necessary reaction time to reach the final DAR below a threshold of 1%. A systematic screening with regards to the three model outputs was done by independently varying the initial mAb concentration between 1.5 to 10 g/L and the drug excess between 5 to 14 molar excess while keeping the other constant.

3.3 Results

First, the UV/Vis measurements for the different payloads and the determination of the payload depletion rates are presented. Second, the bulk of this work demonstrates the selection of suitable conjugation kinetics models. Additionally, the established kinetic models are assessed regarding their ability to predict experimental kinetics for DAR 2 and 8 with high accuracy. An additional test study was conducted to evaluate the model's capability for in-silico screening using the DAR 8 model as example.

3.3.1 Payload depletion

The course of the normalized absorbance for the three payloads is shown in Figure 3-2. For NPM, a rapid decrease in absorbance is visible within 1 h, accompanied by larger error bars, whereas Drug1 exhibits only a minor decrease, and no remarkable change is visible for Drug2. The conjugation experiments using either NPM or Drug1 which was previously hold in conjugation buffer for 1h, showed that for NPM the achieved DAR is strongly reduced as opposed to Drug1 which still reaches a typical DAR value in the range of 1.8 – 1.9 (see Appendix A, Figure S 1). Since the exact mechanism of the payload depletion is unknown, a first-order reaction was assumed to describe

the depletion of NPM and Drug1, similar to Pfister et al. [145]. The depletion rate for NPM could be approximated by plotting the natural logarithm of the absorbance over time (cf. Appendix A, Figure S 2). Hereby, a depletion rate of $k_{\text{drug,NPM}} = 0.041\ s^{-1}$ could be derived using the slope of the linear regression curve. For Drug1 a depletion rate of $k_{\text{drug,Drug1}} = 0.001\ s^{-1}$ was assumed, as this value showed to agree within the modeling workflow. The depletion rate for Drug2 was assumed to be zero.
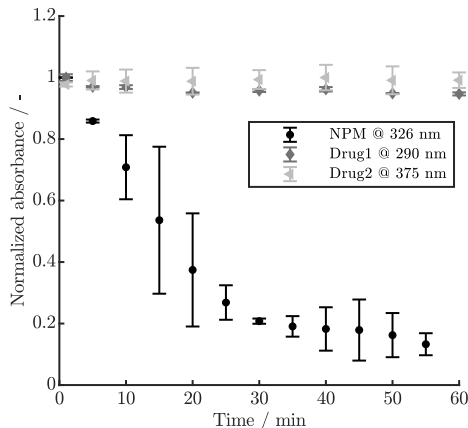

Figure 3-2: Normalized absorbance for the three payloads NPM, Drug1 and Drug2 at different wavelengths over time. The error bars indicate the standard deviation of the duplicate measurements.

3.3.2 Model complexity evaluation

Two types of kinetic models with varying complexity, namely simple and detailed, for either DAR 2 or DAR 8 conjugation kinetics were evaluated with regards to the parameter importance (assessed through the parameter identifiability analysis), parameter confidence and model error. The entire set of estimated parameters including their standard deviation and confidence intervals as well as model error are listed in Table 3.2. For DAR 2, the difference between the two models is the addition of a second rate to account for an independent second conjugation step. The rate k_1 displays small standard deviation and narrow confidence intervals for both model candidates as opposed to k_2 which exhibits large confidence intervals. At the same time, the addition of a second rate did not improve the average prediction performance considerably as can be seen from similar RMSECV and RMSEP values between the two parameter subsets. The parameter importance ranking for the detailed DAR 2 model demonstrated that the rate k_2 has almost zero effect on the model output as opposed to the rate k_1 which has remarkably higher importance as shown in Figure 3-3A. Hence for the further modeling purpose the simple (1k) model was chosen for the modeling of DAR 2 conjugation kinetic.

Table 3.2: Results of the parameter estimation for the kinetic model candidates for either DAR 2 or DAR 8 conjugation. Parameter estimation include parameter estimate ($\widehat{\boldsymbol{\theta}}$), parameter standard deviation ($\boldsymbol{\sigma}$) and lower (lb) and upper bound (ub) of the confidence intervals (all in L (mmol s)$^{-1}$ units). Model error include cross-validation (RMSECV) and test errors (RMSEP) in $\mu mol\ L^1$.

ADC	Model	Parameter estimation					Model error	
		Parameter	$\widehat{\theta}_j$	σ_j	lb	ub	RMSECV	RMSEP
DAR 2	Simple (1k)	k_1	0.251	0.023	0.204	0.298	1.863	3.188
	Detailed (2k)	k_1	0.297	0.027	0.242	0.351	1.808	3.207
		k_2	1.712	6.585	11.515	14.939		
DAR 8	Simple (1k)	k_1	1.733	0.018	1.697	1.768	1.213	1.339
	Detailed (4k)	k_1	1.220	0.003	1.214	1.227	0.620	0.964
		k_2	1.853	0.006	1.841	1.865		
		k_3	5.117	1.015	3.101	7.133		
		$k_{4/5}$	2.312	0.356	1.605	3.020		

With respect to the DAR 8 model, initially a model with five conjugation rates accounting for each conjugation step were used. The parameter importance ranking demonstrated that the last conjugation rate k_5 that accounts exclusively for the over-conjugation step is non-identifiable compared to the other rates, as shown in Figure 3-3B. Consequently, this rate was combined with the rate for the previous conjugation step k_4 (referred as $k_{4/5}$) yielding a lumped version of the model having four kinetic rates. The comparison for the simple and detailed model regarding the parameters and model error is given in Table 3.2. For both DAR 8 models, the confidence intervals are in acceptable ranges, while the intervals relative to the mean for k_1 and k_2 are with values below 1% considerably smaller than for k_3 and k_4 with values around 25%. It could be shown that the detailed model reduces the RMSECV and RMSEP by approximately 49% and 37%, respectively. To further examine the effect of the two models on the prediction performance, a comparison of the two model's predictions is shown versus the experimental data for one example run in Figure 3-3C. The experimental data indicate a more rapid decrease of H0 compared to L0, as well as a sequential formation of the species H1, H2, H3 and H4 in the mentioned order. It can be clearly seen that the detailed (4k) model aligns with the kinetics of all single species more precisely compared to the simple (1k) model. Specifically, the dynamic behavior of H1, H2 and H3 is remarkably better captured by the detailed model. As shown in Table 3.2, the resulting estimated rate for conjugation to light chain (k_1) is slower than the rates representing conjugation to heavy chain (k_2-k_4). For the stepwise conjugation to heavy chain, the second step (k_3) seems to be faster than the first step while the

third/fourth step (k_4) is again slower. The enhanced performance of the detailed model was observed throughout all runs. Therefore, the detailed (4k) model was used for the subsequent modeling of the DAR 8 conjugation reactions.
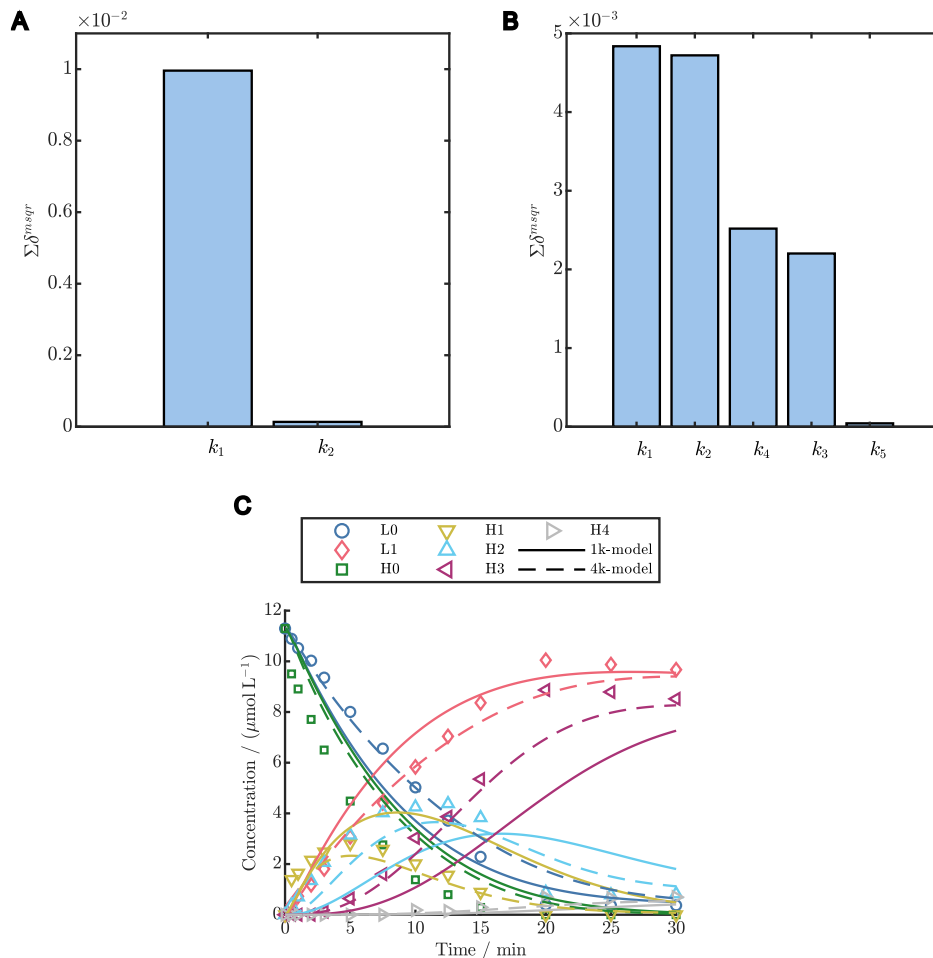

Figure 3-3: Parameter importance ranking according to the summed significance $\sum \boldsymbol{\delta^{msqr}}$ for the detailed DAR 2 model (A) and DAR 8 (B), and (C) comparison of the predictions for the simple (1k) and detailed (4k) DAR 8 kinetic model vs. the experimental data for an example training run with 1.5 g/L ADC2 + 11x NPM and $\boldsymbol{t_f}$ = 30 min.

3.3.3 Modeling site-specific DAR 2 conjugation kinetics

3.3.3.1 Batch conjugation kinetic

The dynamic behavior of the DAR 2 kinetics under various starting conditions in comparison with the predictions from the established kinetic model is depicted for the six training runs in Figure 3-4 (test runs are given in the Appendix A, Figure S 3). The experimental kinetics show a decrease in the concentration of H0 over time concurrent with the formation of H1 and H2. This trend intensifies with increasing drug excess and mAb concentration. The species H2 generally exhibits very low concentrations. Notably, the distribution of the final species depends on the drug excess. Below a drug excess of 3x, the

final concentration of H0 diminishes, while H1 increases proportionally with increasing drug excess, as can be observed for the kinetics in the top row (Figure 3-4A-B). For a drug excess of 3x or higher (Figure 3-4C-F), the final composition remains consistent, which is also characterized by reaching a constant DAR plateau (see Appendix A, Figure S 4). This final DLD contains mainly 84% H1 on average and lower quantities of H0 and H2 at 11% and 5%, respectively. In general, the model accurately predicts the course of the reacting species, exhibiting a minor deviation only during the initial dynamic stage of the reaction. In contrast, the model precisely captures the steady state of each reaction. Overall, the averaged R^2 values over both the training and test set for H0, H1 and H2 were 95.6%, 98.4%, and 71.1%, respectively.
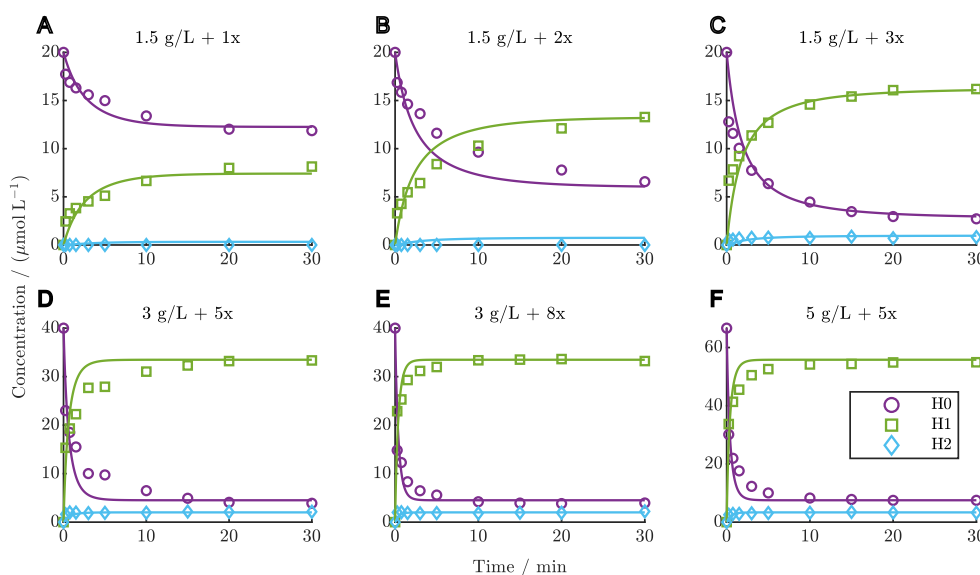

Figure 3-4: Comparison of DAR 2 model predictions vs. experimental data for the six training runs using ADC1 + Drug 1 (A-F).

3.3.3.2 Influence of different payloads

A model re-calibration on the NPM dataset yielded comparable model accuracy with averaged R^2 values of 93.1%, 94.4% and 86.1% for H0, H1 and H2, respectively. Despite the higher depletion rate of NPM, the kinetics revealed a faster conjugation compared to Drug1 under identical initial conditions, as demonstrated for one condition in Figure 3-5 (all NPM kinetics are provided in Appendix A, Figure S 5). This comparison shows that NPM reaches the steady-state more rapidly compared to the other payload, while the distribution of the final species is comparable. The observation was quantified through a comparison of the estimated conjugation rate of the two payloads with $k_{1,Drug1} = 0.251\ L\ (mmol\ s)^{-1}$ and $k_{1,NPM} = 4.840\ L\ (mmol\ s)^{-1}$ indicating a substantial difference in the conjugation rates.
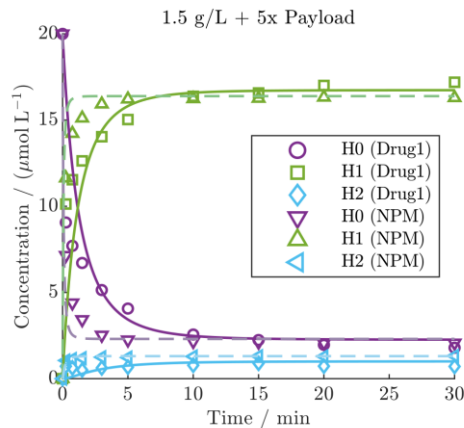
Figure 3-5: Comparison of DAR 2 kinetic with Drug1 or NPM under the same initial condition including the model predictions of the calibrated kinetic models as lines (solid line: Drug1, dashed line: NPM).

3.3.4 Modeling interchain-cysteine conjugation kinetics

3.3.4.1 DAR 8 batch and fed-batch conjugation

The ability of the established model to predict both DAR 8 fed-batch and batch conjugation reactions for the identical ADC was investigated in more detail. The results of three representative kinetics are presented in Figure 3-6 (all kinetics are given in the Appendix A, Figure S 6 and Figure S 7). Compared to experimental batch runs, the data from the fed-batch demonstrates that this drug addition mode results in remarkably slower reaction rates for the individual species within the initial time period of 15 mins. This helps resolve the trajectories of each of the reacting species. In contrast, the samples from the batch kinetic exhibit complete conjugation already for the first time point. Furthermore, the comparison of the final steady states of all runs with a drug excess greater than 8x, suggests that a constant final species distribution is reached, as also characterized by reaching a DAR plateau (shown in Appendix A, Figure S 8). This distribution is characterized by L1 and H3 as the main species, with an average of 96% L1 (from total light chain) and 84% H3 (from total heavy chain). The minor species were L0, H0, H1, H2 and H4 with an average of 4% L0 and 0.01% H0, 0.5% H1, 9% H2 and 6% H4. The utilization of lower drug excesses, such as 6x, does not lead to full conjugation as can be seen in Figure 3-6C. Generally, the kinetic model is able to predict the kinetics of L0, L1, H0, and H3 with high run-averaged R^2 values of 0.97, 0.96, 0.98, and 0.93, respectively. However, it exhibits slightly lower run-averaged R^2 values of 0.86, 0.76, and 0.86 for predicting the kinetics of H1, H2, and H4, respectively, across all runs. This indicates a minor discrepancy between model predictions and experimental data for the species H1, H2, and H4. The estimated kinetic rates for the DAR 8 conjugation of NPM range from approx. 1.2 to 5.1 L $(\text{mmol s})^{-1}$ (cf. Table 3.2). This demonstrates that these rates are within

the same range as observed for the conjugation of the same payload in the case of DAR 2 (cf. section 3.3.3.1).
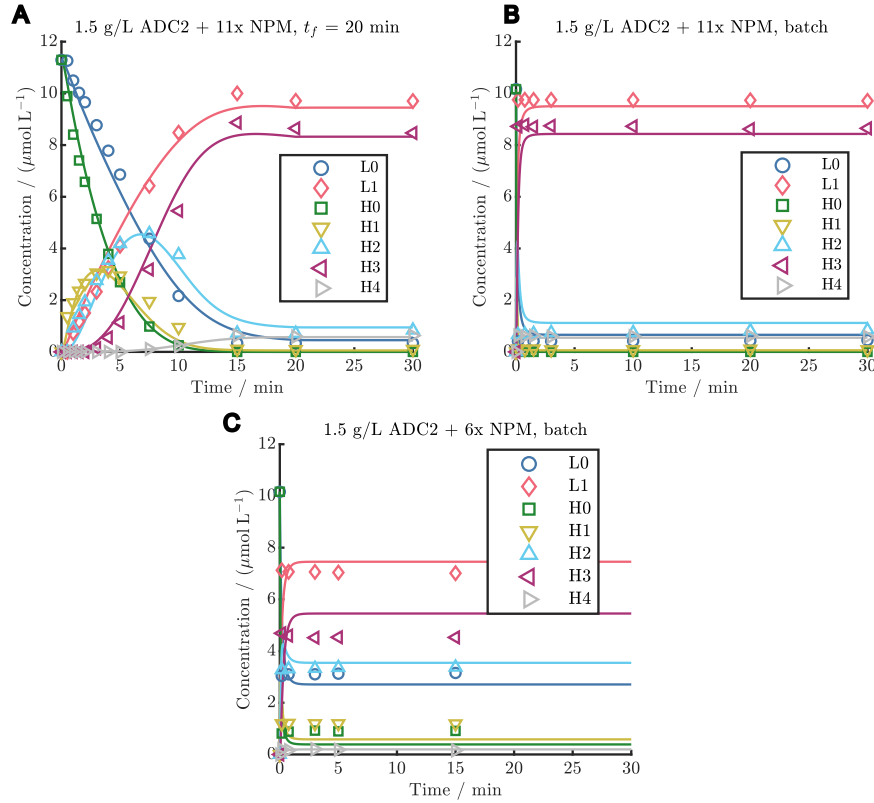

Figure 3-6: Comparison of DAR 8 model predictions vs. experimental data for ADC2 at a mAb concentration of 1.5 g/L for (A) a fed-batch run with 20 min drug feeding time and 11x NPM, and two batch runs with (B) 11x NPM and (C) 6x NPM.

3.3.4.2 In-silico screening for influence of initial mAb and payload concentrations

The selected DAR 8 model was first calibrated to the dataset for ADC3 + Drug2 resulting in similar kinetic rates and model fit with an R^2 of 0.99 averaged for all species (cf. Appendix A, Figure S 9). Afterwards, this model was employed for the in-silico screening for DAR, free unconjugated payload, and reaction time. The results of this screening are depicted in Figure 3-7 for the three outputs. Figure 3-7A illustrates a rapid increase in DAR with escalating drug excess, independent of the mAb concentration, reaching a saturated DAR plateau at approximately 7.7x drug excess. The concentration of unconjugated payload (Figure 3-7B) remains at zero during the concentration range at which the DAR was found to increase, and exhibits a linear rise dependent on both variables as soon as the DAR saturation is reached. The required reaction time to achieve the final DAR (Figure 3-7C) begins with short reaction time of approximately 50 s, increasing with higher drug excess and decreasing mAb concentration to a maximum reaction time of 350 s.

Subsequently, it decreases depending on both variables to reaction times at around 50 s.
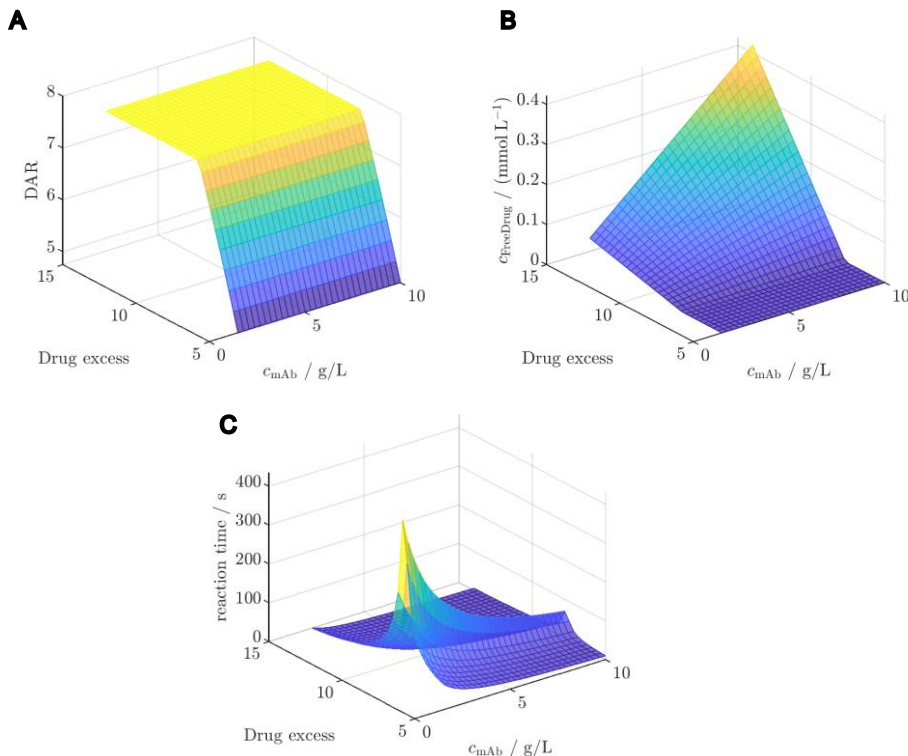

Figure 3-7: Results of the in-silico screening for (A) DAR, (B) free unconjugated payload and (C) reaction time for varying initial mAb concentration c_{mAb} and drug excess.

3.4 Discussion

3.4.1 Kinetic model development

3.4.1.1 Importance of payload depletion rate

Additional measurements to determine the payload depletion rates were necessary due to relatively high parameter uncertainties and strong correlation coefficients with the kinetic rates in models with the depletion rate as parameter to be estimated. The difference in the decrease of the UV/Vis signal for the three payloads, as observed in Figure 3-2, suggests that Drug1 and Drug2 are stable in conjugation buffer over the time studied here, whereas NPM becomes rapidly unavailable for participation in the reaction. Conjugations using pre-mixed NPM or Drug1 in conjugation buffer proved that NPM becomes largely unreactive within typical reaction time (1 h) as opposed to Drug1. This observation is in agreement with Andris et al. [140], who postulate that the depletion of NPM is due to unspecific adsorption to the vessels walls or chemical inactivation. In other studies, it has been similarly proven that polycyclic

aromatic hydrocarbons, such as pyrene, show high affinity to plastic in water due to chemisorption and hydrophobic interactions [146]. NPM may also slowly precipitate due to its low water solubility. The involvement of multiple possible phenomena may result in the larger error bars as observed for the duplicates. For Drug1, a minor inactivation was apparent from a slight decrease of the UV/Vis signal and since the model performance was improved by the addition of a slow depletion rate. In summary, this demonstrates the need for investigating the payload behavior and stability as well as highlights the difference in depletion between the two real drugs, Drug1 and Drug2, and NPM.

3.4.1.2 Model complexity evaluation for DAR 2 and DAR 8

For both conjugation model groups, DAR 2 and DAR 8, the parameter that exclusively accounts for the last conjugation step, had almost zero effect on the model output as shown by low parameter importance values (cf. Figure 3-3). This can be attributed to the consecutive reaction pathway coupled with the fact that the highest conjugated species are a low-concentrated product formed as the last species during the reaction. For instance, for DAR 2 the rate k_1 affects all reacting species, whereas the rate k_2 only partially affects the concentration of H2. In total, this results in minor contributions by the later conjugation rates to the parameter sensitivity. Therefore, the available data is not informative to estimate the rates for the over-conjugation and justifies the removal of the over-conjugation rate leading to a reduced model version with no loss in prediction error indicated by similar RMSE values for DAR 2. The decreasing trend of the influence of the later conjugation rates, as shown in Table 3.2, can similarly explain why the rates k_3 and k_4 in the DAR 8 model, representing the subsequent conjugations, have lower importance and larger confidence intervals. However, here the standard deviation is below 20% relative to the estimated parameter, indicating acceptable estimation quality [75]. The methodology of using the parameter uncertainty and the parameter importance for the selection of proper sub-models can also be found in other works [147].

Especially when applied to model DAR 8 fed-batch conjugation, the detailed model demonstrates enhanced accuracy compared to the simple model, as depicted in Figure 3-3C, and supported by the reduction in RMSE values. The simple 1k model exhibits inadequacies in resolving the species trajectories during the initial stages of the reaction. In contrast, the detailed model excels in capturing the nuances of species trajectories by discerning the rates associated with each sequential conjugation step. Notably, it is observed that the conjugation to the light chain, with a rate constant (k_1) of 1.408 L (mmol s)$^{-1}$, is slightly slower than the first conjugation step of the heavy chain, which has a rate constant (k_2) of 2.111 L (mmol s)$^{-1}$. This discrepancy may be attributed to the presence of only one binding site on the light chain compared to the three

available binding sites on the heavy chain. Additionally, the second conjugation step exhibits an almost twofold increase in kinetic rate compared to the first step, while the third step experiences a reduced kinetic rate. Consistent with findings in Andris et al. [140], an ascending trend in subsequent conjugation steps is noted, associated with the increasing hydrophobicity on the mAb when the payload is bound. The final decrease in the kinetic rate in our case may be linked to the limited availability of one binding site left to react. The estimated conjugation rates ranged between $0.3\text{-}5 \cdot 10^3$ L (mol s)$^{-1}$ which is in the expected range for rate constants of maleimides with thiols [148], and significantly faster than amine conjugation of reactive PEG with rate constants around $1 \cdot 10^{-1}$ L (mol s)$^{-1}$ [149]. One disadvantage of the current model is the lack of resolution for the reactivities of the different binding sites in a single molecule. This would require more sophisticated analytics, such as LC-MS [150], in combination with advanced kinetic models which could enable the forecasting of the reactivity of specific sulfhydryl groups and how this would affect the probability of the subsequent conjugation to the other residues. For instance, Mao et al. [149] recently presented a structure-depended reactivity model for PEGylation which enabled them to estimate the reactivity of individual amines based on molecular descriptors.

3.4.2 Insights from DAR 2 conjugation modeling

3.4.2.1 Modeling accuracy and analytical challenges

As shown in Figure 3-4, the DAR 2 model was able to model the trajectories of the heavy chains accurately, especially for H0 and H1 with R^2 values above 93%. Only the species H2 is not precisely modeled, which can be attributed to its high analytical variance owing to poor HPLC resolution of this species. It was reported that this over-conjugation is located at non-reformed disulfide bond between heavy and light chain [143]. The herein established model, however, does not enable derivation of additional information about this phenomenon as this species makes up a low percentage of the total mixture, thus causing a small impact on the model predictions. Notably, the transient behavior of the distribution of the final species for varying drug excess is also captured by the model. The observation that a consistent composition is reached with around 3x drug excess, despite a stoichiometric requirement of 2x drug excess for the number of binding sites, further exemplifies the occurring payload depletion for Drug1. Additionally, the model showed a systematic discrepancy in the initial stage of the reaction. According to Andris et al. [140], the sequential conjugation in the hinge region results in a step-wise increase in the conjugation rate due to an increase in hydrophobicity in this region. In contrast, the herein established heavy chain model lumps this phenomenon in

one single rate (k_1) because of an inability to account for this phenomenon due to the reducing analytical assay which removes the information about the number of payloads already bound to the intact ADC molecule. In summary, these discrepancies were considered acceptable for the purpose of modeling the DAR 2 conjugation kinetic.

3.4.2.2 Conjugation rates difference for the utilized payloads

Calibrating the model on the two datasets using either Drug1 or NPM could show that the conjugation kinetic rate for NPM is approximately 20 times higher than for Drug1 (see Figure 3-5). This difference may be rather related to molecular size than hydrophobicity, as NPM has a molecular mass that is approximately five times lower than Drug1 and both drugs exhibited a similar elution time in the RP (data not shown). Pfister et al. [145] demonstrated differences in conjugation rate depending on the size of the conjugated molecule, in the case of conjugation of PEG to lysozyme which was described using a core-shell model. In our case, the finding highlights the inequality of two payloads with regards to their conjugation kinetic rate in addition to the said difference in their depletion.

3.4.3 Insights from DAR 8 conjugation modeling

3.4.3.1 Modeling accuracy and exploring drug excess thresholds

The experimental data revealed that there is a drug excess threshold between 8-11x in which a consistent DLD is achieved. Using a drug excess larger than 8x did not alter the percentages in the final DLD or led to higher DAR values (cf. Appendix A, Figure S 8), which suggest that a critical point is reached beyond which further increases in the drug excess does not result in a higher DAR. Drug excesses below this point (e.g. 6x, see Figure 3-6C), lead to not reaching full conjugation which is exemplified in the lower percentages of L1 and H3 in the steady-state. Because NPM was found to simultaneously deplete in the solution, this suggests that some payloads require a drug excess higher than the stoichiometric molar quantity of binding sites to achieve full conjugation. Furthermore, the species distribution in the final steady states remained relatively constant, primarily related to the constant reduction conditions within this kinetic subset. The lower percentages of L0, H0, H1, and H2 likely originate from partially reduced species, with fewer interchain disulfides. In the literature, comparable DLDs were reported with these low levels of under-conjugated species [151], [152], and it was noted that achieving full conjugation is challenging due to the relatively low concentration of reactants [153]. The presence of low percentages of H4 is unique to the herein

utilized IgG1 and is likely to be attributed to the said "mis-alkylation" of the payload [144].

3.4.3.2 Feeding mode comparison

The utilization of fed-batch runs lowered the conjugation rates enabling the elucidation of the reaction mechanism (see Figure 3-6A,B). Feeding times of 10 mins or longer yielded satisfactory analytical resolution of the individual species. The model exhibited slightly lower accuracy for the species H0, H1, H2, and H4, which can be primarily associated to their higher analytical error due to lower concentration ranges. The batch runs emphasized that DAR 8 conjugation reactions are notably faster than DAR 2, with a time-scale of only a few minutes, compared to 5-10 min for DAR 2 reaction. Hereby it became also evident that the utilization of the complex (4k) model and the precise determination of the conjugation rates are only required when modeling fed-batch conjugation reaction. As expected, the model's performance is primarily affected by the initial cysteine distribution when being applied for the batch reaction, as this defines the steady state of the reaction. Using the simple (1k) DAR 8 model for the prediction of the ADC3 batch kinetics yielded an identical average R^2 of 0.99 as the detailed (4k) model. This highlights the importance of accurately determining the initial cysteine distribution for the model precision, when applying the model to batch kinetics.

3.4.3.3 Model adaptability

Moreover, it was demonstrated that the estimated kinetic rates for the same payload molecule are comparable among different ADC types in our study. This suggests the potential adaptability of the utilized modeling approach across different modalities and its potential for transfer to diverse conjugation kinetics without the necessity of prior calibration.

3.4.3.4 In-silico screening for enhanced reaction understanding

The observed behavior in the screening for ADC3 + Drug2 reveals significant insights into the interplay of initial mAb concentration and drug excess on the studied key model outputs, as shown in Figure 3-7. The rapid increase in DAR until a plateau with increasing drug excess, independent of mAb concentration, aligns with the already observed saturation effect. The saturation is independent from the initial mAb concentration as the initial mAb concentration only defines the time point at which the DAR is reached. Notably, the achieved saturated DAR is achieved once the stoichiometric molar drug ratio to mAb is utilized, as Drug2 did not show to deplete and, thus, being complete available for conjugation. Similar results were reported for the DAR 2 reaction for the DAR saturation [140]. The linear increase in unconjugated

payload concentration beyond the DAR plateau phase suggests that, once the saturation point is reached, excess drug molecules remain unbound and contribute to the free payload concentration. Regarding the reaction time, the initial rapid increase is due to the increasing saturation DAR value below the critical drug excess of 7.7x, which requires prolonged reaction times. The maximum reaction times, aligning with DAR saturation, is achieved at low mAb concentration due to slower conjugation rates. After the DAR is saturated, both initial mAb concentration and drug excess lead to faster conjugation rates which corresponds to shorter reaction times. Overall, these findings highlight the benefit of being able to forecast the reaction behavior for various process parameters. Overall, this screening showcases the benefit of a kinetic model due to additional insights into the reaction kinetics and the reaction time-scales which is crucial for optimization of operating conditions. Especially, the knowledge of the optimal drug excess is highly useful since payloads are costly and free payload requires subsequent filtration steps in order to ensure product safety [135].

3.5 Conclusion

In this study, we developed a kinetic modeling methodology that can accurately predict both site-specific conjugations (DAR 2), and conjugations to reduced interchain disulfide bonds (DAR 8) of various payloads. A readily available RP-UHPLC method was used to acquire the kinetic data. For the kinetic model development, we covered multiple facets: First, UV/Vis measurements were performed to determine the payload stability in conjugation buffer alone. Secondly, different kinetic model candidates for the two modalities were formulated and selected. Fed-batch experiments proved to be crucial in resolving the rapid conjugation kinetics in the case of interchain disulfide conjugation. Thirdly, the kinetic model was applied to external datasets containing other payloads. A key outcome was the confirmation of major differences in both the conjugation rates and the stability of the tested pseudo and real payloads. Modeling DAR 2 and DAR 8 conjugation with the same payload demonstrated comparable rates, while for DAR 8 conjugation varying conjugation rates became apparent for the sequential conjugation steps. Following this, the calibrated DAR 8 kinetic model was used for an exemplary in-silico screening to study the reaction behavior across a range of initial conditions. Overall, these results emphasize that this kinetic model framework is highly valuable for augmenting experimental studies, providing enhanced process understanding and optimizing the conjugation process. This approach not only complements traditional DoE methodologies but also addresses the inherent gap in DoE by offering mechanistic insights of the conjugation process,

thereby accelerating process development. In future, it could be extended with other relevant influencing factors, such as pH or salt effects, or tested on other conjugation chemistries.

Acknowledgments

Supplementary Materials

Appendix A contains the Supplementary Materials associated with this chapter:

- S3.1 Detailed Overview of all experimental conditions
- S3.2 Reaction schemes
- S3.3 DAR 8 model ODEs
- S3.4 Conjugation with "inactivated" NPM and Drug1
- S3.5 Payload inactivation for NPM
- S3.6 DAR 2 model predictions for Drug1 and NPM
- S3.7 DAR 8 model predictions using ADC2 for both training and test set

4

Kinetic modeling of the antibody disulfide bond reduction reaction with integrated prediction of the drug load profile for cysteine-conjugated ADCs

Jan Tobias Weggen[1], Pedro González[1], Kimberly Hui[2], Ryan Bean[2], Michaela Wendeler[2], Jürgen Hubbuch[1]

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

[2]Purification Process Sciences, BioPharmaceuticals Development, R&D, AstraZeneca, Gaithersburg, Maryland, USA

Abstract

Antibody-drug conjugates (ADC) constitute a groundbreaking advancement in the field of targeted therapy. In the widely utilized cysteine conjugation, the cytotoxic payload is attached to reduced interchain disulfides which involves a reduction of the native monoclonal antibody (mAb). This reaction needs to be thoroughly understood and controlled as it influences the critical quality attributes (CQAs) of the final ADC product, such as the Drug-to-Antibody Ratio (DAR) and the drug load distribution (DLD). However, existing

methodologies lack a mechanistic description of the relationship between process parameters and CQAs. In this context, kinetic modeling provides comprehensive reaction understanding, facilitating the model-based optimization of reduction reaction parameters and potentially reduce the experimental effort needed to develop a robust process. With this study, we introduce an integrated modeling framework consisting of a reduction kinetic model for the species formed during the mAb reduction reaction in combination with a regression model to quantify the number of conjugated drugs by DAR and DLD. The species formed during reduction will be measured by analytical capillary gel electrophoresis (CGE), and the DAR and DLD will be derived from reversed phase (RP) chromatography. First, we present the development of a reduction kinetic model to describe the impact of reducing agent excess and reaction temperature on the kinetic, by careful investigation of different reaction networks and sets of kinetic rates. Second, we introduce a cross-analytical approach based on multiple linear regression (MLR), wherein CGE data is converted into the RP-derived DAR/DLD. By coupling this with the newly developed reduction kinetic model, an integrated model encompassing the two consecutive reaction steps, reduction and conjugation, is created to predict the final DAR/DLD from initial reduction reaction conditions. The integrated model is finally utilized for an *in silico* screening to analyze the effect of the reduction conditions, TCEP excess, temperature and reaction time, directly on the final ADC product.

4.1 Introduction

Antibody-drug conjugates (ADCs) are a class of innovative therapeutics engineered to deliver potent cytotoxic drugs directly to cancer cells while minimizing damage to healthy tissues [18], [49], [116]. Consisting of monoclonal antibody (mAb), cytotoxic payload, and linker, ADCs offer precise targeting, enhancing efficacy and reducing systemic toxicity compared to traditional chemotherapy. ADCs constitute one of the fastest-growing classes of anticancer drugs, with fourteen ADCs approved by the FDA as of 2023 and hundreds progressing in clinical trials [45]. Driven by persistent research efforts in ADC design, such as more stable linkers, novel payloads and new conjugation strategies, the challenge to develop robust and scalable processes increases [29], [154].

The potency of ADCs largely depends on the type of payload and the number of payload molecules attached, known as the Drug-to-Antibody Ratio (DAR). Controlling the DAR and the drug load distribution (DLD) is crucial for optimizing the ADC therapeutic index, as every species within the DLD exhibits distinct pharmacokinetics and efficacy [52], [57]. A central consideration in this context is the conjugation strategy, which facilitates the connection of the mAb

and linker-payload [34]. Payload conjugation to cysteine residues that are engaged in the four native interchain disulfide bonds is a common conjugation strategy. While this approach is widely used, it can minimally alter the mAb structure and negatively impacts stability, particularly for higher DAR species [155], [156]. With regards to the conjugation process, this approach involves a two-step reaction workflow: an initial mild reduction of the disulfide bonds of the mAb, followed by conjugation of the cytotoxic payload to the reactive cysteine residues. The final DAR for cysteine-linked ADCs ranges from 2 to 8. The reduction reaction is the key step for determining the DAR, as it generates a specific number of binding sites for the payload [45], [56]. The reaction's outcome, including the formation of positional isomers, depends on the choice of reducing agents and various process parameters such as reducing agent concentration, reaction time, temperature and pH [60], [157]. However, comprehensive understanding of the cysteine-based reaction process is limited [136], [158], [159]. Alternative research focus is on improving DAR homogeneity by cross-linking payloads [58] or addition of metal ions [56].

Initiated by the FDA [71], the concept of Quality by Design (QbD) promotes the utilization of modeling techniques to deepen process understanding of unit operations, forecast product behavior, and optimize processes. Consequently, process models have recently gained popularity for a variety of purposes for biopharmaceuticals [123], [160]. To model biochemical reactions, kinetic studies are usually performed to gain knowledge about complex biochemical reaction mechanisms [124]. To construct mathematical process models for these reactions, the most challenging aspects are simultaneously identifying the reaction stoichiometry, inferring the structure of the reaction network, and estimating model parameters [161]. Various solutions have been proposed in the literature to address these challenges, e.g., S-systems [162], target factor analysis [163], automated reaction network generation and identification [164], and chemical reaction neural networks [165].

With regards to modeling of ADC conjugation process, recent studies focused on the conjugation kinetic [140] for site-specific ADCs, or the reduction reaction for cysteine-conjugated ADCs [130]. The latter provides quantitative insights into the reduction mechanisms of interchain disulfide bonds and the use of the kinetic model for process optimization, relying on an complex mathematical approach to resolve positional isomers from hydrophobic interaction chromatography (HIC) and reversed-phase ultrahigh-performance liquid chromatography (RP-UHPLC).

In this study, we introduce the combination of reduction kinetic model and linear regression to model the reduction reaction with directly determining the DAR/DLD at the conjugation endpoint. For the development of these models,

we performed reduction kinetic studies with varying initial conditions, namely the mAb concentration, reducing agent concentrations, as well as the reaction temperatures, while the subsequent conjugation reactions were conducted under the same reaction conditions. Samples were analyzed by both non-reducing capillary gel electrophoresis-sodium dodecyl sulfate (CGE-SDS, or herein shortly CGE) to quantify the reduced species, and RP-UHPLC to determine the DAR/DLD. Based on species formed during reduction, a mechanistic kinetic model for the interchain disulfide bond reduction reaction is developed. Here, various chemical reaction networks (CRNs) with different sets of kinetic rates are proposed and the most suitable CRN was selected. Then, multiple linear regression (MLR) models are employed to predict the final DAR and the percentages of the drug load distribution (DLD-%) from the species formed during the reduction reaction. Finally, the reduction kinetic model and the MLR models are combined, creating an integrated modeling tool enabling *in silico* screening of process parameters in the cysteine-based reaction workflow.

4.2 Material and methods

4.2.1 Experiments

The experiments involved two sequential reaction steps: reduction and conjugation. The reduction conditions were varied to investigate their influence on the partial reduction kinetics of the interchain disulfide bonds. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was used as reducing agent, as this reduces both the two disulfide bonds between the two heavy chains (inter HH) and the bonds between each heavy and light chain (inter HL). The reduction of intrachain disulfides was excluded, given that they are known to remain intact under TCEP treatment [56]. In the subsequent conjugation reaction, maleimide-functionalized payload is conjugated to these reduced disulfide bonds under constant reaction conditions. Non-reducing CGE was used to track the generation of mAb fragments (L (light chain), H (heavy chain), HL, HH, HHL) over time, which are generated as disulfide bonds are reduced. This provided a quantitative measurement of the reduction kinetics. Two IgG1 mAbs, which will be referred to as mAb1 and mAb2, were employed in the scope of the experiments.

4.2.1.1 Chemicals and buffers

The chemicals used in this study were obtained from EMD Millipore, unless otherwise stated. Standard conjugation buffer was 20 mM NaH_2PO_4 (J.T. Baker), 1 mM EDTA in ultrapure water at pH of 7.0. TCEP-HCl (VWR) was used to reduce the mAb. For conjugation, a maleimide-functionalized payload

was dissolved in dimethyl sulfoxide (DMSO, Sigma-Aldrich). For sample treatment for DAR analysis, samples were diluted in denaturing buffer containing guanidine HCL (Thermo Scientific), Tris (Thermo Scientific), EDTA and dithiothreitol (Thermo Scientific) at pH 7.6. RP-UHPLC was performed using acetonitrile (VWR) and HPLC water (VWR), both with 0.1% (v/v) trifluoroacetic acid (Thermo Scientific). For sample treatment for CGE analysis, samples were diluted in 1 M Tris, pH 7.0.

Table 4.1: List of conducted reduction experiments using different mAbs, reduction reaction temperatures, initial mAb concentration and initial TCEP excess. Additionally, the corresponding subset for each run is given.

RunID	mAb	Temperature / °C	c_{mab} / g/L	Molar TCEP excess / -	Subset
1	mAb1	20	15	2	Training
2	mAb1	20	15	3	Training
3	mAb1	20	15	4	Training
4	mAb1	20	15	6	Training
5	mAb1	20	15	8	Training
6	mAb1	20	3.75	4	Training
7	mAb1	20	7.5	4	Training
8	mAb1	20	10	4	Training
9	mAb1	37	15	2	Training
10	mAb1	37	15	4	Training
11	mAb1	37	15	6	Training
12	mAb1	37	15	8	Training
13	mAb1	4	15	2	Training
14	mAb1	4	15	4	Training
15	mAb1	4	15	8	Training
16	mAb1	12	15	4	Test
17	mAb1	12	15	6	Test
18	mAb2	20	15	4	External
19	mAb2	20	20	8	External
20	mAb2	20	20	10	External

4.2.1.2 Reduction kinetic procedure

Initially, the mAb was thawed and diluted to the desired starting mAb concentration in conjugation buffer. A 50 mM TCEP solution was prepared in conjugation buffer and pH adjusted to 7 using NaOH (J.T. Baker). The reduction reaction was started by adding a predetermined volume of the 50 mM TCEP solution to the diluted mAb to reach a desired molar excess of TCEP over mAb. The reaction was allowed to shake on an orbital shaker (Eppendorf) at the specified temperature. Kinetic samples were taken at 5, 10, 20, 30, 60, 120, and 240 min. The reduction reaction was quenched by addition of a 12-fold molar excess of payload (25 mM solution in DMSO) to conjugate all available cysteines, thus preventing reduced disulfides from reforming. DMSO was spiked into the conjugation reaction to reach 10% (v/v) DMSO to improve

payload solubility during conjugation. The conjugation reaction was allowed to proceed on a tube rotator at room temperature for 30 min. In 17 runs using mAb1, the reduction temperature, the initial mAb concentration, and the TCEP excess were varied. Three additional runs with mAb2 were performed. The reduction reaction conditions are summarized in Table 4.1.

4.2.1.3 Sample analytics

4.2.1.3.1 Non-reducing capillary gel electrophoresis

To quantify the partial reduction of the mAb, samples were analyzed with non-reducing CGE-SDS on a PA800+ capillary electrophoresis instrument (AB Sciex) using the assay protocol as published in Cao et al. [143]. Samples were electrokinetically injected into the capillary and separated in reverse polarity. The resulting exported data contained the fractions, denoted as p_i, of intact mAb monomer and its reduced fragments within the reaction mixture. The molar concentrations of each species $\tilde{c}_i$ was converted using the respective mAb concentration of the sample c_{mAb} and the molar mass of the respective species $\widetilde{M}_i$ according to Eq. 1

$$\tilde{c}_i = p_i \cdot \frac{c_{mAb}}{\widetilde{M}_i} \ .$$

(1)

4.2.1.3.2 Reducing reversed-phase ultra-high performance liquid chromatography

To determine the DAR and the DLD, each sample was analyzed using RP-UHPLC. Details about sample treatment procedure and the analytical workflow can be found in Cao et al. [143]. The data for each sample contained the percentages of un- and mono-conjugated light chain, %L0 and %L1, and of un-, mono-, bi- and tri-conjugated heavy chain, %H0, %H1, %H2 and %H3, respectively. The respective DAR was calculated by using the total peak area of all light and heavy chain species, $A_{L,tot}$ and $A_{H,tot}$, and the peak areas of the conjugated species according to Eq. 2

$$DAR = 2 \left(\frac{A_{L1}}{A_{Ltot}} + \frac{A_{H1} + 2A_{H2} + 3A_{H3}}{A_{Htot}} \right) .$$

(2)

4.2.2 Data organization

In this study, two model types were developed and the data were organized differently: (1) For development of the reduction kinetic model, the full reduction kinetic dataset from CGE for mAb1 was manually split based on the reduction temperature into a training subset (4, 20 and 37 °C) with 15 runs and an independent test subset (12 °C) with two runs. The training subset was used for model development and selection, and the test set for independent validation of the reduction kinetic model. The remaining three runs with mAb2 were

employed for external evaluation on the model's applicability on a different mAb molecule. (2) For the development of the MLR models the full dataset comprising both CGE and RP-UHPLC for mAb1 was randomly split into training and test set with a ratio of 75%/25%. The dataset for mAb2 was also used for extern evaluation.

4.2.3 Mechanistic reduction kinetic modeling

The reduction of the interchain disulfide bonds by TCEP results in the occurrence of reduced fragments, HHL, HH, and HL, and single chains, L and H, as measured in the CGE. The development of the mechanistic kinetic reduction model involved the iterative selection of a suitable CRN and a set of kinetic rates accounting for the temperature-dependency of the reaction via the Arrhenius approach. For the first task, different CRNs were proposed and evaluated using the same assumption for kinetic rate discrimination. After selecting the most suitable CRN, the set of kinetic rates was further investigated in more detail to improve model performance.

4.2.3.1 Chemical reaction network construction and kinetic rate discrimination

For the construction of the CRN, special focus was on the two inter HH disulfide bonds localized in the IgG hinge region. As the CGE is unable to differentiate between fragments with two or one intact inter HH disulfide bonds, it was not possible to determine whether the reduction of these bonds occurred stepwise. To assess the impact of the reduced form of these intermediate on the reaction kinetics, two CRN types were constructed: In the *NoInt* CRN, the reduction of the two hinge bonds is considered as single reaction step yielding seven reaction species (= seven CGE species) and six reactions. Conversely, in the *WithInt* CRN, a two-step reduction in the hinge region is included to account for the three non-observable intermediates, IntR, HHLR and HHR (R = reduced form), resulting in a more complex CRN with ten species and eleven reactions. Additionally, a reforming of the reduced species was observed in the data, which coincides with a decreasing DAR value. Therefore, a re-oxidation of the reduced disulfide bonds was assumed in the model assigning a backwards re-oxidation reaction for each forward reduction step, leading to two additional variants of each CRN type, *WithInt+ReOx* (graphically summarized in Figure 4-1A) and *NoInt+ReOx* (given in the Supplementary Materials). To parametrize these CRNs, kinetic rates needed to be assigned to each individual reaction step. As suggested by Chen et al. [70] that TCEP reduction kinetically favors the inter HL over the inter HH bonds, all individual rates belonging to either inter HH or inter HL reduction/re-oxidation were lumped. As HH re-oxidation rates were observed to be substantially slower in comparison to the other kinetic rates, the

respective reaction steps were removed in all CRNs that account for the re-oxidation. Consequently, this resulted in the simple models with two (for CRNs without ReOx) or three kinetic distinct kinetic rates (for CRNs with ReOx, "3k-model"), which were used as baseline rate discrimination to independently analyze the CRN performance.
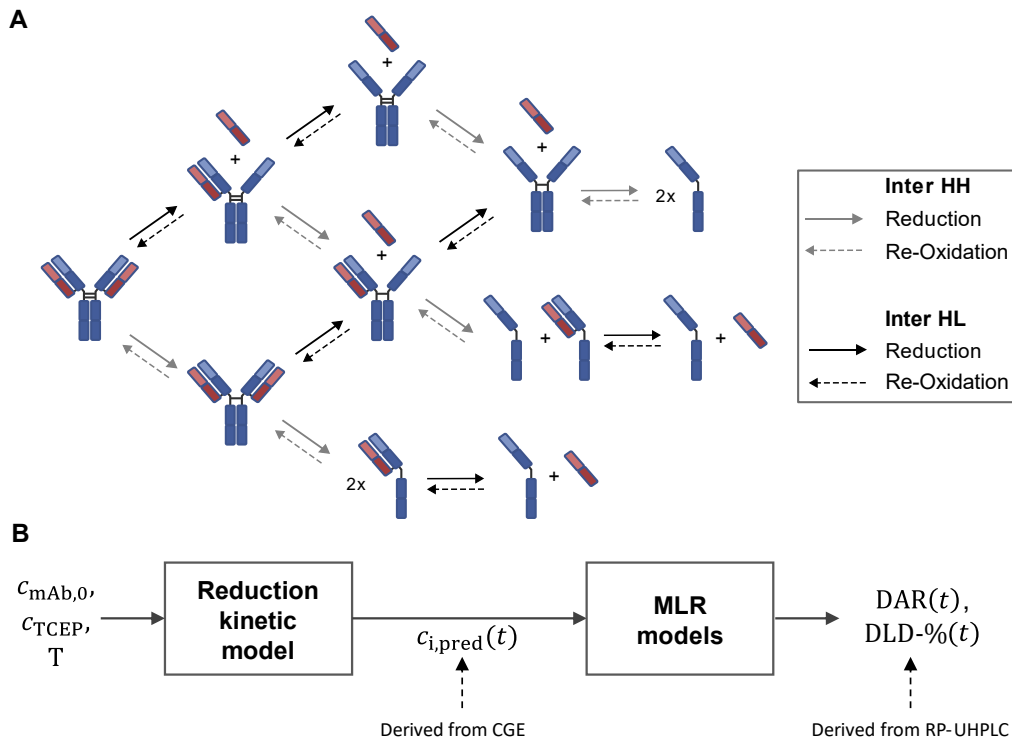

Figure 4-1: (A) Proposed structure of a CRN with intermediates and re-oxidation reaction. Black and gray arrows indicate reactions between inter heavy-heavy (HH) or inter heavy-light (HL) chains. Heavy chains are colored in blue and light chain are colored in red. (B) Schematic representation of the integrated modeling architecture for the sequential combination of the newly developed reduction kinetic modeling with the MLR models predicting the DAR and DLD percentages (DLD-%).

Additionally, two other sets of kinetic rates with more advanced kinetic rate discriminations were evaluated. A second, refined model was defined that specifically discriminates the reduction/re-oxidation belonging to each CGE-observed fragment (e.g. three different rates for the three inter HL reduction of Intact, HHL and HL). Based on the 3k-model, this assumption tripled the number of rates resulting in nine kinetic rates ("9k-model"). Nayak and Richter [130] presented a reduction kinetic model, that indicated a considerable difference between the first and second reduction steps for the inter HH reduction. This finding motivated the design of a third discrimination, which uses a similar approach as the 9k-model, but instead only distinguishes the stepwise inter HH reduction step with two rates, leading in total to eight kinetic rates ("8k-model"). Table 4.2 summarizes the investigated reduction models. A

list of the reactions and the corresponding rates is given in the Supplementary Materials.

Table 4.2: Summary of the tested CRNs and parameter assignments.

No	Intermediates (Int)	Re-Oxidation (Re-Ox)	Rates (k)	Rate Discrimination
1	No	No	2	Groups inter HL/HH reduction
2	No	Yes	3	Groups inter HL and HH reduction/re-oxidation
3	Yes	Yes	3	Groups inter HL and HH reduction/re-oxidation
4	Yes	Yes	9	Individual rates for each fragment
5	Yes	Yes	8	Same as 9k, but for inter HH two rates for the stepwise reduction

4.2.3.2 Temperature modeling

The Arrhenius equation was applied to model the temperature effect of the kinetic rates. This led to double the amount of necessary model parameters in each model, as each rate was split into two Arrhenius parameters. To avoid optimization problems due to high correlation between the Arrhenius parameters, activation energy, E_a, and pre-exponential factor, k_0, it is common to use a reparametrized expression of the Arrhenius equation [166]. The herein used expression is given in Eq. 3

$$k(T) = k_{T_{\text{ref}}} e^{-\frac{E_a}{R}\left(\frac{1}{T} - \frac{1}{T_{\text{ref}}}\right)},$$ (3)

where R denotes the ideal gas constant, T_{ref} is the reference temperature and $k_{T_{\text{ref}}}$ presents the kinetic rate at the reference temperature. $T_{\text{ref}} = 293.15$ K was chosen adequately to minimize the parameter correlation.

4.2.3.3 Parameter estimation

All tested models were implemented as ordinary differential equations (ODE) in MATLAB (The MathWorks Inc.) and numerically solved using the *ode15s* solver. The kinetic rates were estimated based on the nonlinear least squares approach. A custom loss function was utilized that computed the error between the CGE-derived experimental and the model-predicted concentrations. In case of the models with intermediates, the concentrations of the reduced intermediates and the respective non-reduced intermediates were summed up to align with the CGE-derived molar concentrations of the species. To ensure an equal weighting of the six reaction species, the species errors were normalized to the maximum species concentration within the full dataset. As initial parameter values, all E_a were set to 20 kJ/mol, and all $k_{T_{ref}}$ to $1e^{-4}$

$L(\mu mol\,s)^{-1}$. The model parameters were optimized using *lsqnonlin*. The uncertainty of the estimated parameters was assessed via 95% confidence intervals (CI) computed using *nlparci*.

4.2.3.4 Model error, selection and validation

The resulting model performance was quantified by the normalized root mean square error (nRMSE) for each run and for each species computed as given in Eq. 4

$$\text{nRMSE}_i = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(c_i - \hat{c}_i)^2}}{\bar{c}_i} \ , \tag{4}$$

where N is the number of samples, and c_i, $\hat{c}_i$ and $\bar{c}_i$ denote the measured, model predicted, and mean measured concentration of the ith species, respectively. For model selection, the prediction error of the model was assessed based on both cross-validation of the fifteen training runs (nRMSE_{CV}) and on the two test runs (nRMSE_P). A leave-one-run-out scheme was used for the cross-validation, wherein each run was systematically excluded one time and the resulting nRMSE_{CV} values for all rotations were averaged. Additionally, the corrected Akaike's information criterion (AICc) was employed as a balanced measure between the goodness of fit and the model complexity. The AICc is calculated using Eq. 5

$$\text{AICc} = N \cdot \ln\left(\frac{\text{SSE}}{N}\right) + 2\delta + \frac{2\delta(\delta + 1)}{N - \delta - 2} + N \cdot \ln(2\pi) + N \ , \tag{5}$$

where SSE is the sum of squared error between measured and predicted concentration and δ represents the number of model parameters.

4.2.4 Integrated reduction kinetic modeling

As the reduction reaction is the governing step due to its higher reaction time compared to the conjugation reaction [130], a subsequent model was incorporated to directly estimate the final DAR and the DLD-% from each state in the reduction kinetic without explicit consideration of the conjugation kinetic. This sequential combination with the newly developed mechanistic reduction kinetic model allowed then to perform an integrated kinetic modeling of the DAR and the DLD-% with regards to the initial reduction reaction conditions, i.e., $c_{mAb,0}$, TCEP excess, and the temperature. A schematic illustration of this model framework is shown in Figure 4-1B.

4.2.4.1 Final DAR and DLD-% estimation from CGE or reduction model prediction

Multiple linear regression (MLR) models were employed to establish cross-analytical, functional mappings from the reduction state to the final conjugation outcome at each time point. The choice of a linear model was motivated by the robust linear correlation observed between the two analytical datasets (see Supplementary Materials, Figure S4). The regression problem can be expressed using Eq. 6, where the input matrix $\boldsymbol{X}$, contains the concentrations of the six reduction species ($c_{\text{intact}}, c_{\text{HHL}}, c_{\text{HH}}, c_{\text{HL}}, c_{\text{H}}, c_{\text{L}}$) in each column over time, and the response variable $\boldsymbol{y}$ corresponding to the RP-derived DAR or one of the six DLD-% ($\text{L0}, \text{L1}, \text{H0}, \text{H1}, \text{H2}, \text{and } \text{H3}$) at the time point:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \epsilon \tag{6}$$

where $\boldsymbol{\beta}$ represents the regression coefficients and ϵ the error term. To ensure uniform consideration of the entire concentration range, the input concentrations were normalized by the respective mAb concentration. The function *fitlm* in Matlab was used for calibrating the MLR models. In a first step, this approach was evaluated utilizing the complete experimental CGE data as model input. Training and test error between the MLR-predicted and experimental DAR/DLD-% were quantified by the R^2. Subsequently, the MLR models were independently tested on the external dataset of mAb2 to evaluate the potential generalizability of this approach. In a second step, the kinetic model predictions of the newly developed at each experimental time point were employed as model input. A re-calibration of the MLR models was required due to the differences between experimental CGE and model-predicted reduction concentrations. For consistency, the same random training and test split was used in both instances. It is worth noting that this approach is applicable only when a sufficiently high payload excess is utilized to conjugate all reduced disulfide bonds.

4.2.4.2 Integrated in-silico screening and process optimization

To demonstrate the utility of the developed integrated framework in assisting ADC process development, an *in silico* screening of the reduction reaction conditions was performed. Two case scenarios were considered, wherein two out of the four possible process parameters (temperature, TCEP excess, mAb concentration and reduction time) were varied: In the first scenario, both reduction temperature and TCEP excess were varied within intervals of [4, 37] °C and [2, 8], respectively, while the reduction time was fixed to 120 min. In the second scenario, both reaction time and TCEP excess were varied within intervals of [0, 240] min and [2, 8], respectively, with the temperature set at 20 °C. The initial mAb concentration was set to 15 g/L for both scenarios. Utilizing

the calibrated reduction kinetic model, the resulting reduction species concentrations were systematically screened for varying initial conditions. These concentrations were then used to predict the DAR and DLD-% using the previously calibrated MLR models.

In the last part, optimal reaction times were determined across the entire range of temperature and TCEP excess to exemplarily achieve a target DAR of 3. Similar to the *in silico* screening, individual DAR kinetics were simulated under varying temperature and TCEP excess until the desired DAR was reached. Another typical requirement is that the amount of high-loaded species must be reduced [154]. Therefore, the additional constraint was added to ensure the percentage of H2 remained below 12%.

4.3 Results and discussion

4.3.1 Kinetic modeling of interchain disulfide bond reduction reaction

4.3.1.1 Model development: Selection of CRN and parameter set

Five candidate reduction models were evaluated by their model error. Figure 4-2 illustrates the cross-validation and test nRMSE values, respectively, for each model with regards to the individual reduction species, along with the average nRMSE. The first three models represent different CRN variants under the same kinetic rate assumption. When comparing the first two models, the addition of the re-oxidation reactions consistently improves the prediction of each species all species, except for H, resulting in an average nRMSE reduction of 24% and 22% for the CV and the test set, respectively. A particularly strong reduction was found for the intermediates HH and HL as well as for the species L. It is known that re-oxidation can be provoked by the presence of dissolved oxygen [167], [168]. While Tang et al. [129] found that the oxidation of reduced inter HH disulfide bonds tends to be generally more pronounced compared to inter HL, our results instead indicated that inter HH re-oxidation is negligible. This discrepancy suggests either a structural difference among the studied mAbs affecting the reactivity of certain cysteines or other mechanisms, such as preferred de-conjugation of drug molecules bound in the inter HL region, that cause the reforming of these bonds and herein being considered as re-oxidation. A further consistent improvement, except for the prediction of intact mAb, was achieved when including the intermediates with an average nRMSE reduction of 16% and 8% compared to the *NoInt+ReOx* model. The consideration of the reduced intermediates agrees with the structure of the aforementioned reduction kinetic model by Nayak and Richter [130]. Although these intermediates are

not directly distinguishable by the CGE method, their consideration in the CRN contributes to improving the model, particularly concerning the intermediates HH and HL, which can be attributed to higher model flexibility due to the increased number of possible reactions.
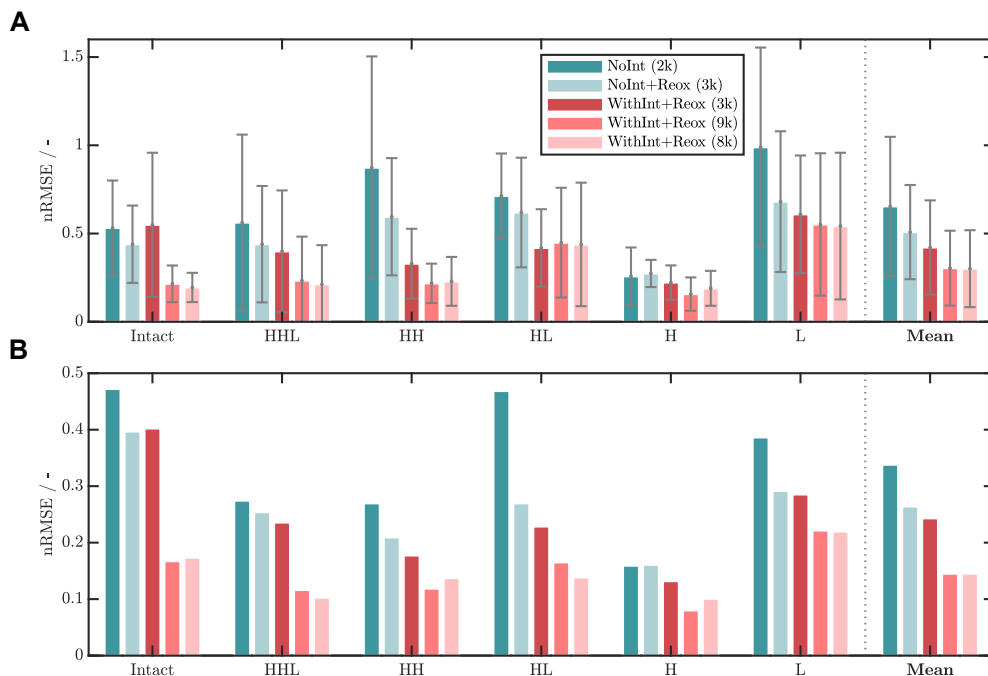

Figure 4-2: Comparison of the five evaluated reduction models with regards to the species-resolved and the mean nRMSE for the cross-validation (A) and the test runs (B). For the cross-validation, the bar represents the average nRMSE and the error bars indicate the standard deviation over all CV rotations.

Therefore, the *WithInt+ReOx* CRN was subsequently chosen for further model finetuning using two additional kinetic rate discriminations, 8k- and 9k-model. In comparison to the *WithInt+ReOx* with three k's, both sets of kinetic rates demonstrated similar error reductions of approximately 27% and 40% for cross-validation and test set, respectively. It was observed that this improvement primarily originates from an enhanced prediction accuracy for Intact, HHL and HH. The lowest AICc value was achieved for the 8k-model with 24.96 in comparison to the 9k-model with 35.84 suggesting a reasonable accurate model fit while minimizing the model complexity. The final use of the 8k-model for the further modeling purpose was justified by its lowest AICc value.

4.3.1.2 Mechanistic insights

Insights into the kinetic mechanisms can be derived from the estimated kinetic rates. Figure 4-3 displays the estimated rates $k_{i,20°C}$ for the 8k-model (estimated E_A values are given in the Supplementary Materials). The kinetic rates are grouped according to the three superordinated reaction steps. The three rates associated with the inter HL disulfide reduction for the three species, Intact (including IntactR), HHL (including HHLR), and HL, showed a decreasing

trend from $6.98 \cdot 10^{-4}$ L (µmol s)$^{-1}$ to $6.81 \cdot 10^{-5}$ L (µmol s)$^{-1}$. These differences might be explained by the different amount of available disulfide bonds (two in case of Intact, one for HHL and HL) and that each reduced disulfide bond might affect the conformational structure of the antibody resulting in a different susceptibility for TCEP reduction. Furthermore, it can be observed that the first inter HL reduction is faster compared to the first inter HH reduction. This is in agreement with Chen et al. [70] who found that the inter HL reduction is kinetically favored over the HH reduction. A reason for this could be the higher solvent accessible surface area, as demonstrated with MD simulations [159]. With regards to the inter HH reduction, the kinetic rate of the second disulfide reduction, k_5, increases 18-fold compared to the first disulfide reduction k_4. This agrees with findings by both Liu et al. [169], who suggest a conformational change in the hinge region once one inter HH bond is reduced resulting in greater exposure, and Nayak and Richter [130], who hypothesize a "zipper"-like effect due to the increased flexibility in the CH2 domain.
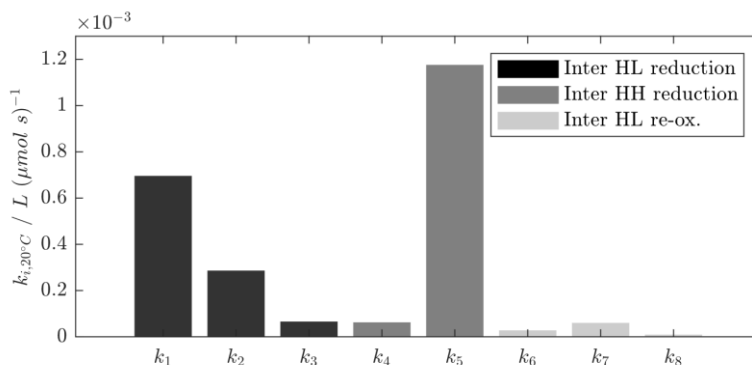

Figure 4-3: Estimated kinetic rates at the reference temperature of 20 °C for the 8k-model.

An observation of the CIs of the estimated parameters (given in the Supplementary Materials, Table S4) , showed that CIs for most of the parameters range between 25% and 60% which indicates an acceptable parameter quality [75]. Considerable wider CIs, especially for the E_a parameters associated with the re-oxidation reactions, were observed suggesting that the data is not informative enough to precisely estimate these parameters. This could have multiple root causes such as the consideration of non-observable intermediates or the analytical error of the CGE method, which tends to increase for lower concentrations [170]. The CIs of the E_a parameters are expected to decrease when more temperatures are added.

4.3.1.3   Kinetic model predictions

A major advantage of a kinetic model is the ability to describe the reacting species' time-course for varying process parameters. To thoroughly assess the model's performance on the reduction kinetic, the experimental vs. predicted

kinetics and with the resulting species nRMSE for each cross-validated run using the 8k-model are plotted in Figure 4-4.
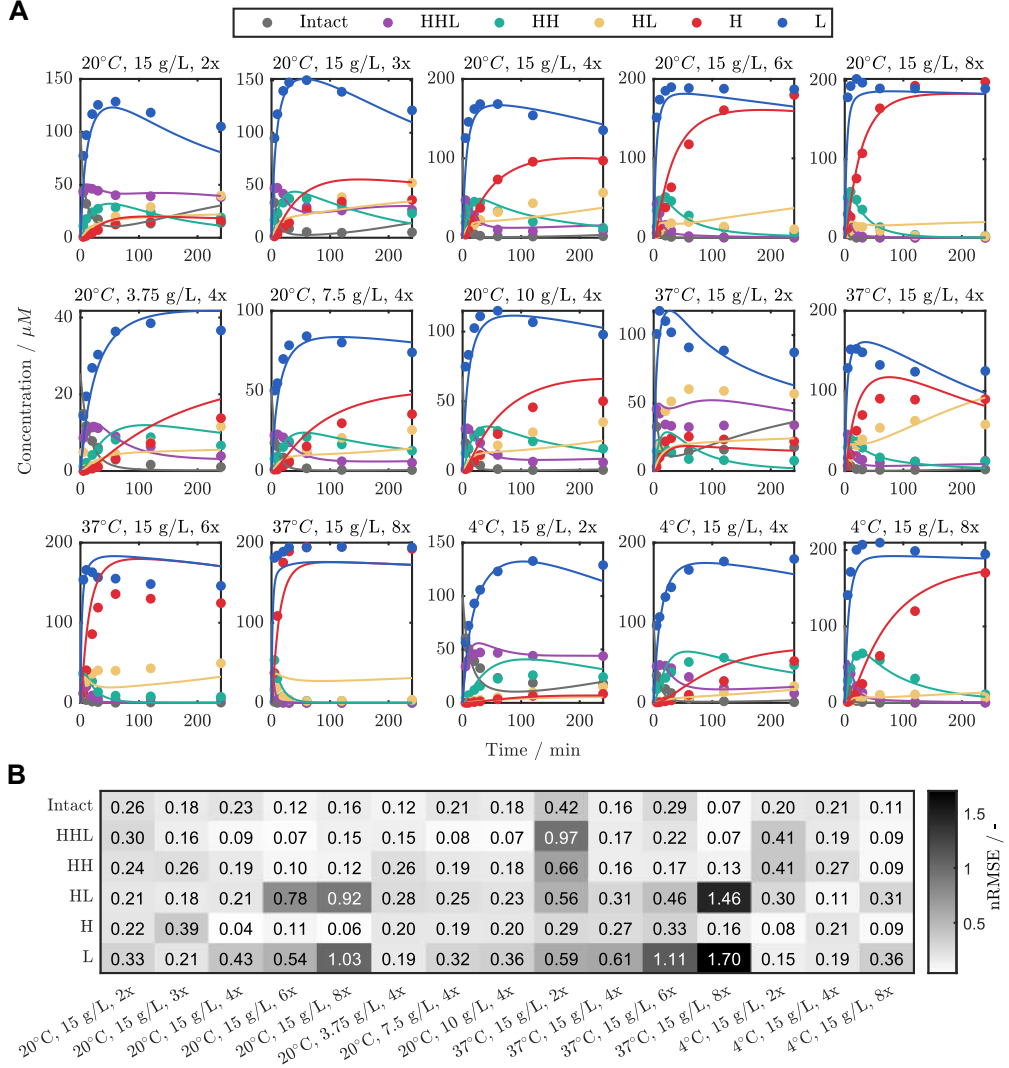

Figure 4-4: (A) Model predictions (lines) vs. experimental CGE data (markers) for the concentrations of the six reduction species for all cross-validation runs. (B) Heatmap of the cross-validation nRMSE resolved for each individual cross-validation run in each run.

According to the CGE data, an increasing TCEP excess generally led to faster generation and higher ratios of H and L in the reduction reaction, while an increase in temperature caused faster kinetics and more pronounced re-oxidation reactions. This dynamic occurred mainly for the species L at lower TCEP excesses. Interestingly, the species HH is generated at higher concentrations with decreasing temperatures, while at the same time the species HL appears to be inversely affected by the temperature. This contradictory influence of the temperature reflects the impact of the varying temperature-dependencies on the individual reaction pathways. A similar behavior for HH/HL was observed in a related process characterization study [158].

73

The model represents the observed trends in the CGE data accurately. When comparing the predictions for the three temperatures, it appears that the model predictions are more accurate at 4 °C and less accurate at 37 °C with average nRMSE values of 0.21 and 0.47, respectively. This decreased performance specifically originated from a mismatch primarily associated with the species HL and L, which can also be observed by comparing the species nRMSE. The decreased performance on the 37 °C kinetics might be caused by non-parametrized effects increasingly occurring at higher temperatures, such as de-conjugation [171] or side reactions between TCEP and maleimide-payloads [172]. These phenomena may also contribute to the generally higher inaccuracy for the species HL and L. However, these effects were not addressed in this study as they could not be proven using the available data. The model shows moderate performance for kinetics at 20 °C with an average nRMSE of 0.28, while the main portion of the error is attributed to the run with 8x TCEP. The runs with lower mAb concentration generally exhibit good performance with an average nRMSE of 0.20.

4.3.1.4 Comparison of mAb1 and mAb2

The external data of mAb2, were used to assess the comparability of the reduction kinetics between the two IgG1 mAbs. When using the same experimental conditions, slight differences in the reduction kinetics of the species were observed (Supplementary Materials, Figure S3). This might be caused by the amino acid sequence causing slight differences in structure or disulfide bond exposure, thus showcasing the need for model re-calibration if applied for another mAb. As a detailed comparison was out of the scope of this study, this was not further analyzed.

4.3.2 End-point DAR and DLD prediction using MLR models

To predict the DAR and the six DLD-% of the conjugation endpoint, MLR models were developed using either experimental CGE data or the reduction model-predicted concentrations as input. Table 4.3 presents the resulting R^2 values for each MLR model concerning the subsets and the model input. It has to be noted, that when using the model predictions as input, mAb2 was not further evaluated due the observed discrepancy in applying the reduction model. With regards to the experimental values, the MLR model exhibited accurate prediction of the DAR, with R^2 values close to 1 for both the training and test set. The models for DLD-% similarly achieved good performance with an average R^2 of 0.973 and 0.961 for the training and test set, respectively. Generally, the prediction accuracy was highest for %H0 and %H3, while the lowest R^2 value was observed for %H2, with an R^2 of 0.893 for the test set.

Table 4.3: Summary of R^2 values for the seven MLR models using either experimental CGE data (training, test and external data (mAb2)) or reduction model-predicted data (train and test) as input. The mean DLD corresponds to the mean of the six DLD species (L0, L1, H0, H1, H2 and H3).

Input	Experimental CGE data			Reduction model-predicted data	
Output	R^2_{train}	R^2_{test}	$R^2_{external}$	R^2_{train}	R^2_{test}
DAR	0.996	0.995	0.981	0.976	0.972
%L0	0.987	0.981	0.940	0.905	0.951
%L1	0.987	0.981	0.940	0.905	0.951
%H0	0.993	0.990	0.922	0.954	0.950
%H1	0.976	0.932	0.961	0.938	0.891
%H2	0.901	0.893	0.937	0.703	0.703
%H3	0.994	0.991	0.947	0.971	0.970
Mean DLD	0.973	0.961	0.941	0.896	0.903

The lower performance with respect to H2 is likely to be caused by an increased analytical error due to relatively low %H2 peak levels. In total, these findings highlight the capability of the MLR models to accurately predict the DAR/DLD-% directly from the CGE data. To the best of our knowledge, no comparable method to link CGE and RP data for ADCs has been presented in literature. A strong correlation between CGE and UHPLC was also shown for other applications [173], [174]. Additionally, good transferability of the proposed method is shown when applying the MLR models to another mAb. Only minor losses in accuracy across all models, except H2, were observed, which is likely to be caused by slight differences in the molecule structure of the mAbs affecting the relationship between CGE and RP. Higher accuracy is expected when training the models on datasets with a larger portion of mAb2 and/or using more advanced regression techniques that conduct an internal mAb-classification. In summary, our model approach presents a novel method to precisely forecast the RP results from CGE data with minimal practical effort.

When the MLR models were trained using the model-predicted concentrations, the accuracy is slightly reduced compared to the models using the experimental data. Still, the DAR value as well as %H0 and %H3 are precisely predicted with R^2 values above 0.95. The R^2 values for the other DLD-% species is slightly decreased, with %H2 exhibiting the lowest R^2. To further investigate the model's predictability uniquely for each run, the observed vs. predicted plots for the DAR and the DLD-% models are shown in Figure 4-5. In general, these

plots confirm that for DAR the approach exhibits relatively higher precision compared to the DLD-%, with the lowest performance for %H2. Moreover, it becomes clear that specific runs, such as Run5, 11 and 12 contribute mainly to the larger error as indicated by their systematic offset from the regression line. This could be attributed to their larger error in the reduction kinetic model concerning certain species being propagated with the MLR model.
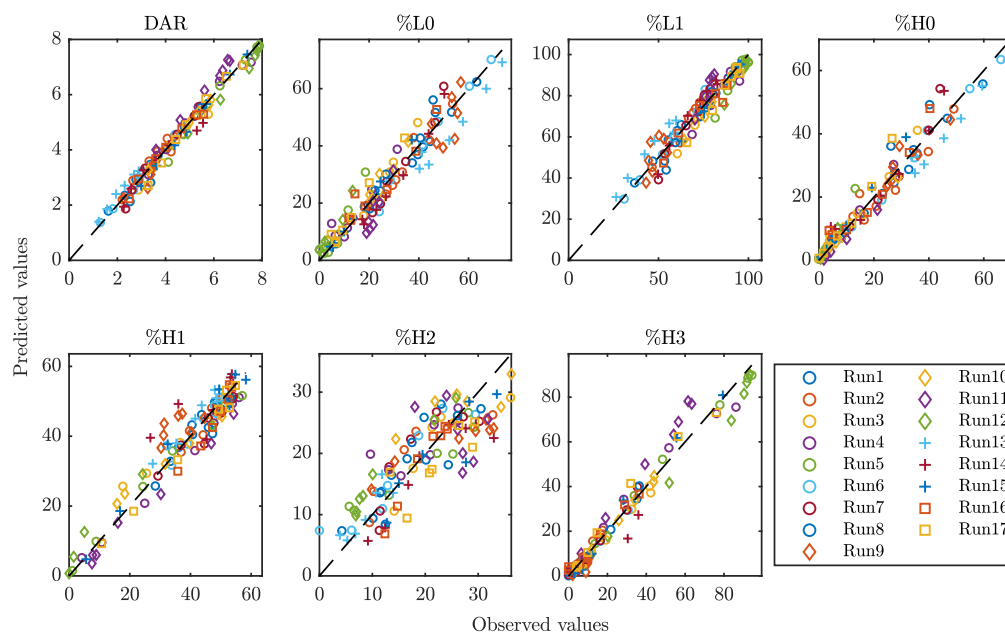

Figure 4-5: Observed vs. predicted DAR values and DLD percentages of the MLR models using the normalized model-predicted concentrations as input. The markers are grouped according to the four temperatures (circles: 20 °C, diamonds: 37 °C, crosses: 4 °C and squares: 12 °C).

In total, these results highlight the capability of the MLR models to accurately predict the DAR/DLD-% also from the newly developed reduction model. For both model inputs, the models achieved better performance on the prediction of DAR compared to DLD-% which can be explained by the fact that the DAR value is calculated from all DLD-%, so that individual errors are compensated making it a more robust feature. A comparison of the model coefficients for both cases, demonstrated that the inputs species are assembled differently in the two model types (data not shown). Combined with the finding that the MLR model had to be trained independently for each instance, this suggests that the model adapts to the kinetic model error and, thus, acts as an internal model correction mechanism.

4.3.3 Integrated in-silico reduction screening

Kinetic models are particularly beneficial for understanding, predicting, and evaluating the effects of reaction parameters due to their ability to represent complex biochemical behavior. The combination of the newly developed

reduction model with the different MLR models allows for the comprehensive evaluation of the effect of the reduction parameters on the resulting DAR and DLD-%. The resulting predictions for the DAR and two exemplary DLD-%, %L1 and %H2, based on the reduction parameter screening for the two case scenarios is provided in Figure 4-6. For constant reaction times (Figure 4-6A), DAR is shown to increase with both increasing TCEP excess and temperature, reaching a plateau at the highest TCEP excess and temperature. With regards to %L1, a similar behavior compared to the DAR is observed with a slightly different influence of the temperature at lower TCEP excesses. In contrast, %H2 exhibits an inverse relationship to both parameters with a maximum around 3.5x TCEP and 37 °C. The root cause for this behavior is the different temperature-dependency of the kinetic rates of inter HL and HH reduction enabling to direct conjugation to heavy or light chain. These results showcase the benefits of the proposed screening approach. While the temperature appears to have minor impact on the DAR, it seems to have a larger impact on the ratio of conjugation to light and heavy chain, especially for lower target DAR values. Knowledge of this trend is especially useful when the goal is to decrease the ratio of certain undesired species in the DLD.
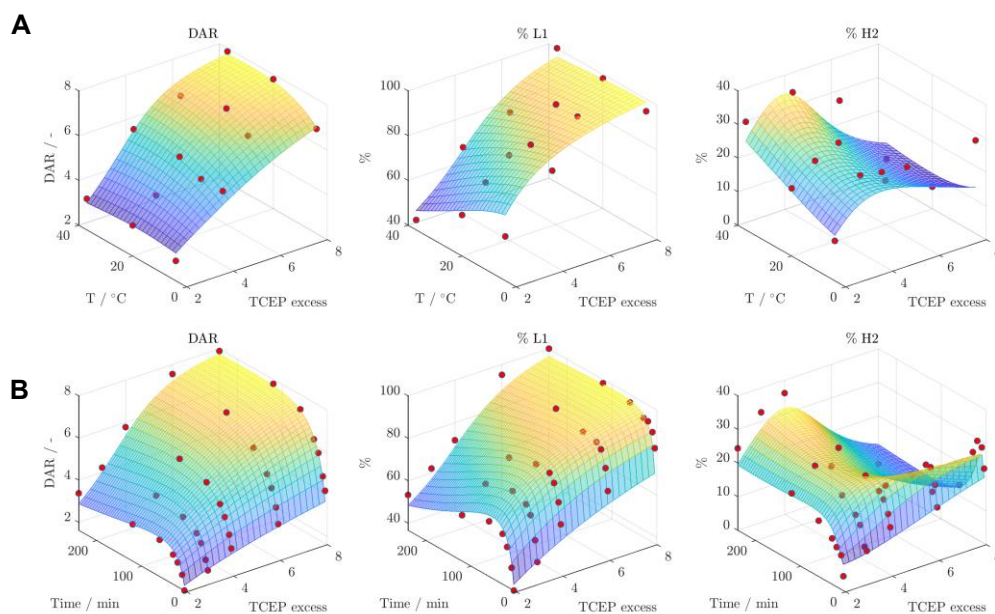

Figure 4-6: Integrated *in silico* screening for two case scenarios: (A) Varying temperature and TCEP excess for constant $c_{mAb} = 15$ g/L and $t = 120$ min, and (B) Varying reaction time and TCEP excess for constant $c_{mAb} = 15$ g/L and $T = 20$ °C. The surface plots represent the model predictions for DAR (left panel), %L1 (middle panel) and %H2 (right panel) including the experimental RP values (red spherics). Blue colors indicate lower values and yellow indicate higher values.

Furthermore, the conjugation results for a constant temperature but for varying TCEP excess and reaction time are given in Figure 4-6B. For both DAR and %L1, the values rapidly increase for longer reaction times to a maximum and then decreases at lower TCEP excesses. This is due to the re-oxidation reaction,

which is particularly important for low TCEP excesses. On the contrary, %H2 has a distinct profile with a maximum at high TCEP and low reaction times, and a plateau at a TCEP excess of 4x. For longer reaction times, %H2 is only slightly affected, as the re-oxidation is not occurring at the inter HH bonds. These results indicate that, in addition to the temperature and the TCEP excess, the reduction time also needs to be carefully chosen for a desired DAR value. Overall, the *in silico* modelled surfaces capture the dynamic in the experimental data (red spheres) across the design space, with a comparably higher scattering for %H2. The comparably higher inaccuracy for %H2 might be due to the lack-of-fit of the reduction model and the lower accuracy of the MLR model for %H2. These results showcase the benefit of the dual modeling approach for gaining complementary process understanding.

4.3.4 Process optimum for desired DAR and narrow DLD

To find appropriate reaction conditions for a desired DAR of 3 while simultaneously keeping the DLD as narrow as possible, the objective was to find the optimal reaction time over the whole design space of temperature and TCEP excess, while minimizing the %H2 according to the integrated model predictions.
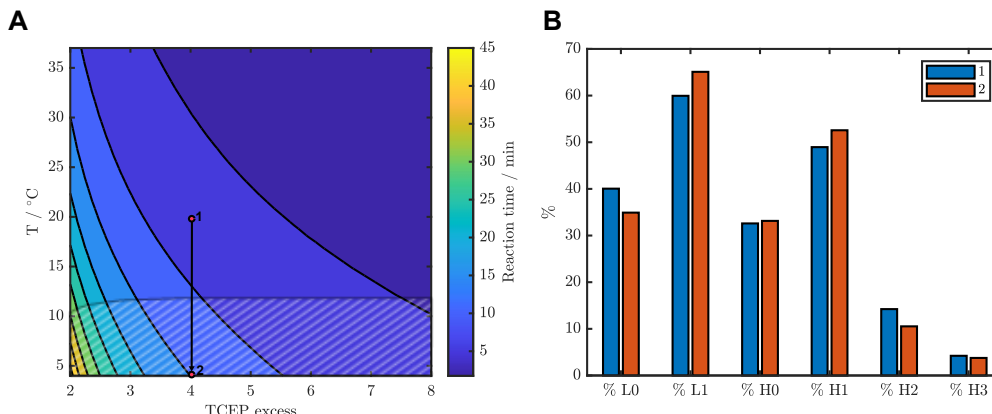

Figure 4-7: Contour plot of the *in silico* determined reaction times for a target DAR of 3 over the whole design space (A) and the resulting predicted DLD for two scenarios (B). The shaded area in (A) depicts the temperature-TCEP region for which the percentage of H2 is below 12%. The scenarios "1" and "2" depict two exemplary reduction conditions for the determined reaction time: Condition 1: T = 20 °C and 4x TCEP, condition 2: T = 4 °C and 4x TCEP.

Figure 4-7A shows the estimated reaction times as contour plot over the temperature-TCEP range. It can be seen that in the region of higher temperature and TCEP excess reaction times are short (around 5 min), while the reaction times are longer towards lower temperature and TCEP excesses. This decreasing trend in the reaction times is mainly driven by the slower reaction rates at lower temperatures and for less reducing agent. The region for

which the required level of H2 is below 12% is mainly below a temperature of approximately 12 °C. To demonstrate the successful adjustment of the final DLD for the same DAR value with changing only the reduction temperature, two scenarios were considered. Figure 4-7B shows the predicted DLD of the two scenarios. While for condition 1, the DLD is wider containing higher levels of L0 and H2, the DLD for the improved condition 2 is narrower with lower levels of L0 and H2. This behavior results from the observed effect of the temperature on the level of H2. These two scenarios highlight the capability of the integrated model to effectively simulate alternative reaction conditions to fine-tune the DLD while achieving the same DAR value. Overall, this screening showcases the benefit of the designed framework of the two models to effectively cover two consecutive reaction steps and gain useful insights in the complexity of reaction behavior with low experimental effort. Although the conjugation reaction conditions were set constant throughout this study, they could also be varied requiring the integration of the reduction model with a conjugation kinetic model [175].

4.4 Conclusion and outlook

This study presents the development of a reduction kinetic model for the partial reduction of the mAb interchain disulfide bonds, along with MLR modeling to directly forecast the resulting conjugation outcome at each reduction state. The dataset included reduction kinetic studies under various reaction conditions, analyzed by both CGE and RP-UHPLC for quantification of the reduced species and DAR/DLD.

We first developed a reduction kinetic model by considering the CRN structure and the set of kinetic rates. Evaluating different possible CRNs revealed that the reduction reaction is driven by intermediates and re-oxidation pathways, necessitating their incorporation to accurately project the reaction dynamics. Further improvements were achieved by isolating kinetic rates. This is particularly important in the sequential inter HH reduction pathway, where an acceleration of the kinetic rate for the reduction of the second inter HH bond was observed upon the reduction of the first inter HH bond, which is in accordance with recent literature findings. Adopting this method to other reducing agents would be highly interesting, as it is known that reducing agents differ in their reduction preference of certain disulfide bonds [157]. Additionally, more sophisticated analytical approaches offering higher resolution regarding the species mass could be utilized to resolve the non-observable intermediates by CGE and validate our mechanistic findings. The temperature dependency modeled by the Arrhenius equation was unable to capture biochemical effects at higher temperatures. Further detailed experimental investigation in this

temperature range or the exploration of empirical approaches to capture the temperature dependency should be considered. Assuming the same reaction pathway for two different mabs of the IgG1 class was proven appropriate, although re-calibration is necessary to achieve same level of accuracy. Comparing the reduction kinetics of different classes of mAbs for developing multi-class models would be interesting for future research.

We second established MLR-models to predict the RP-derived final DAR and DLD-% from the experimental CGE data, exhibiting remarkable performance across all outputs and the entire experimental range. Using the kinetic predictions from the newly developed kinetic model as input for the MLR models enabled direct prediction of the DAR and the DLD-% obtained after the two-step ADC reduction-conjugation workflow. However, a slight decrease in performance was attributed to the incorporation of the lack-of-fit of the reduction model, suggesting potential improvement with a more accurate reduction model.

The hereby created integrated kinetic model provided detailed mechanistic understanding of the complex DAR/DLD-% behavior along the reduction kinetic under varying reduction conditions. Furthermore, this framework enabled the determination of multidimensional reaction conditions to achieve a target DAR and minimize undesired species in the DLD.

In summary, our presented methodology presents an innovative approach to capture the temperature-dependent disulfide bond reduction kinetic with a simultaneously evaluation of the possible conjugation outcome. To further enhance the utility of the reduction model, additional effects, such as pH-sensitivity of the reaction rates could be parametrized, and a detailed comparison with a DoE concerning the predictability and insights gained could be conducted.

Acknowledgments

Supplementary Materials

Appendix B contains the Supplementary Materials associated with this chapter:

- S4.1 Overview of the reaction pathways and the assigned reaction rates
- S4.2 Exemplary ODE
- S4.3 Parameter confidence intervals for the 8k-model

- S4.4 Evaluation of reduction model on mAb2
- S4.5 Pearson correlation matrix between CGE and RP
- S4.6 All DLD-% results for the Integrated in-silico screening

<div style="text-align: right; font-size: 4em;">5</div>

Kinetic studies and CFD-based reaction modeling for insights into the scalability of ADC conjugation reactions

Jan Tobias Weggen[1], Janik Seidel[1], Ryan Bean[2], Michaela Wendeler[2], Jürgen Hubbuch[1]

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

[2]Purification Process Sciences, BioPharmaceuticals Development, R&D, AstraZeneca, Gaithersburg, Maryland, USA

Abstract

The manufacturing of antibody-drug conjugates (ADCs) involves the addition of a cytotoxic small-molecule linker-drug (= payload) to a solution of functionalized antibodies. For the development of robust conjugation processes, initially small-scale reaction tubes are used which requires a lot of manual handling. Scale-up to larger reaction vessels is often knowledge-driven and scale-comparability is solely assessed based on final product quality which does not account for the dynamics of the reaction. In addition, information about the influence of process parameters, such as stirrer speed, temperature, or payload addition rates, is limited due to high material costs. Given these limitations, there is a need for a modeling-based approach to investigate conjugation scale-up. In this work, both experimental kinetic studies and computational fluid

dynamics (CFD) conjugation simulations were performed to understand the influence of scale and mixing parameters. In the experimental part, conjugation kinetics in small-scale reaction tubes with different mixing types were investigated for two ADC systems and compared to larger bench-scale reactions. It was demonstrated that more robust kinetics can be achieved through internal stirrer mixing instead of external mixing devices, such as orbital shakers. In the simulation part, 3D-reactor models were created by coupling CFD-models for three large-scale reaction vessels with a kinetic model for a site-specific conjugation reaction. This enabled to study the kinetics in different vessels, as well as the effect of process parameter variations in-silico. Overall, it was found that for this conjugation type sufficient mixing can be achieved at all scales and the studied parameters cause only deviations during the payload addition period. An additional time-scale analysis demonstrated to aid the assessment of mixing effects during ADC process scale up when mixing times and kinetic rates are known. In summary, this work highlights the benefit of kinetic models for enhanced conjugation process understanding without the need for large-scale experiments.

5.1 Introduction

Antibody-drug conjugates (ADCs) are highly potent biopharmaceuticals that combine the targeting specificity of a monoclonal antibody with the potent cytotoxicity of chemotherapy. In the last decade, ADCs have made considerable progress: In 2021, ten ADCs were approved by the Food and Drug Administration (FDA) and >80 are in clinical trials [118]. For the coupling of the cytotoxic drug/payload to the monoclonal antibody (mAb), functional groups such as lysine residues or free thiols after reduction of interchain disulfide bonds are used [21]. However, these conjugation techniques often cause heterogeneous drug-load profiles and a variety of positional isomers [32], [60]. In the development of scalable and robust ADC processes, a major challenge is the characterization of critical process parameters (CPP) in each synthesis step that impact critical quality attributes (CQAs) [60]. Especially, the drug-to-antibody ratio (DAR), drug load profile and aggregate level are important CQAs as they directly influence the product safety, efficacy and pharmacokinetics [137]. Different site-directed conjugation methods have been developed that aim to synthesize more homogenous ADC products, control the site of attachment, and achieve more stable conjugates [50], [64]. But even for these methods, product-related impurities, such as under- and over-conjugated species or aggregation still occur [53], [65], [143].

At the same time, regulatory agencies increasingly promote the understanding of both product and process already in the development phase according to the

concept of *Quality by Design (QbD)* [71]. As the needed intermediates for ADCs are costly and difficult to handle, scale-down models (SDM) are often used in process development. These models are typically designed by selecting one scale-down parameter to be similar along scales. However, this approach becomes difficult for larger scale differences, because certain factors, such as the power input per volume ratio (P/V), are impractical to be kept constant. [176]–[178]. For ADCs, no systematic approach has been reported and scale-comparability is assessed based on constant ADC quality attributes, such as the DAR or aggregate level [53].

Due to the ongoing digitalization of bioprocesses, the use of process models describing complex biopharmaceutical processes are promoted [179]. Different types of statistical or computational approaches, such design of experiment (DoE), mechanistic or hybrid models and computational fluid dynamics (CFD) were recently applied, also aiming to extrapolate beyond the design space and predict larger scales [121], [160], [180]. Within the last years, computational fluid dynamics (CFD) have gained more attention for (bio)-reactor scale-up due to the ability to provide high resolution results of the complete flow pattern at various scales [181], [182]. The goal is to establish an advanced process model that allows to examine the effects of scale, turbulence, and mixing parameters completely *in-silico*. In the field of biotechnology, CFD was recently applied to study bioreactor mixing performance [83], [84], [183], predict large-scale mixing times and oxygen mass transfer [181], [184]–[186] and explore inhomogeneity effects on the cell metabolism [187]. However, most studies focus on comparably slow bio-chemical processes, such as fermentation, with characteristic times in the range of min to hours, while typical chemical reactions being significantly faster (down to nano-sec). There are only a few cases, where CFD and mechanistic models were coupled to predict the mixing effect on the course of chemical reactions [188]. Due to high computational demand elegant ways to minimize the computational effort by using compartment modeling [189], [190] or surrogate models [191] were also developed.

In the field of ADCs, mostly statistical approaches, such as design of experiments (DoE) are utilized to optimize the process and gain a solid understanding of the CPP-CQA relationship [137]. For scale-up prediction, the use of process models has not been reported yet. Especially the conjugation step is considered to demand adequate mixing and careful considerations of the payload addition method or rate. In addition, single-use reactors are becoming more commonly used in ADC manufacturing [192], [193]. However, their reactor design and mixing geometry is different from conventional stirred vessels which might add further possibilities to affect the product. In an earlier study [194], CFD models were generated to compare different reaction vessels based on mixing times simulations, while the importance for ADC reactions was not

discussed. Recently, a mechanistic kinetic model for a site-directed conjugation was developed by Andris et al. [140], showing that the model could successfully optimize concentrations and reaction times. Since the dataset was limited to small-scale experiments and not compared to large-scale data, evidence of the potential to predict scale-up parameter is still lacking. In addition to the differences in size and geometry, manual handling in small scales vs. a higher degree of automation in larger scales may result in fundamental differences in flow characteristics and mass transfer. These effects might impact the conjugation reaction.

Here, we present a thorough investigation of scale effects and mixing parameters on the course of ADC conjugation reactions by applying experimental kinetic studies and the coupling of CFD models with an ADC conjugation kinetic model. In the experimental kinetic studies, the influence of mixing in typically used small reaction tubes is analyzed for two model ADCs aiming to produce robust conjugation kinetics. One ADC has a target DAR value of 2, and conjugation is achieved by site-specific attachment to inserted cysteine residues. The second ADC has a target DAR value of 8, and the conjugation workflow is based on stochastic conjugation to interchain disulfide bonds after reduction. It is shown, how well this test tube scale (~1 ml) mimics the conjugation reaction in a glass reactor at lab-scale (100 mL) at industrial-relevant concentrations. In the second modeling part, CFD simulations are performed for three differently sized vessels typically used for pilot and large-scale ADC manufacturing (up to 50 L), namely two conventional glass stirred vessels and one single-use vessel. The steady-state results and mixing time simulations are considered for a CFD-based vessel comparison. Subsequently, a kinetic model for the site-directed conjugation reaction is incorporated in the CFD models resulting in a full 3D reactor model and is further used to study *in-silico* how scale and process parameters affect the course of the conjugation reaction.

5.2   Materials and methods

5.2.1   Experimental conjugation kinetic studies with two model ADCs

Two types of kinetic studies with ADC1 and 2 were conducted: (1) Mixing kinetic studies to determine optimal mixing conditions for small-scale conjugation reactions by using either external mixing or internal mixing. (2) Conjugation kinetics with the optimized small-scale conditions vs. lab-scale conjugation to evaluate the scale comparability.

5.2.1.1 Chemicals, ADCs and functionalization steps

Two ADCs were investigated within the experimental part of this study: ADC1 with two engineered cysteines for a site-directed DAR 2 conjugation and ADC2 for a cysteine-linked DAR 8 conjugation. For DAR 2, a functionalized mAb solution was generated through a full reduction with tris(2-carboxyethyl) phosphine hydrochloride (TCEP, EMD Millipore), followed by a buffer exchange using Vivaspin 20 (30 kDa MWCO, Cytiva) and a re-oxidation of the interchain disulfides with (L)-dehydroascorbic acid (DHAA, Sigma-Aldrich). For the ADC2 with a DAR of 8, a mild reduction of the interchain disulfides with TCEP was performed. In both cases, conjugation was carried out with a maleimide-functionalized payload that was dissolved in DMSO (Sigma-Aldrich). All other solutions were prepared with 20 mM sodium phosphate buffer (J.T. Baker), 1 mM EDTA (EMD Millipore), pH 7.0.

5.2.1.2 Conjugation kinetics

For all studies, functionalized mAb solution was prepared by the procedure described above. MAb solutions with ADC1 were diluted to a concentration of 10 mg/mL and conjugated with 5x molar payload excess. In case of ADC2, a lower concentration of 1.5 mg/mL was tested, and conjugations were performed with 11x molar payload excess. As mixing vessel for the small-scale experiments, reaction tubes (1.5 mL Safe-Lock tubes, Eppendorf) were used. Preliminary studies with ADC2 showed that thorough initial payload mixing is required to prevent lower DAR values or inconsistent kinetics, especially when an orbital shaker is used for subsequent mixing (Appendix C, S5.1). To solely assess the influence of the final mixing, the initial payload mixing was conducted by one end-over-end rotation of the tube (tube rotator, VWR). Final mixing was either achieved by "external" mixing (Eppendorf thermomixer C) or internally with a magnetic stir bar (Magnetic stir bar, Merck, Part #23226) placed inside the tube. Two shaking/mixing speeds for each condition were tested. Conjugations with ADC2 were performed in duplicates.

Lab-scale conjugations were performed to compare the kinetics with the optimized small-scale mixing system. The mAb concentrations were set to 5 mg/mL (ADC1) or to 20 mg/mL (ADC2) and the same payload excesses as in the small-scale were used. The stirrer-based mixing was found to lead to more ideal kinetics in the small-scale reactions and was therefore used for the comparison to the lab-scale kinetics. Lab-scale conjugations were conducted with 100 mL mAb solution in a stirred glass reactor (Chemglass, inner diameter = 108 mm, Model CG-1949-x-300), that was also included in the CFD study (later referred to as GST-1). The anchor stirrer (Chemglass, impeller diameter = 81 mm, Model CG-2081-A-04) was installed so that the stirrer was close to

the bottom surface. The stirrer speed was set to 60 rpm and payload solution was manually added with a pipette to the stirred mAb solution.

5.2.1.3 Reference analytics

To obtain data on conjugation kinetics, samples were taken at defined timepoints over 1 h and immediately quenched with N-Acetyl cysteine (Sigma-Aldrich). Each sample was further treated with reducing buffer, incubated at 37°C for 30 min and analyzed using reversed-phase ultra-high-performance liquid chromatography (RP-UHPLC). A detailed description of the applied protocol, method and chromatography system can be found in [143]. The DAR was calculated based on the peak areas of unconjugated/conjugated light and heavy chain peaks.

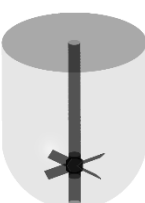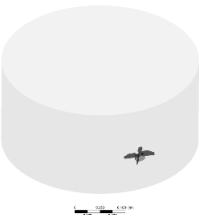## 5.2.2 CFD simulations for large-scale vessels

Multiple CFD simulations were performed, and their individual purpose is shortly described in the following. First, steady-state and transient mixing time simulations were conducted to characterize three industrially relevant mixing vessels. Validation of the CFD models was done based on available mixing times. Next, a calibrated kinetic model for the site-directed conjugation reaction was incorporated in the existing CFD models. This enabled to study the direct impact of mixing geometry and scale on the conjugation kinetic and is referred to as 3D-model. The significance of the 3D-model to accurately describe the conjugation reactions was estimated by comparing the predicted kinetics with the 0D-model (ideal mixing assumption). For one of the studied vessels the influence of varying process parameters was further exemplarily studied. Due to GMP limitations, a validation of the large-scale conjugation kinetic could only be performed for the smallest mixing vessel.

5.2.2.1 Geometries and meshes

The studied geometries comprise three disparate, unbaffled vessels: A 300 mL "lab-scale" glassed stirred tank (GST-1) equipped with an anchor stirrer (Chemglass, Model CG-1949-x-300), a 50 L "large-scale" glass stirred tank (GST-2) equipped with a 45° pitched-blade stirrer (Chemglass, Model CG-1968-81) and a 50 L "large-scale" single-use mixer (SUM) equipped with an eccentric bottom-mounted agitator (Mobius MIX Bag, Merck Millipore). Liquid volumes and stirrer speeds were selected to be comparable to real process conditions. All geometries were designed in ANSYS DesignModeler. An overview of the vessels and the parameters is given in Table 5.1. The exact dimensions are tabularized in the Appendix C, S5.2. For the GST-1, the whole volume was modelled as a single rotating frame having the same rotational speed as the stirrer. For both large-scale vessels (GST-2 and SUM), the fluid domain was divided into two

zones, a cylindrical rotating zone around the impeller and a stationary zone for the remaining volume, to model the stirrer motion using the multiple reference frames (MRF) approach. The water surface was assumed to be flat. The geometries were discretized with Poly-Hexcore meshes using the integrated FLUENT mesher. Five boundary layers of prism cells were applied for the walls (vessel, impeller and shaft) in order to resolve the transition of the flow in the near-wall region. The mesh close to the impeller was further refined because of higher expected gradients. The resulting meshes consisted of approximately 687,000 (GST-1), 26,000 (GST-2) and 38,000 (SUM) mesh elements per liter. The higher mesh cell density for the GST-1 was due to the larger impeller area relative to the volume which had to be refined. Overall, a minimum orthogonality of 0.2 or greater and a maximum skewness less than 0.8 was achieved. To judge sufficient spatial discretization, mesh independency tests were performed for each vessel at the highest stirrer speed based on global average velocity magnitude and turbulence parameters (Appendix C, S5.4). The final meshes and mesh metrics are depicted in the Appendix C, S5.3.

Table 5.1: Geometries and parameters of the studied reaction vessels. *For GST-2, three stirrer speeds were investigated in the parameter study. **For the SUM a lower stirrer speed of 250 rpm was added because mixing time data were available only for this stirrer speed.

3D view	Name	Volume liquid	Impeller type	Speed / rpm	Re_{imp}
	Lab-scale glass stirred tank (GST-1)	300 mL	2 blades, anchor-style, centric	60	6218
	Large-scale glass stirred tank (GST-2)	22 L	4 blades, pitched bladed (45°), centric	60, 80*, 120*	22141, 29522, 44282
	Single-use mixer (SUM)	25 L	4 blades, 15° angle, bottom mounted, eccentric	250**, 400	15765, 25224

5.2.2.2 Steady-state simulations and P/V ratio

Steady-state CFD simulations were performed for the predictions of the stationary flow field. All simulations were run using the finite volume method with pressure-based solvers in ANSYS Fluent v2020 R2. Turbulence models are required since all impeller Reynolds numbers (Re_{imp}) are in the transitional or fully turbulent regime. In a preliminary study, two frequently used Reynolds-averaged Navier Stokes (RANS) models, namely the k-ε-RNG and the Reynolds Stress Model (RSM) model, were compared. Since the k-ε-RNG was found to be more stable and achieve similar results in less computational time, it was applied in this work. The physical properties of the fluid were assumed to be equal to water with 10 % DMSO (a typical media composition for conjugation reactions). The density was set to 1010.5 kg m^{-3} and the dynamic viscosity to 0.00106 kg m^{-1} s^{-1}. The liquid surface was defined as no-shear (free slip), tank and stirrer walls were treated with zero velocity (no-slip) boundary condition. The near-wall region was modeled with standard wall function. The SIMPLEC algorithm was used for pressure-velocity coupling. Further, the second-order upwind scheme for interpolation and Green-Gause node based for gradient determination were used. The simulations were run for at least 10,000 iterations and convergence was judged based on continuity of volume-averaged velocity magnitude, impeller torque and turbulent energy dissipation (ε) as well as scaled residuals. The flow was assumed to be stationary when no considerable deviation of these values was observed (data are shown in the Appendix C, S5.5). The stationary impeller torque (M) was used to calculate the simulated P/V:

$$\frac{P}{V} = \frac{2\pi * N * M}{V} \tag{1}$$

where N is the stirrer speed and V is the liquid volume.

5.2.2.3 Mixing time studies – computational

Tracer simulations using the species transport model were performed for computational mixing time studies. The flow in each vessel was initialized with the respective steady-state result and subsequently "frozen". This approach was expected to be valid, since the flow fields in GST-1 and GST-2 were considerably steady and did not fluctuate (see Appendix C, S5.5). For the SUM, the flow field fluctuated slightly, but a comparison between frozen and dynamic approach showed only minor differences in the simulated mixing times which justifies the use of the frozen approach also in this case. After the initialization, a non-reactive species mimicking a 1 M NaCl solution was added below the water surface at a position analogue to experimental procedure (exact addition positions are described in Appendix C, S5.3). With the defined tracer volume,

a tracer concentration of 0.1 % (v/v) was reached. The tracer diffusion coefficient was specified to D = $1*10^{-9}$ m^2/s. A first-order implicit methods for the temporal discretization was used. Within a preliminary time step analysis, the time step size was gradually reduced until convergence of the simulated mixing time curve was achieved. This analysis was done for the GST-2 at 120 rpm due to the highest average velocity gradients and resulted in a time step size of 0.01 s to be sufficient for all vessels. The Courant number was smaller than unity for most mesh cells to ensure numerical stability and convergence. The simulations were run for up to 300 s. Similar to published literature [82], a homogenization criterion of 95 % was selected to determine mixing times. For a complete representation of the whole vessel, the global mixing indicator M_{global} was chosen which is quantified by the squared deviation of concentrations in the entire fluid domain:

$$M_{global}(t) = 1 - \sqrt{\frac{1}{V} \int \left(\frac{c(t)}{c_\infty} - 1 \right)^2 dV} \qquad (2)$$

where V is the vessel volume, $c(t)$ is the cell concentration over time and c_∞ is the volume-average mean concentration. The time to reach 95 % homogenization ($M_{global} = 0.95$) is the simulated mixing time. For GST-1 (at 60 rpm) experimental mixing time data were performed in the laboratory, whereas for the SUM (at 250 rpm) data were available from the vendor. In both cases, the local tracer concentration at the probe position was taken from the simulation. For the SUM, mixing times for the 99 % criterion were available and therefore evaluated.

5.2.2.4 Mixing time studies – experimental

Salt spiking experiments could be performed for GST-1 at equal volume and stirrer speed to the CFD simulations. The vessel was filled with desalted water as model fluid. The stirrer speed was set and 30 µL of 1 M KCl (Merck KGaA) solution was manually added with a pipette to the top of the liquid surface. The conductivity of the vessel solution was measured externally: A peristaltic pump (Minipuls 3, Gilson, Middleton, USA) was used to continuously pump the solution through PEEK tubing to an in-line pH/C-900 conductivity monitor (Cytiva, Uppsala, Sweden) at a flow rate of 1 mL/min. Since the influence on the volume of the reactor was assumed to be neglectable (< 1 %), the outflowing solution was discarded to prevent flow field disturbances. Analogue to the CFD simulations, the mixing time was determined at 95 % of the final conductivity. The tubing dead volume was determined and measurements were conducted in triplicates. Mixing time data and position for the SUM were available from the vendor.

5.2.2.5 CFD reaction modeling of the ADC conjugation reaction

The ordinary differential equations (ODEs) of the kinetic model describing the DAR 2 conjugation reaction scheme were taken from a previous work [140]. The model consists of two consecutive conjugation steps and a parallel reaction for the payload/drug inactivation:

$$1. \text{conjugation rate} = k_1[mAb][Drug], \qquad k_1 = 0.797 \ (mM * s)^{-1} \qquad (3)$$

$$2. \ \text{conjugation rate} = k_2[mAb_{1Drug}][Drug], \qquad k_2 = 1.476 \ (mM * s)^{-1} \qquad (4)$$

$$\text{Drug sink rate} = k_3[Drug], \qquad k_3 = 0.00155 \ s^{-1} \qquad (5)$$

In short, the model assumes that mAb and payload react to the mono-conjugate (mAb_{1Drug}) and afterwards to the desired bi-conjugate (mAb_{2Drug}). The values of the three calibrated rate constants were taken from previous small-scale experiments using the surrogate payload NPM [140]. In the original model, an initial distribution in % of available cysteines on the mAb is considered which consequently leads to seven ODEs. The percentages of mAbs with two, one and zero activated cysteines were set to 88.59, 8.60 and 2.81 %, which was calculated from the final ratio of mAb, mAb1Drug and mAb2Drug in the validation run. Since this assumption does not affect the time-course of the reaction but increases the computational demand, it was thus only adopted in the reaction simulations for the validation for better agreement with the experimental data and neglected in the remaining CFD simulations.

For the CFD reaction models (3D-model), the reaction is considered as homogeneous liquid reaction system. To predict the course of the ADC conjugation reactions in stirred vessels, simultaneously solving the differential equations of momentum, energy, mass and species is required. The ODEs were implemented as volumetric reactions in the species transport equation in FLUENT. This was realized by adding a rate of production R_i and a source term S_i representing the rate of creation to the mass conservation equation which takes the following differential form for the ith species:

$$\frac{\partial}{\partial t}(\rho Y_i) + \nabla \cdot (\rho \vec{v} Y_i) = -\nabla \vec{J_i} + R_i + S_i \qquad (6)$$

where ρ is the liquid density, v is the fluid viscosity, Y_i is the species mass fraction and $\vec{J_i}$ is the diffusion flux of species i. Constant density was assumed for all species and the increase in volume was neglected due to the rather small volume of added payload of 1.67 % (v/v) with respect to the total volume. The finite-rate model was chosen to calculate the production term. This approach computes the chemical reaction rate of each species directly with neglecting turbulence-chemistry interaction. Backward reactions were set to zero and the liquid temperature was assumed isothermal. For initialization, the entire fluid

domain was homogenously patched with the desired mAb concentration. The payload addition was performed sub-surface. Addition positions can be found in the Appendix C, S5.3. Instant addition was implemented by initializing a spherical domain with a payload concentration of 10 mM. For addition over time a source term in the same region was implemented, which generated a constant amount of payload in each time step for the feeding times. A time step size of 0.01 s was also used for these studies. This was determined by initializing the entire volume for GST-2 as perfectly-mixed (as assumed in the 0D-model) and lowering the time step size until the kinetics converged against the predictions of the 0D-model. The transient reaction simulations were performed with similar settings as the mixing time studies and were run for 300 s since all mAb is conjugated during this period.

5.2.2.6 Validation of CFD reaction modelling

A validation run for the CFD reaction modeling was performed in the GST-1 vessel using 300 mL mAb solution with 1.5 mg/mL (ADC1) and 5x molar payload excess. The stirrer speed was set to 60 rpm and a surrogate payload dissolved in DMSO was added over 1 min with a syringe pump (Nemesys S). Time samples were quenched with NAC and the conjugation kinetic of the intact ADC species was determined with a non-reducing RP-UHPLC method. The same protocol, system and column as described in [140] was used.

5.2.2.7 Vessel comparison and parameter study

Using the 3D-models, the kinetics in three mixing vessels were compared. Furthermore, a parameter study was conducted, exemplarily for GST-2, that covered the variation of three process parameters in typical ranges. The investigated parameters are summarized in Table 5.2. Only one parameter was varied at a time while the other parameters were kept constant at the standard condition of 60 rpm, 5 mg/mL mAb concentration, 5x molar payload excess and 60 s addition time.

Table 5.2: Investigated process parameters within the parameter study conducted for the GST-2. Standard conditions were 60 rpm, 60 s addition time, 5 mg/mL mAb and 5x molar payload excess.

Parameter	Range
Stirrer speed	60, 80 and 120 rpm
Payload addition time	0 (Batch), 60 and 300 s
c_{mAb}	5 and 10 mg/mL

Analogously, the original kinetic model was expanded with a fed-batch term assuming ideal-mixing. This resulted in a classical 0D-model which serves as a reference for the effect of vessel scale and process parameter when comparing

the outcome of 0D- and 3D-model. This deviation between both simulations was quantified using the absolute difference in the DAR value over time:

$$\Delta\text{DAR}(t) = \text{DAR}_{0\text{D}}(t) - \text{DAR}_{3\text{D}}(t) \tag{7}$$

whereas the DAR values were calculated with:

$$\text{DAR}(t) = \frac{c_{\text{mAb,1Drug}}(t) + 2 * c_{\text{mAb,2Drug}}(t)}{c_{\text{mAb,0}}}. \tag{8}$$

In case of the 3D-model, the concentrations of the conjugated species were obtained from the volume-averaged species concentrations at each time step. The 0D-model was simulated in MATLAB R2020a and the differential equations were solved using the *ode45* solver.

5.2.2.8 Time-scale analysis

As an alternative approach to fully modeling the dynamic reaction, one can compare the time-scales of reaction and mixing to receive an expectation regarding the predominating mechanism to be considered. An inhomogeneity of reactant concentration in large-scale reactors operated in fed-batch may be caused by weak distribution of the added reactants. This characteristic time can be quantified with the mixing time [195]. For chemical reactions one can calculate the characteristic time τ_R for a bimolecular reaction according to [188]:

$$\tau_R = \frac{1}{k_i(\widetilde{c_1} + \widetilde{c_2})}, \tag{9}$$

where k_i is the kinetic rate of the ith reaction and $\tilde{c}_j$ is the local concentration of the jth species. If the characteristic reaction time is significantly larger than the mixing time, the reaction can be considered as ideal-mixed, while for larger mixing times in comparison to the reaction time the process becomes mixing-sensitive.

5.3 Results and discussion

The first part of this chapter deals with the experimentally determined conjugation kinetics and the comparability between small- and lab-scale conjugation kinetic. In the second part, the CFD results for the three studied vessels are analyzed involving typical scaling parameters such as P/V, resulting flow fields and mixing times. The chosen process parameters at which the vessels were compared can be found in Table 5.1. Finally, the CFD-simulated ADC conjugation kinetics are compared among the vessels and the influence of process parameters on the reactions is studied for GST-2.

5.3.1 Experimental conjugation kinetic studies

5.3.1.1 Small-scale conjugation kinetic studies

The resulting DAR kinetics from the RP-UHPLC analysis are shown in Figure 5-1A for the DAR 2 species. The DAR increases rapidly for all reactions and the mAb is entirely conjugated after 900 s. Notably, the curves deviate during the initial phase of the conjugation reaction with the orbital shaker at 1000 rpm having the largest offset from the ideal conjugation kinetic. This might be caused by an experimental artefact, but also demonstrate the requirement for proper mixing which cannot be ensured using the orbital shaker. Since conjugation reactions were performed over 1 h and the curves deviate only during the initial reaction period, the final DAR values and the drug load profile (data not shown) were not affected by the mixing type. The achieved average final DAR of all runs is 1.90 which is lower than the theoretical DAR of 2. This can be caused by the previously described pre-inactivation of cysteines [140] that originate from reformed disulfide bridges from reactive thiols formed during the reduction and re-oxidation step [65]. The averaged kinetics for the DAR 8 species are shown in Figure 5-1B. The error bars represent the standard deviation of each duplicate. Similarly, the DAR curve increases rapidly while final DAR values are reached at approx. 300 s. All kinetics show no considerable deviation between the mixing types and final DAR values range between 7.5 and 7.8. Overall, the effect of the mixing on the final DAR for both molecules was rather small if the reaction was run for 1 h. In summary, the conjugation kinetic study demonstrated that internal stirring in reaction tubes may be favored. Using the orbital shaker only results in acceptable kinetics when initial mixing is conducted properly.
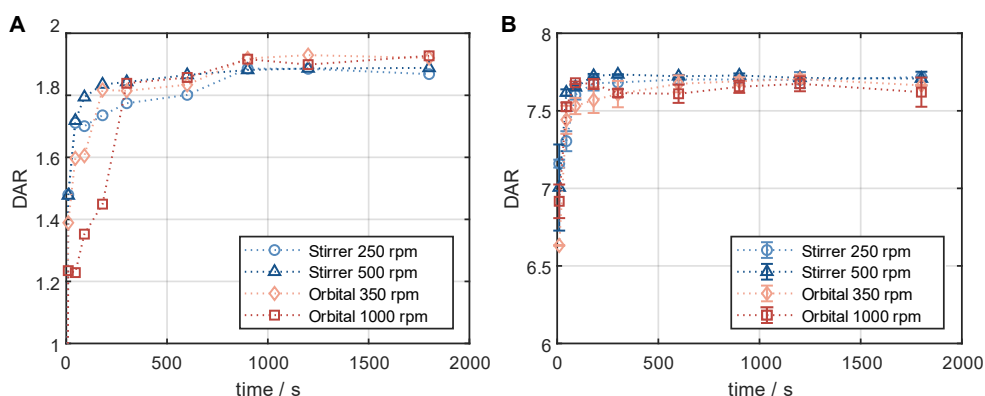

Figure 5-1: Comparison of the RP-UHPLC determined DAR conjugation kinetics performed in small-scale reaction tubes with different mixing systems and speeds. (A) ADC1 at a concentration of 10 mg/mL conjugated with 5x molar payload excess and (B) ADC2 at a concentration of 1.5 mg/mL conjugated with 11x molar payload excess. Runs were performed in duplicates and error bars represent the standard deviation.

5.3.1.2 Conjugation kinetic comparison of reaction tubes and lab-scale stirred vessel

Product quality and DAR of an ADC are usually tested after the conjugation reaction is complete. It was of interest to study changes in conjugation state over the course of the reaction and to compare the small-scale reaction tubes with a larger reactor set-up. Figure 5-2 shows the comparison of the DAR-course in reaction tube vs. lab-scale stirred tank (GST-1) for both the site-specific conjugation to inserted cysteines (left panel) and the stochastic conjugation to reduced interchain disulfide bonds (right panel). For both ADCs the kinetics follow the same course at the two scales. In both cases slightly higher DAR values are achieved in the lab-scale vessel. However, the observed difference is within assay variability. Furthermore, no considerable difference in the drug load profiles was present (data not shown). These experiments demonstrate that the internally stirred reaction tube shows good comparability to the lab-scale vessels for the conjugation chemistries studied.
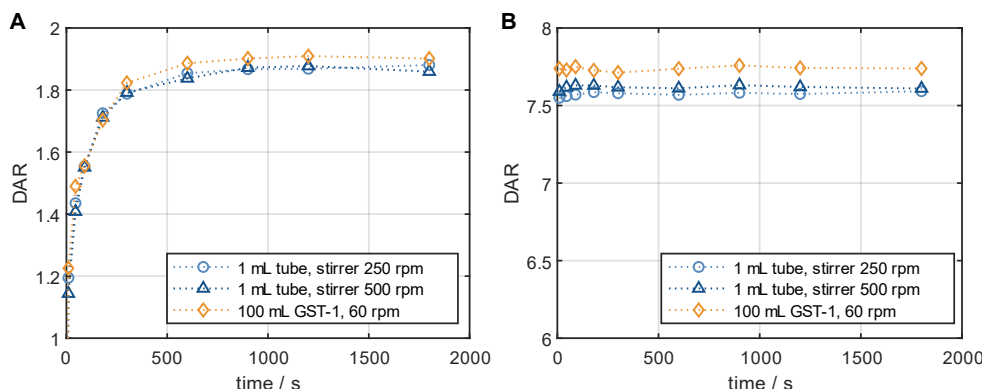

Figure 5-2: Comparison of the small- and lab-scale DAR conjugation kinetics determined by RP-UHPLC. (A) Kinetics for ADC1 at 5 mg/mL and 5x molar payload excess and (B) kinetics for ADC2 at 20 mg/mL and 11x molar payload excess.

5.3.2 CFD simulations for large-scale vessels

5.3.2.1 Large-scale vessel characterization using steady-state and mixing time simulations

Simulations for both glass vessels GST-1 and GST-2 result in comparable P/V values of 3.81 and 2.44 W/m^3, respectively. A tenfold higher P/V value of 23.71 W/m^3 is reached in the SUM due to its higher stirrer speed. Figure 5-3 shows the resulting plots for the velocity magnitude and projected velocity vectors of the steady-state solution for the three studied vessels. The vectors are colored by the magnitude of the axial velocity normalized to the average velocity magnitude in each vessel, in order to examine regions which contribute to the axial transport in the vessel.
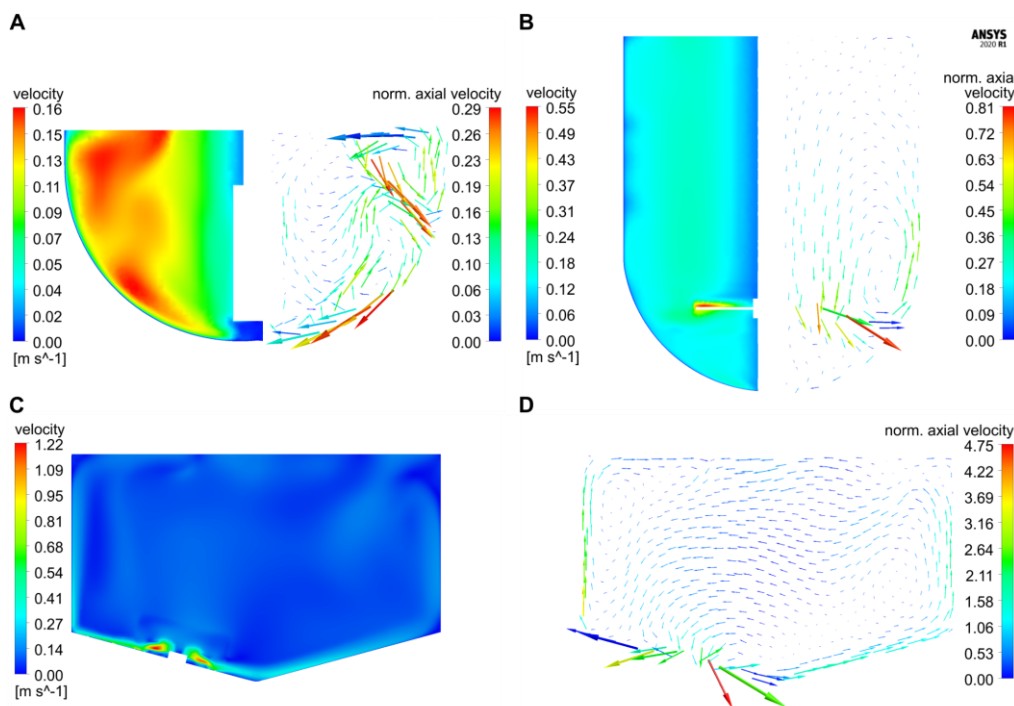
Figure 5-3: Contour plots of the velocity magnitude vs. projected velocity vector plots colored according to the axial velocity magnitude (normalized to the respective average velocity magnitude in each reactor) on vertical cut planes for (A) GST-1 (60 rpm), (B) GST-2 (60 rpm) and (C-D) SUM (400 rpm). Due to the symmetrical design of GST-1 and -2, half of the plots are shown side-by-side. Please note that the vector length corresponds to the velocity magnitude in each reactor but was scaled individually for each plot and, therefore, is not comparable among the reactors.

The comparison of the contour plots demonstrates that in the GST-1 overall a larger part of the bulk has higher relative velocities than the other two reactors. In the GST-2, high velocities were found near the stirrer and medium velocities in the remaining bulk. For both vessels, velocities close to the shaft are lower. In the SUM, high velocities occur near the impeller blades. In contrast to the glass vessels, the majority of the bulk appears to have lower velocity magnitudes compared to the impeller tip speed which is indicated by a larger amount of (light) blue areas in the contour plots. This is due to the impeller discharge towards the bottom of the vessel and the smaller impeller diameter in relation to the vessel diameter. Notably, the average velocity magnitudes of all three reactors are in a similar range. When comparing the vector plots it becomes obvious that the flow direction in the vessels differ strongly. For the two glass reactors, the flow was found to be mainly rotational, but in the GST-1 a larger region appears to contribute more to the axial transport, which is indicates by more vectors having higher axial velocities. For the GST-2, only flow close to the stirrer region contributes to axial transport, whereas the volume above the impeller is mainly rotational with very low axial velocities. This is due to the lack of baffles and the relatively low stirrer speed in comparison to the reactor

volume. In contrast, the vectors for the SUM indicate much higher axial (and also radial) velocities, especially in the area close to the sides of the vessel where flow is directed upwards. Moreover, the average axial velocity is approx. twofold higher than in the other two vessels. The directions of the vectors emphasize that the eccentric position of the impeller also produces a more chaotic and asymmetric flow field with higher gradients in axial/radial direction compared to the glass reactors potentially leading to improved mixing.

The mixing performance of the three reactors was compared based on the CFD mixing time studies. The fastest homogenization is achieved in the GST-1 with a predicted (global) mixing time of 9.4 s due to small volume being relatively well mixed which agrees with the high velocities with axial transport in the bulk. In contrast, the mixing time in the GST-2 is much slower with 32.2 s. Especially, the final homogenization close to 95 % in this vessel is observed to be relatively slow. This is caused by the strong rotational flow and relatively low axial and radial transport which coincides with the findings from the vector plots. The mixing in the SUM is remarkably faster with a mixing time of 17.6 s, although the liquid volume is similar to GST-2. This is caused by the higher stirrer speed and an intensified mixing efficiency due to the eccentric impeller design leading to higher axial/radial transport. The simulated mixing curves are depicted in the Appendix C, S5.6. It is worth mentioning, that using the relation between the reaction times (ranging between 300 – 900 s) and the large-scale mixing times (ca. 10-30 s), one can expect only minor mixing dependency on the reaction at this point.

5.3.2.2   CFD model validation

On the one hand, the developed CFD models were validated by comparing local mixing times. The experimental and simulated mixing times are 9.1 and 8.1 s for GST-1, and 49.0 and 45.2 s for the SUM. This results in an error of 10.0 and 7.8 % for the GST-1 and SUM, respectively, which indicates a potential mismatch between experiments and CFD simulation. This can originate from multiple root causes, such as model simplification through the isotropic flow assumption by the applied RANS model, the MRF technique, frozen flow field or inequality between real and simulated measurement position. Similar errors in the range of 10 % were reported in literature [181], [183] which led to the assumption that the observed deviation are in an acceptable range for the purpose of this study. On the other hand, the validity of the kinetic models to predict large-scale were investigated with a conjugation run in the GST-1, as described in chapter 5.2.2.6. Figure 5-4 presents the predicted kinetics for the three ADC species of both models compared to the reference data from the RP-UPHLC. Both models show a very similar course for all species which is also shown in similar R^2 values of 0.979 and 0.985 for the 0D- and the 3D-model,

respectively. Compared to the reference data both models have an offset between $200 - 500\,\mathrm{s}$ while converging simultaneously to equal species concentrations. These results emphasize the agreement of both model types at this particular scale.
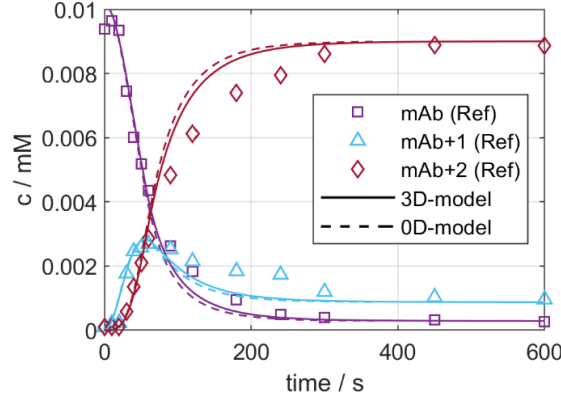

Figure 5-4: Comparison of 0D- and 3D-model predictions of the ADC conjugation kinetic in GST-1 with the experimentally determined kinetic for the first 600 s.

5.3.2.3 3D kinetic modeling of large-scale vessels

In the following, the 3D conjugation kinetics of the three vessels (see chapter 5.2.2.7) are compared. For the examination of the deviation between 3D- and 0D-model, the $\Delta\mathrm{DAR}(t)$ is shown for the three vessels in Figure 5-5.
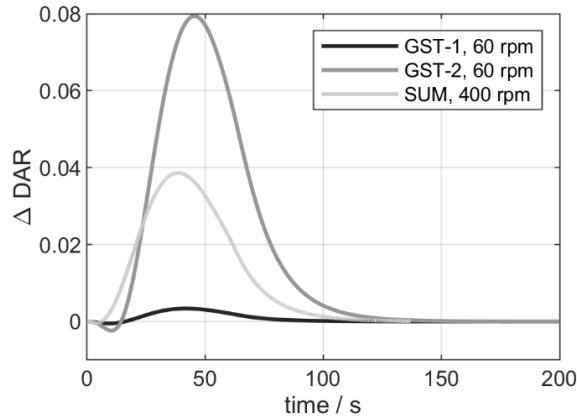

Figure 5-5: Comparison of the $\Delta\mathrm{DAR}$ within the first 200 s for all three studied vessels. Payload addition was simulated over the course of 60 s.

According to this graph, two zones can be distinguished: In the first zone, which is in the beginning of the reaction, the $\Delta\mathrm{DAR}$ curves increase exponentially and reach a maximum value depending on the vessel. Hereby, GST-1 has the smallest deviation (0.0035) and GST-2 the largest deviation (0.08). In the second zone, the three curves converge to $\Delta\mathrm{DAR}(t) = 0$ after around $80 - 130\,\mathrm{s}$. The initial increase in the $\Delta\mathrm{DAR}$ is due to the local availability of added payload in the feed region. The resulting mass transfer limitation leads to actual lower kinetic rates in the remaining bulk which cannot be captured by the 0D-model.

The magnitude of this increase (GST-2 > SUM > GST-1) was found to qualitatively agree with the order of the simulated mixing times.

Additional insights were gained by studying the reaction time-scales at different process times. Local reaction time-scales were exemplarily computed for the first conjugation step at 10 and 60.5 s (immediately after addition is finished at 60 s) according to Eq. 9. The resulting contour plots of the reaction times are shown exemplarily for GST-2 in Figure 5-6.
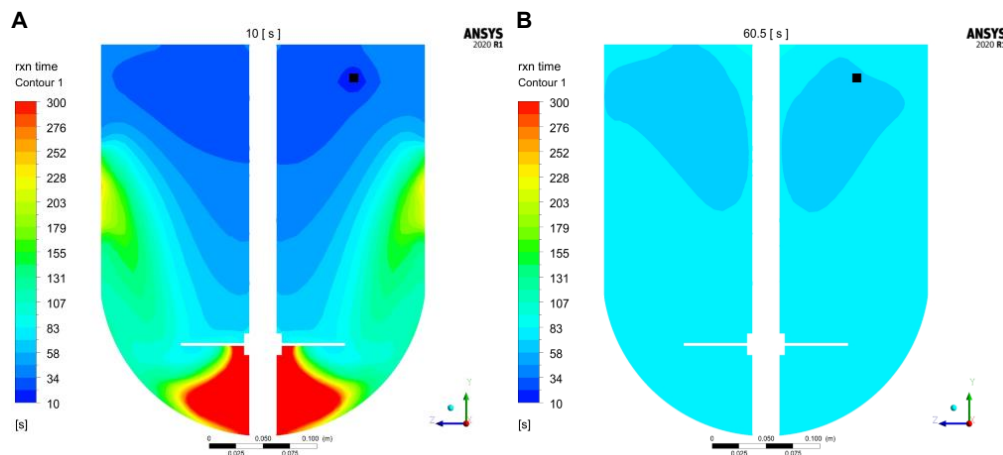

Figure 5-6: Contour plots of reaction time-scales (1. Conjugation step) at vertical cut planes compared for (A) 10 s and (B) 60.5 s exemplary calculated for GST-2. Feeding position is colored in black.

It is noticed that at 10 s the reaction times are lower in the upper part of the vessel. This is equivalent to fast reaction rates, which is due to the freshly added payload in this region. In contrast, the low mass transfer to the region underneath the impeller are found to reduce the reaction rate in this region which is indicated by higher reaction times (green to red areas). This agrees with the observation of the low axial transport downwards to the lower part of the vessel (see Figure 5-3B). In contrast, reaction times are more homogenously distributed after payload feeding is finished (60.5 s) due to a higher degree of homogenization. In summary, this analysis reveals that in the initial phase, reaction times are slightly lower but in a similar magnitude compared to the mixing times which is 32.2 s. Thus, it can be concluded that the speed of homogenization of freshly added payload is responsible for a slight reduction of the conjugation rate in the beginning of the process. However, the calculated maximum ΔDAR values are in a rather irrelevant industrial range, especially, since actual process times are greater than 100 s. The same estimation can be made when calculating the average reaction time by using the mean concentrations of mAb and payload which highlights the benefit of time-scale analysis. Moreover, it should be mentioned that, since the applied conjugation reaction is a consecutive reaction, the deviations in the kinetic of all species are only temporally affected by the vessel mixing and have no large influence on

the final DAR value. In literature, the similarity between a CFD (3D) and an ideal-mixing model (0D) have also been shown by Spann et al. [82] for a fermentation biokinetic model. In this study, the authors presume that the observed local pH changes may not affect the biokinetic and those small differences actually originate from numerical errors in the CFD simulation. In another publication [195], the authors emphasize the advantage of conducting a time-scale analysis in order to determine possible effects of mixing gradients on reactions. It is also contrasted that this analysis does not provide information about the possible effects on relevant CQAs, hence not replacing experimental or in-silico studies.

5.3.2.4 Modeling the influence of process parameters

Additionally, the influence of varying process parameter on the kinetic was studied exemplarily for GST-2. The resulting curves of the ΔDAR are shown in Figure 5-7.
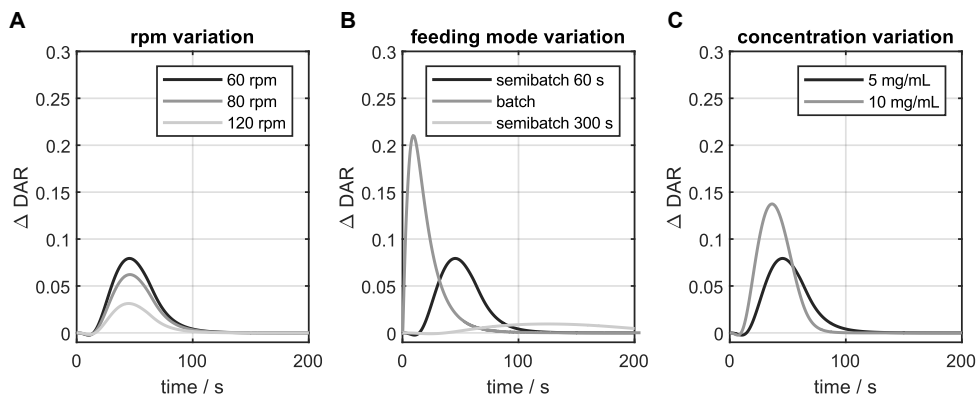

Figure 5-7: Influence of varying process parameters on the ΔDAR of 3D- and 0D-model exemplarily for GST-2: variation of (A) stirrer speed, (B) feeding mode, (C) mAb concentration. The black curve is the reference condition (60 rpm, 60s feeding time, 5 mg/mL + 5x payload excess).

A stepwise increase of the stirrer speed from 60 to 120 rpm leads to a decrease of the model deviation which can be attributed to enhanced mixing performance caused by greater velocity gradients and turbulence. However, increasing the stirrer speeds appears to have a rather small effect on improving the mixing performance and, thus, reducing the ΔDAR value. A larger influence is observed for the investigated feeding modes (see Table 5.2): The batch addition causes the ΔDAR to increase rapidly to around 0.2 compared to semibatch mode. This behavior can be related to a higher amount payload locally available at the same time producing more inhomogeneities which cannot be captured in the 0D-model. Semibatch feeding led to lower maximum ΔDAR. For 300 s feeding the models deviate only marginally which shows that an increase in feeding time is likely to minimize mixing effects or other phenomena due to locally high payload concentrations. Doubling the mAb concentration also resulted in a

short increase of the ΔDAR. Here, the mixing effect is more pronounced due to faster initial reaction rates. As illustrated by the studied parameters, any changes of the process parameters would only influence the time of completion of the reaction but would reach the same endpoint.

In conclusion, the 3D-model indicated deviations from the ideal conjugation kinetic, especially when all payload is added at once (batch mode) or for higher reactant concentrations. For the studied vessel (GST-2), the stirrer speed had only little influence on the course of the kinetic. Other parameters, like feeding position, were also studied but showed even smaller deviations and are therefore not presented here. Overall, the influence of geometry and process parameters were generally small in the case of the studied consecutive two- step conjugation reaction. This is due to the naturally selective conjugation to the targeted sites which has been reported to simplify process development and scale-up [53]. Furthermore, mixing in the studied vessels is adequate due to relatively the slow (bio-chemical) kinetic reaction in contrast to other typical faster chemical reactions. A complete validation of the predicted species time-course would also be necessary. In our case, the results of the CFD kinetic study gave a comprehensive overview of possible parameter influence on the course of the ADC conjugation kinetic. For stochastic conjugation chemistries, the parameter effects may be relevant and the CFD model might be more advantageous. Moreover, the CFD model could be used to predict other effects like shear rate-depended mAb fragmentation or aggregation which was not observed in this study and would require a additional model to be incorporated.

5.4   Conclusion

This work considered different aspects for a better understanding how scale-up and process parameters affect the ADC conjugation reaction for two model ADCs. First, experimental kinetic studies in reaction tubes dealt with the optimization of mixing by using different mixing types. We could show that reaction tubes that are internally mixed using a magnetic stir bar produce consistent conjugation kinetics which are comparable to kinetics in glass reactors. Secondly, different types of CFD simulations were performed for three commonly used vessels in ADC manufacturing. Using steady-state simulations and mixing time studies we could characterize the vessels' mixing performance and were able to describe local mixing effects. We further implemented a DAR 2 conjugation kinetic model in the CFD models leading to a full 3D-model. By using the classical ideal-mixing model (0D) as benchmark we showed that the relation of achieved mixing times and chemical reaction rates governs the implications obtained during scaling. Current ADC conjugation reactions are, however, in a range where mixing performance in commercially available vessels

is adequate for the fast conjugation kinetic. This indicated that the ratio of mixing time and chemical reaction kinetic is a reliable indicator to be considered during scaling for this reaction type. The often-applied P/V value did not show to correspond well with the observed deviations. Furthermore, we studied variations in process parameters (stirrer speed, feeding mode and concentration variation). We found that the parameters affected the conjugation kinetic only little within the first 100 s of the reaction and final DAR values remained constant. This can be attributed to the highly selective conjugation chemistry and the consecutive nature of the reaction. A time-scale analysis demonstrated that conjugation rate inhomogeneities occur in the feed region and only during the addition phase. In the case of the DAR 2 conjugation reaction, the additional insight from the 3D-model were rather not industrially relevant. Therefore, the 0D-reactor models can be applied for predicting large-scale conjugation kinetics and to be used in a digital twin framework. As this study was successfully conducted for a DAR 2 conjugation reaction it has the potential to be adopted to other (conjugation) reactions.

Acknowledgments

Supplementary Materials

Appendix C contains the Supplementary Materials associated with this chapter:

- S5.1 Experimental mixing study for DAR8
- S5.2 CAD Geometries
- S5.3 Generated meshes
- S5.4 Global mesh study
- S5.5 Convergence monitors for steady-state simulations using the final meshes
- S5.6 Simulated mixing times

6

An adaptive soft-sensor for advanced real-time monitoring of an antibody-drug conjugation reaction

Robin Schiemer[1]*, Jan Tobias Weggen[1]*, Katrin Marianne Schmitt[1], Jürgen Hubbuch[1]

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

*These authors contributed equally

Abstract

In the production of antibody-drug conjugates (ADCs), the conjugation reaction is a central step defining the final product composition and, hence, directly affecting product safety and efficacy. To enable real-time monitoring, spectroscopic sensors in combination with multivariate regression models have gained popularity in recent years. The Extended Kalman filter (EKF) can be used as so-called soft-sensor to fuse sensor predictions with long-horizon forecasts by process models. This enables the dynamic update of the current state and provides increased robustness against experimental noise or model errors. Due to the uncertainty associated with sensor and process models in biopharmaceutical applications, the deployment of such soft-sensors is challenging. In this study, we demonstrate the combination of an uncertainty-

aware sensor model with a kinetic reaction model using an EKF to monitor a site-directed ADC conjugation reaction. As the sensor model, a Gaussian process regression model is presented to realize a time-variant determination of the sensor uncertainty. The EKF fuses the time-discrete predictions of the amount of conjugated drug from the sensor model with the time-continuous predictions from the kinetic model. While the ADC species are not distinguishable by on-line recorded UV/Vis spectra, the developed soft-sensor is able to dynamically update all relevant reaction species. It could be shown that the use of time-variant process and sensor noise computation approaches improved the performance of the EKF and achieved a reduction of the prediction error of up to 23 % compared to the kinetic model. The developed framework proved to enhance robustness against noisy sensor measurements or wrong model initialization and was successfully transferred from batch to fed-batch mode. In future, this framework could be implemented for model-based process control and be adopted for other ADC conjugation reaction types.

6.1   Introduction

Antibody-drug conjugates (ADCs) are one of the fastest growing class of biopharmaceuticals using the target specificity of a monoclonal antibody (mAb) to exclusively deliver a cytotoxic drug to tumor cells [196]. Many different conjugation strategies exist consisting of multiple reaction steps while one of the greatest challenges remains in controlling the drug-to-antibody ratio (DAR) [50], [119]. To ensure consistent final product quality, regulatory agencies have been promoting the implementation of advanced monitoring and control strategies according to the concepts of quality by design (QbD) and process analytical technology (PAT) during a decade [91], [197]. In line with the QbD framework, recent publications showed the successful application of model-based techniques in ADC process development, such as molecular dynamics simulations [65], conjugation kinetic models [140] and chromatography models [198], [199]. Mechanistic process models, however, often require simplification of complex phenomena which may lead to plant-model mismatches or poor predictability of the transient behavior of the process [179]. With regards to PAT, spectroscopic sensors have shown to enable real-time measurement of certain critical quality attributes (CQA) [200]. Consequently, these sensors have been deployed as PAT for various unit operations in downstream processing of mAbs [101], [201], [202]. In the field of ADCs, monitoring of the conjugation reaction using UV/Vis spectra [98] and quantification of aggregation using Raman spectroscopy [100] are recently established PAT tools. Partial least squares regression (PLSR) models have commonly been applied in chemometrics [203], [204] and, more recently, Gaussian process regression

(GPR) models have been identified as a competitive alternative to PLSR [108], [205], [206]. Furthermore, GPR models deliver an uncertainty estimate for each evaluated data point and therefore help to perform risk-based dynamic adjustments during the process [206].

The extracted information from sensors can be used to supply process models with direct feedback from the process in near real-time, which enables timely model updates and, hence, more accurate estimates of process- or product-relevant variables [180]. To fuse estimates from two noisy sources, various optimal state estimation algorithms have been developed and recently been applied to chemical [207], [208] and pharmaceutical [209]–[211] processes. The extended Kalman filter (EKF), as one of the most prominently applied algorithms, has been used to combine non-linear process models and sensor measurements for cell culture processes [113], [211] or for capture chromatography [212]. In these studies, the EKF computes a weighted average of both source models, but does not affect the long-horizon predictions of the process model. In many cases, however, only a few states are directly available from measurement probes. This problem can be circumvented by utilizing an EKF formulation that combines the measurement with mass or energy balances. Thus, an EKF can be utilized to update all states in the model which are not available for measurement. This approach is considered a model-based soft-sensor according to [213] and was similarly demonstrated in recent studies for advanced monitoring of bioprocesses [114], [214]–[216].

This study aims to develop a robust and adaptive soft-sensor to monitor the site-directed conjugation reaction of a mAb with a surrogate payload. Therefore, an uncertainty-aware chemometric sensor and the system of equations from a kinetic model are combined using an EKF algorithm. This enables to deliver dynamically updated predictions of all reaction states not available for measurement. The experimental data set encompasses nine batch conjugations and two fed-batch conjugations. Both kinetic model and chemometric sensor model are calibrated on the batch data based on the methodology of previously published works [98], [140]. For the chemometric sensor, a novel GPR model is evaluated against benchmark alternatives and combined with different preprocessing strategies. Subsequently, an EKF is deployed as a soft-sensor fusing the information of the GPR model with the kinetic model. In detail, this approach recursively updates the unmeasurable ADC species predicted from the kinetic model by using the time-discrete GPR predictions of the conjugated drug concentration. The influence of time-variant uncertainties for the GPR and kinetic model are investigated based on the methods given in literature and the optimal configuration is selected. Finally, the robustness and accuracy of the EKF is assessed by comparison to the two underlying models in different scenarios. The described workflow suggests an unified calibration routine for

the underlying models, enables the quantification of model uncertainties and assesses the performance boost of the developed soft-sensor.

6.2 Material and methods

6.2.1 Experiments

6.2.1.1 Chemicals and buffers

The chemicals used in this study were obtained from Merck KGaA, unless otherwise stated. Standard working buffer was made with NaH_2PO_4, dihydrate (VWR International GmbH) and ultrapure water (PURELAB Ultra, ELGA LabWater). The buffer was titrated to pH 7.2 with 4 M NaOH and filtered through a 0.2 µm cellulose acetate membrane filter (Sartorius AG). For the mAb reduction and oxidation step tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and (L)-dehydroascorbic acid (DHAA) was used. To mimic a cytotoxic drug, the non-toxic surrogate payload N-(1-pyrenyl)maleimide (NPM) was used. For quenching of the conjugation reaction, unconjugated free drug was quenched with N-acetyl cysteine (NAC). For the reversed-phase ultra-high performance liquid chromatography (RP-UHPLC), acetonitrile (VWR) and trifluoroacetic acid (TFA, Thermo Scientific) was used.

6.2.1.2 mAb functionalization and conjugation procedure

The reactions were conducted using an engineered IgG1 mAb with two inserted cysteines in the hinge region which was generously provided by AstraZeneca. All chemicals were dissolved in 50 mM sodium phosphate buffer, except NPM which was dissolved in dimethyl sulfoxide (DMSO). In the beginning, the mAb was thawed, diluted and three reaction steps were consecutively performed: a reduction reaction to uncap the engineered cysteines, a dialysis to remove excess TCEP and a re-oxidation reaction of the reduced interchain disulfide bonds. This functionalization procedure was conducted under the same methodology and conditions as previously described in Andris et al. [140]. The subsequent conjugation reaction was performed in a stirred glass beaker (400 rpm) with a reaction volume of 10 mL. Batch conjugations were started by adding a certain volume of 0.68 mM NPM solution with a pipette to reach the desired molar drug excess. In nine runs, the mAb concentration and drug excess were varied, including some runs in duplicates. Two additional fed-batch runs were performed by adding the NPM solution with a syringe pump (Nemesys S, Cetoni GmbH, Germany) constantly over 5 min and 10 min for run 10 and run 11, respectively. All conducted runs are summarized in Table 6.1.

Table 6.1. Overview of conducted conjugation runs using different mAb concentrations and NPM excesses. Batch runs (1-9) were split into training and test set for model calibration, EKF tuning and performance evaluation. Fed-batch runs (10 and 11) were used later in the study to test the transferability of the designed framework.

Experiment	mAb concentration (mg mL^{-1})	Molar drug excess (-)	Number of samples	Subset
Run 1	1	2x	20	Training
Run 2	1	4x	17	Test
Run 3	1.25	3x	21	Training
Run 4	1.5	2x	21	Test
Run 5	1.5	2x	20	Training
Run 6	1.5	3x	18	Test
Run 7	1.5	4x	21	Training
Run 8	1.75	2x	20	Training
Run 9	1.75	3x	20	Training
Run 10 (Fed-batch)	1.5	3x	22	External
Run 11 (Fed-batch)	1.5	3x	20	External

6.2.1.3 UV/Vis on-line monitoring and off-line RP-UHPLC

To monitor the conjugation reaction, the same set-up as in Andris et al. [98] was used. In detail, an on-line loop was installed where a peristaltic pump (Minipuls 3, Gilson, Middleton, USA) pumped the reaction mixture continuously through the loop tubing. The flow rate was adjusted to 1 ml min⁻¹ by measuring the flow rate with a liquid flow meter SLS-1500 (Sensirion AG, Switzerland). To collect UV/Vis spectra every 0.2 s, a DAD-3000 RS diode array detector with a 0.4 mm path length flow cell (both Thermo Fisher Scientific) was installed in the loop. UV/Vis spectra were recorded between 250 and 390 nm at a resolution of 1 nm. Before starting the conjugation reaction, the entire loop was equilibrated with the oxidized mAb solution for up to 15 min and an autozero of the DAD was performed to collect the pure NPM spectra during the reaction.

Off-line samples were taken at certain time points and immediately quenched with a NAC solution. To determine the concentrations of the single ADC species for the time points, all samples were analyzed with RP-UHPLC. The same device, column and analysis method as described in Andris et al. [98] was applied. The RP-UHPLC data will further be treated as reference data for all models.

6.2.2 Data organization

The full data set containing the reaction time-courses from the RP-UHPLC and the spectra for each run was split into training and test subsets. The split was done based on a Kennard-Stone algorithm [217] using the first-derivative mean spectra of each run to ensure that the choice of the calibration set is chosen objectively and spans the full experimental design space. The training subset was used for hyperparameter tuning of the chemometric models, calibration of the kinetic model parameters and tuning of the EKF. The test subset was used for model performance evaluation and selection. The fed-batch runs were evaluated separately. The subsets used for chemometric model development consisted of 122 and 56 reference data points for training/cross-validation and testing, respectively. An overview of all conducted experimental runs is shown in Table 6.1.

6.2.3 Chemometric model development

Initially, all absorbance spectra were averaged over an interval of 5 s and aligned with the RP-UPHLC reference data. Before regression modeling, the spectra were mean-centered and normalized by the maximum absorbance at 250 nm of the training subset. Several chemometric models were calibrated and compared to determine the best performing option. In the following, all mentions of cross-validation refer to a *leave-one-run-out* rotation.

Partial-least squares regression (PLSR) models were calibrated in combination with a Savitzky-Golay derivative filter (SGF) or a variable selection algorithm. The number of latent variables, the window size and order of derivative for the SGF were tuned by employing a cross-validated grid search. The variable selection algorithm was adapted from Mehmood et al. [218] using the regression coefficients of the PLSR as an importance metric and evaluated using cross-validation.

For Gaussian process regression (GPR) models, multiple kernel functions were screened and combined with a SGF or a variable selection algorithm analogue to the one described above. Multiple combinations of a linear, a radial-basis function (RBF) and an automatic relevance determination (ARD) kernel were evaluated. The final implementation of the kernel given in the Appendix D, S6.1. More theoretical information on GPR modeling can further be found in Appendix D, S6.1. The kernel parameters were estimated by expectation maximization as implemented in Pedregosa et al. [219]. The implemented variable selection algorithm was adapted from Paananen et al. [220]. For the exact mathematical procedure the reader is referred to the cited literature. The computed variable sensitivities are used as an importance metric, ranked and selected as performed for the PLSR variable selection algorithm.

All calibrated models were evaluated based on their root mean square error of cross validation ($RMSECV$) and prediction ($RMSEP$). A detailed definition of the error metric used in this study can be found in the Appendix D, S6.4. All chemometric models presented in this study were implemented with recent versions of Python 3.8, NumPy ≥ 1.20 [221], pandas ≥ 1.3 [222], scikit-learn ≥ 1.1 [219] and related packages.

6.2.4 Kinetic model calibration

A previously developed kinetic model for the site-directed conjugation reaction (see Andris et al. [140]) was applied in this study to be adopted in the EKF framework. A detailed explanation of the model and the full ODE system can be found the Appendix D, S6.2. The kinetic rates k_1, k_2 and k_3 were estimated by non-linear least squares optimization using the *lsqnonlin* solver in MATLAB 2019b (The Mathworks Inc.). The initial values were set to 0.8 mM^{-1} s^{-1}, 1.4 mM^{-1} s^{-1} and 0.01 s^{-1} for k_1, k_2 and k_3, respectively. Assuming that measurement errors in each sample are independent and normally distributed with zero mean and variance σ^2, the parameter covariance matrix $\mathbf{C}_k$ is computed according to Eq. 1,

$$\mathbf{C}_k = cov(\mathbf{k}) = \sigma^2 (\mathbf{J^T J})^{-1} \tag{1}$$

where $\mathbf{J}$ is the Jacobian with regard to the kinetic rates $\mathbf{k}$ and σ^2 is the estimated error variance from the non-linear estimation. The error variance σ^2 is approximated by $RMSEP^2$ as suggested in literature [115], [212].

6.2.5 Extended Kalman filter

6.2.5.1 Terminology and concept

As the herein applied models are calibrated using different target variables, the terminology is shortly introduced: The un-, mono- and biconjugated ADC species will be referred to as mAb, Conj1 and Conj2 for the remainder of this study. It is worth mentioning that "mAb" corresponds to the sum of the three subspecies with different amount of active cysteines: mAb$_{2c}$, mAb$_{1c}$ and mAb$_{0c}$, and "Conj1" to the sum of the two sub-species Conj1$_{1c}$ and Conj1$_{0c}$ of the ODE system. While the kinetic model continuously predicts all species, the chemometric model converts the absorbance spectra to the amount of conjugated drug at discrete time points t_k. The same values can also be computed from the single species predictions by $c_{\text{drug,Conj}} = c_{\text{Conj1}} + 2c_{\text{Conj2}}$, where c_{Conj1} and c_{Conj2} are the concentration of mono- and biconjugated mAbs. In order to enable the sensor model to update the individual ADC species in the ODE system, a special measurement function was chosen (see chapter 6.2.5.2). Furthermore, both model outputs are thought to be noisy as they are

subject to errors caused by model assumptions, plant-model mismatch and experimental handling. An EKF framework was designed to combine the noisy model outputs and generate a more accurate estimate than either of the two sources. Since the two used models are not producing estimates for the same set of species, the EKF acts as a soft-sensor updating several unmeasurable species by an indirect measurement of the sum of conjugated drug. For better understanding, the overall concept and data organization is summarized in Figure 6-1.
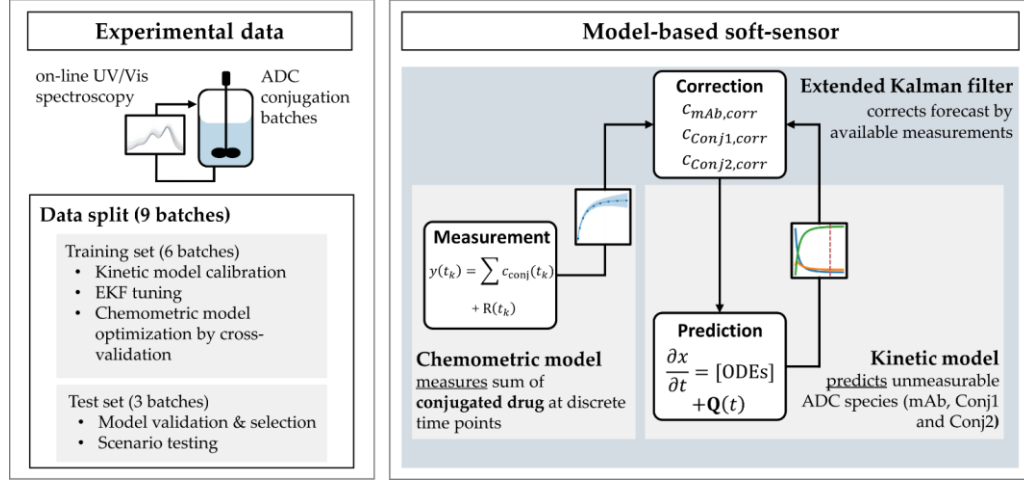

Figure 6-1. Schematic overview of the purposes of the data sets and the methodology of the designed model-based soft-sensor.

6.2.5.2 Mathematical formulation

The theory of the EKF framework is well established and can be found e.g. in Simon [223]. In principle, all species are represented in a state vector $\mathbf{x}$. The seven ordinary differential equations (ODEs) of the kinetic model act as state transition function f propagating the state vector $\mathbf{x}$ through time according to Eq. (2) given deterministic inputs $u(t)$ and the model parameters $\boldsymbol{\theta}$:

$$\frac{\partial \mathbf{x}}{\partial t} = f(\mathbf{x}(t), u(t), \boldsymbol{\theta}) + \mathbf{q}(t) \tag{2}$$

with the state vector x for all seven reacting species in the ODE system defined as:

$$\mathbf{x}(t) = \begin{bmatrix} c_{\mathrm{mAb}_{2c}} & c_{\mathrm{mAb}_{1c}} & c_{\mathrm{mAb}_{0c}} & c_{\mathrm{Conj1}_{1c}} & c_{\mathrm{Conj1}_{0c}} & c_{\mathrm{Conj2}} & c_{\mathrm{NPM}} \end{bmatrix}^{T} \tag{3}$$

where c_{mAb} denotes the concentration of unconjugated mAbs with the respective number of available cysteines (c) and the conjugates with the respective number of conjugated NPM molecules. The free NPM concentration is given as c_{NPM}. Measurements $y(t_k)$ available at discrete time points t_k are related to states vector using the selected measurement function h which is given in Eq. (4). The

process noise $\mathbf{q}(t)$ and the measurement noise $r(t)$ are assumed to be additive, zero-mean Gaussian noise:

$$y(t_k) = h\big(\mathbf{x}(t_k)\big) + r(t_k) = [0\ 0\ 0\ 1\ 1\ 2\ 0] \cdot \mathbf{x}(t_k) + r(t_k) \qquad (4)$$

When a new measurement $y(t_k)$ is available, the measurement function effectively distributes the measured value to the fourth, fifth and sixth reaction state, i.e. the species with available binding sites. This measurement function was chosen as the available chemometric sensor model solely accounts for the amount of conjugated drug which is equivalent to the sum of the respective reaction states (cf. section 6.2.3). A detailed mathematical formulation of the EKF working principle and its elementary functions can be found in the Appendix D, S6.3. The EKF was implemented using the *ExtendedKalmanFilter* class in MATLAB 2019b.

6.2.5.3 Process noise covariance

The process noise $\mathbf{q}(t)$ is parameterized by the covariance matrix $\mathbf{Q}$ which provides a measure of uncertainty and cross-correlation for the states x. For non-linear process models, as used here, the implementation of a time-variant process noise covariance matrix $\mathbf{Q}(t)$ is crucial to capture process dynamics and allow the EKF to flexibly adjust the weighting between the models. In literature Valappil and Georgakis [115], two methods have been suggested to implement a time-variant process noise covariance: A linearized approach that computes a time-dependent process noise covariance matrix $\mathbf{Q}_{\mathrm{lin}}(t)$ based on gradient evaluations with regard to the model parameters. This method uses the Jacobian $\mathbf{J}_{\boldsymbol{\theta},\mathrm{nom}}$ with regard to the nominal model parameters $\boldsymbol{\theta}_{\mathrm{nom}}$, namely the kinetic rates $\mathbf{k}$, as well as their covariance matrix $\mathbf{C}_{\theta}$ and computes $\mathbf{Q}(t)$ according to Eq. (5)

$$\mathbf{Q}_{\mathrm{lin}}(t) = \mathbf{J}_{\theta,nom}\ \mathbf{C}_{\theta}\ \mathbf{J}_{\theta,\mathrm{nom}}^{T}\,. \qquad (5)$$

Alternatively, the Monte Carlo (MC) approach uses linear estimates of the parameter covariance matrix $\mathbf{C}_{\theta}$ to span a normal distribution for all model parameters $\boldsymbol{\theta}_{\mathrm{i}}$ with mean μ and variance σ as $\mathcal{N}_{\theta_i}\big(\mu = \theta_{i,norm}, \sigma^2 = \mathbf{C}_{\theta_i,\theta_i}\big)$. A definite number of n random samples are drawn from the distribution and the sampled parameter values are used to propagate the states $\mathbf{x}$ to the next time step t_{k+1}. The timedependent process noise covariance $\mathbf{Q}_{\mathrm{MC}}$ is then computed as the covariance matrix of the mean-centered residuals $\overline{\mathbf{w}}$ according to Eq. (6)

$$\mathbf{Q}_{\mathrm{MC}}(t) = cov(\overline{\mathbf{w}}) = cov(w_n - \overline{\mathbf{w}}) \qquad (6)$$

with $\overline{\mathbf{w}}$ being the mean of the residuals and the residuals w_n being calculated by

$$w_n = f(t_{k+1}, \boldsymbol{\theta}_n) - f(t_{k+1}, \boldsymbol{\theta}_{\text{norm}}) \tag{7}$$

where $\boldsymbol{\theta}_{\text{norm}}$ and $\boldsymbol{\theta}_n$ denote the nominal and the MC samples of the model parameters, respectively.

6.2.5.4 Measurement noise covariance

The measurement noise covariance matrix $\mathbf{R}$ is the sensor-equivalent of $\mathbf{Q}$. As the chemometric sensor in this study only accounts for the amount of conjugated drug, R is a scalar. In contrast to literature as e.g. Feidl et al. [212] and Narayanan [113], a time-variant measurement noise covariance R_{t_k} may be expressed using the predictive variance of the GPR evaluated at all discrete time points t_k. Further information on the computation of the predictive variance is given in the Appendix D, S6.1.

6.2.5.5 EKF tuning

When updating the reaction states using a new available measurement, the EKF combines the information stored in $\mathbf{Q}$ and R to produce an improved estimate of all states. Since the absolute values of $\mathbf{Q}$ and R depend on the estimation method, the process noise covariance is scaled to $\mathbf{Q}^* = \boldsymbol{k}_\mathbf{Q}\mathbf{Q}$ by a scaling factor $k_\mathbf{Q}$ specific for each method. To ensure an equal weighting of all ADC species i, the factor k_Q is optimized with regards to the normalized error $nRMSE$ averaged over all species. The $nRMSE_i$ for each species is calculated according to Eq. 8

$$nRMSE_i = \frac{RMSE_i}{\bar{c}_i} \tag{8}$$

where $\bar{c}_i$ is the respective mean concentration of each species in each run and $nRMSE_i$ being calculated for each species individually as defined in the Appendix D, S6.4. The total $nRMSE$ is obtained by averaging over all runs and species.

6.2.6 Study design

6.2.6.1 Model-based soft-sensor discrimination

The best performing chemometric model was selected based on the methodology described in section 6.2.3 and a kinetic model was calibrated using the procedures given in section 6.2.4. Based on the selected models, the influence of the different computation approaches for $\mathbf{Q}$ and R were compared. The performance was evaluated based on the prediction error regarding the unmeasurable species in the training and test subsets using the GPR predictions for the conjugated drug concentration. For each approach, the scaling factor $k_\mathbf{Q}$ was tuned independently.

6.2.6.2 Challenge scenarios and transfer to fed-batch

The best performing option was selected and further challenged in three scenarios. In scenario 1, the kinetic model was initialized with randomly sampled mAb and NPM concentrations with a standard deviation of 2% of the nominal mAb/NPM concentration. In scenario 2, the kinetic model was initialized with a randomly sampled cysteine distribution within the experimentally determined standard deviation (see chapter 6.2.4). In the scenario 3, the sensor was fed with noisy absorbance spectra. The spectral data for the test subset was modified by additive Gaussian white noise with a standard deviation of 0.02% of the respective absorption value. Additionally, a random, normally distributed offset was introduced with a standard deviation of 5% of the absorbance at 250 nm. In all scenarios, the random sampling and subsequent model evaluation was conducted for 100 repetitions.

Finally, the investigated models were transferred to fed-batch mode. The kinetic model was extended with a feeding term to account for a continuous addition of free drug solution. The GPR model was supplied with the on-line absorbance spectra which were treated as described in section 6.2.3.

6.3 Results

6.3.1 Chemometric model development

Multiple combinations of PLSR and GPR models with different preprocessing strategies were screened in this study and evaluated with regard to their predictive performance on the training and test set. A summary of the results is presented in Table 6.2. The PLSR model in combination with a SGF achieved the lowest *RMSECV*, while the GPR in combination with variable selection (GPR-VS) achieved the lowest *RMSEP*. Overall, the GPR-VS model performed best with regard to both error metrics and improved accuracy by 18% compared to the PLSR model in combination with variable selection (PLSR-VS). Hence for further evaluation, only the GPR-VS model is considered. The resulting selected variables for the GPR-VS model are shown in Appendix D, S6.5.2.

Table 6.2. Overview of predictive performance of selected chemometric models. Both PLSR and GPR models were combined with SGF and a model-specific VS method.

Model	*RMSECV* (µM)	*RMSEP* (µM)
PLSR-SGF	0.31	0.84
PLSR-VS	0.54	0.51
GPR-SGF	0.61	0.82
GPR-VS	0.43	0.43

The predictive mean and variance for the (cross-validated) training and test set are shown Figure 6-2.
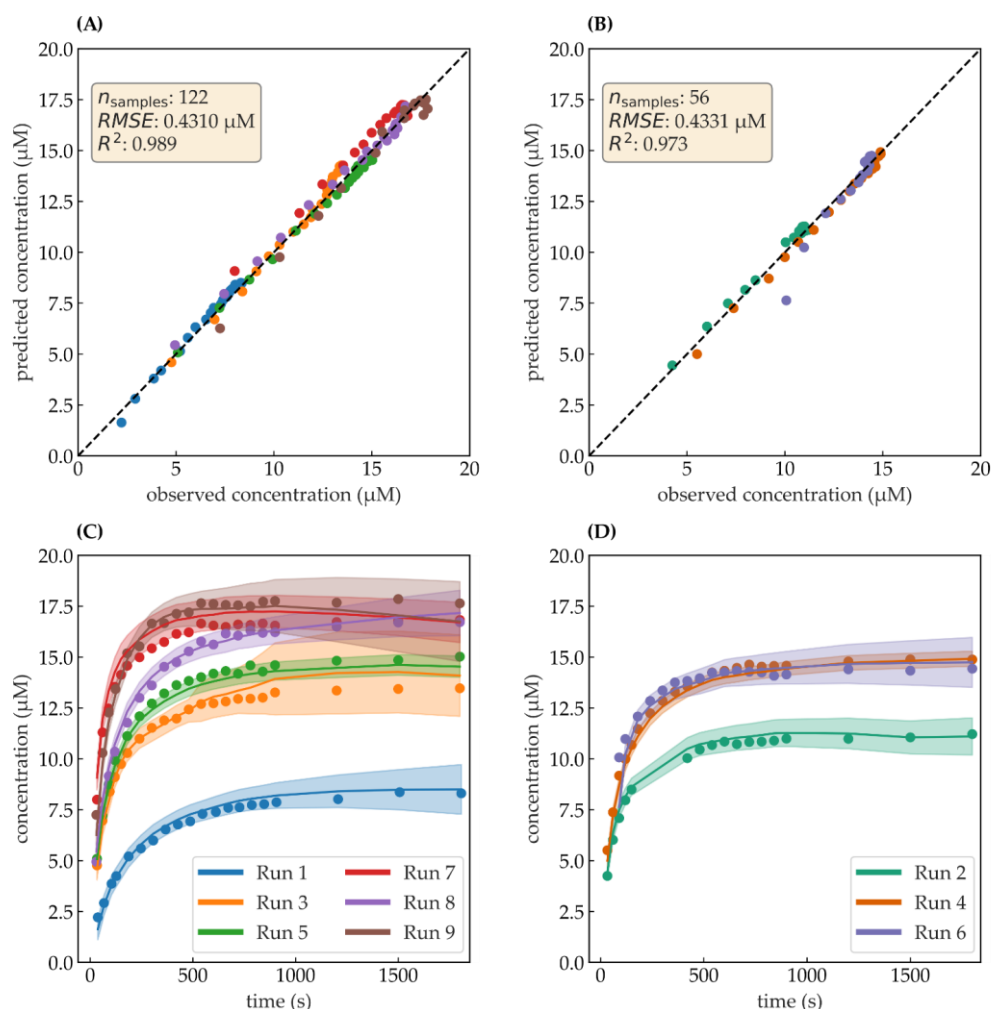

Figure 6-2. Predictions for the GPR-VS model. (A) and (B) show the parity plots for the cross-validation and test subset. The tie line is depicted as dashed line and represents ideal model predictions. The data points are colored by the run numbers. (C) and (D) show the predicted conjugated drug concentration over time. Reference data are illustrated as circles and colored by runs. The GPR-VS predictive mean is showed as solid lines, while the 95% confidence interval is indicated by the colored areas. The 95% confidence interval were approximated by 1.96 standard deviations. Note that the GPR-VS model was only evaluated at the reference data points and the lines are linearly interpolated to guide the eye.

The predictions for the GPR-VS model show good alignment with the reference data for the training and test set with R^2 of 0.968 and 0.973, respectively. For the cross-validation, the predictions for runs 3, 7 and 9 deviate from the reference data towards the end of the reaction. In run 3, an offset of the predictions from the reference data can be observed starting at roughly 500 s. For runs 7 and 9, the predictions reach a maximum at 17.5 μM and decrease towards the end of the reaction to 16.0 μM. This tendency is not present in the reference data. The 95% confidence intervals were approximated by ±1.96

predictive standard deviations σ^* and show a growing uncertainty with time for all runs in all three subsets. For the cross-validation, the confidence intervals of runs 3, 7 and 9 are considerably wider than for the remaining runs indicating an increased uncertainty of the GPR-VS model. For the test subset, the predictions for run 4 deviate only 1 μM from the mean at a 95% confidence as opposed to 2 μM for run 6.

6.3.2 Soft-sensor development

6.3.2.1 Kinetic model calibration

Initially, the end-point compositions of runs with a drug excess of 3x or higher were used to determine the percentages of pre-inactivated cysteines as proposed by Andris et al. [140]. This resulted in 85.98, 10.14 and 3.93% of mAbs with two, one and none available cysteines, respectively. Based on this distribution and the training data, the kinetic rates k_1, k_2 and k_3 were estimated to 0.784 mM^{-1} s^{-1}, 1.625 mM^{-1} s^{-1} and 0.0012 s^{-1}. The parameter covariance matrix was calculated by Eq. 1 for the computation of the process noise covariance for the EKF. The calibrated model resulted in R^2 values for mAb, Conj1 and Conj2 of 94%, 95% and 98% in the training data, and of 93%, 94% and 96% for the test data.

6.3.2.2 Tuning and selection of covariance matrices

In this section, the performance of the EKF with different approaches for the calculation of process noise **Q** and measurement noise R are compared. Each combination was tuned individually by the optimization of the scaling factor k_Q. The effect of the scaling factor on the $nRMSE$ is exemplarily shown for one **Q**/R combination in Figure 6-3.
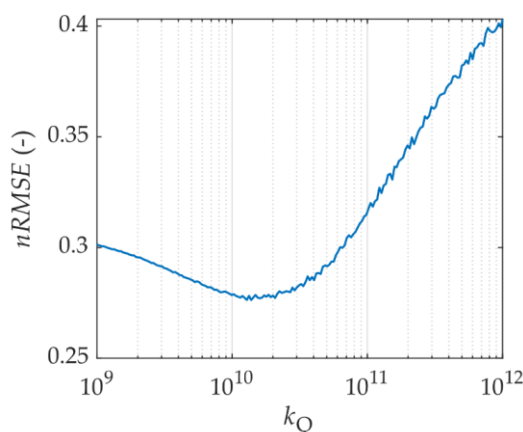

Figure 6-3. EKF prediction error regarding the $nRMSE$ during k_Q tuning

Values in the range of $k_Q=10^9...10^{11}$ result in proper tuning so that the EKF beneficially considers both models in order to minimize the $RMSE$ for all species.

The optimal k_Q varies slightly in each combination of $\mathbf{Q}$ and R as shown in Table 6.3 together with the accomplished *RMSE* reduction by the EKF with regard to the kinetic model predictions. It could be demonstrated that the linearized method does not lead to an overall improvement, while the MC approach achieves an overall *RMSE* reduction in both subsets. Only the prediction error for Conj1 in the training set slightly increases. In summary, this approach yields an averaged *RMSE* reduction of 8.7% and 17.2% for the training and test set, respectively. The time-variant computation of R could further improve the performance by up to 4% for all species compared to the constant setting. In conclusion, the EKF using the $\mathbf{Q}_{MC}$ approach in combination with the R variable setting presents the most valuable combination and is used for the remainder of this study.

Table 6.3. Comparison of the three studied combinations of Q and R regarding their accomplished RMSE reduction compared to the kinetic model in training and test set.

| Q | R | tuned k_Q | Species | *RMSE* reduction (%) | |
				training	test
linearized	variable	9.48	mAb	2.19	2.11
			Conj1	-1.98	3.78
			Conj2	0.19	0.87
MC	variable	10.11	mAb	17.37	11.95
			Conj1	6.67	23.30
			Conj2	21.83	11.07
MC	constant	10.25	mAb	17.78	8.04
			Conj1	8.86	21.04
			Conj2	23.24	7.06

6.3.2.3 EKF soft-sensor performance

The behavior and the soft-sensing capability of the selected EKF combination are exemplarily demonstrated for run 3. Figure 6-4A, B present the predicted conjugated drug concentration as measurable states from the GPR, kinetic model and EKF, as well as the predicted ADC species as the unmeasurable states from the kinetic model and the EKF, respectively. The reference data is overlayed in both figures. During the initial phase until approx. 500 s, it can be observed that the kinetic model predicts the conjugation reaction rate to be faster compared to GPR sensor, which closely follows the reference data. For the unmeasurable states, the same trend is apparent for Conj1 and Conj2. After approximately 200 s, the EKF starts to adjust the trajectory by following the GPR estimates. Towards the end, the kinetic predictions converge to the reference data, while the GPR sensor overestimates the conjugation progress after roughly 600 s. In the same interval, the EKF switches and clearly follows the kinetic model to be aligned with the reference data.
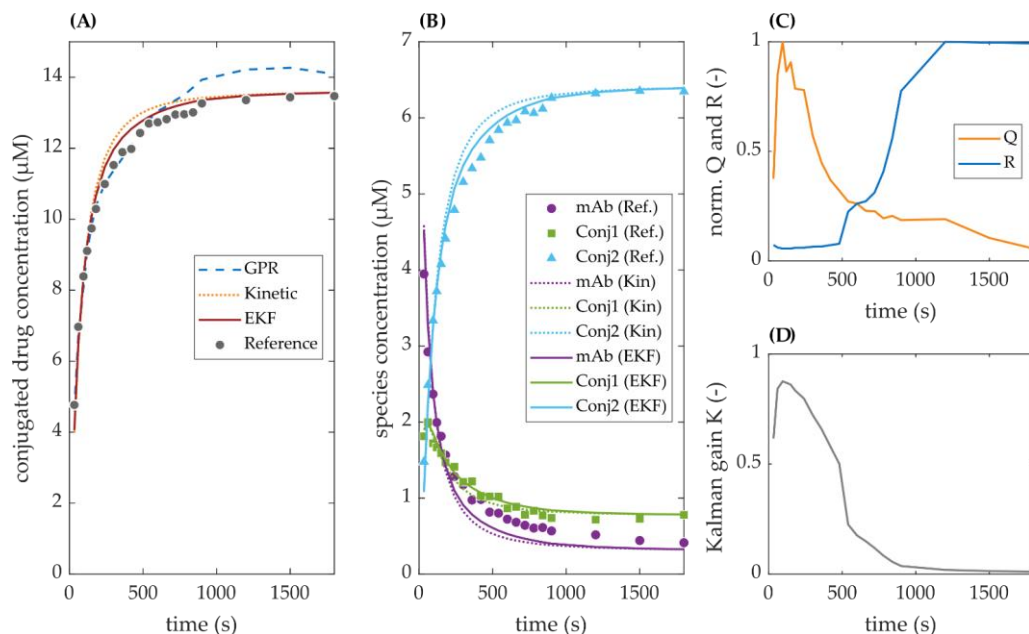
Figure 6-4. Model predictions and calculated process and sensor noise covariance for run 3. (A) Comparison of the conjugated drug concentration as the measured state for the GPR sensor, the kinetic model and the EKF. The reference data from the RP-UHPLC are displayed as circles. (B) Comparison of the predictions for the three ADC species (mAb, Conj1 and Conj2) as the unmeasurable states for kinetic model and the EKF. (C) Time course of the normalized process and sensor noise covariance $\boldsymbol{Q}_{MC}$ and R_{var} and (D) of the Kalman gain K.

This trend can be similarly followed in Figure 6-4C, D, where the normalized time-variant process and sensor noise covariance, and the resulting Kalman gain are displayed. In the beginning, the measurement noise is comparably lower than the process noise regarding all ADC species. This leads the Kalman Gain, and consequently the EKF, to rely more on the GPR predictions with values close to 1. Over the course of the reaction, the process noise is continuously decreasing as the reaction is slowing down and the GPR noise increases conversely relative to the measurement noise. This causes the Kalman gain to shift towards 0 resulting in the EKF relying more strongly on the kinetic model towards the end of the reaction. A similar adjustment dynamic could be observed for all runs in the training and test set and shows the advantage of using time-variant process and measurement noises. The fact that the EKF is converging towards the kinetic model at the end of the reaction, comes along with a decreasing state covariance which may be considered an internal EKF uncertainty. In Appendix D, S6.6, the state covariances for all species are shown for the test subset. For runs, where the Kalman Gain is close to 0 at the end of the reaction, the variance in the EKF estimates is small. Whereas, in Run 4 where the Kalman Gain plateaus at approximately 0.75 due to the high confidence of the GPR sensor, a larger variance is observable. This variance is in agreement with additional data from replicate runs with the same starting conditions which were not used otherwise within this study. Furthermore, the

model and EKF predictions for all training and test runs are depicted in the
Appendix D, S6.6.

6.3.2.4 Soft-sensor challenge scenarios

Three challenging scenarios were created to further investigate the robustness
and accuracy of the designed EKF in comparison to the sole kinetic and sensor
model. The resulting statistical distributions of the *RMSEP* after 100 repetitions
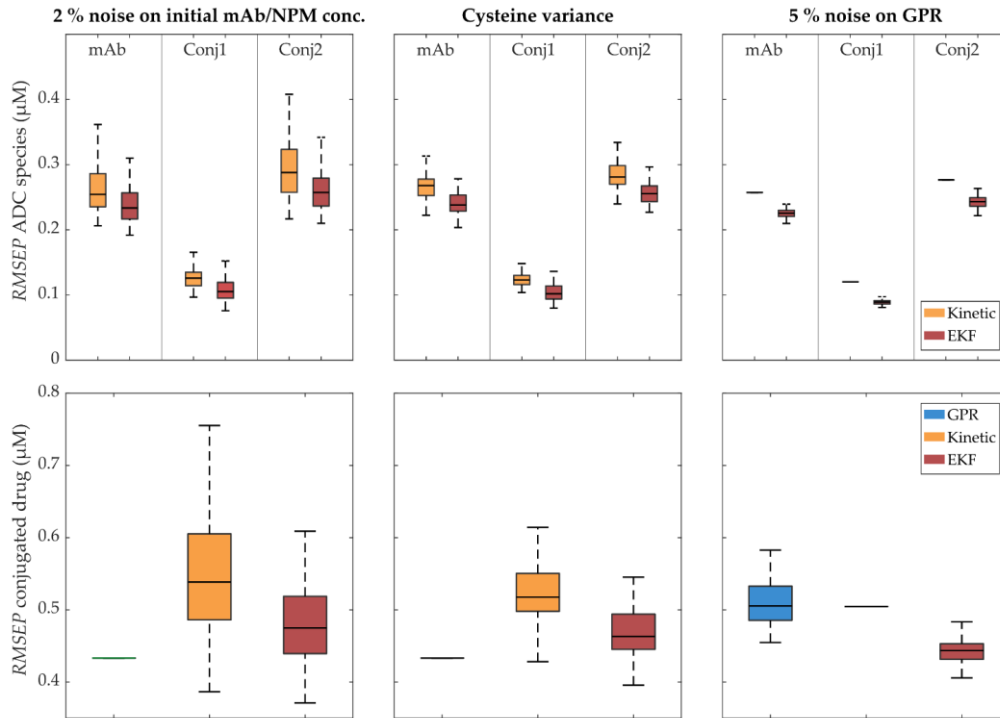are displayed as box plots in Figure 6-5.

Figure 6-5. Comparison of the RMSEP values of the GPR, the kinetic model and
the EKF obtained from 100 simulations for each scenario. The upper row shows
the RMSEP values regarding all ADC species and the lower row the RMSEP
values regarding the conjugated drug concentration.

The upper and lower row present the comparison of the *RMSEP* distributions
with regard to the ADC species and to the conjugated drug, respectively. In the
first column, the effect of noisy initial concentrations of mAb and NPM
(scenario 1) is depicted. By fusing the information from the unmodified GPR
sensor with the faulty initialized kinetic model, the EKF achieves a reduction
of the mean prediction error by 10.7, 21.5 and 18.2% compared to the single
species predictions, mAb, Conj1 and Conj2, from the kinetic model. However,
the interquartile ranges of the displayed boxes are overlapping and their span
is similar for all species indicating that the noisy concentrations are projected
to the EKF estimates. Regarding the conjugated drug, the EKF reduces the
error by 18.0% compared to the kinetic model while not reaching the baseline
of the GPR error entirely. In the second column, the effect of deviation in the

activated cysteines (scenario 2) is shown. All achieved EKF reductions are similar to the results in scenario 1 with overall more dense distributions of the noisy data in both the kinetic model and EKF estimates. The third column demonstrates the effect of noise in the sensor data (scenario 3). With regard to the *RMSEP* of the single species, a reduction in the range of 12%-25% compared to the kinetic model is still accomplished by the EKF. In this case, the EKF effectively reduces the error in the conjugated drug by 13% compared to the noisy GPR as measured by the mean of both distributions.

6.3.2.5 Transfer to fed-batch mode

Finally, the EKF was applied to fed-batch runs in order to test the ability of the EKF to perform under different reaction dynamics. In these runs, the relevant wavelengths exhibit a different behavior compared to the experiments operated in batch mode due to a timely overlay of spectral effects from the conjugation reaction and the NPM feed. A comparison of the spectral data from fed-batch and batch mode can be found in the Appendix D, S6.5.1. To compensate for this changing effect in the absorbance spectra, the GPR was re-calibrated with fed-batch run 10 being part of the training subset.
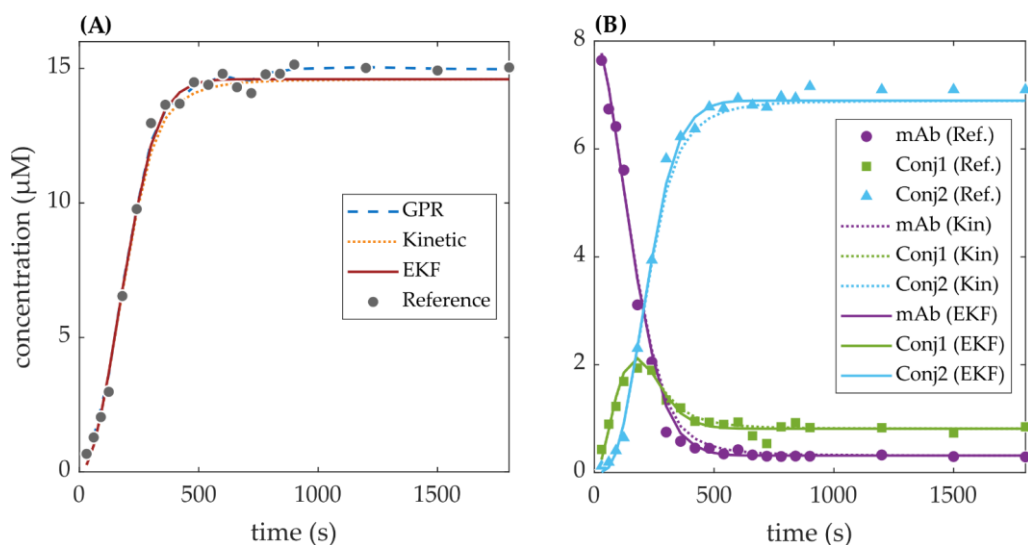

Figure 6-6. Model predictions for fed-batch run 11 for the concentration of conjugated drug (A) and the ADC reaction species (B). To compensate for a changing effect in the absorbance spectra compared to the experiments operated in batch mode, the GPR sensor model was re-calibrated with run 10 being part of the training subset

Figure 6-6 illustrates the model predictions for the fed-batch run 11. The modified kinetic model closely aligns with the reference data until a reaction time of 750 s. Towards the end of the reaction, the modified kinetic model underestimates the conjugated drug concentration, while the GPR sensor follows the reference data. Combining the information of both models, the EKF shows minor improvements during the whole reaction and all species while

converging to the predictions of the modified kinetic model towards the end of the reaction. Thus, the EKF errors for all species are improved compared to the kinetic model by 13.2%, 0.2% and 16.4%. For the conjugated drug, the EKF follows the kinetic model after approx. 750 s and hence results in a relative error increase by 37.9% compared to the GPR model.

6.4   Discussion

6.4.1   Chemometric model development

6.4.1.1   Preprocessing and hyperparameter optimization

The developed GPR model presented in this study uses a composite kernel function and a variable selection algorithm based on the Kullback-Leibler divergence. The model showed elevated predictive performance compared to commonly applied PLSR models and other combinations of kernel functions and preprocessing methods for the GPR (see Table 6.2).

SGF did not improve the predictive capability for either of the two studied regression models. In Andris et al. [98], a baseline correction and subsequent first-derivative filter was found to achieve the best performance, while the herein used GPR model does not rely on any further preprocessing method prior to variable selection. The herein implemented variable selection methods notably improved the predictive accuracy for both the PLSR and the GPR model. This suggests that the absorbance spectra contain uninformative regions which can effectively be removed by the employed algorithms. Particularly wavelengths above 350 nm were found to have a negative impact on the sensor model performances due to the presence of spectral artifacts probably caused by the surrogate drug molecule (cf. Appendix D, S6.5).

The implemented kernel function which is a combination of the linear and the radial-basis function (RBF) kernel has previously been shown to be a viable combination for chemometric purposes in Chen et al. [224]. In other cases, the RBF kernel alone was used for GPR models using near-infrared and Raman spectroscopy data [108], [206], [224]. In general, the authors consider the usage of covariance kernels beneficial to model the non-linear observed spectral shift and baseline drift of the absorbance spectra compared to the linear subspace projections the PLSR models are based upon. It has previously been shown, that peak shifts which may be resembled by closely overlapping absorption maxima of multiple components have a strong influence on the regression vector in PLS modeling which may be a reason for the inferior performance compared to kernel-based models [225], [226].

6.4.1.2   Quantification of spectra-based uncertainty

When the GPR model is evaluated at an unseen spectrum, the kernel function computes the covariance with all available training instances and uses that information to quantify the prediction estimate and the associated uncertainty. The uncertainty estimate is hence influenced by the distribution of available data points along the process as well as the spectral differences among the samples. As emphasized in Figure 6-2, the trajectories of run 4 and 6 closely resemble one another. However, the associated uncertainty estimate is considerably higher for run 6. This observation is in agreement with the experimental data as run 4 is a replicate of run 5 and has been conducted on the same day. Run 6, instead, has a different initial condition than all training runs suggesting higher uncertainty due to unseen spectral variation. Similarly, run 3 and run 9 supposedly exhibit spectral features causing the covariance function to produce a wider confidence interval than the other runs. In fact, the variance increases towards the end of the reaction when the conjugation reaction is considerably slowing down. This may be influenced by an irregular chemical degradation of the NPM molecule which has been observed in experiments with pure NPM solutions (cf. Appendix D, S6.5). However, this effect may not occur for other molecules such as real payloads due to incomparable spectral properties [227].

It should be noted that GPR models are a comparably novel method in chemometrics and the physical validity of the computed confidence intervals has not been extensively studied in the literature. As the confidence intervals are computed with regard to both the experimental variance of the reference and the spectral data, the predictive uncertainty of the GPR can not directly compared to the variance of the reference data alone. Nonetheless, the implemented method provides further insights into the quality of the spectral data and the relation with the training instances and exceeds conventional models in terms of predictive accuracy. The GPR is therefore considered sufficiently accurate to be used as a sensor within the soft-sensor framework. In future, separate studies specifically designed for the validation of predictive uncertainty estimates should be conducted.

6.4.2   Kinetic model calibration

The identified kinetic reaction rates are in agreement with the ones found in Andris et al. [140] with k_1 and k_2 being 1.6 and 9.2% below and above the values in the referenced study. The difference might originate from minor differences in the determined cysteine distribution or the overall experimental or analytical procedure. The distribution of pre-inactivated cysteines is a necessary model parameter to account for actual amount of available bindings sites on the

oxidized mAb which in practice are lower than two. This phenomenon is already discussed in detail in Cao et al. [228].

6.4.3 Soft-sensor development

The selection and appropriate tuning of the process and sensor noise covariance play a crucial role in the development of an optimal state estimator such as an EKF [211]. It has been observed that the EKF mainly especially selects the Conj1 state to compensate for the deviation when k_Q is set too high. This unfavorable adjustment is due to the underlying physical constraints from the kinetic model ODE system and could be effectively minimized by using a normalized *RMSE* for tuning.

Valappil and Georgakis [115] stated that the applicability of the computation method for $\mathbf{Q}(t)$ has to be systemically tested for each individual case. In our study, the MC method outperforms the linearized method in terms of prediction error (see Table 6.3). Remarkably, the EKF using the MC method improves the errors for all ADC species in the test set of at least 11%. It appears that the sampling-based MC method is more adequate and robust in the case of the non-linear ADC reaction. However, it should be noted that the parameter distributions used for the MC method are solely approximations. For increasingly non-linear models, Gaussian distributions may no longer be assumed for all model parameters. Another disadvantage of the MC method may be the diminishing process variance and hence an overconfident EKF state estimate towards the steady-state of the reaction. If applicable, Bayesian inference methods could be used to assess more accurate estimations of the parameter distributions and provide a means for more conservative variance estimation.

The sensor noise covariance R is commonly set to a constant scalar and derived from the calibration error of the sensor model [113], [211], [212]. The usage of the GPR model firstly allows the implementation of a time-varying sensor noise covariance. As the uncertainty derived from the GPR increases with time, the sensor noise resembles this trend and the EKF counteracts to this by relying more on the kinetic model towards the end of the reaction. In comparison to a time-invariant R, the variable approach achieves small improvements, while the setting of $\mathbf{Q}$ majorly contributes to the presented error reduction. However, it should be emphasized that in this case the GPR provides an estimate of the cumulative conjugated drug and does not affect all states individually. It would be hence interesting to evaluate the performance of the variable sensor noise in cases where direct measurements of all states are available. Moreover, another drawback of the implemented EKF are the steady-state constraints, i.e. the experimentally derived cysteine distribution assumed for all batches, which

make the EKF rather stiff towards the end of the reaction. This is a known problem of the EKF [206] when dealing with only a few states being available for measurement as well as for non-linear models. In theory, this problem could be circumvented with other more flexible state estimation algorithms, such as particle filters as implemented in Golabgir and Herwig [210] and Stelzer et al. [229] or unconstrained Kalman filter algorithms as used in Kolås [230] and Simutis and Lübbert [209] which were beyond the scope of this study.

The three challenging scenarios prove that even for a faulty EKF initialization and a noisy sensor model, the designed soft-sensor consistently leads to more accurate predictions compared to the sole chemometric or kinetic models. In cases of large offsets of the kinetic model, the EKF could be averted to rely on the sensor model instead, which is demonstrated in challenge scenario 1 and 2. In case of noisy sensor inputs, the EKF produces less noisy predictions, also regarding the unmeasured species. As discussed before, the predictions of both the process and the sensor model are dependent on the variations in the absorbance spectra or initial conditions. The EKF was shown to be able to compensate foreseeable deviations and hence demonstrated improved robustness which is essential for the successful deployment of the soft-sensor.

Finally, the adaptability of the soft-sensor was investigated by transfer to fed-batch operation mode. Under fed-batch conditions, the EKF also performed adequately while an extended re-tuning was not required. Despite the necessary addition of the fed-batch run 10 to the training data set for the GPR model (see chapter 6.3.2.5), the configuration of the model remained the same. It is therefore recommended to perform multiple fed-batch experiments with varying initial conditions if the chemometric model is intended to be used for this operation mode. In the studied case, the kinetic model aligns perfectly with the reference data due to accurate initial values and precise feeding. Regarding the transferability to other setups or ADC reactions, a re-tuning of the EKF may be necessary as soon as one of the two underlying models' performances change, as the EKF relies on weighting of the two models. However, a thorough investigation this was out of the scope of this study. The scenarios and the fed-batch transfer showed that the EKF is less sensitive to deviations and could be implemented for model-based process control, e.g. for endpoint determination of the conjugation reaction.

6.5 Conclusion

In the ADC conjugation reaction, a cytotoxic payload is conjugated to a previously functionalized mAb with different levels of available binding sites. Current multivariate measurement approaches only enable the on-line

measurement of the amount of conjugated drug as a whole but not the ADC species along the conjugation pathway. In this study, a robust and adaptive model-based soft-sensor for the advanced monitoring of a site-directed ADC conjugation reaction was presented combining a chemometric sensor based on a GPR model with a kinetic reaction model. The devised GPR model exhibited superior accuracy compared to commonly applied PLSR models and provided uncertainty estimates for each measurement. The associated uncertainty is directly influenced by the variation in the spectral data and could provide a means for risk-based monitoring approaches. The deployed soft-sensor EKF fuses the information from the GPR model with the timely forecasts of the kinetic model. The EKF is able to dynamically adjust the predictions of all unmeasurable ADC species throughout the entire reaction using time-variant computation approaches for the process and sensor noise covariances. The designed EKF improved the predictions on average by 8.7% and 17.2% for the training and test set, respectively. The EKF was further shown to be capable to compensate for experimental noise effects, wrong model initialization and poorly determined kinetic model parameters compared the unmodified kinetic or sensor model. In conclusion, the combination of a chemometric sensor and a kinetic model using an EKF-based soft-sensor is highly beneficial for accurate process monitoring and enables a high degree of flexibility with regard to process deviations. This ultimately helps to facilitate robust process control or early end-point determination. The presented methodology further ensures a unified calibration routine for all underlying components and provides knowledge on sensor and model uncertainties.

Acknowledgments

Supplementary Materials

Appendix D contains the Supplementary Materials associated with this chapter:

- S6.1 Gaussian process regression
- S6.2 Kinetic modeling
- S6.3 Extended Kalman Filter
- S6.4 Error metrics

- S6.5 Chemometric model development

- S6.6 Soft-sensor development

7

General Discussion and Conclusion

Having the appropriate digital tools in place for the predictive modeling of processes is key to accelerating the development of more robust and scalable biopharmaceutical manufacturing processes. The goal of this work is to showcase the capability of modeling approaches in the area of process development and manufacturing for antibody-drug conjugates (ADC). The focus was specifically on the conjugation process of two related ADC modalities—site-directed cysteine conjugation and stochastic cysteine-based conjugation. Kinetic modeling was employed as a core method to enhance the scientific process understanding and to demonstrate its application within model-based concepts, intending to align with current Quality by Design (QbD)-demands. The first part of this thesis (Chapters 3 and 4), concentrated on the establishment of mechanistic kinetic models for the reduction and conjugation steps, being the primary reaction steps in the conjugation workflow to control the final ADC CQAs. This provided novel insights into the underlying physical and biochemical mechanisms essential for the successful process development. The second part (Chapter 5 and 6) presented two case studies dealing with the application of an existing kinetic model for the site-directed conjugation reaction: (1) model integration with computational fluid dynamics (CFD) to investigate reaction scalability and influence or large-scale process parameters and (2) combination with a Process Analytical Technology (PAT) sensor to build a soft-sensor to advance real-time monitoring capability.

In the first research project (Chapter 3), conjugation kinetic models were developed for both the site-specific conjugation to engineered cysteines (Drug-to-Antibody Ratio (DAR) 2) and the stochastic conjugation to reduced interchain disulfides (DAR 8) using multiple payloads. To ensure consistency in the modeling workflow and facilitate industrial application, all kinetic

samples were analyzed with a widely applied reversed-phase ultra high-performance liquid chromatography (RP-UHPLC) method. Fed-batch conjugations supported resolving the fast DAR 8 conjugation kinetics, which allowed to gain insights about the consecutive conjugation to heavy chain cysteines. Additional investigations about individual payload stability by ultraviolet/visible (UV/Vis) measurements revealed remarkable differences in the payload stability, which showed to be important to consider for defining the minimum required payload excess. The established kinetic models could also describe the presence of over- and under-conjugated species and quantify differences in the conjugation rate of various payloads. Furthermore, comparable conjugation rates could be determined when the same payload was applied for both DAR 2 and DAR 8 monoclonal antibodies (mAbs). Overall, the modeling results demonstrated that the conjugation kinetics are relatively fast, especially for DAR 8, and that its design mainly involves defining the required payload excess. Finally, the benefits of the kinetic model for model-based process development were highlighted using an *in silico* screening to simulate the process behavior on changing starting concentrations. In the future, this approach may assist in determining the most economic payload excess to reach full conjugation.

The previous study suggested that, especially for conjugation to interchain disulfides, the achievable DAR is determined by the previous reducing step. The second study (Chapter 4) describes the development of an integrated reduction kinetic model for the cysteine-based conjugation workflow, featuring final DAR/Drug load distribution (DLD) prediction for DAR values in between 2 to 8. Experimental reduction kinetics were conducted with varying reaction parameters and were analyzed by both capillary gel electrophoresis (CGE) and RP-UHPLC. As a first model component, a kinetic model for the partial reduction of the interchain disulfides was established based on the CGE-determined reduction species. The model could successfully parametrize the effect of the reaction temperature in between 4 and 37 °C utilizing the Arrhenius approach and account for reduced intermediates as well as backward re-oxidation pathways. From the calibrated kinetic rates, it became evident that reduction rates vary significantly depending on the amount of intact cysteine bonds. Afterwards, multiple linear regression (MLR) models were established to directly predict both the DAR and the percentages of the DLD species without necessitating to account for the conjugation kinetic. Combined with the reduction kinetic model an integrated reduction model could be constructed. This model demonstrated its favorable use for simulating the complex design space, i.e., variations of reduction reaction parameters, such as reaction time, reducing agent excess and temperature, and their effect on the DAR/DLD. Finally, the model showed its capability to optimize the reaction parameters for

a desired DAR value while maximizing the yield of certain DLD species. Another aspect of this study was the transferability of each single model on another mAb, which suggested mAb-dependent reduction rates. In conclusion, a novel modeling approach for the two connected reaction steps involved in cysteine conjugates is established using two analytics and only a handful of experiments.

In Chapter 5, it was explored mixing at different scales and scale-up affect the conjugation kinetic at both small- and large-scales. In small-scale, experimental kinetic studies in reaction tubes were optimized by comparing different internal and external mixers as well as mixing speeds. The optimized mixing parameters showed to result in DAR kinetics that are reproducible and comparable to kinetics in lab-scale stirred vessels. Next, a conjugation kinetic model for DAR 2 was incorporated into validated, steady-state CFD simulations of three commonly used large-scale vessels (two glass vessels of different sizes and one single-use vessel). The characterization of the fluid dynamics allowed to simulate locally resolved conjugation kinetics inside these vessels, which were compared to the predictions of the original perfectly-mixed model. Additionally, for one vessel, potential variations in process parameters, such as stirrer speed, feeding mode and concentration variation, and their effect on the DAR kinetic were estimated. It was concluded that both scale and parameter variations result only in short-term deviation and level out due to the consecutive nature of the reaction. It was further demonstrated that mixing time can be used as a reliable scale indicator to estimate potential scale influence when compared to the reaction time scale. In total, this study suggested that the kinetic model developed using small-scale experimental kinetics can be applied to predict large-scale conjugation kinetic.

The fourth research project (Chapter 6) addressed the limitation of a recently established UV/Vis-based PAT sensor for the site-directed conjugation reaction to measure the conjugated drug concentration, but not the single ADC species. A model-based soft-sensor framework based on an Extended Kalman Filter (EKF) was developed to robustly combine the sensor predictions with the kinetic model and, thus, enable the monitoring of non-observable single species. Since the EKF also accounts for the time-varying models' uncertainty or noise, a Gaussian process regression model was developed as the sensor model. Upon proper tuning and process noise selection, the soft-sensor EKF demonstrated to result in lower model error compared to the uncorrected kinetic model for all nine experimental conjugation runs. Within three challenges scenarios, focusing on incorrect model initialization and noisy sensor predictions, the framework also proved to be robust. Finally, it was also evaluated that the framework is transferable to fed-batch conjugation. In summary, this study presented an

EKF-based soft sensor to enhance process monitoring accuracy and flexibility, which facilitates robust process control and early endpoint determination.

Overall, the herein described models and concepts provide significant advancements in the model-based ADC process development and manufacturing. It became evident that mechanistic kinetic models are essential to understanding the multi-dimensionality of the governing process parameters for single reaction steps as well as how connected reaction steps interrelate. The potential of kinetic models for ADC process digitization could be highlighted by utilizing a conjugation kinetic model for assessing scale-up and advancing real-time monitoring. In the future, this may ultimately lead to increasing process knowledge and helping to ensure defining robust design spaces. This will pave the way towards model-driven process development and implementation of a digital twin of the full conjugation workflow.

Bibliography

[1]   M. Kesik-Brodacka, "Progress in biopharmaceutical development," *Biotechnol. Appl. Biochem.*, vol. 65, no. 3, pp. 306–322, May 2018, doi: https://doi.org/10.1002/bab.1617.

[2]   M. Misra, "Biosimilars: Current perspectives and future implications," *Indian J. Pharmacol.*, vol. 44, no. 1, pp. 12–14, 2012, doi: 10.4103/0253-7613.91859.

[3]   D. E. Gerber, "Targeted therapies: a new generation of cancer treatments," *Am. Fam. Physician*, vol. 77, no. 3, pp. 311–319, 2008.

[4]   A. C. Szkodny and K. H. Lee, "Biopharmaceutical Manufacturing: Historical Perspectives and Future Directions," *Annu. Rev. Chem. Biomol. Eng.*, vol. 13, no. Volume 13, 2022, pp. 141–165, 2022, doi: https://doi.org/10.1146/annurev-chembioeng-092220-125832.

[5]   A. M. Wu and P. D. Senter, "Arming antibodies: Prospects and challenges for immunoconjugates," *Nat. Biotechnol.*, vol. 23, no. 9, pp. 1137–1146, 2005, doi: 10.1038/nbt1141.

[6]   I. Pastan, R. Hassan, D. J. FitzGerald, and R. J. Kreitman, "Immunotoxin therapy of cancer," *Nat. Rev. Cancer*, vol. 6, no. 7, pp. 559–565, 2006, doi: 10.1038/nrc1891.

[7]   C. Dumontet, J. M. Reichert, P. D. Senter, J. M. Lambert, and A. Beck, "Antibody–drug conjugates come of age in oncology," *Nat. Rev. Drug Discov.*, vol. 22, no. 8, pp. 641–661, 2023, doi: 10.1038/s41573-023-00709-2.

[8]   F. Riccardi, M. D. Bo, P. Macor, and G. Toffoli, "A comprehensive overview on antibody-drug conjugates : from the conceptualization to cancer therapy," no. September, pp. 1–21, 2023, doi: 10.3389/fphar.2023.1274088.

[9]   P. Holm, M. Allesø, M. C. Bryder, and R. Holm, "Q8 (R2) Pharmaceutical Development," *ICH Qual. Guidel. an Implement. Guid.*, pp. 535–577, 2017.

[10]   D. C. Hinz, "Process analytical technologies in the pharmaceutical industry: the FDA's PAT initiative," *Anal. Bioanal. Chem.*, vol. 384, no. 5, pp. 1036–1042, 2006, doi: 10.1007/s00216-005-3394-y.

[11]   P. Valent *et al.*, "Paul Ehrlich (1854-1915) and His Contributions to the Foundation and Birth of Translational Medicine," *J. Innate Immun.*, vol. 8, no. 2, pp. 111–120, 2016, doi: 10.1159/000443526.

[12]   K. Strebhardt and A. Ullrich, "Paul Ehrlich's magic bullet concept: 100

Years of progress," *Nat. Rev. Cancer*, vol. 8, no. 6, pp. 473–480, 2008, doi: 10.1038/nrc2394.

[13]   R. S. Schwartz, "Paul Ehrlich's Magic Bullets," *N. Engl. J. Med.*, vol. 350, no. 11, pp. 1079–1080, 2004, doi: 10.1056/NEJMp048021.

[14]   A. Gilman and F. S. Philips, "The Biological Actions and Therapeutic Applications of the B-Chloroethyl Amines and Sulfides," *Science (80-.).*, vol. 103, no. 2675, pp. 409–436, 1946, doi: 10.1126/science.103.2675.409.

[15]   S. Farber, L. K. Diamond, R. D. Mercer, R. F. Sylvester, and J. A. Wolff, "Temporary Remissions in Acute Leukemia in Children Produced by Folic Acid Antagonist, 4-Aminopteroyl-Glutamic Acid (Aminopterin)," *N. Engl. J. Med.*, vol. 238, no. 23, pp. 787–793, 1948, doi: 10.1056/NEJM194806032382301.

[16]   M. Visentin, R. Zhao, and I. D. Goldman, "The Antifolates," *Hematol. Oncol. Clin. North Am.*, vol. 26, no. 3, pp. 629–648, 2012, doi: https://doi.org/10.1016/j.hoc.2012.02.002.

[17]   L. Falzone, S. Salomone, and M. Libra, "Evolution of Cancer Pharmacological Treatments at the Turn of the Third Millennium," *Front. Pharmacol.*, vol. 9, 2018, doi: 10.3389/fphar.2018.01300.

[18]   P. Khongorzul, C. J. Ling, F. U. Khan, A. U. Ihsan, and J. Zhang, "Antibody-drug conjugates: A comprehensive review," *Mol. Cancer Res.*, vol. 18, no. 1, pp. 3–19, 2020, doi: 10.1158/1541-7786.MCR-19-0582.

[19]   G. MATHE, L. O. C. TRAN BA, and J. BERNARD, "Effet sur la leucémie 1210 de la souris d'un combinaison par diazotation d'A-méthoptèrine et de gamma-globulines de hamsters porteurs de cette leucèmie par hétérogreffe," *C. R. Hebd. Seances Acad. Sci.*, vol. 246, no. 10, pp. 1626–1628, Mar. 1958.

[20]   G. S. Hamilton, "Antibody-drug conjugates for cancer therapy: The technological and regulatory challenges of developing drug-biologic hybrids," *Biologicals*, vol. 43, no. 5, pp. 318–332, 2015, doi: 10.1016/j.biologicals.2015.05.006.

[21]   K. Tsuchikama and Z. An, "Antibody-drug conjugates: recent advances in conjugation and linker chemistries," *Protein Cell*, vol. 9, no. 1, pp. 33–46, 2018, doi: 10.1007/s13238-016-0323-0.

[22]   G. Köhler and C. Milstein, "Continuous cultures of fused cells secreting antibody of predefined specificity. 1975.," *J. Immunol.*, vol. 174, no. 5, pp. 2453–2455, Mar. 2005.

[23]   J. K. H. Liu, "The history of monoclonal antibody development e Progress , remaining challenges and future innovations," *Ann. Med. Surg.*, vol. 3, no. 4, pp. 113–116, 2014, doi: 10.1016/j.amsu.2014.09.001.

[24]   D. Schrama, R. A. Reisfeld, and J. C. Becker, "Antibody targeted drugs

as cancer therapeutics," *Nat. Rev. Drug Discov.*, vol. 5, no. 2, pp. 147–159, 2006, doi: 10.1038/nrd1957.

[25]   C. H. Ford *et al.*, "Localisation and toxicity study of a vindesine-anti-CEA conjugate in patients with advanced cancer," *Br. J. Cancer*, vol. 47, no. 1, pp. 35–42, 1983.

[26]   G. Pietersz and K. Krauer, "Antibody-Targeted Drugs for the Therapy of Cancer," *J. Drug Target.*, vol. 2, no. 3, pp. 183–215, Jan. 1994, doi: 10.3109/10611869408996804.

[27]   E. L. Sievers *et al.*, "Efficacy and safety of gemtuzumab ozogamicin in patients with CD33-positive acute myeloid leukemia in first relapse.," *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 19, no. 13, pp. 3244–3254, Jul. 2001, doi: 10.1200/JCO.2001.19.13.3244.

[28]   B. ten Cate *et al.*, "A novel AML-selective TRAIL fusion protein that is superior to Gemtuzumab Ozogamicin in terms of in vitro selectivity, activity and stability.," *Leukemia*, vol. 23, no. 8, pp. 1389–1397, Aug. 2009, doi: 10.1038/leu.2009.34.

[29]   J. M. Sasso *et al.*, "The Evolving Landscape of Antibody-Drug Conjugates: In Depth Analysis of Recent Research Progress," *Bioconjug. Chem.*, vol. 34, no. 11, pp. 1951–2000, 2023, doi: 10.1021/acs.bioconjchem.3c00374.

[30]   P. J. Carter and G. A. Lazar, "Next generation antibody drugs: Pursuit of the 'high-hanging fruit,'" *Nat. Rev. Drug Discov.*, vol. 17, no. 3, pp. 197–223, 2018, doi: 10.1038/nrd.2017.227.

[31]   M. Damelin, W. Zhong, J. Myers, and P. Sapra, "Evolving Strategies for Target Selection for Antibody-Drug Conjugates," *Pharm. Res.*, vol. 32, no. 11, pp. 3494–3507, 2015, doi: 10.1007/s11095-015-1624-3.

[32]   N. Kommineni, P. Pandi, N. Chella, A. J. Domb, and W. Khan, "Antibody drug conjugates: Development, characterization, and regulatory considerations," *Polym. Adv. Technol.*, vol. 31, no. 6, pp. 1177–1193, 2020, doi: 10.1002/pat.4789.

[33]   M. Nejadmoghaddam, A. Minai-tehrani, and R. Ghahremanzadeh, "Antibody-Drug Conjugates : Possibilities and Challenges," vol. 11, no. 1, 2019.

[34]   M. R. Gordon, M. Canakci, L. Li, J. Zhuang, B. Osborne, and S. Thayumanavan, "Field Guide to Challenges and Opportunities in Antibody-Drug Conjugates for Chemists," *Bioconjug. Chem.*, vol. 26, no. 11, pp. 2198–2215, 2015, doi: 10.1021/acs.bioconjchem.5b00399.

[35]   H. L. Perez *et al.*, "Antibody-drug conjugates: Current status and future directions," *Drug Discov. Today*, vol. 19, no. 7, pp. 869–881, 2014, doi: 10.1016/j.drudis.2013.11.004.

[36]   B. E. C. G. De Goeij *et al.*, "Efficient payload delivery by a bispecific antibody-drug conjugate targeting HER2 and CD63," *Mol. Cancer Ther.*, vol. 15, no. 11, pp. 2688–2697, 2016, doi: 10.1158/1535-7163.MCT-16-0364.

[37]   J. Lu, F. Jiang, A. Lu, and G. Zhang, "Linkers having a crucial role in antibody–drug conjugates," *Int. J. Mol. Sci.*, vol. 17, no. 4, 2016, doi: 10.3390/ijms17040561.

[38]   N. Jain, S. W. Smith, S. Ghone, and B. Tomczuk, "Current ADC Linker Chemistry," *Pharm. Res.*, vol. 32, no. 11, pp. 3526–3540, 2015, doi: 10.1007/s11095-015-1657-7.

[39]   G. M. Dubowchik *et al.*, "Cathepsin B-Labile Dipeptide Linkers for Lysosomal Release of Doxorubicin from Internalizing Immunoconjugates: Model Studies of Enzymatic Drug Release and Antigen-Specific In Vitro Anticancer Activity," *Bioconjug. Chem.*, vol. 13, no. 4, pp. 855–869, Jul. 2002, doi: 10.1021/bc025536j.

[40]   M. Ritchie, L. Tchistiakova, and N. Scott, "Implications of receptor-mediated endocytosis and intracellular trafficking dynamics in the development of antibody drug conjugates," *MAbs*, vol. 5, no. 1, pp. 13–21, Jan. 2013, doi: 10.4161/mabs.22854.

[41]   T. Tedeschini *et al.*, "Polyethylene glycol-based linkers as hydrophilicity reservoir for antibody-drug conjugates," *J. Control. Release*, vol. 337, pp. 431–447, 2021, doi: https://doi.org/10.1016/j.jconrel.2021.07.041.

[42]   R. Y. Zhao *et al.*, "Synthesis and Evaluation of Hydrophilic Linkers for Antibody–Maytansinoid Conjugates," *J. Med. Chem.*, vol. 54, no. 10, pp. 3606–3623, May 2011, doi: 10.1021/jm2002958.

[43]   L. Shefet-Carasso and I. Benhar, "Antibody-targeted drugs and drug resistance - Challenges and solutions," *Drug Resist. Updat.*, vol. 18, pp. 36–46, 2015, doi: 10.1016/j.drup.2014.11.001.

[44]   D. Leung, J. Wurst, T. Liu, R. Martinez, A. Datta-Mannan, and Y. Feng, "Antibody Conjugates-Recent Advances and Future Innovations," *Antibodies*, vol. 9, no. 1, p. 2, 2020, doi: 10.3390/antib9010002.

[45]   V. Metrangolo and L. H. Engelholm, "Antibody–Drug Conjugates: The Dynamic Evolution from Conventional to Next-Generation Constructs," *Cancers (Basel).*, vol. 16, no. 2, pp. 1–20, 2024, doi: 10.3390/cancers16020447.

[46]   A. S. Ratnayake *et al.*, "Natural Product Bis-Intercalator Depsipeptides as a New Class of Payloads for Antibody–Drug Conjugates," *Bioconjug. Chem.*, vol. 30, no. 1, pp. 200–209, Jan. 2019, doi: 10.1021/acs.bioconjchem.8b00843.

[47]   L. Conilh, L. Sadilkova, W. Viricel, and C. Dumontet, "Payload

diversification: a key step in the development of antibody–drug conjugates," *J. Hematol. Oncol.*, vol. 16, no. 1, p. 3, 2023, doi: 10.1186/s13045-022-01397-y.

[48]   C. M. Yamazaki *et al.*, "Antibody-drug conjugates with dual payloads for combating breast tumor heterogeneity and drug resistance," *Nat. Commun.*, vol. 12, no. 1, pp. 1–13, 2021, doi: 10.1038/s41467-021-23793-7.

[49]   Z. Fu, S. Li, S. Han, C. Shi, and Y. Zhang, "Antibody drug conjugate: the 'biological missile' for targeted cancer therapy," *Signal Transduct. Target. Ther.*, vol. 7, no. 1, 2022, doi: 10.1038/s41392-022-00947-7.

[50]   S. Panowski, S. Bhakta, H. Raab, P. Polakis, and J. R. Junutula, "Site-specific antibody drug conjugates for cancer therapy," *MAbs*, vol. 6, no. 1, pp. 34–45, 2014, doi: 10.4161/mabs.27022.

[51]   Y. Matsuda, M. Leung, T. Okuzumi, and B. Mendelsohn, "A purification strategy utilizing hydrophobic interaction chromatography to obtain homogeneous species from a site-specific antibody drug conjugate produced by ajicaptm first generation," *Antibodies*, vol. 9, no. 2, 2020, doi: 10.3390/antib9020016.

[52]   C. R. Behrens and B. Liu, "Methods for site-specific drug conjugation to antibodies," *MAbs*, vol. 6, no. 1, pp. 46–53, 2014, doi: 10.4161/mabs.26632.

[53]   M. H. Hutchinson, R. S. Hendricks, X. X. Lin, and D. A. Olsson, "Process Development and Manufacturing of Antibody-Drug Conjugates," *Biopharm. Process. Dev. Des. Implement. Manuf. Process.*, pp. 813–836, 2018, doi: 10.1016/B978-0-08-100623-8.00041-4.

[54]   H. Yao, F. Jiang, A. Lu, and G. Zhang, "Methods to design and synthesize antibody-drug conjugates (ADCs)," *Int. J. Mol. Sci.*, vol. 17, no. 2, 2016, doi: 10.3390/ijms17020194.

[55]   A. M. Sochaj, K. W. Świderska, and J. Otlewski, "Current methods for the synthesis of homogeneous antibody-drug conjugates," *Biotechnol. Adv.*, vol. 33, no. 6, pp. 775–784, 2015, doi: 10.1016/j.biotechadv.2015.05.001.

[56]   J. You, J. Zhang, J. Wang, and M. Jin, "Cysteine-Based Coupling: Challenges and Solutions," *Bioconjug. Chem.*, vol. 32, no. 8, pp. 1525–1534, 2021, doi: 10.1021/acs.bioconjchem.1c00213.

[57]   A. V. Kamath and S. Iyer, "Preclinical Pharmacokinetic Considerations for the Development of Antibody Drug Conjugates," *Pharm. Res.*, vol. 32, no. 11, pp. 3470–3479, 2015, doi: 10.1007/s11095-014-1584-z.

[58]   C. R. Behrens *et al.*, "Antibody-Drug Conjugates (ADCs) Derived from Interchain Cysteine Cross-Linking Demonstrate Improved Homogeneity

and Other Pharmacological Properties over Conventional Heterogeneous ADCs," *Mol. Pharm.*, vol. 12, no. 11, pp. 3986–3998, 2015, doi: 10.1021/acs.molpharmaceut.5b00432.

[59]   G. Badescu *et al.*, "Bridging Disulfides for Stable and Defined Antibody Drug Conjugates," *Bioconjug. Chem.*, vol. 25, no. 6, pp. 1124–1136, Jun. 2014, doi: 10.1021/bc500148x.

[60]   Y. Matsuda and B. A. Mendelsohn, "An overview of process development for antibody-drug conjugates produced by chemical conjugation technology," *Expert Opin. Biol. Ther.*, vol. 21, no. 7, pp. 963–975, 2021, doi: 10.1080/14712598.2021.1846714.

[61]   K. Bechtold-Peters *et al.*, "CMC Regulatory Considerations for Antibody-Drug Conjugates," *J. Pharm. Sci.*, vol. 112, no. 12, pp. 2965–2980, 2023, doi: 10.1016/j.xphs.2023.09.007.

[62]   J. R. Junutula *et al.*, "Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index," *Nat. Biotechnol.*, vol. 26, no. 8, pp. 925–932, 2008, doi: 10.1038/nbt.1480.

[63]   B.-Q. Shen *et al.*, "Conjugation site modulates the in vivo stability and therapeutic activity of antibody-drug conjugates," *Nat. Biotechnol.*, vol. 30, no. 2, pp. 184–189, 2012, doi: 10.1038/nbt.2108.

[64]   D. Y. Jackson, "Processes for Constructing Homogeneous Antibody Drug Conjugates," *Org. Process Res. Dev.*, vol. 20, no. 5, pp. 852–866, 2016, doi: 10.1021/acs.oprd.6b00067.

[65]   R. G. E. Coumans *et al.*, "A Platform for the Generation of Site-Specific Antibody-Drug Conjugates That Allows for Selective Reduction of Engineered Cysteines," *Bioconjug. Chem.*, vol. 31, no. 9, pp. 2136–2146, 2020, doi: 10.1021/acs.bioconjchem.0c00337.

[66]   A. Wakankar, Y. Chen, Y. Gokarn, and F. S. Jacobson, "Analytical methods for physicochemical characterization of antibody drug conjugates," *MAbs*, vol. 3, no. 2, pp. 161–172, 2011, doi: 10.4161/mabs.3.2.14960.

[67]   A. Wagh, H. Song, M. Zeng, L. Tao, and T. K. Das, "Challenges and new frontiers in analytical characterization of antibody-drug conjugates," *MAbs*, vol. 10, no. 2, pp. 222–243, 2018, doi: 10.1080/19420862.2017.1412025.

[68]   B. Bobály, S. Fleury-Souverain, A. Beck, J. L. Veuthey, D. Guillarme, and S. Fekete, "Current possibilities of liquid chromatography for the characterization of antibody-drug conjugates," *J. Pharm. Biomed. Anal.*, vol. 147, pp. 493–505, 2018, doi: 10.1016/j.jpba.2017.06.022.

[69]   L. N. Le, J. M. R. Moore, J. Ouyang, X. Chen, M. D. H. Nguyen, and W. J. Galush, "Profiling antibody drug conjugate positional isomers: A

system-of-equations approach," *Anal. Chem.*, vol. 84, no. 17, pp. 7479–7486, 2012, doi: 10.1021/ac301568f.

[70]   H. Chen *et al.*, "In-Depth Structure and Function Characterization of Heterogeneous Interchain Cysteine-Conjugated Antibody − Drug Conjugates," 2023, doi: 10.1021/acsptsci.3c00235.

[71]   I. C. H. ICH, "Q8 (R2) Pharmaceutical Development," *ICH Q9 (quality risk Manag. ICH Q*, vol. 10, 2008.

[72]   C. G. Hill and T. W. Root, *Introduction to Chemical Engineering Kinetics and Reactor Design.* Johne Wiley & Sons, 2014.

[73]   G. Maria, "A review of algorithms and trends in kinetic model identification for chemical and biochemical systems," *Chem. Biochem. Eng. Q.*, vol. 18, no. 3, pp. 195–222, 2004.

[74]   J. Wakefield, "Linear Models BT - Bayesian and Frequentist Regression Methods," J. Wakefield, Ed. New York, NY: Springer New York, 2013, pp. 195–252.

[75]   G. Sin and K. Gernaey, "Data Handling and Parameter Estimation," in *Experimental Methods in Wastewater Treatment*, M. C.M. van Loosdrecht, P. H. Nielsen, C. M. Lopez-Vazquez, and D. Brdjanovic, Eds. United Kingdom: IWA Publishing, 2016, pp. 201–234.

[76]   G. A. F. Seber and C. J. Wild, *Nonlinear Regression.* John Wiley & Sons, 1989.

[77]   L. M. Helleckes, M. Osthege, W. Wiechert, E. Von Lieres, and M. Oldiges, *Bayesian calibration, process modeling and uncertainty quantification in biotechnology*, vol. 18, no. 3. 2022.

[78]   P. Loskot, K. Atitey, and L. Mihaylova, "Comprehensive review of models and methods for inferences in bio-chemical reaction networks," *Front. Genet.*, vol. 10, no. JUN, 2019, doi: 10.3389/fgene.2019.00549.

[79]   K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference: a practical information-theoretic approach.* Springer, 2002.

[80]   J. Wutz *et al.*, "Predictability of kLa in stirred tank reactors under multiple operating conditions using an Euler–Lagrange approach," *Eng. Life Sci.*, vol. 16, no. 7, pp. 633–642, Oct. 2016, doi: https://doi.org/10.1002/elsc.201500135.

[81]   C. Haringa *et al.*, "Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: An analysis of organism lifelines," *Eng. Life Sci.*, vol. 16, no. 7, pp. 652–663, 2016, doi: 10.1002/elsc.201600061.

[82]   R. Spann *et al.*, "CFD predicted pH gradients in lactic acid bacteria cultivations," *Biotechnol. Bioeng.*, vol. 116, no. 4, pp. 769–780, 2019, doi: 10.1002/bit.26868.

[83]   J. Wutz, B. Waterkotte, K. Heitmann, and T. Wucherpfennig, "Computational fluid dynamics (CFD) as a tool for industrial UF/DF tank optimization," *Biochem. Eng. J.*, vol. 160, no. April, p. 107617, 2020, doi: 10.1016/j.bej.2020.107617.

[84]   Z. Xing *et al.*, "A CFD model for predicting protein aggregation in low-pH virial inactivation for mAb production," *Biotechnol. Bioeng.*, vol. 117, no. 11, pp. 3400–3412, 2020, doi: 10.1002/bit.27505.

[85]   K. G. Desai *et al.*, "Mixing of a mAb Formulation in a New Magnetically Coupled Single-Use Mixing System: Key Learnings of Preliminary Experimental and Computational Evaluation," *J. Pharm. Sci.*, vol. 108, no. 12, pp. 3932–3937, 2019, doi: 10.1016/j.xphs.2019.09.001.

[86]   J. H. Ferziger, M. Perić, and R. L. Street, *Computational Methods for Fluid Dynamics, Fourth Edition.* Springer International Publishing, 2019.

[87]   S. Schmalzriedt, M. Jenne, K. Mauch, and M. Reuss, "Integration of physiology and fluid dynamics.," in *Advances in biochemical engineering/biotechnology*, vol. 80, U. von Stockar, L. A. M. van der Wielen, A. Bruggink, J. M. S. Cabral, S.-O. Enfors, P. Fernandes, M. Jenne, K. Mauch, D. M. F. Prazeres, M. Reuss, S. Schmalzriedt, D. Stark, U. von Stockar, A. J. J. Straathof, and L. A. M. van der Wielen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 19–68.

[88]   J. B. Joshi *et al.*, "CFD simulation of stirred tanks: Comparison of turbulence models. Part I: Radial flow impellers," *Can. J. Chem. Eng.*, vol. 89, no. 1, pp. 23–82, 2011, doi: 10.1002/cjce.20446.

[89]   K. A. Bakeev, *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries: Second Edition.* John Wiley & Sons, Ltd, 2010.

[90]   M. Schiestl, T. Stangler, C. Torella, T. Čepeljnik, H. Toll, and R. Grau, "Acceptable changes in quality attributes of glycosylated biopharmaceuticals," *Nat. Biotechnol.*, vol. 29, no. 4, pp. 310–312, 2011, doi: 10.1038/nbt.1839.

[91]   FDA, "Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacuring, and Quality Assurance," 2004. [Online]. Available: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatory Information/Guidances/ucm070305.pdf.

[92]   A. S. Rathore, R. Bhambure, and V. Ghare, "Process analytical technology (PAT) for biopharmaceutical products," *Anal. Bioanal. Chem.*, vol. 398, no. 1, pp. 137–154, 2010, doi: 10.1007/s00216-010-3781-x.

[93]   M. Rüdt, T. Briskot, and J. Hubbuch, "Advances in downstream

processing of biologics – Spectroscopy: An emerging process analytical technology," *J. Chromatogr. A*, vol. 1490, pp. 2–9, 2017, doi: 10.1016/j.chroma.2016.11.010.

[94]   L. Rolinger *et al.*, "Multi-attribute PAT for UF/DF of Proteins— Monitoring Concentration, particle sizes, and Buffer Exchange," *Anal. Bioanal. Chem.*, vol. 412, no. 9, pp. 2123–2136, Apr. 2020, doi: 10.1007/s00216-019-02318-8.

[95]   M. A. Hough, "Choosing the optimal spectroscopic toolkit to understand protein function," *Biosci. Rep.*, vol. 37, no. 3, p. BSR20160378, Jun. 2017, doi: 10.1042/BSR20160378.

[96]   M. Rüdt, T. Briskot, and J. Hubbuch, "Advances in downstream processing of biologics – Spectroscopy: An emerging process analytical technology," *J. Chromatogr. A*, vol. 1490, pp. 2–9, 2017, doi: 10.1016/j.chroma.2016.11.010.

[97]   Y. Tang *et al.*, "Real-Time Analysis on Drug-Antibody Ratio of Antibody-Drug Conjugates for Synthesis, Process Optimization, and Quality Control," *Sci. Rep.*, vol. 7, no. 1, pp. 2–11, 2017, doi: 10.1038/s41598-017-08151-2.

[98]   S. Andris, M. Rüdt, J. Rogalla, M. Wendeler, and J. Hubbuch, "Monitoring of antibody-drug conjugation reactions with UV/Vis spectroscopy," *J. Biotechnol.*, vol. 288, no. October, pp. 15–22, 2018, doi: 10.1016/j.jbiotec.2018.10.003.

[99]   A. L. de Faria e Silva and A. G. Ryder, "Analyzing protein conjugation reactions for antibody-drug conjugate synthesis using polarized excitation emission matrix spectroscopy," *Biotechnol. Bioeng.*, vol. 119, no. 12, pp. 3432–3446, 2022, doi: https://doi.org/10.1002/bit.28229.

[100]   C. Zhang, J. S. Springall, X. Wang, and I. Barman, "Rapid, quantitative determination of aggregation and particle formation for antibody drug conjugate therapeutics with label-free Raman spectroscopy," *Anal. Chim. Acta*, vol. 1081, pp. 138–145, 2019, doi: 10.1016/j.aca.2019.07.007.

[101]   L. Rolinger, M. Rüdt, and J. Hubbuch, "A critical review of recent trends, and a future perspective of optical spectroscopy as PAT in biopharmaceutical downstream processing," *Anal. Bioanal. Chem.*, vol. 412, no. 9, pp. 2047–2064, 2020, doi: 10.1007/s00216-020-02407-z.

[102]   L. M. Rolinger, "Smart PAT concepts for the downstream process of biologics — Application of multimodal soft sensors, machine learning, and data fusion," Karlsruher Institut für Technologie (KIT), 2022.

[103]   S. M. Mercier, B. Diepenbroek, M. C. F. Dalm, R. H. Wijffels, and M. Streefland, "Multivariate data analysis as a PAT tool for early bioprocess development data," *J. Biotechnol.*, vol. 167, no. 3, pp. 262–270, 2013, doi: 10.1016/j.jbiotec.2013.07.006.

[104] R. Gautam, S. Vanga, F. Ariese, and S. Umapathy, "Review of multidimensional data processing approaches for Raman and infrared spectroscopy," *EPJ Tech. Instrum.*, vol. 2, no. 1, p. 8, 2015, doi: 10.1140/epjti/s40485-015-0018-6.

[105] C. Cui and T. Fearn, "Comparison of partial least squares regression, least squares support vector machines, and Gaussian process regression for a near infrared calibration," *J. Near Infrared Spectrosc.*, vol. 25, no. 1, pp. 5–14, 2017, doi: 10.1177/0967033516678515.

[106] C. E. Rasmussen and C. K. I. Williams, "Gaussian Processes for Machine Learning." The MIT Press, Nov. 23, 2005, doi: 10.7551/mitpress/3206.001.0001.

[107] A. Dürauer and T. Scharl, "Sensors and chemometrics in downstream processing," no. March, pp. 1–18, 2023, doi: 10.1002/bit.28499.

[108] C. Cui and T. Fearn, "Comparison of partial least squares regression, least squares support vector machines, and Gaussian process regression for a near infrared calibration," *J. Near Infrared Spectrosc.*, vol. 25, no. 1, pp. 5–14, Feb. 2017, doi: 10.1177/0967033516678515.

[109] V. Lyubenova, G. Kostov, and R. Denkova-Kostova, "Model-Based Monitoring of Biotechnological Processes—A Review," *Processes*, vol. 9, no. 6. 2021, doi: 10.3390/pr9060908.

[110] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven Soft Sensors in the process industry," *Comput. Chem. Eng.*, vol. 33, no. 4, pp. 795–814, 2009, doi: 10.1016/j.compchemeng.2008.12.012.

[111] R. Alexander, G. Campani, S. Dinh, and F. V. Lima, "Challenges and opportunities on nonlinear state estimation of chemical and biochemical processes," *Processes*, vol. 8, no. 11, pp. 1–27, 2020, doi: 10.3390/pr8111462.

[112] D. Haßkerl, S. Subramanian, S. Markert, S. Kaiser, and S. Engell, "Multi-rate state estimation applied to a pilot-scale reactive distillation process," *Chem. Eng. Sci.*, vol. 185, pp. 256–281, 2018, doi: https://doi.org/10.1016/j.ces.2018.04.018.

[113] H. Narayanan *et al.*, "Hybrid-EKF: Hybrid model coupled with extended Kalman filter for real-time monitoring and control of mammalian cell culture," *Biotechnol. Bioeng.*, vol. 117, no. 9, pp. 2703–2714, 2020, doi: 10.1002/bit.27437.

[114] A. Tuveri, F. Pérez-García, P. A. Lira-Parada, L. Imsland, and N. Bar, "Sensor fusion based on Extended and Unscented Kalman Filter for bioprocess monitoring," *J. Process Control*, vol. 106, pp. 195–207, 2021, doi: 10.1016/j.jprocont.2021.09.005.

[115] J. Valappil and C. Georgakis, "Systematic estimation of state noise

statistics for extended Kalman filters," *AIChE J.*, vol. 46, no. 2, pp. 292–308, 2000, doi: 10.1002/aic.690460209.

[116] C. H. Chau, P. S. Steeg, and W. D. Figg, "Antibody–drug conjugates for cancer," *Lancet*, vol. 394, no. 10200, pp. 793–804, 2019, doi: 10.1016/S0140-6736(19)31774-X.

[117] S. Ponziani *et al.*, "Antibody-drug conjugates: The new frontier of chemotherapy," *Int. J. Mol. Sci.*, vol. 21, no. 15, pp. 1–28, 2020, doi: 10.3390/ijms21155510.

[118] A. Q. Dean, S. Luo, J. D. Twomey, and B. Zhang, "Targeting cancer with antibody-drug conjugates: Promises and challenges," *MAbs*, vol. 13, no. 1, 2021, doi: 10.1080/19420862.2021.1951427.

[119] N. Joubert, A. Beck, C. Dumontet, and C. Denevault-sabourin, "Antibody – Drug Conjugates : The Last Decade," no. April, pp. 1–31, 2020.

[120] A. S. Rathore and H. Winkle, "Quality by design for biopharmaceuticals," *Nat. Biotechnol.*, vol. 27, no. 1, pp. 26–34, 2009, doi: 10.1038/nbt0109-26.

[121] D. Roush *et al.*, "Toward in silico CMC: An industrial collaborative approach to model-based process development," *Biotechnol. Bioeng.*, 2020, doi: 10.1002/bit.27520.

[122] D. K. Babi *et al.*, "Opportunities and challenges for model utilization in the biopharmaceutical industry: current versus future state," *Curr. Opin. Chem. Eng.*, vol. 36, p. 100813, 2022, doi: 10.1016/j.coche.2022.100813.

[123] H. Narayanan *et al.*, "Bioprocessing in the Digital Age: The Role of Process Models," *Biotechnol. J.*, vol. 15, no. 1, pp. 1–10, 2020, doi: 10.1002/biot.201900172.

[124] C. J. Taylor *et al.*, "A Brief Introduction to Chemical Reaction Optimization," *Chem. Rev.*, 2022, doi: 10.1021/acs.chemrev.2c00798.

[125] N. I. Vollmer, J. L. S. P. Driessen, C. K. Yamakawa, K. V. Gernaey, S. I. Mussatto, and G. Sin, "Model development for the optimization of operational conditions of the pretreatment of wheat straw," *Chem. Eng. J.*, vol. 430, no. P4, p. 133106, 2022, doi: 10.1016/j.cej.2021.133106.

[126] S. Kyriakopoulos *et al.*, "Kinetic Modeling of Mammalian Cell Culture Bioprocessing: The Quest to Advance Biomanufacturing," *Biotechnol. J.*, vol. 13, no. 3, pp. 1–11, 2018, doi: 10.1002/biot.201700229.

[127] I. W. Ashworth, L. C. Chan, B. G. Cox, I. M. McFarlane, and A. R. Phillips, "Understanding the Alkylation of a Phenol by 1-(3-Chloropropyl)pyrrolidine: Evidence for the Intermediacy of an Azetidinium Ion," *J. Org. Chem.*, vol. 84, no. 8, pp. 4754–4762, 2019, doi: 10.1021/acs.joc.8b02458.

[128] D. Pfister, N. Ulmer, A. Klaue, O. Ingold, and M. Morbidelli, "Modeling the Kinetics of Protein Conjugation Reactions," *Chemie-Ingenieur-Technik*, vol. 88, no. 11, pp. 1598–1608, 2016, doi: 10.1002/cite.201600046.

[129] P. Tang *et al.*, "Optimization and kinetic modeling of interchain disulfide bond reoxidation of monoclonal antibodies in bioprocesses," *MAbs*, vol. 12, no. 1, 2020, doi: 10.1080/19420862.2020.1829336.

[130] S. Nayak and S. M. Richter, "Kinetic Studies of the Partial Reduction and Conjugation Reactions in an Antibody-Drug Conjugate (ADC) Synthesis," 2023, doi: 10.1021/acs.oprd.3c00264.

[131] G. Sin, K. V. Gernaey, and A. E. Lantz, "Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis," *Biotechnol. Prog.*, vol. 25, no. 4, pp. 1043–1053, 2009, doi: 10.1002/btpr.166.

[132] E. Anane *et al.*, "Output uncertainty of dynamic growth models: Effect of uncertain parameter estimates on model reliability," *Biochem. Eng. J.*, vol. 150, no. February, p. 107247, 2019, doi: 10.1016/j.bej.2019.107247.

[133] A. D. Rodman and D. I. Gerogiorgis, "Parameter estimation and sensitivity analysis for dynamic modelling and simulation of beer fermentation," *Comput. Chem. Eng.*, vol. 136, p. 106665, 2020, doi: 10.1016/j.compchemeng.2019.106665.

[134] M. Abdollahpour-Alitappeh *et al.*, "Antibody–drug conjugates (ADCs) for cancer therapy: Strategies, challenges, and successes," *J. Cell. Physiol.*, vol. 234, no. 5, pp. 5628–5642, 2019, doi: https://doi.org/10.1002/jcp.27419.

[135] L. Fernandez-Cerezo *et al.*, "Strategies for UF/DF-Based Impurity Removal in the Post-conjugation Purification of Antibody-Drug Conjugates," *Org. Process Res. Dev.*, vol. 27, no. 7, pp. 1258–1268, 2023, doi: 10.1021/acs.oprd.3c00137.

[136] Y. Matsuda, Z. Tawfiq, M. Leung, and B. A. Mendelsohn, "Insight into Temperature Dependency and Design of Experiments towards Process Development for Cysteine-Based Antibody-Drug Conjugates," *ChemistrySelect*, vol. 5, no. 28, pp. 8435–8439, 2020, doi: 10.1002/slct.202001822.

[137] B. Stump and J. Steinmann, "Conjugation Process Development and Scale-Up," *Methods Mol. Biol.*, vol. 1045, pp. 235–248, 2013, doi: 10.1007/978-1-62703-541-5_14.

[138] S. Andris, M. Wendeler, X. Wang, and J. Hubbuch, "Multi-step high-throughput conjugation platform for the development of antibody-drug conjugates," *J. Biotechnol.*, vol. 278, no. May, pp. 48–55, 2018, doi: 10.1016/j.jbiotec.2018.05.004.

[139] Y. Nakahara, B. A. Mendelsohn, and Y. Matsuda, "Antibody-Drug Conjugate Synthesis Using Continuous Flow Microreactor Technology," *Org. Process Res. Dev.*, vol. 26, no. 9, pp. 2766–2770, 2022, doi: 10.1021/acs.oprd.2c00217.

[140] S. Andris, J. Seidel, and J. Hubbuch, "Kinetic reaction modeling for antibody-drug conjugate process development," *J. Biotechnol.*, vol. 306, no. September, pp. 71–80, 2019, doi: 10.1016/j.jbiotec.2019.09.013.

[141] J. T. Weggen, J. Seidel, R. Bean, M. Wendeler, and J. Hubbuch, "Kinetic studies and CFD-based reaction modeling for insights into the scalability of ADC conjugation reactions," *Front. Bioeng. Biotechnol.*, vol. 11, no. April, pp. 1–12, 2023, doi: 10.3389/fbioe.2023.1123842.

[142] R. Schiemer, J. T. Weggen, K. M. Schmitt, and J. Hubbuch, "An adaptive soft-sensor for advanced real-time monitoring of an antibody-drug conjugation reaction," *Biotechnol. Bioeng.*, vol. 120, no. 7, pp. 1914–1928, 2023, doi: 10.1002/bit.28428.

[143] M. Cao *et al.*, "Site-specific antibody-drug conjugate heterogeneity characterization and heterogeneity root cause analysis," *MAbs*, vol. 11, no. 6, pp. 1064–1076, 2019, doi: 10.1080/19420862.2019.1624127.

[144] J. Paulech, N. Solis, and S. J. Cordwell, "Characterization of reaction conditions providing rapid and specific cysteine alkylation for peptide-based mass spectrometry," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1834, no. 1, pp. 372–379, 2013, doi: 10.1016/j.bbapap.2012.08.002.

[145] D. Pfister, E. Bourgeaux, and M. Morbidelli, "Kinetic modeling of protein PEGylation," *Chem. Eng. Sci.*, vol. 137, pp. 816–827, 2015, doi: 10.1016/j.ces.2015.07.031.

[146] Y. H. Fung *et al.*, "Different sorption behaviours of pyrene onto polyethylene microplastics in a binary system with water and a ternary system with water and sediment," *Environ. Technol. Innov.*, vol. 30, p. 103086, 2023, doi: https://doi.org/10.1016/j.eti.2023.103086.

[147] F. C. C. Montes, K. Gernaey, and G. Sin, "Dynamic Plantwide Modeling, Uncertainty, and Sensitivity Analysis of a Pharmaceutical Upstream Synthesis: Ibuprofen Case Study," *Ind. Eng. Chem. Res.*, vol. 57, no. 30, pp. 10026–10037, 2018, doi: 10.1021/acs.iecr.8b00465.

[148] N. J. Smith *et al.*, "Fast, irreversible modification of cysteines through strain releasing conjugate additions of cyclopropenyl ketones," *Org. Biomol. Chem.*, vol. 16, no. 12, pp. 2164–2169, 2018, doi: 10.1039/c8ob00166a.

[149] L. Mao, A. J. Russell, and S. Carmali, "Moving Protein PEGylation from an Art to a Data Science," *Bioconjug. Chem.*, vol. 33, no. 9, pp. 1643–1653, 2022, doi: 10.1021/acs.bioconjchem.2c00262.

[150] M.-C. Janin-Bussat, M. Dillenbourg, N. Corvaia, A. Beck, and C. Klinguer-Hamour, "Characterization of antibody drug conjugate positional isomers at cysteine residues by peptide mapping LC–MS analysis," *J. Chromatogr. B*, vol. 981–982, pp. 9–13, 2015, doi: https://doi.org/10.1016/j.jchromb.2014.12.017.

[151] M. Källsten, R. Hartmann, K. Artemenko, S. B. Lind, F. Lehmann, and J. Bergquist, "Qualitative analysis of antibody-drug conjugates (ADCs): An experimental comparison of analytical techniques of cysteine-linked ADCs," *Analyst*, vol. 143, no. 22, pp. 5487–5496, 2018, doi: 10.1039/c8an01178h.

[152] J. Jones, L. Pack, J. H. Hunter, and J. F. Valliere-Douglass, "Native size-exclusion chromatography-mass spectrometry: suitability for antibody–drug conjugate drug-to-antibody ratio quantitation across a range of chemotypes and drug-loading levels," *MAbs*, vol. 12, no. 1, 2020, doi: 10.1080/19420862.2019.1682895.

[153] K. Li, Z. Zhang, Z. J. Lin, H. Shi, and Y. Ma, "Accurate determination of drug-to-antibody ratio of interchain cysteine-linked antibody–drug conjugates by LC-HRMS," *Anal. Bioanal. Chem.*, vol. 412, no. 4, pp. 833–840, 2020, doi: 10.1007/s00216-019-02280-5.

[154] A. S. Prashad, B. Nolting, V. Patel, A. Xu, B. Arve, and L. Letendre, "From R&D to Clinical Supplies," *Org. Process Res. Dev.*, vol. 21, no. 4, pp. 590–600, 2017, doi: 10.1021/acs.oprd.7b00020.

[155] J. Guo *et al.*, "Characterization and Higher-Order Structure Assessment of an Interchain Cysteine-Based ADC: Impact of Drug Loading and Distribution on the Mechanism of Aggregation," *Bioconjug. Chem.*, vol. 27, no. 3, pp. 604–615, 2016, doi: 10.1021/acs.bioconjchem.5b00603.

[156] Y. T. Adem, K. A. Schwarz, E. Duenas, T. W. Patapoff, W. J. Galush, and O. Esue, "Auristatin antibody drug conjugate physical instability and the role of drug payload," *Bioconjug. Chem.*, vol. 25, no. 4, pp. 656–664, 2014, doi: 10.1021/bc400439x.

[157] M. M. C. Sun *et al.*, "Reduction-alkylation strategies for the modification of specific monoclonal antibody bisulfides," *Bioconjug. Chem.*, vol. 16, no. 5, pp. 1282–1290, 2005, doi: 10.1021/bc050201y.

[158] D. V Nadkarni *et al.*, "Process Development and Structural Characterization of an Anti-Notch 3 Antibody–Drug Conjugate," *Org. Process Res. Dev.*, vol. 22, no. 3, pp. 286–295, Mar. 2018, doi: 10.1021/acs.oprd.7b00337.

[159] Y. Song, H. Cai, Z. Tan, N. Mussa, and Z. J. Li, "Mechanistic insights into inter-chain disulfide bond reduction of IgG1 and IgG4 antibodies," *Appl. Microbiol. Biotechnol.*, vol. 106, no. 3, pp. 1057–1066, 2022, doi: 10.1007/s00253-022-11778-5.

[160] J. Smiatek, A. Jung, and E. Bluhmki, "Towards a Digital Bioprocess Replica: Computational Approaches in Biopharmaceutical Development and Manufacturing," *Trends Biotechnol.*, vol. 38, no. 10, pp. 1141–1153, 2020, doi: 10.1016/j.tibtech.2020.05.008.

[161] I. C. Chou and E. O. Voit, "Recent developments in parameter estimation and structure identification of biochemical and genomic systems," *Math. Biosci.*, vol. 219, no. 2, pp. 57–83, 2009, doi: 10.1016/j.mbs.2009.03.002.

[162] T. Forster, D. Vázquez, M. N. Cruz-Bournazou, A. Butté, and G. Guillén-Gosálbez, "Modeling of bioprocesses via MINLP-based symbolic regression of S-system formalisms," *Comput. Chem. Eng.*, vol. 170, no. December 2022, 2023, doi: 10.1016/j.compchemeng.2022.108108.

[163] J. Fromer, C. Georgakis, and J. Mustakis, "Toward the Identification of Stoichiometric Models for Complex Reaction Mixtures," *Ind. Eng. Chem. Res.*, vol. 62, no. 5, pp. 2225–2239, 2023, doi: 10.1021/acs.iecr.2c01231.

[164] C. J. Taylor *et al.*, "An automated computational approach to kinetic model discrimination and parameter estimation," *React. Chem. Eng.*, vol. 6, no. 8, pp. 1404–1411, 2021, doi: 10.1039/d1re00098e.

[165] W. Ji and S. Deng, "Autonomous Discovery of Unknown Reaction Pathways from Data by Chemical Reaction Neural Network," *J. Phys. Chem. A*, vol. 125, no. 4, pp. 1082–1092, 2021, doi: 10.1021/acs.jpca.0c09316.

[166] M. Schwaab and J. C. Pinto, "Optimum reference temperature for reparameterization of the Arrhenius equation. Part 1: Problems involving one kinetic constant," *Chem. Eng. Sci.*, vol. 62, no. 10, pp. 2750–2764, 2007, doi: 10.1016/j.ces.2007.02.020.

[167] K. M. Hutterer *et al.*, "Monoclonal antibody disulfide reduction during manufacturing Untangling process effects from product effects," *MAbs*, vol. 5, no. 4, pp. 608–613, 2013, doi: 10.4161/mabs.24725.

[168] T. Wang, Y. D. Liu, B. Cai, G. Huang, and G. C. Flynn, "Investigation of antibody disulfide reduction and re-oxidation and impact to biological activities," *J. Pharm. Biomed. Anal.*, vol. 102, pp. 519–528, 2015, doi: 10.1016/j.jpba.2014.10.023.

[169] H. Liu, C. Chumsae, G. Gaza-Bulseco, K. Hurkmans, and C. H. Radziejewski, "Ranking the susceptibility of disulfide bonds in human IgG1 antibodies by reduction, differential alkylation, and LC-MS analysis," *Anal. Chem.*, vol. 82, no. 12, pp. 5219–5226, 2010, doi: 10.1021/ac100575n.

[170] C. E. Sänger–van de Griend, "CE-SDS method development, validation, and best practice—An overview," *Electrophoresis*, vol. 40, no. 18–19, pp. 2361–2374, 2019, doi: 10.1002/elps.201900094.

[171] T. Chen *et al.*, "Chemical de-conjugation for investigating the stability of small molecule drugs in antibody-drug conjugates," *J. Pharm. Biomed. Anal.*, vol. 117, pp. 304–310, 2016, doi: 10.1016/j.jpba.2015.09.015.

[172] T. Kantner, B. Alkhawaja, and A. G. Watts, "In Situ Quenching of Trialkylphosphine Reducing Agents Using Water-Soluble PEG-Azides Improves Maleimide Conjugation to Proteins," *ACS Omega*, vol. 2, no. 9, pp. 5785–5791, 2017, doi: 10.1021/acsomega.7b01094.

[173] K. D. Altria and S. D. Filbey, "Quantitative pharmaceutical analysis by capillary electrophoresis," *J. Liq. Chromatogr.*, vol. 16, no. 11, pp. 2281–2292, 1993, doi: 10.1080/10826079308020985.

[174] A. E. Mahan *et al.*, "A method for high-throughput, sensitive analysis of IgG Fc and Fab glycosylation by capillary electrophoresis," *J. Immunol. Methods*, vol. 417, pp. 34–44, 2015, doi: 10.1016/j.jim.2014.12.004.

[175] J. T. Weggen, R. Bean, K. Hui, M. Wendeler, and J. Hubbuch, "Kinetic models towards an enhanced understanding of diverse ADC conjugation reactions," *Front. Bioeng. Biotechnol.*, vol. 12, p. 1403644, 2024, doi: 10.3389/fbioe.2024.1403644.

[176] M. P. C. Marques, J. M. S. Cabral, and P. Fernandes, "Bioprocess scale-up: Quest for the parameters to be used as criterion to move from microreactors to lab-scale," *J. Chem. Technol. Biotechnol.*, vol. 85, no. 9, pp. 1184–1198, 2010, doi: 10.1002/jctb.2387.

[177] T. Tajsoleiman, L. Mears, U. Krühne, K. V. Gernaey, and S. Cornelissen, "An Industrial Perspective on Scale-Down Challenges Using Miniaturized Bioreactors," *Trends Biotechnol.*, vol. 37, no. 7, pp. 697–706, 2019, doi: 10.1016/j.tibtech.2019.01.002.

[178] I. Montes-Serrano, P. Satzer, A. Jungbauer, and A. Dürauer, "Characterization of hydrodynamics and volumetric power input in microtiter plates for the scale-up of downstream operations," *Biotechnol. Bioeng.*, no. July, pp. 1–12, 2021, doi: 10.1002/bit.27983.

[179] C. L. Gargalo *et al.*, "Towards the Development of Digital Twins for the Bio-manufacturing Industry," 2020, doi: 10.1007/10_2020_142.

[180] P. Sinner, S. Daume, C. Herwig, and J. Kager, "Usage of Digital Twins Along a Typical Process Development Cycle," in *Digital Twins: Tools and Concepts for Smart Biomanufacturing*, C. Herwig, R. Pörtner, and J. Möller, Eds. Cham: Springer International Publishing, 2021, pp. 71–96.

[181] J. Scully *et al.*, "Beyond heuristics: CFD-based novel multiparameter scale-up for geometrically disparate bioreactors demonstrated at industrial 2kL–10kL scales," *Biotechnol. Bioeng.*, vol. 117, no. 6, pp. 1710–1723, Jun. 2020, doi: 10.1002/bit.27323.

[182] S. Werner, S. C. Kaiser, M. Kraume, and D. Eibl, "Computational fluid dynamics as a modern tool for engineering characterization of bioreactors," *Pharm. Bioprocess.*, vol. 2, no. 1, pp. 85–99, 2014, doi: 10.4155/pbp.13.60.

[183] M. C. Martinetz *et al.*, "Hybrid Approach for Mixing Time Characterization and Scale-Up in Geometrical Nonsimilar Stirred Vessels Equipped with Eccentric Multi-Impeller Systems—An Industrial Perspective," *Processes*, vol. 9, no. 5, p. 880, 2021, doi: 10.3390/pr9050880.

[184] J. Wutz, R. Steiner, K. Assfalg, and T. Wucherpfennig, "Establishment of a CFD-based kLa model in microtiter plates to support CHO cell culture scale-up during clone selection," *Biotechnol. Prog.*, vol. 34, no. 5, pp. 1120–1128, 2018, doi: 10.1002/btpr.2707.

[185] C. Bach *et al.*, "Evaluation of mixing and mass transfer in a stirred pilot scale bioreactor utilizing CFD," *Chem. Eng. Sci.*, vol. 171, pp. 19–26, 2017, doi: 10.1016/j.ces.2017.05.001.

[186] G. Nadal-Rey *et al.*, "Computational fluid dynamics modelling of hydrodynamics, mixing and oxygen transfer in industrial bioreactors with Newtonian broths," *Biochem. Eng. J.*, vol. 177, no. October 2021, p. 108265, 2022, doi: 10.1016/j.bej.2021.108265.

[187] C. Haringa, A. T. Deshmukh, R. F. Mudde, and H. J. Noorman, "Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic S. cerevisiae fermentation," *Chem. Eng. Sci.*, vol. 170, pp. 653–669, 2017, doi: 10.1016/j.ces.2017.01.014.

[188] L. Vicum, S. Ottiger, M. Mazzotti, Ł. Makowski, and J. Bałdyga, "Multi-scale modeling of a reactive mixing process in a semibatch stirred tank," *Chem. Eng. Sci.*, vol. 59, no. 8–9, pp. 1767–1781, 2004, doi: 10.1016/j.ces.2004.01.032.

[189] S. Yang, S. Kiang, P. Farzan, and M. Ierapetritou, "Optimization of reaction selectivity using CFD-based compartmental modeling and surrogate-based optimization," *Processes*, vol. 7, no. 1, 2019, doi: 10.3390/pr7010009.

[190] M. Öner *et al.*, *Scale-up Modeling of a Pharmaceutical Crystallization Process via Compartmentalization Approach*, vol. 44. Elsevier Masson SAS, 2018.

[191] U. Kaya, S. Gopireddy, N. Urbanetz, I. Nopens, and J. Verwaeren, "Predicting the Hydrodynamic Properties of a Bioreactor: Conditional Density Estimation as a Surrogate Model for CFD Simulations," *Chem. Eng. Res. Des.*, 2022.

[192] O. Marcq, "Outlook on Next Generation Technologies and Strategy Considerations for ADC Process Development and Manufacturing," in

Innovations for Next-Generation Antibody-Drug Conjugates, M. Damelin, Ed. Cham: Springer International Publishing, 2018, pp. 113–161.

[193] D. R. Schmidhalter, S. Elzner, and R. Schmid, "Progress in the Development of Single-Use Solutions in Antibody–Drug Conjugate (ADC) Manufacturing," in *Single-Use Technology in Biopharmaceutical Manufacture*, John Wiley & Sons, Ltd, 2019, pp. 303–310.

[194] X. Hu, E. Bortell, F. W. Kotch, A. Xu, B. Arve, and S. Freese, "Development of Commercial-Ready Processes for Antibody Drug Conjugates," *Org. Process Res. Dev.*, vol. 21, no. 4, pp. 601–610, 2017, doi: 10.1021/acs.oprd.7b00023.

[195] G. Nadal-Rey, D. D. McClure, J. M. Kavanagh, S. Cornelissen, D. F. Fletcher, and K. V. Gernaey, "Understanding gradients in industrial bioreactors," *Biotechnol. Adv.*, vol. 46, no. November 2020, p. 107660, 2021, doi: 10.1016/j.biotechadv.2020.107660.

[196] A. Thomas, B. A. Teicher, and R. Hassan, "Antibody–drug conjugates for cancer therapy," *Lancet Oncol.*, vol. 17, no. 6, pp. e254–e262, 2016, doi: 10.1016/S1470-2045(16)30030-4.

[197] ICH, "Q8(R2) Pharmaceutical Development," 2009. [Online]. Available: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific _guideline/2009/09/WC500002872.pdf.

[198] S. Andris and J. Hubbuch, "Modeling of hydrophobic interaction chromatography for the separation of antibody-drug conjugates and its application towards quality by design," *J. Biotechnol.*, vol. 317, no. May, pp. 48–58, 2020, doi: 10.1016/j.jbiotec.2020.04.018.

[199] W. R. Keller and M. Wendeler, "Using multimodal chromatography for post-conjugation antibody-drug conjugate purification: A methodology from high throughput screening to in-silico process development," *J. Chromatogr. A*, vol. 1653, p. 462378, 2021, doi: 10.1016/j.chroma.2021.462378.

[200] A. S. Rathore, R. Bhambure, and V. Ghare, "Process analytical technology (PAT) for biopharmaceutical products," *Anal. Bioanal. Chem.*, vol. 398, no. 1, pp. 137–154, 2010, doi: 10.1007/s00216-010-3781-x.

[201] S. Großhans *et al.*, "In-line Fourier-transform infrared spectroscopy as a versatile process analytical technology for preparative protein chromatography," *J. Chromatogr. A*, vol. 1547, pp. 37–44, 2018, doi: 10.1016/j.chroma.2018.03.005.

[202] B. Wei *et al.*, "Multi-attribute Raman spectroscopy (MARS) for monitoring product quality attributes in formulated monoclonal antibody therapeutics," *MAbs*, vol. 14, no. 1, 2022, doi:

10.1080/19420862.2021.2007564.

[203] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemom. Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, Oct. 2001, doi: 10.1016/S0169-7439(01)00155-1.

[204] M. Andersson, "A comparison of nine PLS1 algorithms," *J. Chemom. A J. Chemom. Soc.*, vol. 23, no. 10, pp. 518–529, 2009, doi: https://doi.org/10.1002/cem.1248.

[205] T. Chen, J. Morris, and E. Martin, "Gaussian process regression for multivariate spectroscopic calibration," *Chemom. Intell. Lab. Syst.*, vol. 87, no. 1, pp. 59–71, 2007, doi: 10.1016/j.chemolab.2006.09.004.

[206] A. Tulsyan, H. Khodabandehlou, T. Wang, G. Schorner, M. Coufal, and C. Undey, "Spectroscopic models for real-time monitoring of cell culture processes using spatiotemporal just-in-time Gaussian processes," *AIChE J.*, vol. 67, no. 5, 2021, doi: 10.1002/aic.17210.

[207] E. L. Haseltine and J. B. Rawlings, "Critical evaluation of extended Kalman filtering and moving-horizon estimation," *Ind. Eng. Chem. Res.*, vol. 44, no. 8, pp. 2451–2460, 2005, doi: 10.1021/ie034308l.

[208] R. Schneider and C. Georgakis, "How to NOT make the extended kalman filter fail," *Ind. Eng. Chem. Res.*, vol. 52, no. 9, pp. 3354–3362, 2013, doi: 10.1021/ie300415d.

[209] R. Simutis and A. Lübbert, "Hybrid approach to state estimation for bioprocess control," *Bioengineering*, vol. 4, no. 1, Mar. 2017, doi: 10.3390/bioengineering4010021.

[210] A. Golabgir and C. Herwig, "Combining Mechanistic Modeling and Raman Spectroscopy for Real-Time Monitoring of Fed-Batch Penicillin Production," *Chemie-Ingenieur-Technik*, vol. 88, no. 6, pp. 764–776, 2016, doi: 10.1002/cite.201500101.

[211] M. Elsheikh, R. Hille, A. Tatulea-Codrean, and S. Krämer, "A Comparative Review of Multi-Rate Moving Horizon Estimation Schemes for Bioprocess Applications," *Comput. Chem. Eng.*, vol. 146, p. 107219, 2021, doi: 10.1016/j.compchemeng.2020.107219.

[212] Feidl *et al.*, "Combining Mechanistic Modeling and Raman Spectroscopy for Monitoring Antibody Chromatographic Purification," *Processes*, vol. 7, no. 10, p. 683, Oct. 2019, doi: 10.3390/pr7100683.

[213] R. Luttmann *et al.*, "Soft sensors in bioprocessing: A status report and recommendations," *Biotechnol. J.*, vol. 7, no. 8, pp. 1040–1048, 2012, doi: https://doi.org/10.1002/biot.201100506.

[214] P. Cabaneros Lopez *et al.*, "Towards a digital twin: a hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation," *Biofuels, Bioprod. Biorefining*, vol. 14, no. 5, pp. 1046–

1060, 2020, doi: 10.1002/bbb.2108.

[215] K. Ohadi, R. L. Legge, and H. M. Budman, "Development of a Soft-Sensor Based on Multi-Wavelength Fluorescence Spectroscopy and a Dynamic Metabolic Model for Monitoring Mammalian Cell Cultures," *Biotechnol. Bioeng*, vol. 112, pp. 197–208, 2015, doi: 10.1002/bit.25339/abstract.

[216] R. Spann, C. Roca, D. Kold, A. Eliasson Lantz, K. V. Gernaey, and G. Sin, "A probabilistic model-based soft sensor to monitor lactic acid bacteria fermentations," *Biochem. Eng. J.*, vol. 135, pp. 49–60, 2018, doi: 10.1016/j.bej.2018.03.016.

[217] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, pp. 137–148, 1969, doi: https://doi.org/10.2307/1266770.

[218] T. Mehmood, K. H. Liland, L. Snipen, and S. Sæbø, "A review of variable selection methods in Partial Least Squares Regression," *Chemom. Intell. Lab. Syst.*, vol. 118, pp. 62–69, Aug. 2012, doi: 10.1016/j.chemolab.2012.07.010.

[219] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in {P}ython," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[220] T. Paananen, J. Piironen, M. R. Andersen, and A. Vehtari, "Variable selection for Gaussian processes via sensitivity analysis of the posterior predictive distribution," *Proc. 22nd Int. Conf. Artif. Intell. Stat.*, vol. PLMR, no. 89, pp. 1743–1752, Dec. 2019, [Online]. Available: http://arxiv.org/abs/1712.08048.

[221] C. R. Harris *et al.*, "Array programming with {NumPy}," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.

[222] T. pandas development team, "pandas-dev/pandas: Pandas." Zenodo, 2020, doi: 10.5281/zenodo.3509134.

[223] D. Simon, "Nonlinear Kalman filtering," in *Optimal State Estimation*, John Wiley & Sons, Ltd, 2006, pp. 393–431.

[224] T. Chen, J. Morris, and E. Martin, "Gaussian process regression for multivariate spectroscopic calibration," *Chemom. Intell. Lab. Syst.*, vol. 87, no. 1, pp. 59–71, 2007, doi: 10.1016/j.chemolab.2006.09.004.

[225] K. Kjeldahl and R. Bro, "Some common misunderstandings in chemometrics," *J. Chemom.*, vol. 24, no. 7–8, pp. 558–564, 2010, doi: 10.1002/cem.1346.

[226] H. Witjes, M. Van Den Brink, W. J. Melssen, and L. M. C. Buydens, "Automatic correction of peak shifts in Raman spectra before PLS regression," *Chemom. Intell. Lab. Syst.*, vol. 52, no. 1, pp. 105–116, 2000, doi: 10.1016/S0169-7439(00)00085-X.

[227] J. Ouyang, "Drug-to-antibody ratio (DAR) and drug load distribution by hydrophobic interaction chromatography and reversed phase high-performance liquid chromatography," in *Antibody-Drug Conjugates*, Springer, 2013, pp. 275–283.

[228] M. Cao *et al.*, "Site-specific antibody-drug conjugate heterogeneity characterization and heterogeneity root cause analysis," *MAbs*, vol. 11, no. 6, pp. 1064–1076, 2019, doi: 10.1080/19420862.2019.1624127.

[229] I. V. Stelzer, J. Kager, and C. Herwig, "Comparison of Particle Filter and Extended Kalman Filter Algorithms for Monitoring of Bioprocesses," in *Proceedings of the 27th European Symposium on Computer Aided Process Engineering – ESCAPE 27*, 2017, vol. 40, pp. 1483–1488, doi: 10.1016/B978-0-444-63965-3.50249-X.

[230] S. Kolås, B. A. Foss, and T. S. Schei, "Constrained nonlinear state estimation based on the UKF approach," *Comput. Chem. Eng.*, vol. 33, no. 8, pp. 1386–1401, 2009, doi: 10.1016/j.compchemeng.2009.01.012.

Abbreviations

ADC	antibody-drug conjugate
AIC	Akaike information criterion
API	active pharmaceutical ingredient
ARD	automatic relevance determination
CE-SDS	capillary electrophoresis sodium dodecyl sulfate
CFD	computational fluid dynamics
CGE	capillary gel electrophoresis
CMC	chemistry, manufacturing, and control
CPP	critical process parameter
CQA	critical quality attribute
DAD	diode array detector
DAR	drug-to-antibody ratio
DHAA	dehydroascorbic acid
DLD	drug load distribution
DMSO	Dimethyl sulfoxide
DTT	dithiothreitol
DoE	Design of Experiment
DSP	downstream processing
EDTA	Ethylenediaminetetraacetic acid
EKF	Extended Kalman Filter
FDA	U.S. Food and Drug Administration
GPR	Gaussian process regression
GMP	good manufacturing practice
GST	glass-stirred tank
H	heavy chain
HH	heavy-heavy chain
HHL	heavy-heavy-light chain
HIC	hydrophobic interaction chromatography
HL	heavy-light chain
(U)HPLC	(ultra) high-performance liquid chromatography
HTS	high-throughput screening
IgG	Immunoglobulin G
L	light chain
LS	least squares
mAb	monoclonal antibody
NAC	N-acetyl cysteine
NPM	N-(1-pyrenyl)-maleimide
MC	Monte Carlo

MOA	mechanism of action
ML	machine learning
MLE	maximum likelihood estimation
MLR	multiple linear regression
MRF	multiple reference frames
MS	mass spectrometry
NHS	N-hydroxysuccinimide
nRMSE	normalized root mean squared error
ODE	ordinary differential equation
PAT	process analytical technology
PCA	principial component analysis
PEG	polyethylene glycol
PLS(R)	partial least squares regression
QbD	Quality by Design
QTPP	quality target product profile
RANS	Reynolds-averaged Navier Stokes
RBF	radial-basis function
RMSE	root mean squared error
RMSECV	root mean squared error of cross validation
RMSEP	root mean squared error of prediction
RSM	Reynolds stress model
RP	reversed phase
SGF	Savitzky-Golay derivative filter
SH	sulfhydryl
SUM	single-use mixer
TCEP	tris(2-carboxyethyl) phosphine
TFA	trifluoroacetic acid
UV/Vis	ultraviolet/visible
VS	variable selection

Appendix A Supplementary Material for Chapter 3

S3.1. Detailed Overview of all experimental conditions

Table S 1: Experimental starting conditions for the conjugation kinetics using two different conjugation types with three payloads and in different feeding modes.

Type	ADC	Payload	c_{mAb} / (mg mL^{-1})	Molar drug excess	Feeding time / s	Replicates	Subset
DAR 2	ADC1	Drug1	1.5	1x	-	1	Train
DAR 2	ADC1	Drug1	1.5	2x	-	1	Train
DAR 2	ADC1	Drug1	1.5	3x	-	2	Train
DAR 2	ADC1	Drug1	1.5	5x	-	2	Test
DAR 2	ADC1	Drug1	3	5x	-	2	Train
DAR 2	ADC1	Drug1	3	8x	-	2	Train
DAR 2	ADC1	Drug1	5	5x	-	2	Train
DAR 2	ADC1	Drug1	10	5x	-	2	Test
DAR 2	ADC1	NPM	1.5	3x	-	2	External
DAR 2	ADC1	NPM	1.5	5x	-	2	External
DAR 2	ADC1	NPM	3	3x	-	2	External
DAR 2	ADC1	NPM	3	5x	-	2	External
DAR 8	ADC2	NPM	1.5	11x	20	1	Train
DAR 8	ADC2	NPM	1.5	11x	30	1	Train
DAR 8	ADC2	NPM	3	11x	20	1	Train
DAR 8	ADC2	NPM	1.5	8x	20	2	Train
DAR 8	ADC2	NPM	1.5	13x	20	1	Test
DAR 8	ADC2	NPM	1.5	6x	20	2	Train
DAR 8	ADC2	NPM	2.25	11x	20	1	Test
DAR 8	ADC2	NPM	1.5	11x	10	1	Test
DAR 8	ADC2	NPM	1.5	11x	-	2	Train
DAR 8	ADC2	NPM	1.5	6x	-	1	Train
DAR 8	ADC3	Drug2	1.5	11x	-	4	External
DAR 8	ADC3	Drug2	1.5	14x	-	2	External
DAR 8	ADC3	Drug2	20	11x	-	2	External
DAR 8	ADC3	Drug2	20	14x	-	2	External

S3.2. Reaction schemes

For DAR 2 and detailed reaction scheme:

Heavy chains with two cysteines: $\quad H0_{2c} \underset{+Drug}{\overset{k_1}{\rightarrow}} H1_{2c} \underset{+Drug}{\overset{k_2}{\rightarrow}} H2_{2c}$

Heavy chain with one cysteine: $\quad H0_{1c} \underset{+Drug}{\overset{k_1}{\rightarrow}} H1_{1c}$

Drug inactivation: $\quad Drug \overset{k_{drug}}{\longrightarrow} Drug_{intact}$
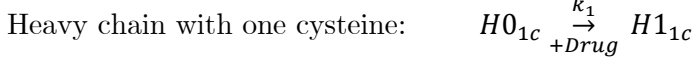
For DAR 8 and detailed reaction scheme:

Light chain with one cysteine: $\quad L0_{1c} \underset{+Drug}{\overset{k_1}{\rightarrow}} L1_{1c}$

Heavy chain with four cysteines: $\quad H0_{4c} \underset{+Drug}{\overset{k_2}{\rightarrow}} H1_{4c} \underset{+Drug}{\overset{k_3}{\rightarrow}} H2_{4c} \underset{+Drug}{\overset{k_4}{\rightarrow}} H3_{4c} \underset{+Drug}{\overset{k_5}{\rightarrow}} H4_{4c}$

Heavy chain with three cysteines: $\quad H0_{3c} \underset{+Drug}{\overset{k_2}{\rightarrow}} H1_{3c} \underset{+Drug}{\overset{k_3}{\rightarrow}} H2_{3c} \underset{+Drug}{\overset{k_4}{\rightarrow}} H3_{3c}$

Heavy chain with two cysteines: $\quad H0_{2c} \underset{+Drug}{\overset{k_2}{\rightarrow}} H1_{2c} \underset{+Drug}{\overset{k_3}{\rightarrow}} H2_{2c}$

Heavy chain with one cysteine: $\quad H0_{1c} \underset{+Drug}{\overset{k_2}{\rightarrow}} H1_{1c}$

Drug inactivation: $\quad Drug \overset{k_{drug}}{\longrightarrow} Drug_{intact}$
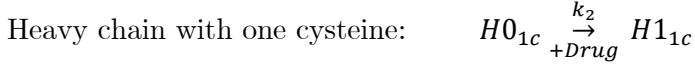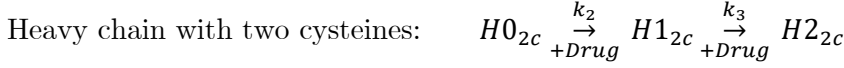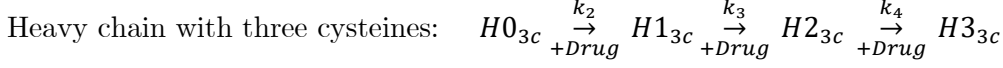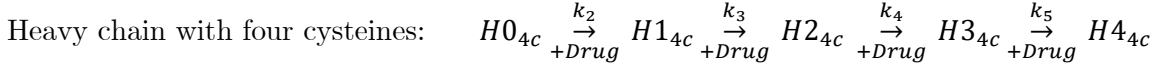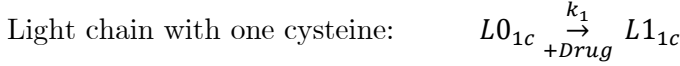
S3.3. DAR 8 model ODEs

The ODEs for a fed-batch conjugation and the detailed model with 4ks are given with:

$$\frac{dc_{L0_{1c}}}{dt} = -k_1 \, c_{L0_{1c}} \, c_{drug} - \frac{c_{L0_{1c}}}{V_l} q_{in} \tag{1}$$

$$\frac{dc_{L0_{0c}}}{dt} = -\frac{c_{L0_{0c}}}{V_l} q_{in} \tag{2}$$

$$\frac{dc_{H0_{4c}}}{dt} = -k_2 \, c_{H0_{4c}} \, c_{drug} - \frac{c_{H0_{4c}}}{V_l} q_{in} \tag{3}$$

$$\frac{dc_{H0_{3c}}}{dt} = -k_2 \, c_{H0_{3c}} \, c_{drug} - \frac{c_{H0_{3c}}}{V_l} q_{in} \tag{4}$$

$$\frac{dc_{H0_{2c}}}{dt} = -k_2 \, c_{H0_{2c}} \, c_{drug} - \frac{c_{H0_{2c}}}{V_l} q_{in} \tag{5}$$

$$\frac{dc_{H0_{1c}}}{dt} = -k_2 \, c_{H0_{1c}} \, c_{drug} - \frac{c_{H0_{1c}}}{V_l} q_{in} \tag{6}$$

$$\frac{dc_{H0_{0c}}}{dt} = -\frac{c_{H0_{0c}}}{V_l} q_{in} \tag{7}$$

$$\frac{dc_{L1_{1c}}}{dt} = k_1\, c_{L0_{1c}}\, c_{drug} - \frac{c_{L1_{1c}}}{V_l} q_{in} \tag{8}$$

$$\frac{dc_{H1_{4c}}}{dt} = -k_3\, c_{H1_{4c}}\, c_{drug} + k_2\, c_{H0_{4c}}\, c_{drug} - \frac{c_{H1_{4c}}}{V_l} q_{in} \tag{9}$$

$$\frac{dc_{L1_{1c}}}{dt} = k_1\, c_{L0_{1c}}\, c_{drug} - \frac{c_{L1_{1c}}}{V_l} q_{in} \tag{8}$$

$$\frac{dc_{H1_{4c}}}{dt} = -k_3\, c_{H1_{4c}}\, c_{drug} + k_2\, c_{H0_{4c}}\, c_{drug} - \frac{c_{H1_{4c}}}{V_l} q_{in} \tag{9}$$

$$\frac{dc_{H1_{3c}}}{dt} = -k_3\, c_{H1_{3c}}\, c_{drug} + k_2\, c_{H0_{3c}}\, c_{drug} - \frac{c_{H1_{3c}}}{V_l} q_{in} \tag{10}$$

$$\frac{dc_{H1_{2c}}}{dt} = -k_3\, c_{H1_{2c}}\, c_{drug} + k_2\, c_{H0_{2c}}\, c_{drug} - \frac{c_{H1_{2c}}}{V_l} q_{in} \tag{11}$$

$$\frac{dc_{H1_{1c}}}{dt} = k_2\, c_{H0_{1c}}\, c_{drug} - \frac{c_{H1_{2c}}}{V_l} q_{in} \tag{12}$$

$$\frac{dc_{H2_{4c}}}{dt} = -k_4\, c_{H2_{4c}}\, c_{drug} + k_3\, c_{H1_{4c}}\, c_{drug} - \frac{c_{H2_{4c}}}{V_l} q_{in} \tag{13}$$

$$\frac{dc_{H2_{3c}}}{dt} = -k_4\, c_{H2_{3c}}\, c_{drug} + k_3\, c_{H1_{3c}}\, c_{drug} - \frac{c_{H2_{3c}}}{V_l} q_{in} \tag{14}$$

$$\frac{dc_{H2_{2c}}}{dt} = k_3\, c_{H2_{1c}}\, c_{drug} - \frac{c_{H2_{2c}}}{V_l} q_{in} \tag{15}$$

$$\frac{dc_{H3_{4c}}}{dt} = -k_4\, c_{H3_{4c}}\, c_{drug} + k_4\, c_{H2_{4c}}\, c_{drug} - \frac{c_{H3_{4c}}}{V_l} \tag{16}$$

$$\frac{dc_{H3_{3c}}}{dt} = k_4\, c_{H3_{3c}}\, c_{drug} - \frac{c_{H3_{3c}}}{V_l} q_{in} \tag{17}$$

$$\frac{dc_{H4_{4c}}}{dt} = k_4\, c_{H3_{4c}}\, c_{drug} - \frac{c_{H4_{4c}}}{V_l} q_{in} \tag{18}$$

$$\frac{dc_{drug}}{dt} = -k_1\, c_{L0_{1c}}\, c_{drug} - k_2\, c_{H0_{4c}}\, c_{drug} - k_2\, c_{H0_{3c}}\, c_{drug}$$
$$-k_2\, c_{H0_{2c}}\, c_{drug} - k_2\, c_{H0_{1c}}\, c_{drug} - k_3\, c_{H1_{4c}}\, c_{drug}$$
$$-k_3\, c_{H1_{3c}}\, c_{drug} - k_3\, c_{H1_{2c}}\, c_{drug} - k_4\, c_{H2_{4c}}\, c_{drug}$$
$$-k_4\, c_{H2_{3c}}\, c_{drug} - k_4\, c_{H3_{4c}}\, c_{drug} - k_{drug}\, c_{drug} + \frac{c_{drug,in}}{V_l} q_{in} \tag{19}$$

$$\frac{dV_l}{dt} = q_{in} \tag{20}$$

Here, c_H and c_L represent the molar concentration of the heavy or light chain, respectively, where the number indicates the number of conjugated drugs and the index indicates the number of initial available cysteines, c_{drug} the molar concentration of payload, $c_{drug,in}$ the molar feed concentration of payload, k_j the reaction rate for the jth conjugation step and k_{drug} the depletion rate of the specific payload, V_l the reaction volume and q_{in} the feed flow rate. For Eq. (20) the liquid volume of the sample is subtracted outside the ODE-solver as sample is only taken at discrete time points. For the simple model with 1k the kinetic rates are simplified respectively. For modeling batch reaction, the last term in Eq. (1)-(19) representing the drug feeding is removed and Eq (20) is not used. This leads to 19 ODEs for the DAR 8 batch reaction.

S3.4. Conjugation with "inactivated" NPM and Drug1
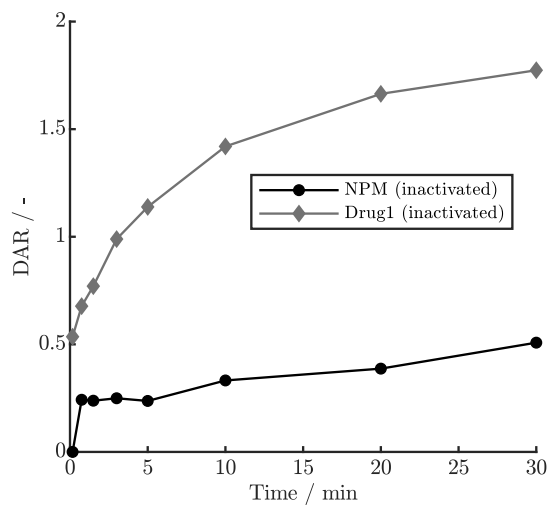
Figure S 1: DAR kinetics for 1.5 g/L mAb concentration using NPM and Drug1 that were pre-mixed in conjugation buffer for one hour.
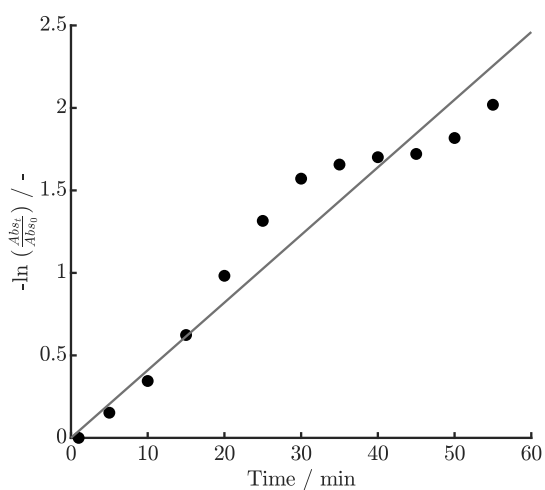
S3.5. Payload inactivation for NPM

Figure S 2: Linearized absorption plot of the absorbance data with a linear regression curve for NPM ($R^2 = 0.926$).
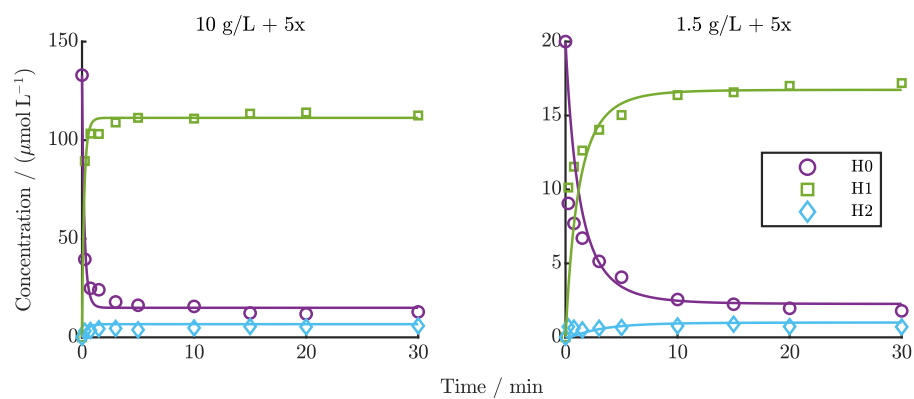
S3.6. DAR 2 model predictions for Drug1 and NPM

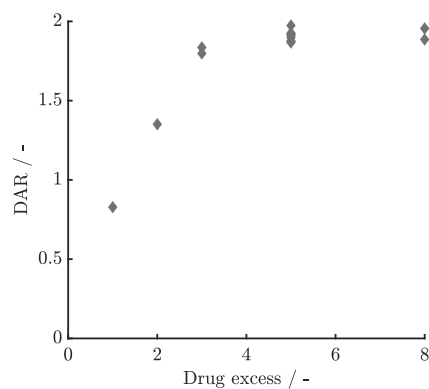Figure S 3: Comparison of DAR 2 model predictions vs. experimental data for test runs using ADC1 + Drug 1.

Figure S 4: Final DAR values over drug excess for all ADC1 runs.
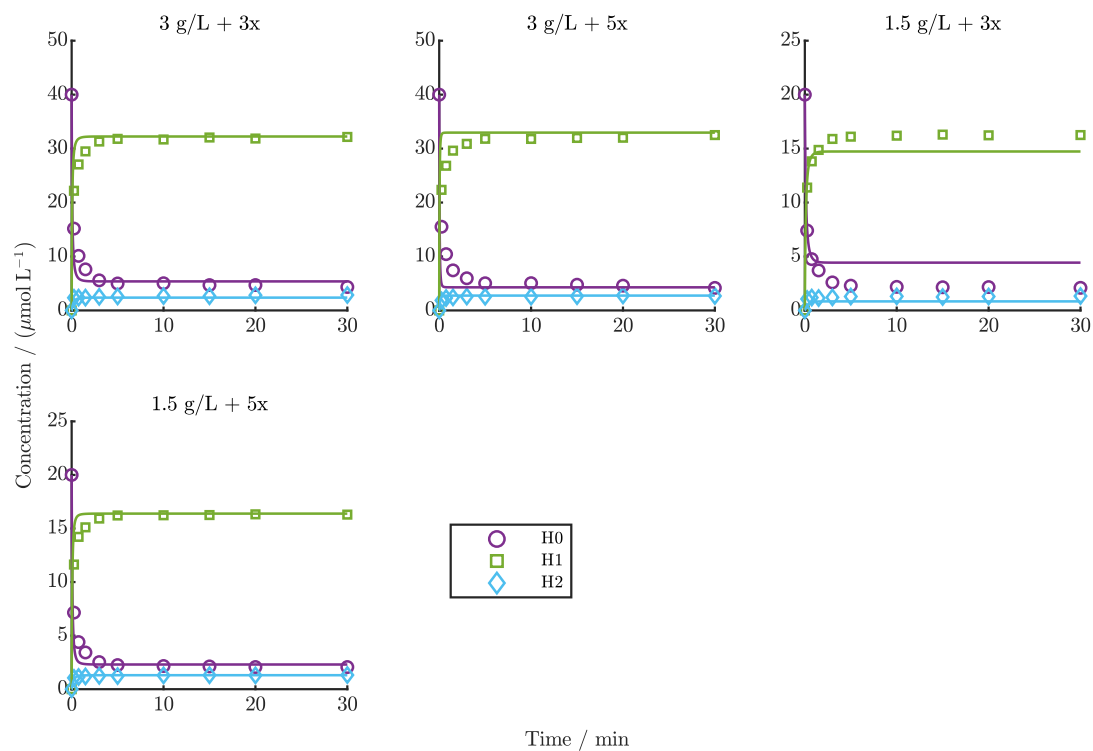
Figure S 5: Comparison of DAR 2 model predictions vs. experimental data for external runs using ADC1 + NPM. Model was re-calibrated and the determined payload depletion rate for NPM was set constant.

S3.7. DAR 8 model predictions using ADC2 for both training and test set
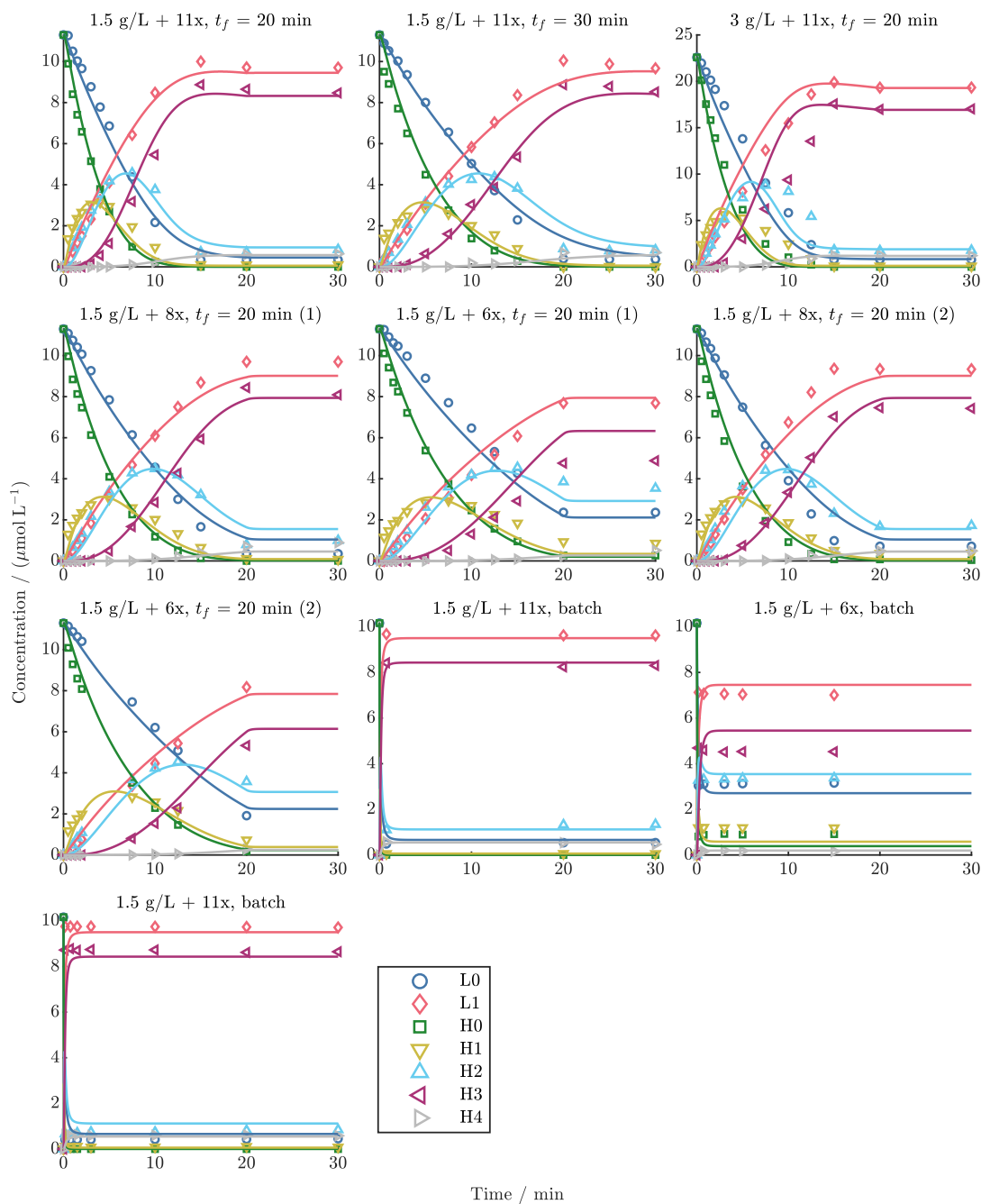
Figure S 6: Comparison of DAR 8 model predictions vs. experimental data for all ten training runs of ADC2 + NPM. For duplicate runs, the replicate number is indicated in the brackets.
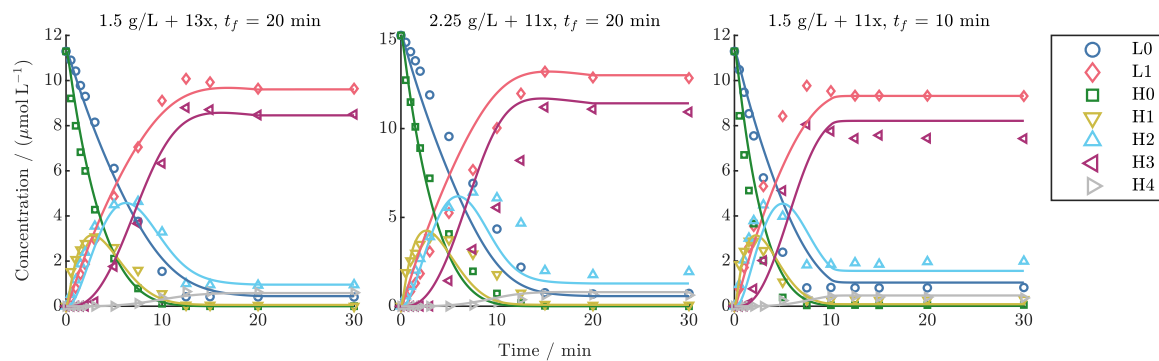
Figure S 7: Comparison of DAR 8 model predictions vs. experimental data for all three test runs of ADC2 + NPM.
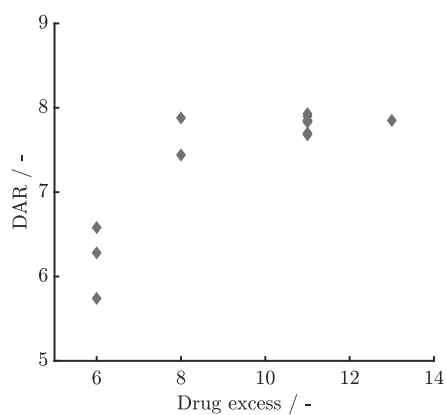
Figure S 8: Final DAR values over drug excess for all ADC2 runs.

S3.8. DAR 8 model predictions using ADC3
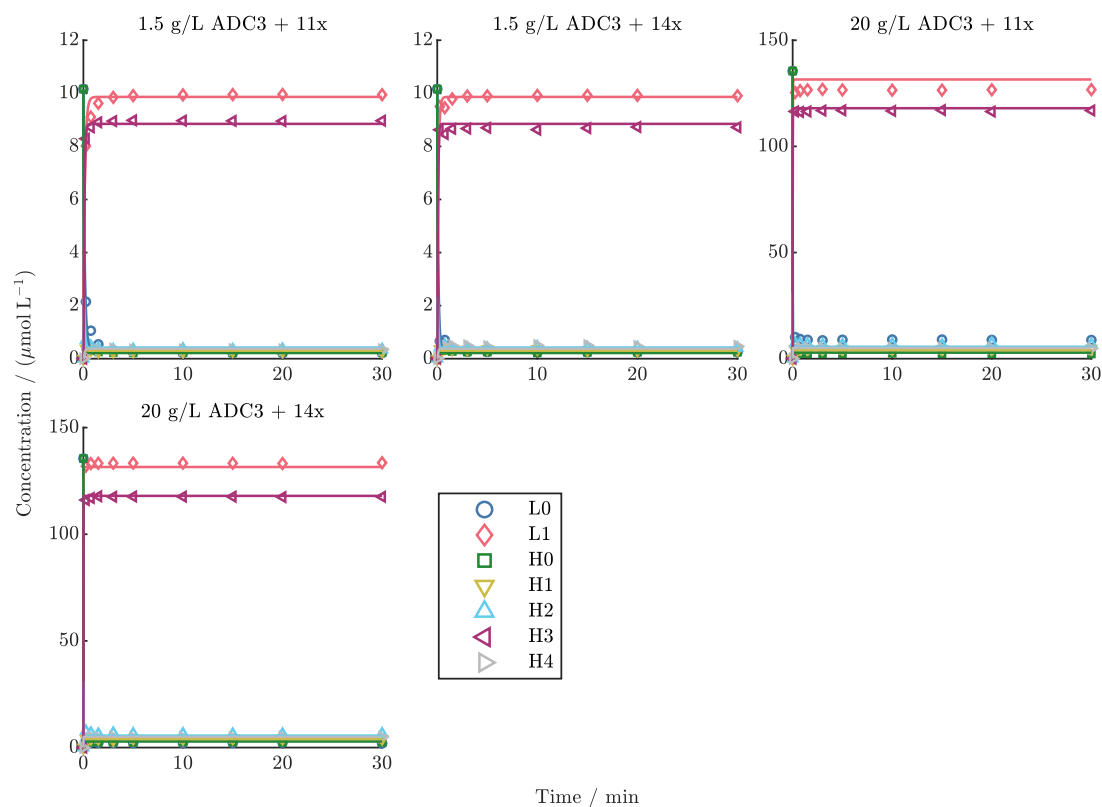
Figure S 9: Comparison of DAR 8 kinetic model predictions vs. experimental data for a mixed kinetic dataset of ADC3 conjugated with Drug2. All runs were performed in duplicates, except condition 1.5 g/L ADC3 + 11x was performed in quadruplicates.

Appendix B Supplementary Material for Chapter 4

S4.1. Overview of the reaction pathways and the assigned reaction rates

All CRN types and sets of kinetic rates for modeling the reduction kinetic are given in this section Figure S 10 shows the reaction pathways for the proposed CRN without intermediates and with re-oxidation (*NoInt-ReOx*). The re-oxidation reactions are the backwards reactions (reaction number 7-12). A list of all reactions with the differently assigned kinetic rates is given in Table S 2. The model with re-oxidation (4k model), was further refined by removing the pathways for inter HH re-oxidation, i.e., setting the rates for reactions 7-9 to zero (3k-model). For a detailed explanation see manuscript.
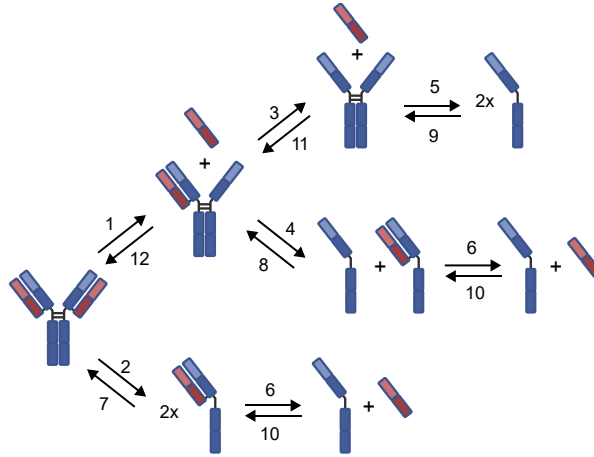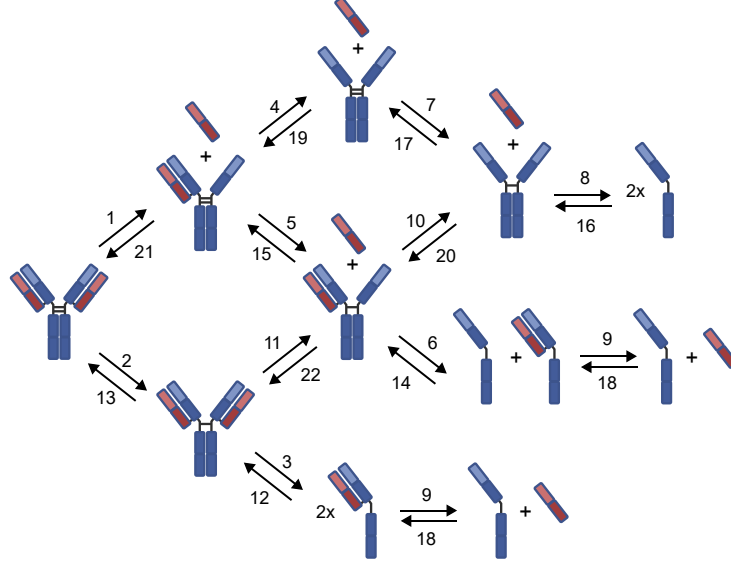

Figure S 10: CRN without intermediates and with re-oxidation (*NoInt+ReOx*) with the reaction numbers. The light chains are indicated in red and heavy chains in blue.

Table S 2: Overview of the reactions in the *NoInt* CRN and the assigned kinetic rates for the 2k-model (not including re-oxidation) and 4k/3k-model (including re-oxidation). The 3k-model sets the reactions for the inter HH re-oxidation steps to zero.

No	Reduction pathway	k_i (2k)	k_i (4k/3k)	No	Re-oxidation pathway	k_i (2k)	k_i (4k/3k)
1	$Int + TCEP \rightarrow HHL + L$	1	1	7	$2HL \rightarrow Int$	-	(4)
2	$Int + 2TCEP \rightarrow 2HL$	2	2	8	$HL + H \rightarrow HHL$	-	(4)
3	$HHL + TCEP \rightarrow HH + L$	1	1	9	$2H \rightarrow HH$	-	(4)
4	$HHL + 2 \cdot TCEP \rightarrow H + HL$	2	2	10	$H + L \rightarrow HL$	-	3
5	$HH + 2TCEP \rightarrow 2H$	2	2	11	$HH + L \rightarrow HHL$	-	3
6	$HL + TCEP \rightarrow H + L$	1	1	12	$HHL + L \rightarrow Int$	-	3

In contrast, Figure S 11 shows the reaction pathways for the proposed CRN with intermediates and re-oxidation (*WithInt-ReOx*). The re-oxidation reactions are the backwards reactions (reaction number 12-22). A list of all reactions with the differently assigned kinetic rates for this CRN is given in Table S 3. The model with re-oxidation (5k model), was further refined by removing the pathways for inter HH re-oxidation, i.e., setting the rates for reactions 12-17 to zero (3k-model). The 8k- and 9k-model further discriminate the rates (for detailed explanation see manuscript).

Figure S 11: CRN with intermediates and re-oxidation (*WithInt+Re-Ox*) with the reaction numbers. The light chains are indicated in red and heavy chains in blue.

Table S 3: Overview of the reactions in the *WithInt+ReOx* CRN and the assigned kinetic rates for the different models.

No	Reduction pathway	k_i (5k, 3k)	k_i (8k)	k_i (9k)	No	Re-oxidation pathway	k_i (5k, 3k)	k_i (8k)	k_i (9k)
1	$Int + TCEP \rightarrow HHL + L$	1	1	1	12	$2HL \rightarrow IntR$	(4)	-	-
2	$Int + TCEP \rightarrow IntR$	2	4	4	13	$IntR \rightarrow Int$	(5)	-	-
3	$IntR + TCEP \rightarrow 2HL$	2	5	4	14	$HL + H \rightarrow HHLR$	(4)	-	-
4	$HHL + TCEP \rightarrow HH + L$	1	2	2	15	$HHLR \rightarrow HHL$	(5)	-	-
5	$HHL + TCEP \rightarrow HHLR$	2	4	5	16	$2H \rightarrow HHR$	(4)	-	-
6	$HHLR + TCEP \rightarrow H + HL$	2	5	5	17	$HHR \rightarrow HH$	(5)	-	-
7	$HH + TCEP \rightarrow HHR$	2	4	6	18	$H + L \rightarrow HL$	3	8	9
8	$HHR + TCEP \rightarrow 2H$	2	5	6	19	$HH + L \rightarrow HHL$	3	7	8
9	$HL + TCEP \rightarrow H + L$	1	3	3	20	$HHR + L \rightarrow HHLR$	3	7	8
10	$HHLR + TCEP \rightarrow HHR + L$	1	2	2	21	$HHL + L \rightarrow Int$	3	6	7
11	$IntR + TCEP \rightarrow HHLR + L$	1	1	1	22	$HHLR + L \rightarrow IntR$	3	6	7

S4.2. Exemplary ODE

The resulting differential equations for each of the ten reacting species in the *Within+Re-Ox* CRN is given in the following as mole balances:

$$\frac{\partial y_{Intact}}{\partial t} = -r_1 - r_2 + r_{13} + r_{21}$$

$$\frac{\partial y_{HHL}}{\partial t} = r_1 - r_4 - r_5 + r_{15} + r_{19} - r_{21}$$

$$\frac{\partial y_{HH}}{\partial t} = r_4 - r_7 + r_{17} - r_{19}$$

$$\frac{\partial y_{HL}}{\partial t} = 2r_3 + r_6 - r_9 - r_{12} - r_{14} + r_{18}$$

$$\frac{\partial y_{H}}{\partial t} = r_6 + 2r_8 + r_9 - r_{14} - r_{16} - r_{18}$$

$$\frac{\partial y_{L}}{\partial t} = r_1 + r_4 + r_9 + r_{10} + r_{11} - r_{18} - r_{19} - r_{20} - r_{21} - r_{22}$$

$$\frac{\partial y_{Intact,r}}{\partial t} = r_2 - r_3 - r_{11} + r_{12} - r_{13} + r_{22}$$

$$\frac{\partial y_{HHL,r}}{\partial t} = r_5 - r_6 - r_{10} + r_{11} + r_{14} - r_{15} + r_{20} - r_{22}$$

$$\frac{\partial y_{HH,r}}{\partial t} = r_7 - r_8 + r_{10} + r_{16} - r_{17} - r_{20}$$

$$\frac{\partial y_{TCEP}}{\partial t} = -r_1 - r_2 - r_3 - r_4 - r_5 - r_6 - r_7 - r_8 - r_9 - r_{10} - r_{11}.$$

For example, for the simple 5k-model, the rate laws for the reduction reactions $(r_1 - r_{11})$ and the re-oxidation reactions $(r_{12} - r_{22})$ in the Arrhenius reparametrized form is summarized in Table S 4. The rates $k_1 - k_5$ are the activation energies E_a and the rates $k_6 - k_{10}$ as the reaction rates at the reference temperature.

Table S 4: Overview of the kinetic rate laws in the reparametrized Arrhenius form for the 5k-model.

Reduction reactions:	Re-oxidation reactions:
$r_1 = k_6 \cdot exp\left(-\frac{k_1}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{Intact} \cdot y_{TCEP}$	$r_{12} = k_9 \cdot exp\left(-\frac{k_4}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot (y_{HL})^2$
$r_2 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{Intact}$ $\cdot y_{TCEP}$	$r_{13} = k_{10} \cdot exp\left(-\frac{k_5}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{Intactr}$
$r_3 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr} \cdot y_{TCEP}$	$r_{14} = k_9 \cdot exp\left(-\frac{k_4}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HL} \cdot y_H$
$r_4 = k_6 \cdot exp\left(-\frac{k_1}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHL} \cdot y_{TCEP}$	$r_{15} = k_{10} \cdot exp\left(-\frac{k_5}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHLr}$
$r_5 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHL} \cdot y_{TCEP}$	$r_{16} = k_9 \cdot exp\left(-\frac{k_4}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot (y_H)^2$
$r_6 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr} \cdot y_{TCEP}$	$r_{17} = k_{10} \cdot exp\left(-\frac{k_5}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr}$
$r_7 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HH} \cdot y_{TCEP}$	$r_{18} = k_8 \cdot exp\left(-\frac{k_3}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_H \cdot y_L$

$$r_8 = k_7 \cdot exp\left(-\frac{k_2}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr} \cdot y_{TCEP} \qquad r_{19} = k_8 \cdot exp\left(-\frac{k_3}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HH} \cdot y_L$$

$$r_9 = k_6 \cdot exp\left(-\frac{k_1}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HL} \cdot y_{TCEP} \qquad r_{20} = k_8 \cdot exp\left(-\frac{k_3}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr} \cdot y_L$$

$$r_{10} = k_6 \cdot exp\left(-\frac{k_1}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHLr} \qquad r_{21} = k_8 \cdot exp\left(-\frac{k_3}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHL} \cdot y_L$$
$$\cdot y_{TCEP}$$

$$r_{11} = k_6 \cdot exp\left(-\frac{k_1}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHr} \qquad r_{22} = k_8 \cdot exp\left(-\frac{k_3}{R} \cdot \left(\frac{1}{T_K} - \frac{1}{T_{ref}}\right)\right) \cdot y_{HHLr} \cdot y_L$$
$$\cdot y_{TCEP}$$

S4.3. Parameter confidence intervals for the 8k-model

Table S 5: Estimated parameters and confidence intervals for the 8k-model. The parameters k1-8 represent the kinetic rates at the reference temperature and k9-16 the corresponding activation energies.

Parameter	Estimate	95% CI		in %
k1	6.98E-04	5.81E-04	8.15E-04	16.74
k2	2.89E-04	2.40E-04	3.37E-04	16.75
k3	6.81E-05	2.41E-05	1.12E-04	64.65
k4	6.47E-05	4.55E-05	8.39E-05	29.68
k5	1.18E-03	-4.27E-04	2.78E-03	136.27
k6	3.00E-05	1.37E-05	4.64E-05	54.51
k7	6.23E-05	3.62E-05	8.84E-05	41.89
k8	1.15E-05	4.68E-06	1.83E-05	59.24
k9	24.30	14.63	33.97	39.80
k10	25.65	15.57	35.72	39.29
k11	26.72	12.50	40.93	53.20
k12	43.05	28.78	57.32	33.14
k13	99.99	98.75	101.23	1.24
k14	4.43	-26.63	35.48	701.46
k15	25.47	-0.51	51.45	101.98
k16	30.17	14.28	46.05	52.65

S4.4. Evaluation of reduction model on mAb2
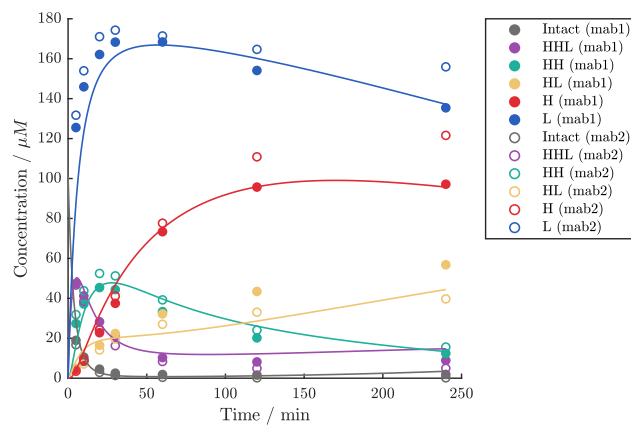
Figure S 12: Comparison of the experimental reduction kinetic of two runs (Run3 and Run18) for mAb1 (filled markers) and mAb2 (empty markers) with the same experimental conditions. The lines indicate the predictions of the reduction model calibrated on the mAb1 data.

S4.5. Pearson correlation matrix between CGE and RP

To illustrate the cross-analytical relationship between the six experimental CGE concentrations and the six RP-UHPLC species as well as the DAR value, a correlation matrix was plotted as shown in Figure S 13. It can be seen, that each RP species exhibits a distinct, individual correlation with the CGE species as indicated by the Pearson correlation coefficents r. For example, the DAR is strongly correlated with HHL ($r = -0.94$) and H ($r = 0.90$), whereas L0 is strongly correlated L ($r = -0.97$). With the exception of H2, all RP species demonstrated at least one Pearson correlation coefficent greater than ± 0.9. It is worth noting, that these correlations changed when instead using the model-predicted CGE concentrations.
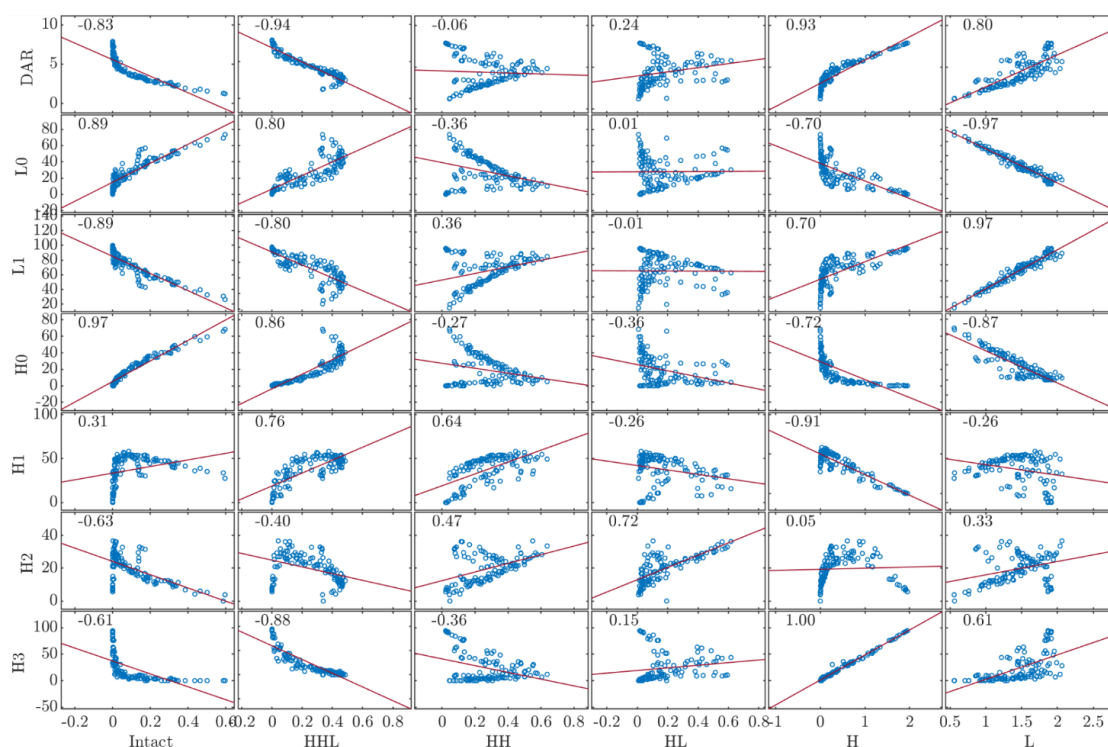
Figure S 13: Pearson correlation matrix between each pair of variables in the CGE data (Intact, HHL, HH, HL, H and L) and the RP data including the DAR (DAR, L0, L1, H0, H1, H2 and H3). The reference line displays the Pearson correlation coefficient.

S4.6. All DLD-% results for the Integrated in-silico screening
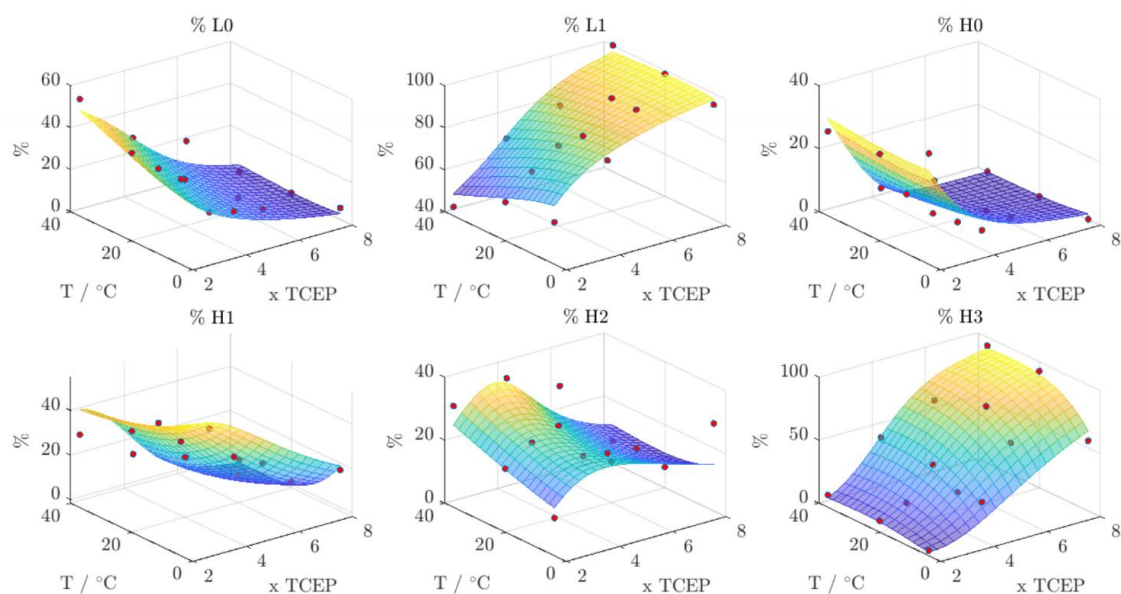
Figure S 14: Results for all DLD-% in the in-silico screening for varying temperature and TCEP excess for constant c_{mAb}=15 g/L and t=120 min. The surface plots represent the model predictions all DLD-% including the experimental RP values (red spherics). Blue colors indicate lower values and yellow indicate higher values.
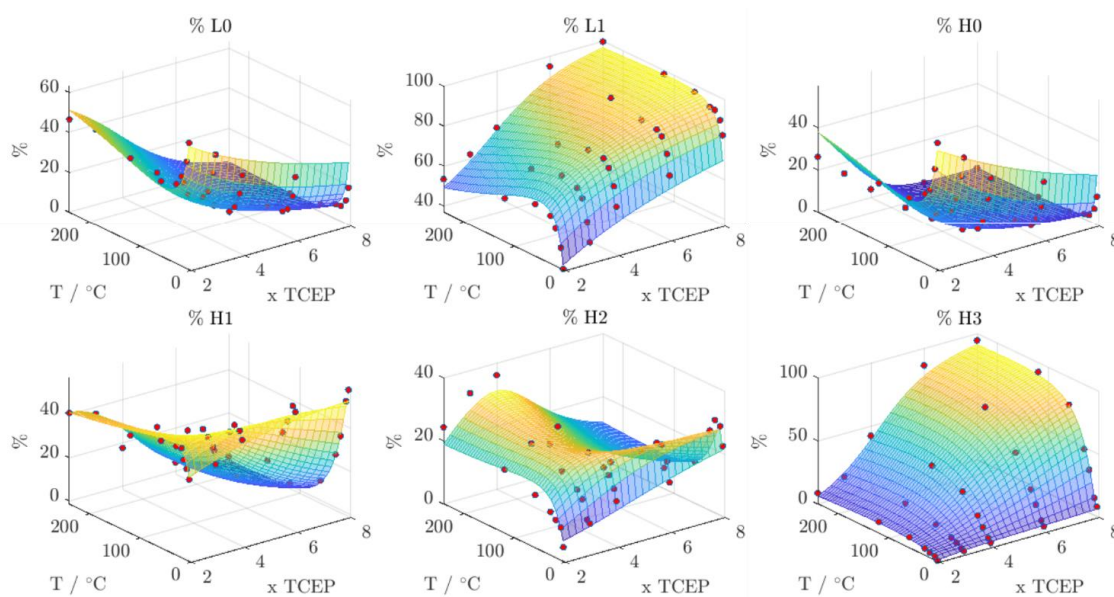
Figure S 15: Results for all DLD-% in the in-silico screening for varying reaction time and TCEP excess for constant c_{mAb}=15 g/L and T=20 °C. The surface plots represent the model predictions for all DLD-% including the experimental RP values (red spherics). Blue colors indicate lower values and yellow indicate higher values.

Appendix C   Supplementary Material for Chapter 5

S5.1. Experimental mixing study for DAR8
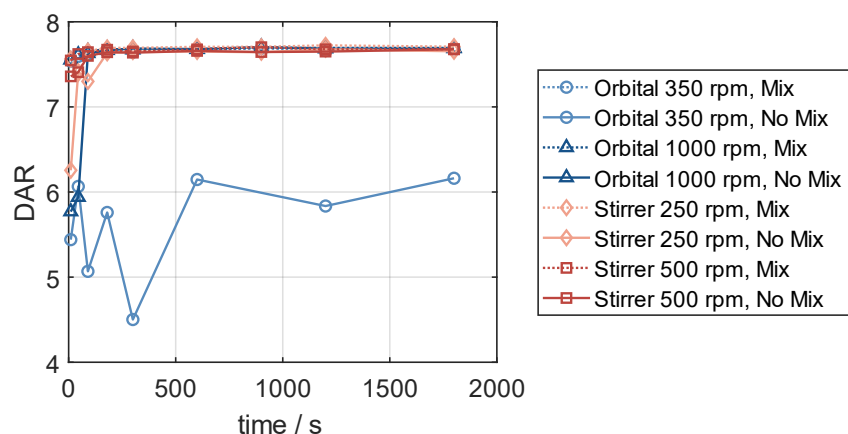
Figure S 16 DAR kinetics for different mixing modes with and without initial payload mixing performed. Reaction conditions were 1.5 g/L mAb and 14x molar payload excess.
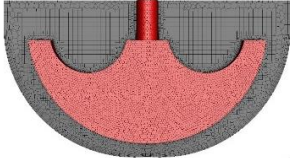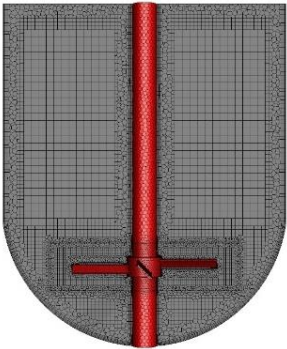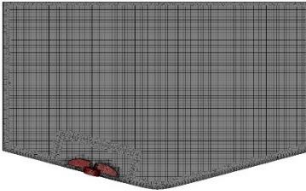
S5.2. CAD Geometries

Table S 6 Vessel, stirrer and shaft dimensions of the CAD models.

	GST-1	GST-2	SUM
Inner diameter vessel / mm	108	300	384
Vessel bottom type	Curved spherical	Curved spherical	Cone
Height of the curved bottom / mm	45	130	52.7
Impeller diameter / mm	81	152	65
Impeller width / mm	35	20	11
Impeller angle	0°	45°	15°
Impeller clearance / mm	10	85	4.8
Shaft diameter / mm	7.2	25.4	-

S5.3. Generated meshes

Table S 7 Mesh metrics for the three final meshes and feeding positions in this study.

Reactor	300 mL GST-1	22 L GST-2	25 L SUM
Stirrer speed	60 rpm	120 rpm	400 rpm
Mesh count	206123	582731	940652
Ortho min (average)	0.21 (0.94)	0.2 (0.96)	0.20 (0.98)
Skew max	0.79	0.8	0.80
Poly-Hexcore meshes			
Tracer/ Payload addition positions	2.5 mm below surface and 2 cm distance to wall	1.5 cm below surface centered in between wall and tank middle axis	2.5 cm below surface and 3 cm distance to wall

S5.4. Global mesh study
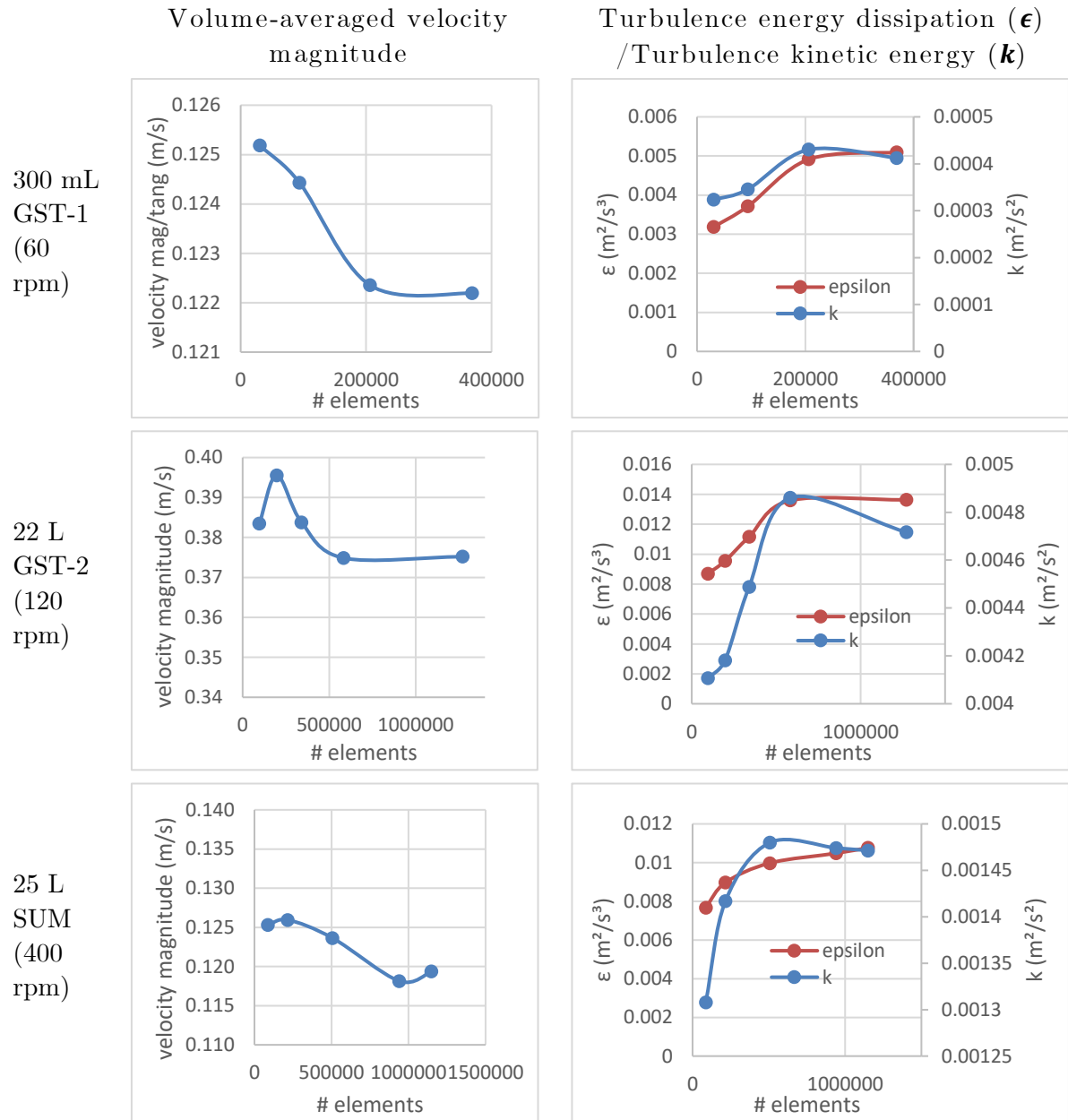
Figure S 17 Mesh independency test for the three vessels (at highest rpm) with regards to volume-averaged velocity magnitude and turbulent energy dissipation and turbulence kinetic energy.

S5.5. Convergence monitors for steady-state simulations using the final meshes
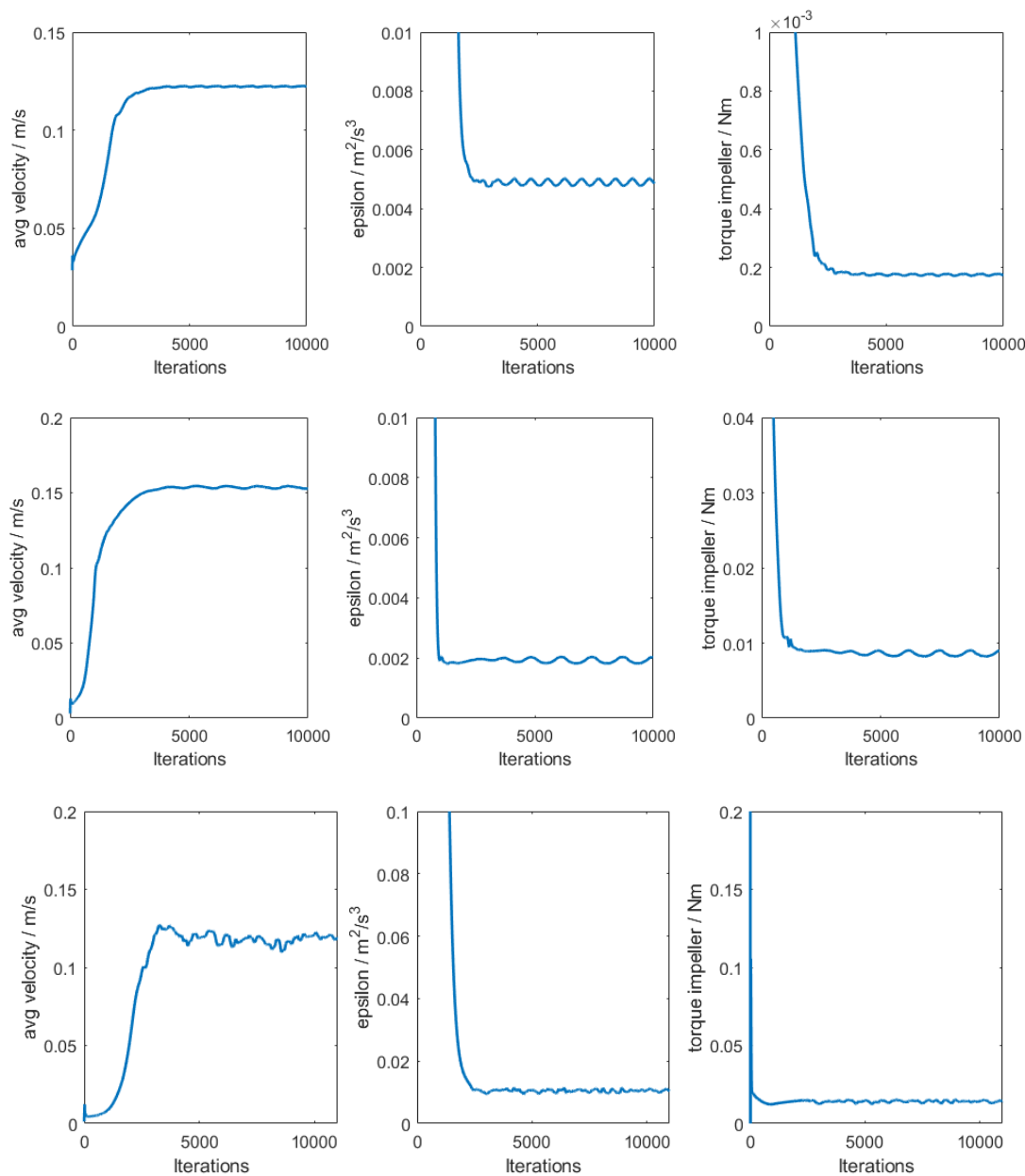
Figure S 18 Monitored values of averaged velocity magnitude, turbulent energy dissipation and impeller torque for the three vessels over the iterations during the steady-state simulations. Top: GST-1 (60 rpm), middle: GST-2 (60 rpm) and bottom: SUM (400 rpm).

S5.6. Simulated mixing times
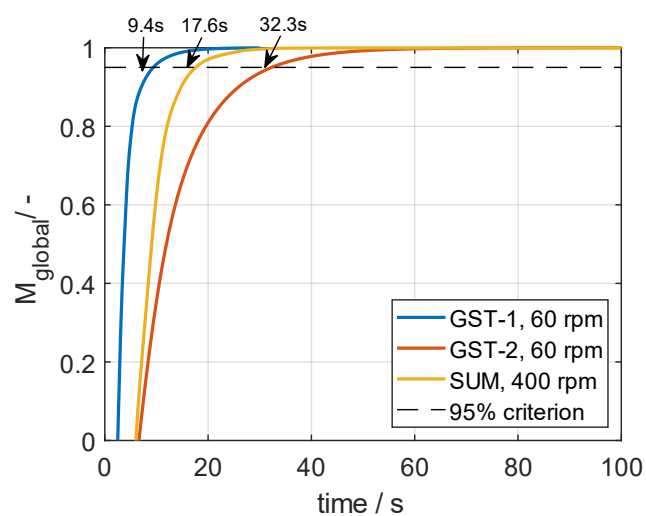
Figure S 19 Simulated <u>global</u> mixing curve for all three studied vessels with characteristic mixing times determined for $M_{global} = 0.95$.

Appendix D Supplementary Material for Chapter 6

S6.1. Gaussian process regression

Mathematical formulation

In Gaussian process regression for spectroscopic data, a functional relationship between an input spectrum $\mathbf{x}$ with M wavelengths with an output variable y is established. The functional relationship is created by a Gaussian process (GP) which can be defined as a collection of random variables such that any finite subset exhibits a joint Gaussian distribution [105], [106]. A GP is mathematically defined by its mean function m and covariance function C and can be written as

$$P(y|\mathbf{X}) = N\big(m(\mathbf{x}), C(\mathbf{x}, \mathbf{x})\big) \tag{1}$$

assuming a Gaussian posterior $P(y|\mathbf{X})$ with $\mathbf{X}$ being the training data consisting of N samples of input spectra. Considering mean-centered input data, the mean function of the GP can be set to $m(\mathbf{x}) = 0$, leaving the covariance function C as the only configurable of the GP. The covariance function $C(\mathbf{x}, \mathbf{x}^*)$ computes a distance measure of two arbitrary input spectra $\mathbf{x}$ and $\mathbf{x}^*$. The choice of C, hence, affects the computed relationships between different input spectra. In this study, a composite covariance function was used which is constructed from the sum of the so-called linear and radial-basis function (RBF) kernels, such that

$$C(\mathbf{x}, \mathbf{x}^*) = a_0 + \sigma_0^2 \sum_{i=1}^{N} \mathbf{x}^T \mathbf{x}^* + \sigma_1^2 \exp\left(-\frac{||\mathbf{x} - \mathbf{x}^*||_2^2}{2l}\right) + \sigma_e^2 \tag{2}$$

with a_0, σ_0^2, σ_1^2 and l being the constant offset, the variance of the linear kernel, the variance of the RBF kernel and the length-scale, respectively. The GP covariance function also assumes Gaussian additive noise with variance σ_e^2. All parameters are considered hyperparameters and need to be inferred from the training data.

For parameter inference, the log-likelihood function is maximized with regard to the hyperparameters. According to [106], the log-likelihood can be calculated as

$$L = -\frac{1}{2}(\text{logdet}\Sigma - \mathbf{y}^T \Sigma^{-1} \mathbf{y} - N \log 2\pi) \tag{3}$$

with Σ being the covariance matrix as computed by Eq. 2 using all spectra available in the training data and $\mathbf{y}$ the corresponding output values.

Using the estimated set of hyperparameters, the predictive posterior distribution is given by the mean y^* and variance σ^{*2} as

$$y^* = \mathbf{k}_{GP}^T(\mathbf{x}^*)\Sigma^{-1}\mathbf{y} \tag{4}$$

$$\sigma^{*2} = C(\mathbf{x}, \mathbf{x}^*) - \mathbf{k}_{GP}^T(\mathbf{x}^*)\Sigma^{-1}\mathbf{k}_{GP}(\mathbf{x}^*), \tag{5}$$

where $\mathbf{k}_{GP} = [C(\mathbf{x_1}, \mathbf{x}^*), \dots, C(\mathbf{x}_N, \mathbf{x}^*)]^T$.

Hyperparameter optimization

The kernel parameters were estimated by expectation maximization as implemented in [219]. The resulting hyperparameter values for the GPR-VS model were $a_0 = 1e\text{-}5$, $\sigma_0 = 143$, $\sigma_1 = 0.105$, $\sigma_e = 0.0133$ and $l = 0.0536$.

S6.2. Kinetic modeling

The herein applied kinetic model assumes that the reactive cysteines of the oxidized mAb conjugate irreversibly with a drug in a two-stage consecutive reaction. The two reaction steps for the first and second conjugation are described by second-order rate laws having two kinetic rates k_1 and k_2. The inactivation of free drug is described by a first-order kinetic with the rate constant k_3. In theory, a free drug molecule is expected to conjugate with each of the two reactive cysteine residues which were inserted in a single mAb. In practice, a residual level of un- and mono-conjugated mAb are present after the conjugation reaction. To account for this, the model assumes an initial distribution of reactive cysteine residues for the oxidized mAb intermediate. This distribution is experimentally determined based on the final composition of conjugated species in runs with a NPM excess of 3x or higher. The kinetic model for the consecutive conjugation reaction of a mixture of mAbs with zero, one or two binding sites with the surrogate payload NPM is described by a set of seven ordinary differential equations (ODEs):

$$\frac{dc_{mAb_{2c}}}{dt} = -k_1 \cdot c_{mAb_{2c}} \cdot c_{NPM} \tag{6a}$$

$$\frac{dc_{mAb_{1c}}}{dt} = -k_1 \cdot c_{mAb_{1c}} \cdot c_{NPM} \tag{6b}$$

$$\frac{dc_{mAb_{0c}}}{dt} = 0 \tag{6c}$$

$$\frac{dc_{Conj1_{2c}}}{dt} = k_1 \cdot c_{mAb_{2c}} \cdot c_{NPM} - k_2 \cdot c_{Conj1_{1c}} \cdot c_{NPM} \tag{6d}$$

$$\frac{dc_{Conj1_{1c}}}{dt} = k_1 \cdot c_{mAb_{1c}} \cdot c_{NPM} \tag{6e}$$

$$\frac{dc_{Conj2}}{dt} = k_2 \cdot c_{Conj1_{1c}} \cdot c_{NPM} \tag{6f}$$

$$\frac{dc_{NPM}}{dt} = -k_1 \cdot c_{mAb_{2c}} \cdot c_{NPM} - k_1 \cdot c_{mAb_{1c}} \cdot c_{NPM} - k_2 \cdot c_{Conj1_{1c}} \cdot c_{NPM} - k_3 \cdot c_{NPM} \tag{6g}$$

where $c_{mAb_{2c}}$, $c_{mAb_{1c}}$ and $c_{mAb_{0c}}$ denote the concentration of unconjugated mAbs with two, one and zero available binding sites, respectively. The free NPM concentration is given as c_{NPM}. Further, the concentrations $c_{Conj1_{2c}}$, $c_{Conj1_{1c}}$, c_{Conj2} define the concentration of Conj1 with one attached NPM molecules and one binding sites, concentration of Conj1 with one attached NPM molecule and no binding site, and the concentration of Conj2 with two attached NPM molecules, respectively.

S6.3. Extended Kalman Filter

The working principle of the EKF can be described by a recursive execution of a prediction and correction step. During the prediction step, the state function is solved for a discrete time interval $\Delta t = t_k - t_{k-1}$ to propagate the state vector $\mathbf{x}$ to the next time step according to Eq. 7. To describe the covariance in each state, the EKF uses the state covariance matrix $\mathbf{P}$ which is propagated to the next time step in parallel according to Eq. 8

$$\mathbf{x}^-(t_k) = \mathbf{x}^+(t_{k-1}) + \int_{t_{k-1}}^{t_k} (f(\mathbf{x}, u(\tau), \boldsymbol{\theta}) d\tau \tag{7}$$

$$\mathbf{P}^-(t_k) = \mathbf{P}^+(t_{k-1}) + \int_{t_{k-1}}^{t_k} \left(\mathbf{Z}(\tau)\mathbf{P}(\tau) + \mathbf{P}(\tau)\mathbf{Z}^T(\tau) + \mathbf{Q}(\tau) \right) d\tau \tag{8}$$

where the superscripts of the state vector and state covariance matrix (+) and (-) indicate the corrected values at the previous time instance and the predicted values at the current time instance, respectively. The process noise covariance matrix is denoted as $\mathbf{Q}$. To estimate the change in the state covariance, the local derivatives $\mathbf{Z}(t)$ of the state function with regard to all model states are introduced as

$$\mathbf{Z}(t_k) = \left(\frac{\partial f}{\partial \mathbf{x}} \right)_{\mathbf{x}(t_k), \boldsymbol{\theta}}. \tag{9}$$

For the first iteration, the reaction states and state covariance matrix require initialization. In this study, all reaction states and the corresponding covariance matrix were initialized as suggested in [208].

During the correction step, the predicted estimate is fused with the available measurement $y(t_k)$ using the Kalman Gain $\mathbf{K}(t_k)$. The Kalman Gain takes into account the state covariance $\mathbf{P}^-(t_k)$, the measurement noise $\mathbf{R}(t_k)$ and the local derivatives $\mathbf{C}(t)$ of the measurement function with regard to all model states and is computed according to

$$\mathbf{K}(t_k) = \mathbf{P}^-(t_k)\mathbf{C}(t_k)^T \left(\mathbf{C}(t_k)\mathbf{P}^-(t_k)\mathbf{C}(t_k)^T + \mathbf{R}(t_k) \right)^{-1} \tag{10}$$

where $\mathbf{C}(t_k)$ is defined as

$$\mathbf{C}(t_k) = \left(\frac{\partial h}{\partial x} \right)_{\mathbf{x}(t_k), \boldsymbol{\theta}}. \tag{11}$$

The corrected states vector and the corrected state covariance are then computed according to

$$\mathbf{x}^+(t_k) = \mathbf{x}^-(t_k) + \mathbf{K}(t_k)\left(y(t_k) - h\big(\mathbf{x}^-(t_k)\big)\right) \tag{12}$$

$$\mathbf{P}^+(t_k) = \big(1 - \mathbf{K}(t_k)\mathbf{C}(t_k)\big)\mathbf{P}^-(t_k) \tag{13}$$

S6.4. Error metrics

The root mean square error ($RMSE$) is defined as

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}(y_i - \hat{y}_i)^2}{n}}, \tag{14}$$

where $\hat{y}_i$ denotes the predicted concentration values, y_i the reference data from RP-UPHLC and n the sample size. To evaluate model equality, the $RMSECV$ and the $RMSEP$ are calculated based on the Eq. 14. To compute the $RMSECV$, the $RMSE$ for the validation data of each iteration are averaged. For the $RMSEP$, Eq. 14 is applied using only the test data. To optimize PLSR model using cross-validation, a scaled sum or squared errors SSE_{scaled} was introduced according to [203] which is given as

$$SSE_{\text{scaled}} = \frac{\sum_{n-1}^{N}(y_i - \hat{y}_i)^2}{n - n_{\text{PLS}} - 1} \tag{15}$$

with n_{PLS} being the number of latent variables used in the PLSR model.

S6.5. Chemometric model development

Spectroscopic data

This section compiles spectroscopic data used for the closer evaluation of the presented results in this study. Figure S 20 presents an overview of all absorbance spectra used in this study. From the raw absorbance spectra, the location of the peak maxima of the two main peak bands of NPM were extracted and are displayed over time side-by-side with the raw spectra. Figure S 21 displays the timely absorbance spectra for the surrogate drug without adding any mAb to the solution.

Variable Selection

The selected variables of the GPR-VS model which were optimized based on the cross-validation error are displayed in Figure S 22. The selected variables are indicated by gray vertical lines which accumulate around the absorbance peaks between 250-280 nm and 320-350 nm. This is also the spectral area where the two peak shifts of the NPM molecule occur during the conjugation reaction as shown described above. The same bathochromic red shift of the NPM spectrum has previously been reported [98].

S6.6. Soft-sensor development

This section compiles the individual soft-sensor predictions for all conjugation runs used in this study. Figure S 23 and Figure S 25 show the model predictions of the spectroscopic sensor model, the kinetic model and the EKF soft-sensor for the conjugated drug concentration and all individual species in the test and training subset.
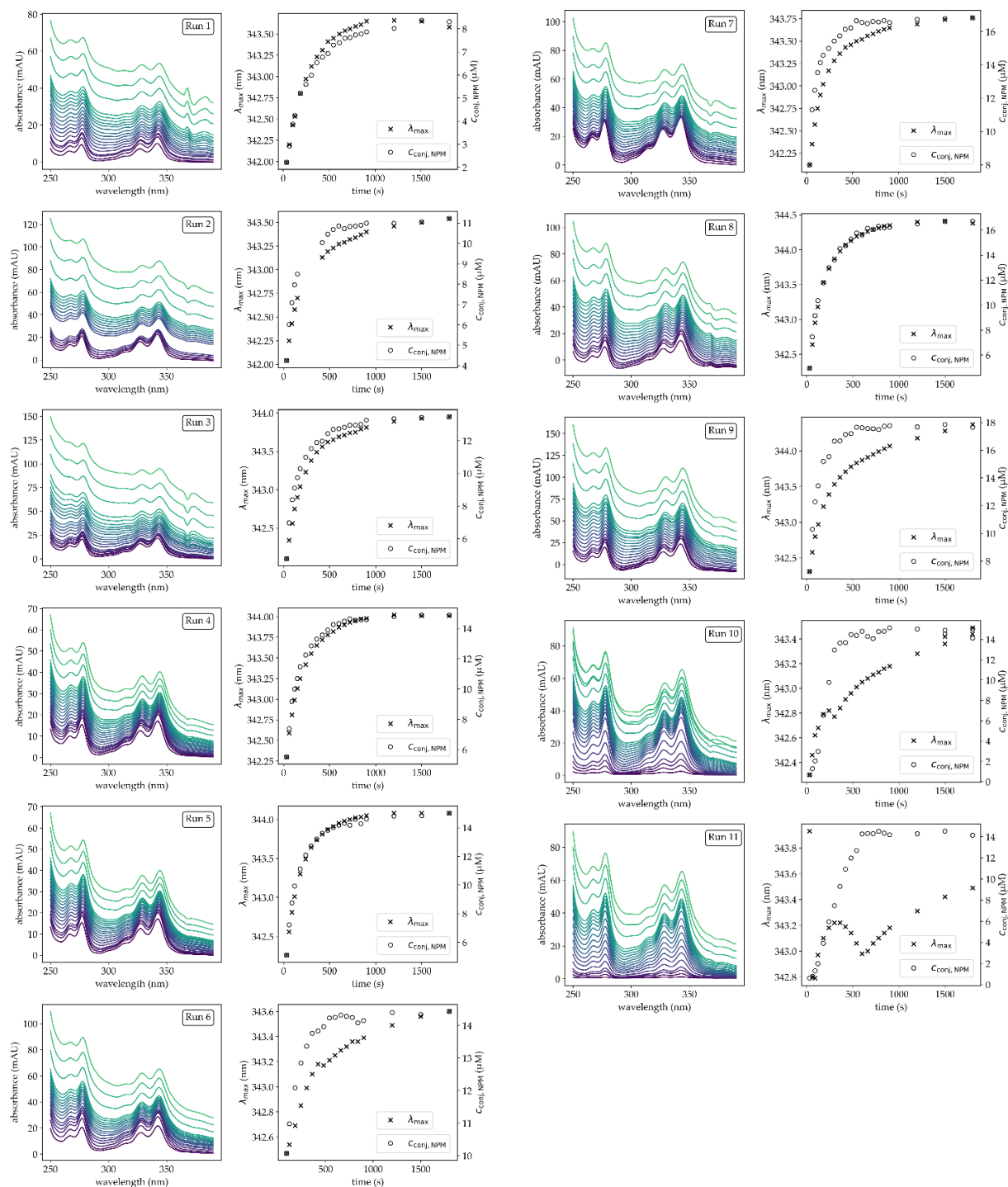

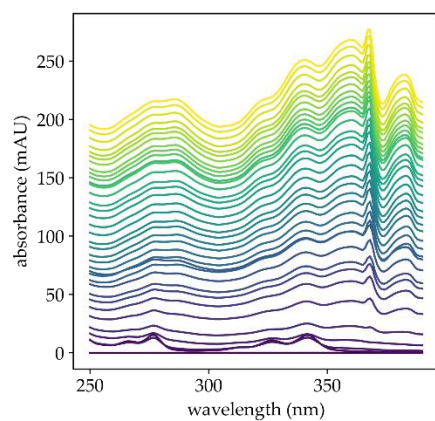Figure S 20 Absorbance spectra and detected shift of the location of the peak maximum for all runs used in this study.

Figure S 21 Raw absorbance spectra for the surrogate drug without addition of mAb solution. The surrogate drug was continuously stirred and circulated through the on-line loop for 30 min. Only spectra for intervals of 25 s are shown.
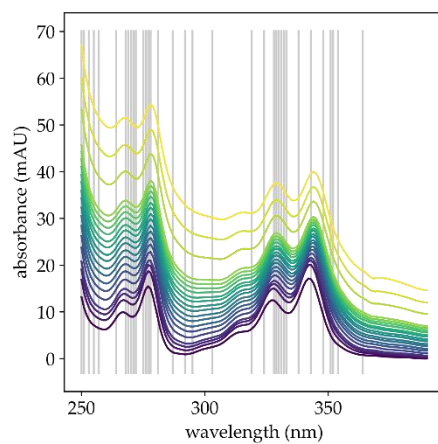

Figure S 22 Selected variables for GPR model in combination with variable selection algorithm. Selected variables are indicated by vertical gray lines. The absorbance spectra of run 5 are shown exemplarily and are colored by the reaction time from purple (0 min) to yellow (30 min).
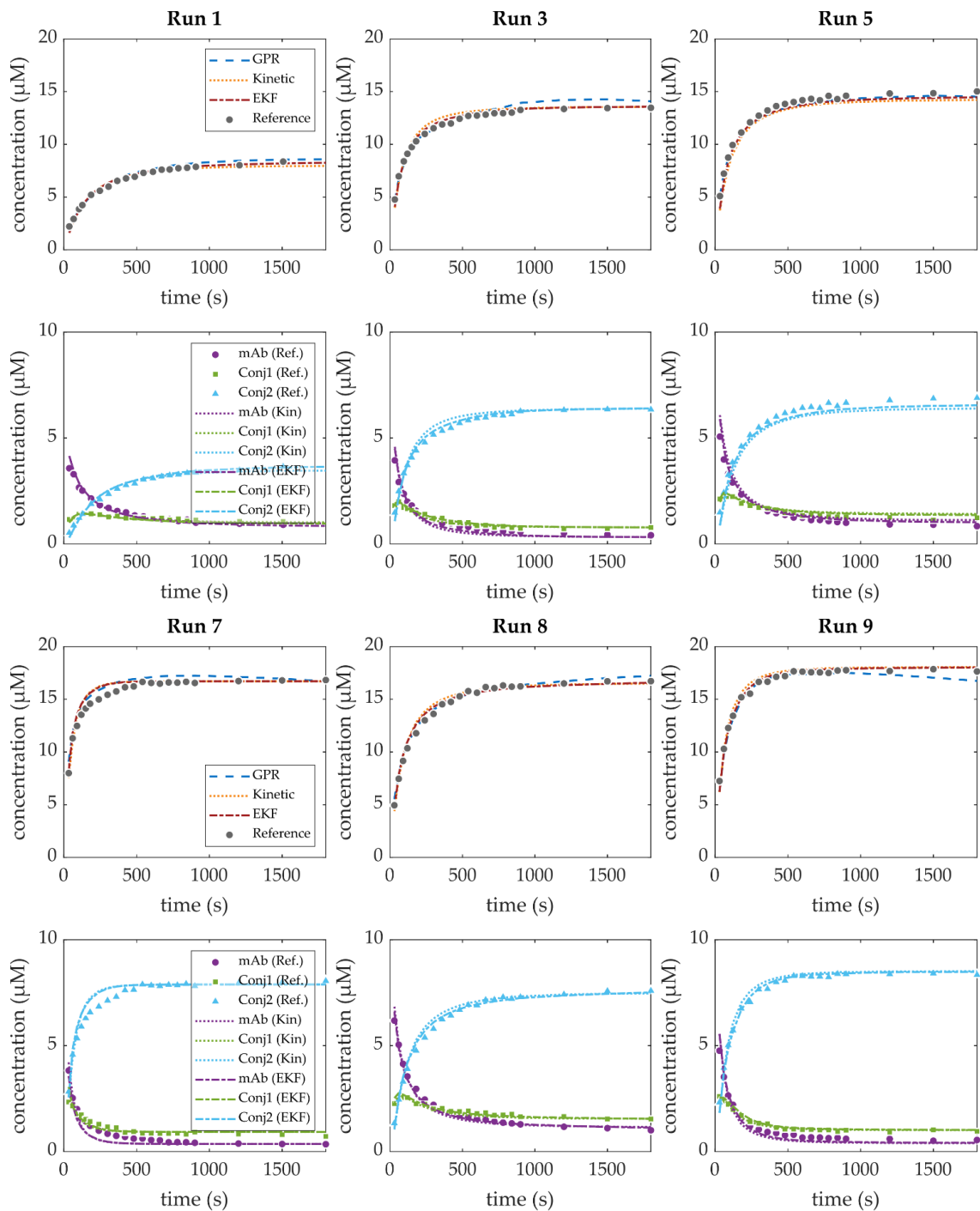
Figure S 23 Model and EKF predictions for all training runs. For each run, the first row shows the predictions of conjugated drug concentration and the second row the ADC species concentration.
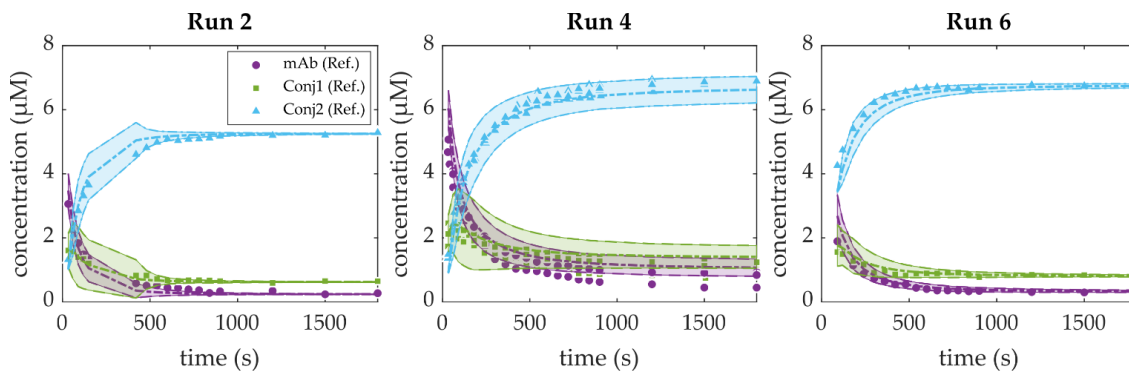
Figure S 24 EKF predictions and uncertainty estimates for all test runs derived from the state covariance matrix. The uncertainty bands were estimated from $\pm 2\sigma_{x(t)}$ for the corresponding states in the EKF state space. For Run 4, additional reference data from replicate runs, which were not used otherwise within this study, were added to compare the experimental and soft-sensor variance.
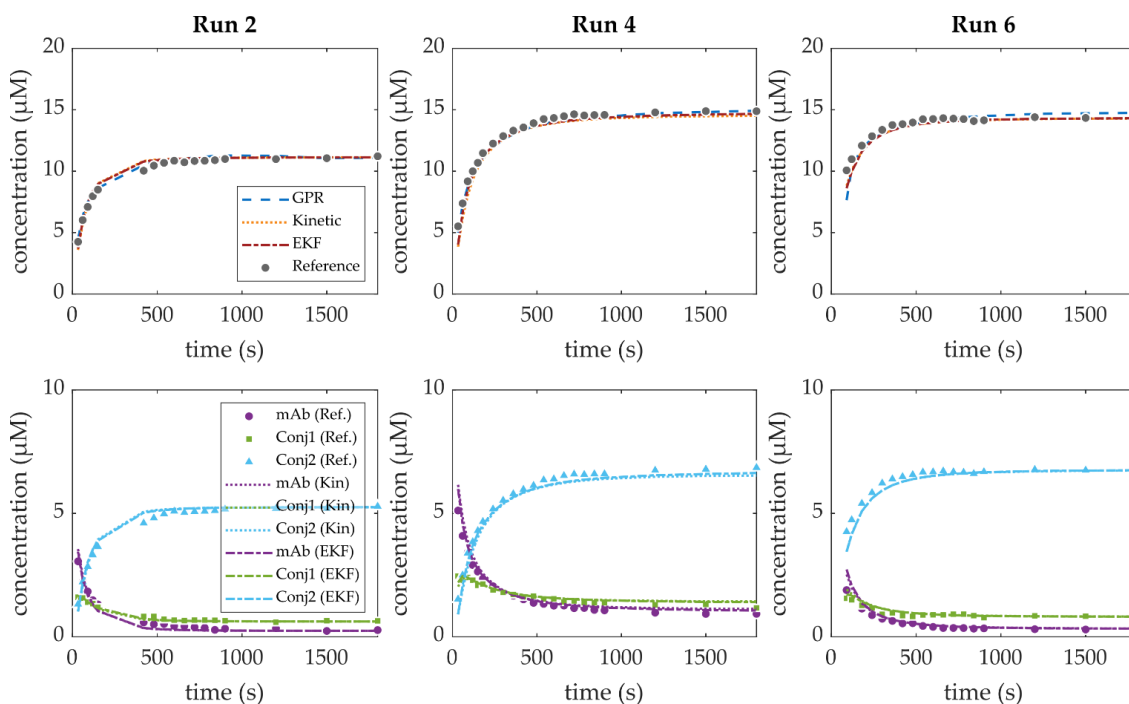

Figure S 25 Model and EKF predictions for all test runs. For each run, the first row shows the predictions of conjugated drug concentration and the second row the ADC species concentration.