

Aus dem Institut für Molekular- und Zellbiologie der Hochschule Mannheim

Leiter: Prof. Dr. rer. nat. Mathias Hafner

Mechanistic modeling and scale-up investigation of a preparative two-column purification step for a biopharmaceutical polypeptide

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)

der

Medizinischen Fakultät Mannheim

der Ruprecht-Karls-Universität

zu

Heidelberg

vorgelegt von

Jonas Martin Maria Koch

aus

Achern

2023

Dekan: Prof. Dr. med. Sergij Goerd

Referent: Prof. Dr. rer. nat. Mathias Hafner

TABLE OF CONTENTS

	Page
LIST OF ABBREVIATIONS	1
LIST OF SYMBOLS.....	2
GREEK LETTERS	6
1 INTRODUCTION.....	8
1.1 The role of biopharmaceuticals in modern medicine	8
1.2 Purification of biopharmaceuticals in the pharmaceutical industry.....	10
1.2.1 Liquid chromatography in downstream processing.....	11
1.2.2 Mechanistic modeling of non-linear range liquid chromatography ...	13
1.3 Protein-Ligand interactions in modern medicine.....	14
1.4 Theoretical considerations.....	16
1.4.1 Model derivation for ion exchange chromatography	16
1.4.2 Model derivation for reversed-phase chromatography	23
1.5 Aim of the work.....	30
2 MATERIALS AND METHODS.....	31
2.1 Chemicals.....	31
2.2 Electronic equipment.....	31
2.3 Computer software	32
2.4 Preparative chromatography resins and columns	33
2.5 Analytical chromatography resins and columns	34
2.6 Miscellaneous items	34
2.7 Polypeptides and sample preparation	35
2.8 Chromatographic experiments	35
2.8.1 CEX resin column experiments	35
2.8.2 RP column experiments.....	37
2.8.3 RP-HPLC experiments	37
2.9 Column characterization.....	37
2.9.1 Determination of porosities	37
2.9.2 Determination of the HETP, asymmetry factors, and plate numbers	38

2.9.3	Determination of polypeptide exclusion factors	39
2.9.4	Determination of ionic capacities	39
2.9.5	Determination of axial dispersion coefficients.....	40
2.10	Modeling and simulation.....	40
2.10.1	Model calibration and simulation with CEX column 1	40
2.10.2	Modeling and simulation with CEX columns 2-4.....	42
2.10.3	Model calibration of the RP column	43
3	RESULTS.....	44
3.1	Column characterization.....	44
3.1.1	CEX resin columns	44
3.1.2	RP resin column	48
3.2	CEX model development.....	50
3.2.1	CEX model calibration at small scale (CEX column 1)	50
3.2.2	CEX model calibration at small scale (TOYOPEARL S-650S)	74
3.2.3	Scale-up investigation of the CEX column 1 and the CEX model....	81
3.3	RP model development	90
3.3.1	RP model calibration at small scale (RP column 1)	90
3.4	CEX model and RP model coupled simulations sampling study	98
4	DISCUSSION.....	101
4.1	Column characterization.....	101
4.1.1	CEX resin columns	101
4.1.2	RP resin column	103
4.2	CEX model development.....	105
4.2.1	CEX model calibration at small scale (CEX column 1)	105
4.2.2	CEX model calibration at small scale (TOYOPEARL S-650S)	114
4.2.3	Scale-up investigation of the CEX column 1 and the CEX model..	117
4.3	RP model development	123
4.3.1	RP model calibration at small scale (RP column 1)	123
4.4	CEX model and RP model coupled simulations sampling study	128
5	CONCLUSION	130
6	ABSTRACT	133
7	REFERENCES.....	134
8	APPENDIX.....	146

8.1	Calculation of the column model.....	146
8.2	Calculation of the SAS isotherm.....	150
8.3	CEX column 1 modeling and simulation	154
8.4	CEX column scale-up modeling.....	156
8.5	CEX model and RP model coupled simulations sampling study	157
9	CURRICULUM VITAE	159
10	ACKNOWLEDGEMENT	163

LIST OF ABBREVIATIONS**AC**, Affinity chromatography**AU**, Absorbance units**CEX**, Cation exchange**CHO**, Chinese Ovary Cells**COVID-19**, Coronavirus disease 2019**CPA**, Colloidal particle adsorption**CV**, Column volume**DNA**, Deoxyribonucleic acid**DoE**, Design of Experiments**ECV**, Extra column volume**EMA**, European Medicines Agency**EMG**, Exponentially modified gaussian**FDA**, Food and Drug Administration**FPLC**, Fast protein liquid chromatography**HCAb**, Heavy Chain-only Antibody**HCP**, Host-cell protein**HIC**, Hydrophobic interaction chromatography**HPLC**, High-performance liquid chromatography**HTS**, High-throughput screening**IEC**, Ion-exchange chromatography**ITC**, Isothermal titration calorimetry**LGE**, Linear gradient elution**mAb**, Monoclonal Antibody**MMC**, Mixed-mode chromatography**mRNA**, messenger ribonucleic acid**miRNA**, small interfering ribonucleic acid**NaCl**, Sodium chloride**NaOH**, Sodium hydroxide**ODE**, Ordinary differential equation**PDE**, Partial differential equation**PEG**, Polyethylene glycol**pKa**, Acid dissociation constant**pH**, Potentia Hydrogenii**QbD**, Quality by Design**RNA**, Ribonucleic acid**RNase**, Ribonuclease**RP**, Reversed-phase**RPC**, Reversed-phase chromatography**SAS**, Self-association**sdAbs**, single-domain antibodies**SDM**, Scale down model**siRNA**, micro ribonucleic acid**SMA**, Steric mass action**UV**, Ultraviolet

LIST OF SYMBOLS

1	Index for protein monomer layer
2	Index for protein dimer layer
\hat{a}	Activity for the adsorbed species
a	Activity for the solute species
A_c	Column cross-sectional area
A_i	Slope of the isotherm
A_s	Asymmetry factor
$b_{\varepsilon_{o,j,k}}$	Factor of the concentric annular column in the entire scale-up column
B	Parameter from the Debye-Hückel theory
c	Molar density
c_{NaOH}	Sodium hydroxide concentration
C_f	Compression factor
c_{tot}	Total molarity of the mobile phase
$c_{s/i}$	Mobile phase counterion/protein concentration
$c_{i,in/out}$	Liquid concentrations of species i accessing/leaving a section n
$c_{s,elu}$	Counterion concentration at the elution point of the protein peak
$c_{s,init}$	Initial salt concentration
$c_{salt,initial}$	Initial salt gradient concentration
$c_{salt,final}$	Final salt gradient concentration
$\Delta G_{i/s}^0$	Practical Gibbs energy change of association of the protein/salt species
d_p	Particle diameter
D	Index for dimer layer
D_{ax}	Axial dispersion coefficient

E	Binary interaction parameter
F	Faraday number
$GH_{salt/M}$	Normalized gradient slope for the salt/organic modifier gradient
H	Index for reversed-phase model
$HETP$	Height equivalent of a theoretical plate
I	Ionic strength
i	Index for protein species i
j	Index for columns with the same resin and different dimensions
j_{ax}	Diffusive flux in axial direction
k_B	Boltzmann constant
k_{eff}	Effective mass transfer coefficient
k_d	Exclusion factor of the protein
$K_{D,i}$	Thermodynamic equilibrium constant for the dimerization process
K_i	Thermodynamic equilibrium constant for the monomer process
K_p	Protein dependent parameter
K_s	Salt-dependent parameter
L	Ligand
L_c	Column length
L_{in}	Convective flow in the segment
M	Organic modifier, Index for organic modifier
n	Any section of the column in the finite difference axial dispersion model
N	Plate number of the column
N_A	Avogadro number
$N_{acidic/basic}$	Number of acidic/basic amino acid-like group
p	The number of concentric annular columns

P	Protein
pKa_{acidic}	pka value of the acidic amino acid-like group
pKa_{basic}	pka value of the basic amino acid-like group
q_i^*	Overall adsorbed protein concentration in the accessible pore volume
$q_{s/i}$	Adsorbed counterion/protein concentration
$Q_{i/s}$	Dimensionless variable for the solid phase protein/salt concentration
r_i	Rate of loss for the protein species
R	Universal gas constant
S	Counterion
s	Index for counterion species
T	Absolute temperature
u_{int}	Interstitial mobile phase velocity
V	Volume of the section in the finite difference axial dispersion model
V_0	Retention volume Blue Dextran peak
V_g	Gradient volume
V_{NaOH}	Applied sodium hydroxide titrant volume
V_R	Retention volume of the protein pulse
V_{salt}	Retention volume salt peak
W	Index for water
x	Column axial coordinate
$x_{M,elu}$	Eluting organic modifier concentration
$x_{H,i/i/s/M}$	Mole fractions for the solute protein/salt/organic modifier species
y	Mole fractions for the adsorbed species
$\dot{y}_{i/s}$	Dimensionless variable for the liquid phase protein/salt concentration
z_i	Charge of the protein species

LIST OF SYMBOLS

z_L	Charge of the ligand
z_S	Charge of the counterion

GREEK LETTERS

α	Number of total protein species
α_H	Parameter in the Wilson equation
β_i	Hydrophobic interaction parameter for the protein species i
β_H	Parameter in the Wilson equation
δ	Parameter in the Wilson equation
ε	Permittivity of the solvent
ε_0	Interstitial porosity
ε_F	Permittivity of the free space
ε_p	Particle porosity
ε_r	Relative permittivity
ε_t	Total porosity
ζ_i	The displacement of the modulator molecule (organic modifier)
η_i	Protein/salt-specific parameter
$\hat{\gamma}$	Activity coefficient for the adsorbed species
$\tilde{\gamma}$	Asymmetric activity coefficient for the adsorbed species
γ	Activity coefficient for the solute species
θ	Available ligands involved in the formation of the association complex
θ_H	Parameter in the Wilson equation
θ_{p_i}	Protein/salt-specific parameter
κ	The inverse of the Debye length
Λ	Ligand density
Λ_{pack}	Ligand density based on packed bed
Λ_{pore}	Ligand density based on pore volume

$\Lambda_{sediment}$	Ligand density based on resin sediment
φ	Parameter in the Wilson equation
$\mu_{EMG,salt}$	First moment of the salt peak
$\mu_{EMG,Dextran}$	First moment of the Blue Dextran peak
$\hat{\mu}_{i/s}^0$	Reference potential for the bound protein/salt species
$\mu_{i/s}$	Reference potential for the solute protein/salt species
μ_{s-i}	Salting in term of the effect from the modulator salt on the protein activity
μ_{s-o}	Salting out term of the effect from the modulator salt on the protein activity
ϑ_i	The displacement of the protein species
ν_i	Characteristic charge of the protein species i
ξ	Shielding factor
σ	Charge ratio of the ligand and the counterion
$\sigma_{EMG,salt}$	Standard deviation of the salt peak
$\sigma_{EMG,salt}^2$	Second moment of the salt peak
$\sigma_{EMG,Dextran}^2$	Second moment of the Blue Dextran peak
τ	Number of proteins in solution
τ_p	Protein/salt-specific parameter
χ	Parameter in the Wilson equation
ψ	Soluble ligands
ω	Number of proteins bound to the resin
ω_H	Parameter in the Wilson equation

1 INTRODUCTION

1.1 The role of biopharmaceuticals in modern medicine

Biopharmaceuticals are mostly therapeutic recombinant proteins produced by using biotechnological techniques ¹ and represent some of the best accomplishments of modern science ², especially in modern medicine. According to the Food and Drug Administration (FDA), biotechnology is defined as “the application of biological systems and organisms to technical and industrial processes” ³. The first commercially produced biopharmaceutical using biotechnological techniques was insulin in 1982. Irving Johnson and colleagues at Eli Lilly established the recombinant DNA technology and produced human insulin in *E. coli* ⁴. The biotechnological production of insulin was a significant advance in the treatment of diabetes, as the standard treatment at the time was the injection of relatively crude insulin derived from the pancreas of pigs and cattle ⁵. This milestone was the starting point for the industrial production of therapeutic recombinant proteins in modern medicine. Nowadays, biopharmaceuticals constitute a significant part of all medical drugs, and the number is constantly increasing. Recombinant proteins as active pharmaceutical ingredients represent the majority of commercially available biopharmaceuticals ². Antibodies represent the major class of biopharmaceutical products ⁶ and continue to dominate biopharmaceutical approvals ⁷. Currently, over 100 antibody-based therapeutics are approved for treating various diseases ⁸. Antibodies cover a wide range of therapeutic usage. They are mainly used to treat cancer, autoimmunity, and chronic inflammatory diseases, with a broader extension to infectious, hematologic, neurological, ophthalmological, metabolic, and musculoskeletal diseases ⁹.

Over the last decades, numerous protein expression platforms have been developed and implemented for the biotechnological production of protein-based biopharmaceuticals. Protein-based biopharmaceuticals can be expressed in yeast cells ¹⁰, bacterial cells ⁵, mammalian cells ¹¹, insect cells ¹², and algae cells ¹³. Each expression system has its advantages and disadvantages. For example, procaryotic systems are mainly used for the industrial production of proteins with less complex structures or peptides due to insufficient protein secretion and the lack of the ability for complex posttranslational modifications. For proteins with a more complex structure, such as monoclonal Antibodies (mAbs), eukaryotic expression systems, for example, Chinese Ovary Cells

(CHO), were used ¹⁴. However, there is an aspiration to express mAbs in plant-based systems since mammalian cells require high production costs and high initial investments ⁶.

Besides the broadly applicable mAbs, there have been numerous efforts in recent years to establish new classes of molecules as therapeutic drugs to aim for new biological targets or existing targets in a new way. One class is single-domain antibodies (sdAbs) (e.g. Nanobody[®] molecules). sdAbs originate in camels and sharks, which produce heavy chain-only antibodies (HCAbs). The sdAb is then the isolated variable domain of the HCAb ¹⁵. sdAbs are resilient molecules with excellent target affinity and specificity ¹⁵. Furthermore, the expression of recombinant sdAbs takes place in the periplasm of microbial cells and is, therefore, not as costly as the production of conventional mAbs in eucaryotic cells. Several companies have sdAb-based therapeutic drugs in the pipeline, and some clinical trials are already ongoing ¹⁵. A type of sdAb, the Nanobody[®] molecules, were initially developed by the company Ablynx, which Sanofi acquired in 2018 ¹⁶. Later on, the first ever Nanobody[®] therapeutic drug Caplacizumab (Cabliivi[®] from Ablynx), which prevents a rare blood-clotting disorder, was approved by the FDA and the European Medicines Agency (EMA) in 2018/2019 ¹⁵.

Another recent class of therapeutic protein molecules are the so-called Synthorin[™] molecules. Synthorin[™] molecules are proteins expressed with non-natural amino acids to gain new biological activity and specificity ¹⁷. Protein expression is performed with a semi-synthetic organism based on an *E. coli* strain and through an expanded genetic alphabet with a novel DNA base pair ¹⁷. Synthorx Inc. developed the Synthorin[™] platform, and Sanofi acquired it in 2019 ¹⁸. Currently, the leading compound from the Synthorin[™] platform THOR-707, a PEGylated recombinant human IL-2 molecule intended for the treatment of a variety of solid tumors, is tested in phase I/II clinical trials ^{17, 19}.

Another famous – at least since the COVID-19 pandemic – and promising new therapeutic drug class are mRNA molecules (oligonucleotides) ²⁰. mRNA-based therapeutics are based on the approach that mRNA molecules containing the sequence for a therapeutic protein are transfected into cells, where the mRNA is then translated ²¹. However, the transfection of mRNA into cells is challenging since the negatively charged mRNA cannot readily pass the anionic cell membrane, and the mRNA is

susceptible to ubiquitous RNases. Furthermore, it is also possible that mRNA induces an undesired immune response²¹. However, to overcome these hurdles, nanoparticle-based delivery systems are used²². Currently, several mRNA-based therapeutics as vaccines against infectious diseases and cancer and as therapeutics against genetic and metabolic disorders are under pre-clinical and clinical trials²¹. The first approved and the most famous mRNA-based biopharmaceuticals were the mRNA-based vaccines against COVID-19 in 2020/2021 (Comirnaty® from BioNTech/Pfizer and Spikevax® from Moderna). These vaccines probably saved millions of lives during the COVID-19 pandemic²³ and demonstrated biopharmaceuticals' significant role in modern medicine and beyond.

Many other already established molecule classes are being explored to overcome existing challenges. For example, peptide-based therapeutics are highly potent. However, peptide-based therapeutics often have poor plasma stability and membrane permeability²⁴. Therefore, a new generation of peptide-derived molecules as therapeutics is explored, for example, the so-called peptidomimetics, cyclic peptides, hydrocarbon-stapled peptides, and hybrid peptide macrocycles²⁴. Furthermore, in the class of oligonucleotides, miRNA and siRNA molecules, which interfere during the translation process of the cell, are explored as new therapeutics²⁴.

To summarize, our ever-growing understanding of biological systems in modern medicine constantly provides new biological targets and therapeutic approaches for treating diseases that are currently difficult or impossible to treat²⁴. Therefore, biopharmaceuticals will continue to play a significant role in modern medicine.

1.2 Purification of biopharmaceuticals in the pharmaceutical industry

The manufacturing of biopharmaceuticals is divided into upstream and downstream processes. Cell growth and the formation of the substrate to the desired product contribute to the upstream process. At the same time, the downstream process includes all steps that contribute to the purification of the product. The aim during the purification step is to reduce the process volume, decrease contaminants, and increase the purity of the desired product²⁵. In general, a downstream process for biopharmaceuticals consists of four sections: preparation, which includes the cell harvest and disruption; capture, which consists of a first initial purification for cell debris removal and product

extraction; base purification, for product renaturation and concentration; and polishing, for the final purification ²⁶. A downstream process usually consists of a sequence of different unit operations, such as centrifugation, microfiltration, ultrafiltration, chromatography, precipitation, distillation, or diafiltration ¹. Downstream process design significantly impacts the overall manufacturing cost since product-related and process-related impurities in administered therapeutics may seriously affect clinical safety and efficacy ²⁵. Meeting the mandatory standards in protein purification during downstream processing can be challenging. Efficient recovery and purification of the desired product are essential to the production process. Therefore, more and more computer-aided tools, such as mathematical model-based approaches for downstream process development, are commonly used in the industry ²⁷. There are three categories of mathematical models: Statistical models, mechanistic models, and hybrid models ²⁸. Statistical models, for example, the Design of Experiments (DoE) approach, are based on statistical data and are widely used in process development. In contrast, the increasingly popular mechanistic models consider physicochemical fundamentals with comparable parameters. Therefore, an increased process understanding can be achieved relative to statistical models. Hybrid models combine statistical and mechanistic models and benefit from both advantages. Since the introduction of the Quality by Design (QbD) concept by the FDA, increased attention has been dedicated to process understanding using model-based approaches, especially to mechanistic models ²⁹.

1.2.1 Liquid chromatography in downstream processing

One of the key unit operations in the downstream processing of biopharmaceuticals and the favorable method for analyzing protein mixtures is liquid chromatography ³⁰ due to its high-resolution capacity ³¹. Therefore, Chromatography continues to be the backbone of downstream processing ³². Since this work focuses on liquid chromatography, especially on the mechanistic modeling of preparative liquid chromatography, the current status and new methods for process development in the field of liquid chromatography are outlined in the following.

Chromatography is based on protein-ligand interactions, whereas different ligands result in various types of chromatography. The most commonly used types of chromatography for the purification of biopharmaceuticals at an industrial scale are affinity chromatography (AC), ion-exchange chromatography (IEC), mixed-mode chromatography (MMC), hydrophobic-interaction chromatography (HIC), and reversed-phase

chromatography (RPC)³⁰. AC with protein A resins are mainly used for an initial capture step in antibody purification³³. IEC is commonly employed as a polishing step to separate product-related impurities and to remove host cell protein (HCP) and DNA in mAb production^{34, 35}. MMC is often used as a capture step^{36, 37} or polishing step^{38, 39}. HIC has been utilized in mAb⁴⁰ and plasmid DNA⁴¹ purification. Like HIC, RPC is based on hydrophobic interactions. However, RPC is mainly used to purify small proteins and peptides, which tolerate harsher conditions due to strong hydrophobic interactions with the resin⁴². Each type of chromatography has different advantages and disadvantages for each molecule class and can therefore be used to target a specific biopharmaceutical molecule for purification.

Nowadays, mini-columns and columns in a high-throughput format are commonly used for the initial development and optimization of chromatographic process steps since they allow a fast and material-saving investigation. High throughput columns are available in a 96-well plate format, so different resins or conditions can be tested in parallel within a short period⁴³. The data obtained with high-throughput screening (HTS) methods can then be analyzed using the DoE approach to describe the influence of process parameters statistically⁴⁴. Alternatively, the data can be described using a mechanistic model⁴⁵⁻⁴⁷. After initial process development at small column scale, the findings are transferred to scale down model (SDM) scale. A SDM represents the chromatographic step in the process scale and can be used for product quality prediction and extrapolating operating conditions^{48 49}. SDM columns typically have the same bed height as the process scale column but smaller diameters to save materials and keep the residence time constant. The flow rate in different scales is then changed in proportion to the column's cross-sectional area, which is also known as the linear scale-up method in liquid chromatography⁵⁰⁻⁵². The concept of timescales is referred to the residence time and is more flexible since it allows scale-up with variable column lengths by keeping the residence time constant. Despite different column lengths, similar separation performance can be achieved since the separation performance (number of theoretical plates) of the columns is kept constant with constant residence time within a specific limit^{50, 52, 53}. The concept of timescales seems to have found successful application in the industry over the last decade⁵². However, the scale-up of chromatographic processes is still challenging due to column packing quality inconsistency resulting from different pressure drops and packing pressure across different column scales^{52, 54}, which significantly influences the separation performance and the column

characteristics⁵⁵. Nevertheless, column inconsistencies across different scales are often neglected in scaling up chromatographic processes, and the same column parameters and packing qualities are assumed^{45, 47, 56-60}. Furthermore, packing inhomogeneities within a column can result in a radial packing density profile, which in turn results in a radial velocity profile that can significantly affect the elution profile of the target molecule. Several publications have shown that radial packing inhomogeneities can strongly influence the peak shape in the linear and non-linear range⁶¹⁻⁷². Considering all these influences during scale-up is challenging and the implementation in a mechanistic model is even more complicated. Therefore, when developing a mechanistic model suitable for scale-up, the effects of packing quality within different column scales and radial packing inhomogeneities within a column scale should be kept in mind.

1.2.2 Mechanistic modeling of non-linear range liquid chromatography

Since the FDA is placing more and more emphasis on model-based process understanding, there has been increased emphasis on implementing mechanistic models for liquid chromatography steps²⁹. In the biopharmaceutical industry, non-linear chromatography (preparative chromatography) is typically performed since acceptable space-time yields are required. However, the mechanistic description of preparative liquid chromatography is still challenging⁷³. In preparative chromatography, unusual elution behavior often is observed, which can have different causes. For example, oligomer formation of the product on the resin^{74, 75}, complex pH-dependent intraparticle diffusion effects⁷⁶, or resin surface-induced conformational changes of the protein⁷⁷ result in peak shoulders and unexpected peak shapes of the protein elution peak. Furthermore, non-ideal thermodynamic conditions may require activity coefficients for the mechanistic description of preparative chromatography⁷⁸. Nevertheless, a large number of literature is available where high loading situations in liquid chromatography have been described using model-based approaches. Creasy et al. used an empirical interpolation method for the description of high loading situations in IEX with a mAb⁷⁹⁻⁸¹. However, an empirical isotherm does not give any physiochemical insight into the adsorption process. Huuk et al. observed complex peak shapes with shoulders in high loading situations and applied the steric mass action (SMA) isotherm⁸² extended with an asymmetrical activity coefficient for the description of thermodynamic non-idealities introduced by Mollerup⁸³. Diedrich et al. also observed a complex high loading elution

behavior and successfully predicted the elution profiles by implementing a multi-state SMA isotherm, which describes multiple binding orientations of the protein ⁸⁴. Briskot et al. applied the non-stoichiometric colloidal particle adsorption (CPA) model, which considers protein-protein interactions in the bound state, for the description of high loading situations ⁸⁵⁻⁸⁷. Mollerup applied the self-association (SAS) isotherm to predict an anti-Langmuirian elution behavior in the non-linear range ⁸⁸. Khalaf et al. observed a complex anti-Langmuirian to Langmuirian elution behavior with increasing load on a polyelectrolyte brush type CEX resin and successfully described the elution behavior using an isotherm, which accounts for multilayer formation ⁸⁹. This selection of literature gives an insight into the research efforts being made for the model-based description of preparative chromatography and the observed complex elution behaviors. In general, the goal in modeling a complex elution behavior is to first understand what is going on in the system and then try to find a mechanistic model that is able to describe these mechanisms.

1.3 Protein-Ligand interactions in modern medicine

Despite liquid chromatography during downstream processing or analytical processes, protein-ligand interactions are a fundamental feature in most biological processes and ubiquitous in modern medicine ⁹⁰. Proteins play a central role in biological systems. They are part of an extensive interaction network with other proteins, small molecules, or nucleotides ⁹¹. The function of a protein molecule inside a biological system involves dynamic binding or unbinding, competing, or complex formation with other molecules during immune recognition, cellular metabolism, regulation of gene expression, and cell signaling processes ⁹⁰, whereas dysfunction often results in disease ⁹¹. Therefore, the ability to understand the detailed mechanism of a protein-ligand binding behavior is a crucial part ⁹⁰. One method to explore and understand the function of proteins inside a biological system is to measure and analyze their interaction with other molecules ⁹¹. This is the point where modern drug design becomes applicable. Biopharmaceutical drugs (primarily proteins) are designed to interact with a target molecule (ligand) and either enhance or inhibit its function. Modern drug design is based on computer-aided approaches, where various structure-based models screen drug substances on their potential pharmacology affinity, selectivity, and stability ⁹². Recent

literature indicates that model-based approaches play an essential role in the development of drugs⁹²⁻⁹⁸.

Protein-ligand interactions in modern drug design are modeled using mechanistic binding isotherms to obtain information about the binding enthalpy and stoichiometry. A standard method to obtain isothermal binding parameters is the isothermal titration calorimetry (ITC)^{99, 100}. With this method, thermodynamically-based drug design is possible, whereby the enthalpy change based on the resulting heat during the binding process of a protein-ligand complex is measured¹⁰¹. These measured thermodynamic parameters can be used in model-based approaches to screen potential affinity, selectivity, and stability to other ligands.

These few examples show that protein-ligand interactions play a significant role in modern drug design. Furthermore, computer-aided approaches combined with thermodynamic binding models derived from experimentally determined data or structure-based models are commonly used for designing new biopharmaceutical drugs.

In the end, mechanistic modeling in the field of liquid chromatography and the field of modern drug design has a large number of intersections. The mechanistic binding models developed with chromatographic methods could be transferred and used in modern drug design approaches, as the models describe mechanistic interactions of protein-ligand or protein-protein complexes.

1.4 Theoretical considerations

This work focuses on the mechanistic description of the chromatographic elution of a polypeptide in ion exchange and reversed-phase chromatography. The following sections show the model derivation for each chromatography type.

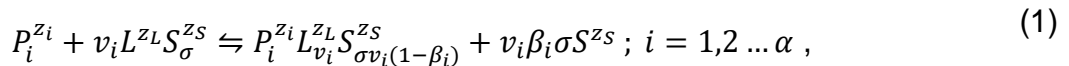
1.4.1 Model derivation for ion exchange chromatography

This section describes the detailed derivation of the mechanistic model used for the description of IEC in this work.

1.4.1.1 Association schemes and equilibrium constants

The adsorption isotherm used in this work is an extended version of the SAS isotherm derived by Mollerup^{88, 102}. The SAS isotherm describes the additional association of a protein molecule with an already ligand-adsorbed protein molecule. Thus, with the SAS isotherm, mechanistic modeling of oligomerization processes in liquid chromatography can be done.

The overall association scheme of proteins that associate with immobilized ligands is given by¹⁰²:



where a protein P of species i with the charge z_i , associates with v_i ligands L carrying the charge z_L by displacing $v_i \beta_i$ counterions S carrying the charge z_S . α denotes the number of total protein species. In pure ion exchange chromatography without hydrophobic interactions $\beta_i = 1$ is valid. The charge ratio of the ligand and the counterion is given by $\sigma = z_L/z_S$. In our case, the ratio is equal to one since the ligand and the counterion charge are equal to one. The stoichiometric coefficient also referred to as characteristic charge, is given by $v_i = z_i/z_S$. Therefore, the association scheme can be simplified to:



Consequently, the thermodynamic equilibrium constant for the monomer layer is given by⁷⁴:

$$K_i = \frac{\hat{a}_{1,i} a_s^{v_i}}{a_i \hat{a}_s^{v_i}} = \frac{y_{1,i} x_s^{v_i} \hat{\gamma}_{1,i} \gamma_s^{v_i}}{x_i y_s^{v_i} \gamma_i \hat{\gamma}_s^{v_i}} = \frac{q_{1,i} (c_s)^{v_i} \hat{\gamma}_{1,i} (\gamma_s)^{v_i}}{c_i (q_s)^{v_i} \gamma_i (\hat{\gamma}_s)^{v_i}}, \quad (3)$$

where $\hat{a}_{1,i}$ and \hat{a}_s are the activities of the adsorbed protein and counterion species, while a_i and a_s are the activities of the species in solution. $y_{1,i}$ and $\hat{\gamma}_{1,i}$ or else y_s and $\hat{\gamma}_s$ are the corresponding mole fractions and activity coefficients for the respective adsorbed protein and counterion species. Furthermore, x_i and γ_i or else x_s and γ_s are the corresponding mole fractions and activity coefficients for the protein and counterion species in solution. $q_{1,i}$ denotes the adsorbed and c_i the mobile phase protein concentration, which is given by $q_{1,i} = y_i c$ and $c_i = x_i c$ with the molar density c . Moreover, c_s denotes the mobile phase and q_s the adsorbed concentration of the counterion, which is given by $c_s = x_s c$ and $q_s = y_s c$. Note that q_s is equal to the concentration of the ligands available for binding.

It is known that the characteristic charge v_i depends on the protein charge and, thus, on the pH. The relationship to describe the pH dependency of the characteristic charge is given by ¹⁰³⁻¹⁰⁵:

$$v_i = -\frac{N_{acidic}}{1 + 10^{pKa_{acidic} - pH}} + \frac{N_{basic}}{1 + 10^{pH - pKa_{basic}}}, \quad (4)$$

where N_{acidic} and N_{basic} denotes the number of the acidic and basic amino acid-like groups involved in binding with the corresponding pKa value pKa_{acidic} and pKa_{basic} . Since the pH range in this work is relatively small, from pH 3.3 to pH 4.3, only one acidic and one basic amino acid-like group are necessary.

The pH dependence of the equilibrium constant is defined according to Mollerup's thermodynamic model using the characteristic charge v_i ⁸³:

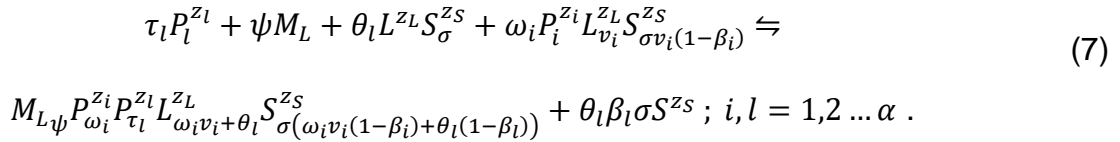
$$RT \ln K_i = -(\hat{\mu}_i^0 - \mu_i^0) + v_i(\hat{\mu}_s^0 - \mu_s^0) = -\Delta G_i^0 + v_i \Delta G_s^0, \quad (5)$$

where $\hat{\mu}_i^0$ and μ_i^0 are the reference potentials of the bound protein and solute protein species, respectively. $\hat{\mu}_s^0$ and μ_s^0 are the reference potentials of the counterion. ΔG_i^0 and ΔG_s^0 are the differences in the reference potentials when the protein or the counterion changes its state from bound to solute.

Assuming that the adsorbed species behave ideally ($\hat{\gamma}_{1,i} = 1$ and $\hat{\gamma}_s = 1$), and by replacing the activity coefficients of the species in solution with asymmetric activity coefficients $\tilde{\gamma}_{i,s}$, the isotherm for the monomer layer can be defined as:

$$\frac{q_{1,i}}{c_i} = K_i \left(\frac{q_s}{c_s} \right)^{v_i} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}}. \quad (6)$$

The overall association scheme of proteins that associate with immobilized proteins is given by ¹⁰²:



With one protein in solution ($\tau_l = 1$), no soluble ligands ($\psi = 0$), only one protein species that is bound to the resin ligand ($\omega_i = 1$), dimer formation is by self-association with one protein species ($i = l$), assuming that no available ligands are involved in the formation of the association complex ($\theta_l = 0$), and the assumptions made above, the association scheme can be simplified to:



Consequently, the thermodynamic equilibrium constant for the dimerization process is given by ⁷⁴:

$$K_{D,i} = \frac{\hat{a}_{2,i}}{a_i \hat{a}_{1,i}} = \frac{y_{2,i}}{x_i y_{1,i}} \frac{\hat{\gamma}_{2,i}}{\gamma_i \hat{\gamma}_{1,i}} = \frac{q_{2,i} c}{c_i q_{1,i}} \frac{\hat{\gamma}_{2,i}}{\gamma_i \hat{\gamma}_{1,i}}, \quad (9)$$

where $\hat{a}_{2,i}$ denotes the activity of the adsorbed protein dimer layer with the corresponding mole fraction and the activity coefficient $y_{2,i}$ and $\hat{\gamma}_{2,i}$. The adsorbed protein dimer layer concentration is given by $q_{2,i} = y_{2,i} c$. Assuming that the adsorbed dimer species behaves ideally ($\hat{\gamma}_{2,i} = 1$), and the assumptions made above, the isotherm for the dimer layer is defined as:

$$\frac{q_{2,i}}{c_i} = q_{1,i} \frac{K_{D,i}}{c} \tilde{\gamma}_i. \quad (10)$$

Since it is practically impossible to determine the monomer and dimer layer concentration individually in the overall association process, the monomer and dimer layer concentration is replaced by the total protein concentration q_i given by the following relationship ⁸⁸:

$$q_i = q_{1,i} + 2q_{2,i}. \quad (11)$$

Consequently, the overall isotherm for the described protein dimerization process in ion exchange chromatography is given by:

$$\frac{q_i}{c_i} = K_i \left(\frac{q_s}{c_s} \right)^{v_i} \left(\frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} + 2 \frac{K_{D,i}}{c} c_i \frac{\tilde{\gamma}_i^2}{\tilde{\gamma}_s^{v_i}} \right). \quad (12)$$

1.4.1.2 Material balance

The material balance of the ligands available for binding is defined as ⁷⁴:

$$q_s = \Lambda - \sum_{i=1}^{\alpha} [(v_i + \xi_i)q_{1,i} - (v_i + \xi_{D,i})q_{2,i}]. \quad (13)$$

With $q_{1,i}^{max} = \Lambda/(v_i + \xi_i)$ and $q_{2,i}^{max} = \Lambda/(v_i + \xi_{D,i})$, in which $q_{1,2,i}^{max}$ are the maximum binding concentrations of the monomer and dimer species, ξ_i , and $\xi_{D,i}$ are the shielding factors of the monomer and dimer species, and Λ denotes the ligand density.

Assuming that $(v_i + \xi_{D,i}) = 2(v_i + \xi_i)$, the maximal dimer concentration is half the amount of the monomer concentration $q_{2,i}^{max} = q_{1,i}^{max}/2$ ⁸⁸. Thus, the material balance simplifies to:

$$q_s = \Lambda - \sum_{i=1}^{\alpha} (v_i + \xi_i)q_i. \quad (14)$$

Combining Equations (12) and (14) results in the overall isotherm:

$$\frac{q_i}{c_i} = K_i \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} \left(\frac{\Lambda}{c_s}\right)^{v_i} \left[1 - \sum_{j=1}^{\alpha} \left(\frac{v_j + \xi_j}{\Lambda}\right) q_j\right]^{v_i} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{\gamma}_i\right). \quad (15)$$

1.4.1.3 Activity coefficients

Two activity coefficients were considered in the overall adsorption isotherm: The asymmetric activity coefficient of the salt ion and protein species in solution. Due to the assumed ideal behavior, the activity coefficients of the bound salt ion and protein species were set to one. The asymmetric activity coefficient of the salt ion species in solution is described by the Davies equation ¹⁰⁶ extended with a temperature-dependent term introduced by Debye and Hückel ¹⁰⁷:

$$\log \tilde{\gamma}_s = -B \left(\frac{\sqrt{I}}{1 + \sqrt{I}} - 0.3 \cdot I \right), \quad (16)$$

$$B = 1.82 \cdot 10^6 \cdot (\varepsilon T)^{-\frac{3}{2}}, \quad (17)$$

where B is a parameter adopted from the Debye-Hückel theory with the permittivity of the solvent ε and the ionic strength of the solution I . Since sodium chloride was used as the mobile phase modulator in this work, the ionic strength is equal to the molar concentration of the salt. Note that the influence of the constant concentration of the short aliphatic alcohol was considered by applying the appropriate dielectric constant of the water/alcohol mixture.

Mollerup describes the asymmetric activity coefficient of the protein in solution with the following relationship ⁸³:

$$\ln \tilde{\gamma}_i = K_s c_s + K_p c_i, \quad (18)$$

where K_s and K_p are parameters in the salt and protein species' activity coefficient model. They depend on the charge of the protein and thus are pH-dependent ⁸³. This work introduced an empirical pH dependency since the model should describe the elution behavior at various pH values. The empiric pH dependency of the parameters K_s

and K_p , using the characteristic charge and the introduced salt and protein dependent constants K_s^* and K_p^* , was described with the following relationship:

$$K_s = \frac{K_s^*}{v_i^2}, \quad (19)$$

$$K_p = \frac{K_p^*}{10^{v_i}}. \quad (20)$$

1.4.1.4 Linear gradient elution experiments

Linear gradient elution (LGE) experiments were performed for the data determination in the linear and non-linear range of the isotherm. Yamamoto et al. ^{108, 109} developed a protein retention model in IEC based on counterion-induced LGE experiments to determine linear range parameters. In the linear range of the isotherm, the correlation between the normalized gradient slope GH_{salt} and the eluting salt concentration is:

$$\int_{c_{s,init}}^{c_{s,elu}} c_s^{v_i} dc_s = GH_{salt} \frac{1}{L_c} K_i \left(\frac{\Lambda}{z_s \cdot c_{s,elu}} \right)^{v_i} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} \int_0^{L_c} dx, \quad (21)$$

where L_c is the column length, and $c_{s,elu}$ is the eluting and $c_{s,init}$ the initial salt concentration. Differentiation results in the following correlation for the normalized gradient slope and the slope of the isotherm A_i ¹¹⁰:

$$\frac{dGH_{salt}}{dc_{s,elu}} = \frac{1}{k_{d,i}A_i + k_{d,i} - 1} = \left[k_{d,i}K_i \left(\frac{\Lambda}{c_{s,elu}} \right)^{v_i} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} + k_{d,i} - 1 \right]^{-1}, \quad (22)$$

where $k_{d,i}$ denotes the exclusion factor of the protein. The normalized gradient slope is defined according to Pedersen et al. ¹¹¹:

$$GH_{salt} = g_{salt} \cdot (V_c(1 - \varepsilon_0)\varepsilon_p) = \frac{c_{salt,final} - c_{salt,initial}}{V_g} \cdot (V_c(1 - \varepsilon_0)\varepsilon_p), \quad (23)$$

where V_g is the gradient volume, $c_{salt,initial}$ is the initial, $c_{salt,final}$ is the final salt gradient concentration, ε_0 is the interstitial porosity, and ε_p the intraparticle porosity.

1.4.1.5 Column simulation model

For the simulation of the protein elution, the mobile phase salt, and pH elution profiles, a lumped rate model was applied ¹¹²:

$$\frac{\partial c_i}{\partial t} + u_{int} \cdot \frac{\partial c_i}{\partial x} + \frac{(1 - \varepsilon_0)\varepsilon_p k_{d,i}}{\varepsilon_0} \cdot \frac{\partial q_i^*}{\partial t} = D_{ax} \cdot \frac{\partial^2 c_i}{\partial x^2}, \quad (24)$$

where u_{int} is the interstitial mobile phase velocity, D_{ax} denotes the axial dispersion coefficient, q_i^* is the overall adsorbed protein concentration in the accessible pore volume, and x is the axial column coordinate. For the calculation of c_s the salt was assumed to be an inert modulator, with local equilibrium between c_s and q_s^* , and with $D_{ax,s} = 0$ and $k_{d,s} = 1$. The mobile phase pH was treated like a common mobile phase modulator with the same assumption for the salt in solution. A linear driving force approximation, which considers the mass transfer in the solid phase to be dominant, was used to describe mass transfer limitations ^{112, 113}:

$$\frac{\partial q_i^*}{\partial t} = k_{eff,i} \frac{6}{d_p} (q_{eq,i}^* - q_i^*), \quad (25)$$

where $k_{eff,i}$ describes the effective mass transfer coefficient, d_p denotes the particle diameter of the resin and $q_{eq,i}^*$ denotes the hypothetical adsorbed protein concentration at equilibrium with the mobile phase. Note that $k_{eff,i}$ is a lumped rate parameter describing the transfer between q_i^* and $q_{eq,i}^*$ ¹¹². The column simulation model is discretized by a first- and/or second order central finite-difference method, as shown by Ingham et al. ¹¹⁴. The system of ordinary differential equations (ODE) was solved by applying the fourth-order Rosenbrock (stiff) method in Berkeley Madonna (V8.3.18). Model computation was done by using at least eighty increments. The adsorption equilibrium is modeled according to the SAS formalism:

$$q_{eq,i}^* = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} \left(\frac{q_s}{c_s}\right)^{v_i-1} c_i \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{Y}_i\right)}{c_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{q_s}{c_s}\right)^{v_j-1} \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} c_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} c_j \tilde{Y}_j\right) \right]} \quad (26)$$

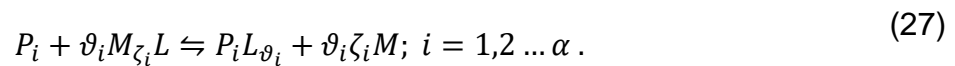
The formalism was solved using a Newton iteration, as shown by Frederiksen ¹¹⁵. A detailed description of how the column simulation model and the adsorption isotherm are calculated using the ODE solver software Berkeley Madonna (V8.3.18) is shown in Appendix 8.1 and 8.2.

1.4.2 Model derivation for reversed-phase chromatography

This section describes in detail the derivation of the mechanistic model used for the description of RPC in this work.

1.4.2.1 Association schemes and equilibrium constants

In this work, an adapted version of the HIC adsorption isotherm derived from Arkell et al. was used ¹¹⁶⁻¹¹⁹. For a detailed derivation of the adsorption model, the publications from Arkell et al. ¹¹⁶⁻¹¹⁹ are recommended. The adsorption mechanism for the hydrophobic interaction is described by the displacement of $\vartheta_i \zeta_i$ adsorbed modulator molecules M induced by the protein P_i species i . Whereas the protein forms the complex $P_i L_{\vartheta_i}$ by reversibly binding to ϑ_i ligands L . The association scheme can be written as ¹¹⁶:



The association scheme is equivalent to the scheme used for IEC (see Equation (2)) based on the SMA model. Instead of electrostatic interactions modulated by salt ions, hydrophobic interactions are modulated by an organic modifier. It is assumed that the stationary phase is saturated with the organic modifier M . Consequently, the equilibrium constant for the hydrophobic protein-ligand interaction can be written as ¹¹⁶:

$$K_{H,i} = \frac{\hat{a}_{H,i} a_M^{\vartheta_i \zeta_i}}{a_{H,i} \hat{a}_M^{\vartheta_i}} = \frac{y_{H,i} x_M^{\vartheta_i \zeta_i} \hat{\gamma}_{H,i} \gamma_M^{\vartheta_i \zeta_i}}{x_{H,i} y_M^{\vartheta_i} \gamma_{H,i} \hat{\gamma}_M^{\vartheta_i}} = \frac{q_i x_M^{\vartheta_i \zeta_i} c_{ref,m} \hat{\gamma}_{H,i} \gamma_M^{\vartheta_i \zeta_i}}{c_i y_M^{\vartheta_i} c_{ref,s} \gamma_{H,i} \hat{\gamma}_M^{\vartheta_i}}, \quad (28)$$

where $\hat{a}_{H,i}$ and \hat{a}_M are the activities of the adsorbed protein and displaced species M , while $a_{H,i}$ and a_M are the activities of the species in solution. $y_{H,i}$ and $\hat{\gamma}_{H,i}$ or else y_M and $\hat{\gamma}_M$ are the corresponding mole fractions and activity coefficients for the respective adsorbed protein species and the species M . Furthermore, $x_{H,i}$ and $\gamma_{H,i}$ or else x_M and γ_M are the corresponding mole fractions and activity coefficients for the protein species and the species M in solution. The adsorption isotherm is then given by ¹¹⁶:

$$\frac{q_i}{c_i} = K_{H,i} \left(\frac{y_M \hat{\gamma}_M}{x_M^{\zeta_i} \gamma_M^{\zeta_i}} \right)^{\vartheta_i} \frac{\gamma_{H,i}}{\hat{\gamma}_{H,i}}. \quad (29)$$

1.4.2.2 Material balance

The molar fraction of free ligands y_M can be calculated from the material balance ¹¹⁶:

$$y_M = \frac{\Lambda}{c_{tot}} \left[1 - \sum_{i=1}^{\alpha} \frac{(\vartheta_i + \xi_i) q_i}{\Lambda} \right], \quad (30)$$

where c_{tot} is the total molarity of the mobile phase. In the linear range of the adsorption isotherm, the material balance can be simplified to:

$$\lim_{q_i \rightarrow 0} y_M = \frac{\Lambda}{c_{tot}}. \quad (31)$$

With the assumption that the activity coefficients of the ligands available for binding $\hat{\gamma}_M$ and the bound protein species $\hat{\gamma}_{H,i}$ are reasonably constant ¹²⁰, the adsorption isotherm in the linear range can be written as ¹¹⁶:

$$\frac{q_i}{c_i} = A_{H,i} = K_{H,i} \left(\frac{y_M \hat{\gamma}_M}{x_M^{\zeta_i} \gamma_M^{\zeta_i}} \right)^{\vartheta_i} \frac{\gamma_{H,i}}{\hat{\gamma}_{H,i}} \approx \frac{A_{0,i}}{(x_M \gamma_M)^{\vartheta_i \zeta_i}} \left(\frac{\Lambda}{c_{tot}} \right)^{\vartheta_i} \gamma_{H,i}, \quad (32)$$

where $A_{0,i}$ is a constant for a certain protein-ligand-modular complex. Applying the natural logarithm results in ¹¹⁶:

$$\ln(A_{H,i}) = \ln(A_{0,i}) + \vartheta_i \ln\left(\frac{\Lambda}{c_{tot} x_M^{\zeta_i} \gamma_M^{\zeta_i}}\right) + \ln(\gamma_{H,i}). \quad (33)$$

Consequently, the general adsorption isotherm in the linear and non-linear range is given by the following relationship:

$$\frac{q_i}{c_i} = A_{0,i} \left(\frac{\Lambda}{c_{tot}}\right)^{\vartheta_i} \frac{1}{(x_M \gamma_M)^{\vartheta_i \zeta_i}} \left[1 - \sum_{i=1}^{\alpha} \frac{(\vartheta_i + \xi_i) q_i}{\Lambda}\right]^{\vartheta_i} \gamma_{H,i}. \quad (34)$$

1.4.2.3 Activity coefficients

In this model, several activity coefficients describing the effect of the modulator salt and the organic modifier on the protein activity and the effect of the mobile phase composition on the activity coefficient of the organic modifier were used. For the description of the effect of the mobile phase composition on the activity coefficient on the organic modifier, the Wilson equation¹²¹ was chosen. The activity coefficient of the organic modifier can then be written as¹¹⁶:

$$\ln(\gamma_M) = -\ln(x_M + E_{M,W} x_W) + x_W \left(\frac{E_{M,W}}{x_M + E_{M,W} x_W} - \frac{E_{W,M}}{x_W + E_{W,M} x_M} \right), \quad (35)$$

where $E_{M,W}$ and $E_{W,M}$ are the binary interaction parameters for the organic modifier and water. In this work, the parameters $E_{M,W} = 0.7380$ and $E_{W,M} = 0.2523$ as fitted for the water-organic modifier system by Arkell et al.¹¹⁶ were used.

The effect of the organic modifier on the protein activity can be written with a simplified version of a Wilson equation¹²¹ for a ternary system (organic modifier-water-protein) as described by Arkell et al.¹¹⁶:

$$\ln(\gamma_{H,i,M}) = \ln(\omega_H) + \ln(\varphi + x_M) + \frac{\alpha_H x_M + \beta_H}{\chi x_M^2 + \delta x_M + E_{W,M}}, \quad (36)$$

with:

$$\omega_H = (E_{M,P} - E_{W,P}) \exp\left(\frac{\theta_H}{\chi} - 1\right), \quad (37)$$

$$\theta_H = E_{P,W}(E_{W,M} - 1) + E_{P,M}(E_{M,W} - 1), \quad (38)$$

$$\varphi = \frac{E_{W,P}}{E_{M,P} - E_{W,P}}, \quad (39)$$

$$\alpha_H = E_{P,W}(1 - 2E_{W,M}) + E_{P,M} - \frac{\theta_H}{\chi}, \quad (40)$$

$$\beta_H = E_{W,M}\left(E_{P,W} - \frac{\theta_H}{\chi}\right), \quad (41)$$

$$\chi = (E_{M,W} - 1)(1 - E_{W,M}), \quad (42)$$

$$\delta = 1 + E_{W,M}(E_{M,W} - 2), \quad (43)$$

where the indices W and P refer to water and protein.

The effect of the modulator salt on the protein activity is described by the, from Mollerup et al. ¹²⁰ derived, salting-in and salting-out potential according to Kirkwood's theories on the electrostatics of macromolecules in solution. The effect of the modulator salt on the protein activity can then be written as ^{116, 120}:

$$\ln(\gamma_{H,i,salt}) = \frac{\mu_{S-i}}{RT} + \frac{\mu_{S-o}}{RT}, \quad (44)$$

with the relationship for the salting-in term ^{116, 120}:

$$\begin{aligned} \frac{\mu_{s-i}}{RT} = & -\frac{3N_A}{32\pi\epsilon RT} \left[\eta_i \tau_{p_i} \left(1 - \frac{\ln(1 + \tau_{p_i} \kappa^2)}{\tau_{p_i} \kappa^2} \right) \right. \\ & \left. + \frac{(z_i F)^2}{\kappa^2 \epsilon RT} \sum_{k=1}^{\alpha} c_k \eta_k \tau_{p_k} \left(\frac{\ln(1 + \tau_{p_k} \kappa^2)}{\tau_{p_k} \kappa^2} - \frac{1}{1 + \tau_{p_k} \kappa^2} \right) \right], \end{aligned} \quad (45)$$

and with the relationship for the salting-out term ^{116, 120}:

$$\frac{\mu_{s-o}}{RT} = \frac{3}{2} \kappa^2 \tau_{p_i} \theta_{p_i}, \quad (46)$$

where F is the Faraday number, N_A the Avogadro number, κ the inverse of the Debye length, and η_i , τ_{p_i} , and θ_{p_i} are protein- and salt-specific parameters. The permittivity of the mobile phase ϵ was set to be constant during the organic modifier gradients due to the marginal change in the gradient. The inverse Debye length was rearranged to the following expression:

$$\frac{1}{\kappa} = \sqrt{\frac{\epsilon_F \epsilon_r k_B T}{2N_A e^2}} \cdot \frac{1}{\sqrt{I}}, \quad (47)$$

$$\frac{1}{\kappa^2} = \frac{\epsilon_F \epsilon_r k_B T}{2N_A e^2} \cdot \frac{1}{I}, \quad (48)$$

where ϵ_F is the permittivity of free space, ϵ_r is the relative permittivity, k_B is the Boltzmann constant, and e is the elementary charge.

Combining Equations (33), (36), and (44) results in the general isotherm applicable in the linear range ¹¹⁶:

$$\ln(A_{H,i}) = \ln(A_{0,i}) + \vartheta_i \ln\left(\frac{\Lambda}{c_{tot} x_M^{\zeta_i} \gamma_M^{\zeta_i}}\right) + \ln(\gamma_{H,i,M}) + \ln(\gamma_{H,i,salt}), \quad (49)$$

$$\begin{aligned} \ln(A_{H,i}) = \ln(A_{0,i}\omega) + \vartheta_i \ln\left(\frac{\Lambda}{c_{tot} x_M^{\zeta_i} \gamma_M^{\zeta_i}}\right) + \frac{\mu_{s-i}}{RT} + \frac{\mu_{s-o}}{RT} \\ - \ln(\varphi + x_M) - \frac{\alpha x_M + \beta}{\chi x_M^2 + \delta x_M + E_{W,M}}. \end{aligned} \quad (50)$$

In order to simplify the derived model, some assumptions were made, as stated in the work of Arkell et al. ¹¹⁶. Arkell et al. suggested, based on their experimental data, that $\ln(A_{H,i})$ is linear on the molar fraction of the organic modifier. Although a different experimentally framework was used, this assumption was adopted in this work. Therefore, as explained by Arkell et al. ¹¹⁶, the terms $\ln(\varphi + x_M)$ and $\beta_H/(\chi x_M^2 + \delta x_M + E_{W,M})$ were assumed to be constant and integrated in the first term of Equation (50). Furthermore, the salting-out term (Equation (46)) was omitted since no salting-out effect was observed. In the linear range, the second term of the salting-in term was neglected since c_i approaches zero. Additionally, the salting-in term was simplified in order to linearize the dependency on κ^2 and the solute protein concentration c_i . The terms in the salting-in potential then simplify to ¹¹⁶:

$$\frac{3N_A \eta_i \tau_i}{32\pi \epsilon RT} \left(1 - \frac{\ln(1 + \tau_i \kappa^2)}{\tau_i \kappa^2}\right) \approx \frac{3N_A}{64\pi \epsilon RT} \kappa^2 (\eta \tau^2)_i, \quad (51)$$

$$\begin{aligned} \frac{3N_A}{32\pi \kappa^2} \left(\frac{z_i F}{\epsilon RT}\right)^2 \sum_{k=1}^{\alpha} c_k \eta_k \tau_{p_k} \left(\frac{\ln(1 + \tau_k \kappa^2)}{\tau_k \kappa^2} - \frac{1}{1 + \tau_{p_k} \kappa^2}\right) \\ \approx \frac{3N_A}{64\pi} \left(\frac{z_{P,i} F}{\epsilon RT}\right)^2 \sum_{k=1}^{\alpha} c_k (\eta \tau^2)_k. \end{aligned} \quad (52)$$

Note that in this work, the term $(\eta \tau^2)_k$ was fitted to the linear range data.

Consequently, the final form of the isotherm in the linear range can be written as ¹¹⁶:

$$\ln(A_{H,i}) = \ln(A'_{0,i}) - \vartheta_i \zeta_i \ln(x_M \gamma_M) - \frac{3N_A}{64\pi\epsilon RT} \kappa^2 (\eta\tau^2)_i - \frac{\alpha_H x_M}{\chi x_M^2 + \delta x_M + E_{W,M}} \quad (53)$$

With the relationship:

$$A'_{0,i} = A_{0,i} \left(\frac{\Lambda}{c_{tot}} \right)^{\vartheta_i}, \quad (54)$$

the explicit isotherm valid for the linear and non-linear range is then given by:

$$\frac{q_i}{c_i} = A_{0,i} \left(\frac{\Lambda}{c_{tot}} \right)^{\vartheta_i} \frac{1}{(x_M \gamma_M)^{\vartheta_i \zeta_i}} \left[1 - \sum_{i=1}^{\alpha} \frac{(\vartheta_i + \xi_i) q_i}{\Lambda} \right]^{\vartheta_i} \gamma_{H,i,M} \gamma_{H,i,salt}, \quad (55)$$

$$\begin{aligned} \frac{q_i}{c_i} = & A_{0,i} \left(\frac{\Lambda}{c_{tot}} \right)^{\vartheta_i} \frac{1}{(x_M \gamma_M)^{\vartheta_i \zeta_i}} \left[1 - \sum_{i=1}^{\alpha} \frac{(\vartheta_i + \xi_i) q_i}{\Lambda} \right]^{\vartheta_i} \\ & \exp \left(- \frac{\alpha_H x_M}{\chi x_M^2 + \delta x_M + E_{W,M}} \right) \exp \left(- \frac{3N_A}{64\pi\epsilon RT} \kappa^2 (\eta\tau^2)_i \right. \\ & \left. - \frac{3N_A}{64\pi} \left(\frac{z_i F}{\epsilon RT} \right)^2 \sum_{k=1}^{\alpha} c_k (\eta\tau^2)_k \right). \end{aligned} \quad (56)$$

1.4.2.4 Linear gradient elution experiments

Similar to the model calibration for the IEC columns, linear range parameters were determined by applying LGE experiments at low loading conditions. The relationship between the normalized gradient slope and the organic modifier elution concentration is given by:

$$\frac{dGH_M}{dx_M} = \frac{1}{k_d A_{H,i} + k_d - 1}. \quad (57)$$

For the simulation of the protein elution and the mobile phase salt elution profiles, the same lumped rate model, as shown in section 1.4.1.5, was applied.

1.5 Aim of the work

The objective of this work was to develop an *in silico* model for a preparative chromatographic purification process consisting of a cation exchange chromatography step followed by a reversed-phase chromatography step using mechanistic modeling. First, the uncommon elution behavior of the main product, a polypeptide, on the cation exchange resin at analytical and preparative loading range should be investigated and described with an appropriate mechanistic model. The same mechanistic modeling approach should be applied to the general elution behavior of the polypeptide on the reversed-phase resin. Calibration experiments will be performed on a small scale with mini-columns to save time and materials. The model is calibrated using linear gradient elution experiments in the linear and non-linear range of the adsorption isotherm. Consequently, the scale-up ability of the developed mechanistic model calibrated with a small-scale column and the small-scale calibration column itself should be investigated for a transfer to large benchtop column scale. For the investigation and the modeling of the separation performance, product- and process-related impurities should also be modeled. The implementation of the impurities, in the end, will result in a representative mechanistic multicomponent model of each preparative chromatographic step.

After successful model development on both chromatographic steps, both models should be coupled to a sequential two-column model. In conclusion, a sampling study will investigate appropriate loading conditions for the initial load on the first chromatographic step and the influence on theoretical pools and step yields and recoveries.

2 MATERIALS AND METHODS

2.1 Chemicals

The chemicals used in this work are listed in Table 1.

Table 1: Chemicals used in this work.

Chemical	Grade/Purity	Manufacturer
Acetonitrile	≥ 99.9 % LiChrosolv®	Merck KGaA
Blue Dextran (void volume marker for gel filtration columns)	2000 kDa	Sigma-Aldrich
Ethanol	≥ 99.9 % LiChrosolv®	Merck KGaA
Hydrochloric acid (HCl), 1 mol/L	≥ 99.5 % Titripur®	Merck KGaA
Hydrochloric acid (HCl), 37 %	NORMAPUR	VWR
Lactic acid (90 %)	EMPROVE®	Merck KGaA
MES	≥ 99.5 %	AppliChem GmbH
Sodium chloride (NaCl)	EMSURE®	Merck KGaA
Sodium hydroxide (NaOH)	EMSURE®	Merck KGaA
Sodium hydroxide (NaOH), 1 mol/L	≥ 99.5 % Titripur®	Merck KGaA
Sodium hydroxide (NaOH), 5 mol/L	≥ 99.5 % Titripur®	Merck KGaA
Technical buffer solution pH 2, 4.01, 7, and 10	InLab® Solutions	Mettler Toledo GmbH
Reference electrode Filling solution	ROSS™	Thermo Scientific

2.2 Electronic equipment

The electronic equipment used in this work is listed in Table 2.

Table 2: Electronic equipment used in this work.

Instrument	Manufacturer	Description
ÄKTA Purifier 10	GE Healthcare Life Sciences (Now Cytiva)	FPLC
Alliance Waters s2695 and 2489	Waters	HPLC
Balance (CP12001 S)	Sartorius	Balance
Balance (CP225D)	Sartorius	Analytical balance
Balance (type 1413)	Sartorius	Balance
Knick Portamess® incl. WTW TetraCon® conductivity meter	Knick / WTW	Conductivity meter
Orion VersaStar Pro (Advanced Electrochemistry Meter) incl. 8157BNUMD Triode	Thermo Scientific	pH meter
PURELAB® Chorus 1	Elga Veolia	Water purification system
Stirrer/Heater	Heidolph	Magnetic stirrer
Vortex Mixer	VWR	Vortex Mixer

2.3 Computer software

The software used in this work is listed in Table 3.

Table 3: Computer software used in this work.

Software	Manufacturer	Description
Berkeley Madonna (V8.3.18)	Berkeley Madonna, University of California	Differential equation solver
MATLAB R2021a	MathWorks	Numerical computing software
MS Office 2021	Microsoft	Office software
Peakfit v4.12	SeaSolve Software Inc.	Peak convolution software

Software	Manufacturer	Description
UNICORN 5.31	GE Healthcare Life Sciences (Now Cytiva)	ÄKTA system control and evaluation software

2.4 Preparative chromatography resins and columns

In this work, two strong CEX resins and a silica-based reversed-phase resin were used. The first CEX resin was a polymer-based strong cationic exchange resin with an SO_3^- functional group. The second CEX resin was TOYOPEARL GigaCap S-650S (Tosoh), also with an SO_3^- functional group. Each resin was purchased in various pre-packed columns with different column scales. In addition, one custom-packed column was used. The chromatography columns with the corresponding resin are listed in Table 4. The columns with the first CEX resin were named in the following CEX column 1-4, and the RP column is named RP column 1. The CEX column 1 (1 mL scale) was used as the model calibration column. Based on the determined model parameters with CEX column 1, scale-up to larger column volumes with the CEX column 2-4 was done. The column dead volume was assumed to be 30.4 μL for the CEX column 1 and 55.7 μL for the CEX column 2-3¹²². For the custom-packed CEX column 4, the dead volume was determined only with the system and additionally with an empty column. The second CEX resin (TOYOPEARL GigaCap S-650S) was purchased in a 1 mL pre-packed column, named in the following TOYOPEARL S-650S column.

Table 4: Preparative chromatography columns used in this work.

Column	Dimension (mm)	Volume (mL)	Product/manufacturer
CEX column 1	50 x 5	1.0	Pre-packed OPUS® Mini-Chrom® (Repligen)
CEX column 2	50 x 11.3	5.01	Pre-packed OPUS® Mini-Chrom® (Repligen)
CEX column 3	100 x 11.3	10.03	Pre-packed OPUS® Mini-Chrom® (Repligen)

Column	Dimension (mm)	Volume (mL)	Product/manufacturer
CEX column 4	359 x 10	28.20	Custom-packed in Super Compact 600-10 (Götec Labortechnik)
TOYOPEARL S-650S	50 x 5	1.0	Pre-packed OPUS® Mini-Chrom® (Repligen)
RP column 1	50 x 4.6	0.83	Pre-packed

2.5 Analytical chromatography resins and columns

In this work, a pre-packed analytical RP-HPLC column was used for the quantitative and qualitative determination of the sample proteins. The analytical column is listed in Table 5.

Table 5: Analytical column used in this work.

Column	Dimension (mm)	Volume (mL)	Product/manufacturer
Analytical C ₁₈ RP column, 2-3 µm particle diameter	250 x 3	1.77	Merck Millipore

2.6 Miscellaneous items

All other miscellaneous items used in this work are listed in Table 6.

Table 6: Miscellaneous items used in this work.

Item	Manufacturer	Description
20 mL syringe Luer lock	Thermo Scientific	Plastic Syringe
33 mm, PES, 0.2 and 0.45 µm Filter	Thermo Scientific	Syringe filter
500 mL bottle top filter	Corning	Filter Cup

Item	Manufacturer	Description
Nitrile Gloves	Kimtech	Nitrile Gloves
Parafilm®	American National Can	Sealing
Rapid flow filter Cup 250 mL	Thermo Scientific	Sterile filter Cup
Sample tube, 1.5, 2, 15, and 50 mL	Thermo Scientific	Sample tube
Superloop, 50, and 150 mL	GE Healthcare Life Sciences (Now Cytiva)	Sample loop for ÄKTA systems
UV cell, 2, and 10 mm	GE Healthcare Life Sciences (Now Cytiva)	UV cell for ÄKTA systems

2.7 Polypeptides and sample preparation

The sample for all chromatographic experiments was a recombinant lyophilized polypeptide, named in the following as polypeptide, two product-related impurities, named in the following as variant A and variant B, and one process-related impurity, named variant C. The purity, based on total protein content, was $\geq 99\%$. For the lyophilized polypeptide and the impurities, stock solutions with 10 mg/mL and 1 mg/mL were prepared and filtered with a sterile filter (0.2 μm). The sample concentration was prepared using the stock solution and an appropriate dilution buffer (mostly Buffer A from the chromatographic experiment). Before injection, the sample was filtered with a particle filter (0.45 μm).

2.8 Chromatographic experiments

2.8.1 CEX resin column experiments

The binding and eluting buffer for the linear and non-linear range experiments contained 0.05 mol/L lactic acid and a constant concentration of a short-chain aliphatic alcohol. The eluting buffer contained 0.2 mol/L sodium chloride, and all experiments were performed between pH 3.3 and pH 4.3.

For the LGE calibration experiments in the linear range for the CEX column 1 (calibration column), five different gradient slopes (from $8 \times 10^{-3} \text{ mol}_{\text{NaCl}}/\text{L}\times\text{CV}^{-1}$ to $1.4 \times 10^{-3} \text{ mol}_{\text{NaCl}}/\text{L}\times\text{CV}^{-1}$) at pH 3.3, 3.5, 3.7, 3.9, 4.1, and 4.3 were chosen. The sample load in the linear range was 0.05 mg/mL_{CV} with a 0.1 mg/mL sample concentration for the polypeptide and the impurity variants. Non-linear experiments were performed using different sample loads up to 40 mg/mL_{CV} with a polypeptide concentration of 1 mg/mL. Chromatographic runs with a load up to 0.75 mg/mL_{CV} were performed with a 10 mm UV detector cell. Chromatographic runs with a higher load were performed with a 2 mm UV detector cell. The load for the CEX columns 2-4 was 0.05 mg/mL_{CV} for the linear range experiments and 1.5, 5, and 10 mg/mL_{CV} for the non-linear range experiments. Note that no experiments in the non-linear range were performed with the impurities.

For the isocratic elution experiments for the CEX column 1 with the polypeptide species, 200 μL sample solution was injected with a buffer mixture of 60 % and 75 % Buffer B. The column load was 0.05, 0.5, 1, 2, 3, 4 mg/mL_{CV} (60 % Buffer B runs) and 0.05, 0.5, 1, 2, 3, 4, 5.5, 6.5, 7.5, 8 mg/mL_{CV} (75 % Buffer B runs). The sample solution for each run contained a different polypeptide concentration of up to 20 mg/mL (60 % Buffer B runs). For the 75 % Buffer B runs, the polypeptide sample was diluted in Buffer A (without salt) up to a sample concentration of 40 mg/mL. A higher sample concentration leads to precipitation of the polypeptide in the buffer system. Therefore, column loads above 8 mg/mL_{CV} could not be tested.

For the CEX column 1 validation experiment, a dual gradient from pH 4.3 to pH 3.4 and from 0 mol/L to 0.2 mol/L sodium chloride with a column load of 7.5 mg/mL_{CV} was performed. A replacement column for CEX column 1 with the same column dimensions and individual column parameters (ϵ_t : 0.736, ϵ_p : 0.594, Λ_{pack} : 0.105 mol/L) was used.

For the calibration experiments in the linear range for the TOYOPEARL S-650S column, five different gradient slopes (from $8 \times 10^{-3} \text{ mol}_{\text{NaCl}}/\text{L}\times\text{CV}^{-1}$ to $1.4 \times 10^{-3} \text{ mol}_{\text{NaCl}}/\text{L}\times\text{CV}^{-1}$) at pH 3.7 were chosen. The sample load in the linear range was 0.05 mg/mL_{CV} with a 0.1 mg/mL sample concentration for the polypeptide. Non-linear experiments were performed using different sample loads up to 60 mg/mL_{CV} (0.25, 0.5, 0.75, 1.5, 5, 10, 15, 25, 40, and 60 mg/mL_{CV}) with a polypeptide concentration of 1 mg/mL.

For each pH value, a buffer calibration was done. The concentration of the eluting buffer was increased stepwise, and with the knowledge of the counterion concentration, the conductivity signal was recalculated to molar concentrations. The pH values determined with the chromatography system were adjusted to adapt them to the more accurate external pH measurements.

2.8.2 RP column experiments

The binding and eluting buffer for the RP column 1 experiments consisted of 0.1 mol/L of a buffer substance buffering in the range of $\text{pH } 2.8 \pm 1$, 0.01 mol/L of a salt modulator, and a concentration gradient of an organic modifier. All experiments were performed between pH 3.3 and pH 4.1.

For the linear range calibration experiments, five different gradient slopes between 0.5 and $2 \text{ \%}_{\text{organic-modifier}}/\text{L} \times \text{CV}^{-1}$ were chosen. The sample load for the linear range experiments was 0.05 mg/mL_{CV} with a 0.1 mg/mL sample concentration for the polypeptide and the impurity variants. Sample load in the non-linear range was up to 25 mg/mL_{CV} with a 1 mg/mL sample concentration for the polypeptide. All linear range chromatographic runs with the RP column 1 were done with a 10 mm UV cell, and all non-linear range experiments were recorded with a 2 mm UV cell.

2.8.3 RP-HPLC experiments

Fraction analysis was done with RP-HPLC using an alliance HPLC system from Waters. For the analysis, a standard RP-HPLC protocol was used. The composition for Buffer A was: 0.05 mol/L NaH₂PO₄, 0.3 mol/L NaCl, and 25 % (v/v) acetonitrile. For Buffer B: 0.05 mol/L NaH₂PO₄, 0.05 mol/L NaCl, and 65 % (v/v) acetonitrile. All buffers were adjusted to pH 2.5. A gradient from 100 % to 0 % Buffer A in 30 minutes was performed. The injection volume was 15 μL , the flow rate was 0.55 mL, and the HPLC analysis was performed at 35 °C column temperature. The sample was detected at a wavelength of 214 nm.

2.9 Column characterization

2.9.1 Determination of porosities

The porosities for the CEX resin columns were determined by applying pulse experiments. With the retention volume of a pore diffusing tracer V_{salt} , the retention volume

of a non-pore diffusing tracer V_0 and concerning the dead volumes of the system, the porosities of the applied column can be calculated using the following relationships ¹²³:

$$\varepsilon_0 = \frac{V_0}{V_C}, \quad (58)$$

$$\varepsilon_p = \frac{\varepsilon_t - \varepsilon_0}{1 - \varepsilon_0} = \frac{\frac{V_{salt}}{V_C} - \varepsilon_0}{1 - \varepsilon_0}, \quad (59)$$

where ε_t denotes the total porosity and V_C the column volume. Dead volumes of the system were determined by applying tracer injections without the column and additionally with an empty column for the CEX column 4 (28.2 mL scale). For the CEX columns, a pore diffusing tracer, a 1 mol/L NaCl solution, and a non-pore diffusing tracer, a 1 g/L Dextran solution (Blue Dextran from Merck Millipore, 2000 kDa mean molecular weight), were used. For the CEX column 1 (1 mL scale), 10 μ L sample solution; for the CEX columns 2 and 3 (5 mL and 10 mL scale), 50 μ L; and for the CEX column 4 (28.2 mL scale), 100 μ L sample solution were injected, each with a 0.05 mol/L MES running buffer including 0.4 mol/L NaCl at pH 6 with a flow rate of 100 cm/h.

The total porosity for the RP column 1 was determined by applying salt pulse experiments with sodium chloride. The binding buffer from section 2.8.2 was used. Due to the difficulties of interstitial porosity determination with reversed-phase resins ¹²⁴, the interstitial porosities for the RP column 1 were estimated using literature data.

2.9.2 Determination of the HETP, asymmetry factors, and plate numbers

The salt pulse experiments for the CEX columns and the RP column were also evaluated in order to determine the height equivalent of a theoretical plate *HETP* and the asymmetry factor A_s of the column. The *HETP* value and the plate number N is given by:

$$HETP = \frac{\sigma_{EMG,salt}^2}{\mu_{EMG,salt}^2} L_c, \quad (60)$$

$$N = \frac{L_c}{HETP}, \quad (61)$$

where $\mu_{EMG,salt}$ and $\sigma_{EMG,salt}^2$ are the first and the second moment of the salt peak and L_c is the column length. The asymmetry factor (at 10 % peak height) was determined

with the chromatography system software UNICORN 5.31 from GE Healthcare Life Sciences (now Cytiva).

2.9.3 Determination of polypeptide exclusion factors

The exclusion factor $k_{d,i}$ of the polypeptide for the CEX columns was determined by applying pulse experiments under non-binding conditions with the buffer system used for the chromatographic experiments, including 1 mol/L sodium chloride. With the retention volume of the protein pulse V_R , the exclusion factor is given by the following relationship ¹¹¹:

$$k_{d,i} = \frac{\frac{V_R}{V_C} - \varepsilon_0}{(1 - \varepsilon_0) \cdot \varepsilon_p} \quad (62)$$

Due to the local porosity changes and differences during measuring retention volumes in RPC at non-binding conditions ^{125, 126}, the polypeptide exclusion factor for the RP column was estimated using literature data. It was assumed that the exclusion factors for the impurity variants are the same as for the polypeptide since their molecular weight is in the same range.

2.9.4 Determination of ionic capacities

The ionic capacities of the CEX resin columns were determined according to Tugcu et al. ¹²⁷. The column was first flushed with a 0.5 mol/L HCl solution (Merck Millipore) and then rinsed with MilliQ water. Then the column was titrated with a 0.1 mol/L NaOH solution (Merck Millipore) at a flow rate of 100 cm/h until the conductivity signal was rising and constant. The ionic capacities, also referred to as ligand densities, can be calculated using the following relationships ¹²⁸:

$$\Lambda_{pack} = \frac{V_{NaOH} \cdot c_{NaOH}}{V_C} \quad (63)$$

$$\Lambda_{pore} = \frac{\Lambda_{pack}}{(1 - \varepsilon_0) \cdot \varepsilon_p} \quad (64)$$

$$\Lambda_{sediment} = \frac{\Lambda_{pack}}{1/(1 - C_f)} \quad (65)$$

where V_{NaOH} is the applied titrant volume of the titration curve, c_{NaOH} is the concentration of the NaOH solution and C_f is the compression factor of the resin ($C_f = 0.1$ for all

CEX columns). The applied titrant volume was determined with the peak maximum of the differentiated titration curve evaluated with the software UNICORN 5.31.

The ligand density of the RP column 1 was calculated based on pore volume using resin-specific information from the manufacturer.

2.9.5 Determination of axial dispersion coefficients

The axial dispersion coefficient for the CEX resin columns was determined by applying Blue Dextran pulse experiments. The same experimental conditions as used in the porosity determination were applied. The axial dispersion coefficient of the column was then calculated using the following relationship ¹²⁸:

$$D_{ax} = \frac{\sigma_{EMG,Dextran}^2 L_c u_{int}}{\mu_{EMG,Dextran}^2 2}. \quad (66)$$

The pulse experiments' first and second moments were determined from an exponentially modified Gaussian (EMG) function ¹²⁹ fit computed with the software PeakFit (v4.12, SeaSolve Software Inc). Since no large non-pore diffusing tracer suitable for RP columns was available, the axial dispersion coefficient of the RP column 1 was fitted simultaneously with the mass transfer coefficients at analytical loading conditions.

2.10 Modeling and simulation

2.10.1 Model calibration and simulation with CEX column 1

In this work, the mechanistic model was calibrated at small column scale (CEX column 1). All model parameters were determined with the CEX column 1 and seen to be constant at larger column scales (CEX column 2-4). Model parameter estimation in the linear range of the isotherm was done by applying Yamamoto's approach ^{108, 109}. For the determination of the salt and pH dependency of the model, a set of salt-induced LGE experiments with different gradient slopes at different pH values in the linear range of the isotherm is required. The eluting salt concentration from the LGE experiments was determined by evaluating the first moment of a fitted EMG peak ¹²⁹ (PeakFit v4.12, SeaSolve Software Inc). With the equations derived in section 1.4.1.4, a GH_{salt} vs. $c_{s,elut}$ plot was created, and the linear range isotherm parameters were fitted. Model

calculations and simulations were done using the differential equation solver Berkeley Madonna (V8.3.18).

Parameters in the linear range of the isotherm were fitted using the following fit strategy leading to a robust fit with a reproducible solution: In the first step, the linear range experimental data were estimated with Equation (22) to the variables v_i and K_i without the influence of the asymmetric activity coefficient of the protein $\tilde{\gamma}_i$. A unique solution for the fitted variables for each pH value occurs. In the second step, the pH dependency of the equilibrium constant was estimated using the parameters $\Delta G_i^0/RT$ and $\Delta G_s^0/RT$. As initial guess values, the slope and the y-intercept of the $\ln K_i$ vs. v_i plot from the previous step were used. The characteristic charge was estimated individually for each pH value using the previous step values $\pm 10\%$. In the third step, the parameter K_s^* from the asymmetric activity coefficient was determined by varying the parameters ξ_i , K_s^* and $K_{D,i}/c$ in high load simulations (25 mg/mL_{CV}) using the previously determined v_i , $\Delta G_i^0/RT$ and $\Delta G_s^0/RT$ for the respective pH. K_p^* was set to be zero due to the observed marginal influence in this step. The parameter fitting results in a robust and reproducible solution for the determined model parameters. In the last step, all parameters in the linear range of the isotherm were fitted again, including the parameter K_s^* from the asymmetric activity coefficient model. The term K_s^*/v_i^2 was varied $\pm 10\%$ for the pH value of the high loading experiment from the previous step applying the determined values for K_s^* and v_i . Simultaneously, the pH dependency of the characteristic charge was fitted with two amino acid-like groups, considering the number of the involved groups N_{acidic} and N_{basic} and the pKa value pKa_{acidic} of the acid group. The pKa value of the basic amino acid-like group pKa_{basic} was assumed to be 11 due to the experimental pH range from pH 3.3 to pH 4.3. Note that the term $K_p \times c_i$ approaches zero due to the low protein concentration in the linear range of the isotherm. The last fit results in robust and reproducible values for the model parameters determined in the linear range of the isotherm (K_s^* , $\Delta G_i^0/RT$, $\Delta G_s^0/RT$, N_{acidic} , N_{basic} and pKa_{acidic}).

Parameter estimation in the non-linear range of the isotherm was done by applying the inverse fitting method. Simulated chromatograms were fitted to experimentally determined chromatograms by varying the parameters in the non-linear range (ξ_i , $K_{D,i}/c$ and K_p^*). A maximum of four chromatograms with different sample loads were fitted

simultaneously. Ideally, all non-linear range calibration experiments would be computed simultaneously, but this was not possible due to software and computation time limitations.

The effective mass transfer parameter $k_{eff,i}$ was fitted individually for each load and pH value. The mass transfer parameter only showed an influence in simulations with a load lower than 1.5 mg/mL_{CV}. A marginal load and pH dependency was observed, which was not further investigated due to the insignificant influence of $k_{eff,i}$ at higher loads (≥ 1.5 mg/mL_{CV}).

2.10.2 Modeling and simulation with CEX columns 2-4

The model parameters in the linear (K_s^* , $\Delta G_i^0/RT$, $\Delta G_s^0/RT$, N_{acidic} , N_{basic} and pKa_{acidic} , pKa_{basic}) and non-linear (ξ_i , $K_{D,i}/c$, and K_p^*) range, determined at small column scale (CEX column 1), were seen to be constant at larger column scale for the CEX column 2-4. Scale-up simulations were done in two approaches. In the first empirical approach, the shielding factor and the mass transfer coefficient $k_{eff,i}$ were simultaneously fitted for each scale-up column. It should be noted that $k_{eff,i}$ had a negligible effect on the overall peak shape at columns loads higher than 1.5 mg/mL_{CV} so the peak shape is largely determined by the shielding factor in high loading situations. The experimental loading dependencies for each column scale's effective mass transfer coefficient are shown in Appendix 8.4 in Table 20. The second approach considers packing inhomogeneities while keeping all model parameters constant, including the shielding factor. The same mass transfer coefficient values from the first empirical approach were used.

2.10.2.1 Modeling of column packing inhomogeneity

For the consideration of the effects of inhomogeneous packing, the different scaled CEX columns (CEX columns 2-4) were divided into p homogeneously packed concentric annular columns with equal face areas. It was assumed that no radial dispersion takes place¹³⁰ and that the axial dispersion coefficient and the linear velocity are constant in each concentric annular column. Ten different interstitial porosities ($\varepsilon_{0,j,k}$) for each column scale j , with a variation of about $\pm 20\%$ around the experimentally determined values, were used to simulate the elution profile in a concentric annular column. Note that a change in the radial interstitial porosity also results in a change in the radial interstitial velocity since the convection term in the column model (Equation (24)) is

driven by u_{int} . The resulting overall elution profile for all concentric annular columns was calculated by summing the concentration profiles eluted from each concentric annular column with different $\varepsilon_{0,j,k}$, using the following relationship:

$$c_{i,sum} = \sum_{k=\varepsilon_{0,j,min}}^{\varepsilon_{0,j,max}} c_{i,\varepsilon_{0,j,k}} \left(\frac{b_{\varepsilon_{0,j,k}}}{p} \right), \quad (67)$$

where p ($p = 50$) is the number of concentric annular columns and $b_{\varepsilon_{0,j,k}}$ is the factor indicating how often the concentric annular column, with a defined interstitial porosity, is represented in the entire scale-up column. $b_{\varepsilon_{0,j,k}}$ was determined by using Equation (67) and fitting the number of the concentric annular columns with a defined interstitial porosity to the experimental elution profile. The fits were performed using the solver function in Microsoft® Excel® and minimizing the squared difference between the experimental and calculated elution profile.

2.10.3 Model calibration of the RP column

Model calibration in the linear range of the adsorption isotherm for the RP model was done, as for the CEX model, by applying Yamamoto's approach^{108, 109}. LGE experiments at different pH values with analytical loading conditions were performed. The relationship between the normalized gradient slope GH_M and the eluting organic modifier concentration $x_{M,elu}$ (see Equation (57)) was used to determine the linear range parameters ($(\eta\tau^2)_i$, $A'_{0,i}$, α_H , and $\vartheta_i\zeta_i$). Note that the same model assumptions as made by Arkell et al.¹¹⁶ were adopted in this work. The detailed model derivation and the corresponding assumption made for the model are described in section 1.4.2.

For the model calibration in the non-linear range of the adsorption isotherm, the inverse fitting method was applied. The simulated UV signal was fitted to the experimentally determined UV signal by varying the non-linear range parameters (ξ_i , $A'_{0,i,n}$, ϑ_i). During model calibration, it has been observed that no good fit could be achieved without adjusting one linear parameter in the non-linear range. Therefore, the parameter $A'_{0,i}$, fitted in the linear range, was refitted in the non-linear range resulting in the parameter $A'_{0,i,n}$. The effective mass transfer parameter $k_{eff,i}$ was fitted individually for each load.

3 RESULTS

Parts of this work have been summarized in three different publications. Two of these papers have already been published during the doctoral project in two peer-reviewed scientific publications, and one paper is currently in draft mode. The results and the corresponding discussion can be found in the references from Koch et al. ^{73, 131, 132}.

3.1 Column characterization

A crucial point in the modeling and simulation of liquid column chromatography is the column characterization, which includes determining the packing quality and column-specific parameters, such as porosities, protein exclusion factors, and ligand densities. The following chapters show the results of the column characterization from the different chromatographic columns used in this work.

3.1.1 CEX resin columns

In this work, two strong cationic exchange resins were used. One resin was available in four different column scales (CEX column 1-4), and the second resin (TOYOPEARL GigaCap S-650S) was purchased in a pre-packed 1 mL column (TOYOPEARL S-650S column). Table 7 summarizes the determined column-specific parameters for each CEX column used in this work. The interstitial and total porosities were obtained from salt pulse injections. The particle porosity was calculated according to Equation (59). The axial dispersion was obtained from moment analysis of Blue Dextran injections (Equation (66)), the ionic capacities from column back titration experiments (Equation (63)), the polypeptide exclusion factor from polypeptide pulse injections under non-binding conditions (Equation (62)), and the extra column volume (ECV) to column volume (CV) ratio from system death volume determining salt pulse injections. In Figure 1, the column-specific parameters for the CEX columns 1-4 are shown to visualize the scalability of the columns packed with the same resin. The interstitial porosity ranges from 0.347 to 0.369 and is within a theoretically plausible range. The total and the particle porosity decrease with increasing column scale, whereby the porosity values of the 1 mL column (CEX column 1) are relatively high compared to the values of the CEX columns 2-4. The ligand density based on pore volume is for the CEX column 1 (1 mL), the CEX column 2 (5 mL), and the CEX column 3 (10 mL) in the same range

and only varies around 3×10^{-3} mol/L. For the 28.2 mL column scale, the ligand density based on pore volume is 0.275 mol/L and slightly higher but still in the same range. The ligand density based on the packed bed decreases with increasing column scale. The $k_{a,i}$ values for the polypeptide show a decreasing dependency with increasing column scale and the axial dispersion coefficient D_{ax} shows different values for each column scale.

Comparing the column-specific parameters determined for the TOYOPEARL S-650S column with the values of CEX column 1, it is noticeable that the values are in the same range, except for the ionic capacity, which is more than twice as high for the TOYOPEARL S-650S column.

The salt pulse evaluation, representing the packing quality for each column¹³³, is summarized in Table 8. The first moment (peak center) and the standard deviation of the salt pulses were corrected by the first moment and the standard deviation of the system ECV to consider ECV effects. Figure 2 visualizes a comparison of the CEX columns with the same resin and different dimensions for the investigation of the scalability of these columns. The first moments of the peaks for the CEX columns 1-4 increase linearly with increasing column volume. The standard deviations of the salt peaks were determined as a measure of peak broadening. The standard deviations tend to increase with increasing column volume. The black and orange data points in Figure 2 represent the evaluation in which only the ECV effects of the system were considered since no empty pre-packed columns were available. The blue data points represent the values for the CEX column 4 (28.2 mL) corrected by the system values, including an empty column. The standard deviation for the CEX column 4 decreases significantly when the empty column is also considered, although the first moment is not influenced. The determined *HETP* values and plate numbers for each column are shown in Figure 2B. Furthermore, the asymmetry factor for each column and the ratio between the ECV and the CV is shown in Figure 2C. An intensive parameters discussion with a view of the scalability of the CEX column 1 to larger column volumes can be found in section 4.1.1.

RESULTS

Table 7: Overview of the determined column-specific parameters for the CEX resin columns used in this work.

Parameters	CEX column 1	CEX column 2	CEX column 3	CEX column 4	TOYOPEARL S-650S
Dimension [mm]	50 x 5	50 x 11.3	100 x 11.3	359 x 10	50 x 5
Column vol- ume [mL]	0.98	5.01	10.03	28.20	0.98
Interstitial po- rosity ε_0 [-]	0.349	0.347	0.356	0.369	0.338
Total porosity ε_t [-]	0.776	0.732	0.728	0.717	0.790
Particle po- rosity ε_p [-]	0.656	0.589	0.578	0.552	0.682
Axial disper- sion D_{ax} [cm²/s]	2.1×10^{-4}	3.0×10^{-5}	2.3×10^{-5}	1.4×10^{-4}	3×10^{-4}
Ionic capacity Λ_{pack} [mol/L]	0.110	0.099	0.097	0.096	0.280
Ionic capacity $\Lambda_{sed.}$ [mol/L]	0.099	0.089	0.087	0.086	0.252
Ionic capacity Λ_{pore} [mol/L]	0.258	0.257	0.260	0.275	0.620
Exclusion factor k_d [-]	0.819	0.815	0.808	0.792	0.898
ECV/CV [-]	0.21	0.06	0.03	0.01 0.03*	0.21

* Corrected value with an empty column.

RESULTS

Table 8: Overview of the determined packing quality parameters for the CEX resin columns used in this work.

Parameters	CEX column 1	CEX column 2	CEX column 3	CEX column 4	TOYOPEARL S-650S
$\mu_{EMG,salt}$ [mL]	0.762	3.668	7.304	20.789 20.223*	0.765
$\sigma_{EMG,salt}$ [mL]	0.0334	0.1667	0.3373	0.3305 0.2306*	0.0344
<i>HETP</i> [cm]	0.0095	0.0103	0.0213	0.0091 0.0047*	0.0101
Plate number <i>N</i> [-]	521	484	469	3956 7693*	495
Asymmetry factor A_s [-]	1.087	1.277	1.565	0.990	0.95

* Corrected value with an empty column.

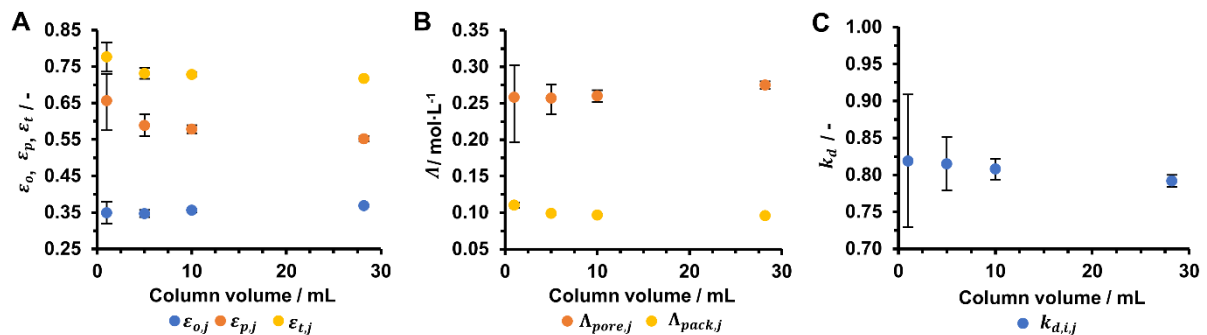


Figure 1: Determined porosities (A), ligand densities (B), and polypeptide exclusion factors (C) of the four different scaled columns (CEX column 1-4) with the same resin used in this work. Each data point contains error bars.

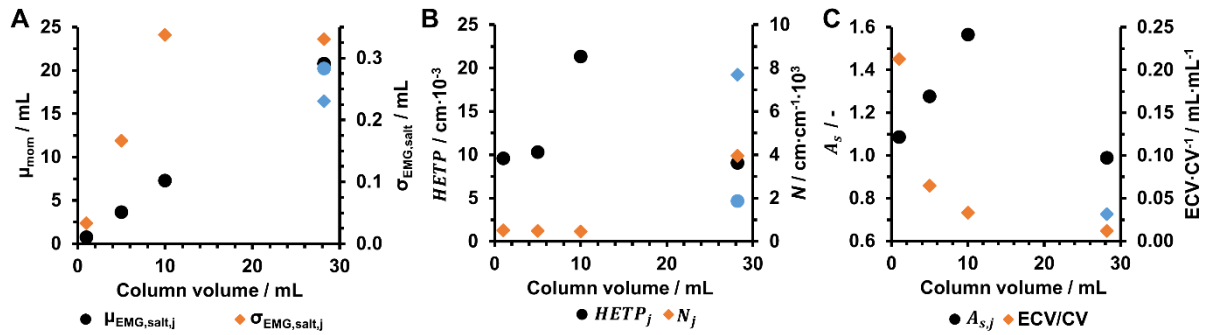


Figure 2: Determined first moment and standard deviation (A), and HETP and plate number (B) of the salt pulse peak and the column as a function of column volume of the four different scaled CEX columns (CEX column 1-4) with the same resin used in this work. (C) shows the asymmetry factor for each column and the ratio between the ECV and the CV as a function of column volume. The blue data points show the corrected values for CEX column 4, determined with an empty column.

3.1.2 RP resin column

In this work, one RP resin column (RP column 1) was used and purchased in a pre-packed steel column with similar dimensions as the CEX column 1. All determined column-specific and packing quality parameters of the RP column 1 are summarized in Table 9. A classification of the plausibility and a discussion of the parameters can be found in section 4.1.2.

Table 9: Overview of the determined column-specific and packing quality parameters for the RP column 1.

Parameters	RP column 1
Dimension [mm]	50 x 4.6
Column volume [mL]	0.831
Interstitial porosity ε_0 [-]	0.35*
Total porosity ε_t [-]	0.490
Particle porosity ε_p [-]	0.215
Axial dispersion D_{ax} [cm ² /s]	0.001
Ligand density Λ_{pore} [mol/L]	2.66×10^{-3}
Exclusion factor k_d [-]	0.75*
ECV/CV [-]	0.25
HETP [cm]	0.05744
$\mu_{EMG,salt}$ [mL]	0.408
$\sigma_{EMG,salt}$ [mL]	0.00659
Plate number N [-]	1741
Asymmetry factor A_s [-]	1.70

* Estimated value based on literature data.

3.2 CEX model development

This section presents the model development results for the chromatographic CEX columns 1-4 and the TOYOPEARL S-650S column. Model development for the CEX columns 1-4 was done at small column scale (CEX column 1) and was then later scaled up to the scale-up columns CEX column 2-4.

3.2.1 CEX model calibration at small scale (CEX column 1)

3.2.1.1 Investigation of the polypeptide elution behavior

Intending to model a chromatographic step, which should be valid in the linear and non-linear loading range, it is essential to be aware of the general elution behavior of the sample. With that knowledge, appropriate adsorption isotherms, which are able to describe the investigated elution behavior, can be selected and modified for the modeling task. Therefore, several loading series with the polypeptide and with column loads from 0.05 mg/mL_{CV} (linear range of the isotherm) up to 25 mg/mL_{CV} (non-linear range of the isotherm) at different pH values were performed by applying linear salt gradient elution experiments (Figure 3). Additionally, two isocratic loading series at pH 3.7 with the polypeptide and column loads from 0.05 mg/mL_{CV} up to 8 mg/mL_{CV} were performed (Figure 4 and Figure 5). In Figure 4, the polypeptide was diluted in 60 % Buffer B, and in Figure 5, in 0 % Buffer B. Note that no higher loadings than 8 mg/mL_{CV} could be tested isocratically since no higher sample concentration could be achieved without precipitation of the polypeptide in the sample buffer. The loading series show that the polypeptide exhibits no typical Langmuirian adsorption behavior. In a typical Langmuirian elution behavior, the elution peak of the polypeptide would shift to the left-hand side with a steep rise in the front of the peak and a diffuse rear with increasing load. An anti-Langmuirian elution behavior would be present if the elution peak shifts to the right-hand side with a diffuse front region of the peak and a steep rear with increasing load¹³⁴. However, there is a right-to-left movement of the elution peak in the gradient elution and the isocratic experiments with increasing load (anti-Langmuirian to Langmuirian elution behavior). Furthermore, a strong influence of the pH on the elution peak can be seen. Increasing the pH leads to an earlier elution and a narrower and taller peak. Note that the high loading experiments in Figure 3 were recorded with a 2 mm UV detector cell and then recalculated to a 10 mm UV detector cell. The absorption signal of all experiments carried out is still in the linear range of the UV detector cell,

although absorption signals over 3000 mAU are shown. The isocratic experiments in Figure 4 and Figure 5 were recorded with a 10 mm UV detector cell.

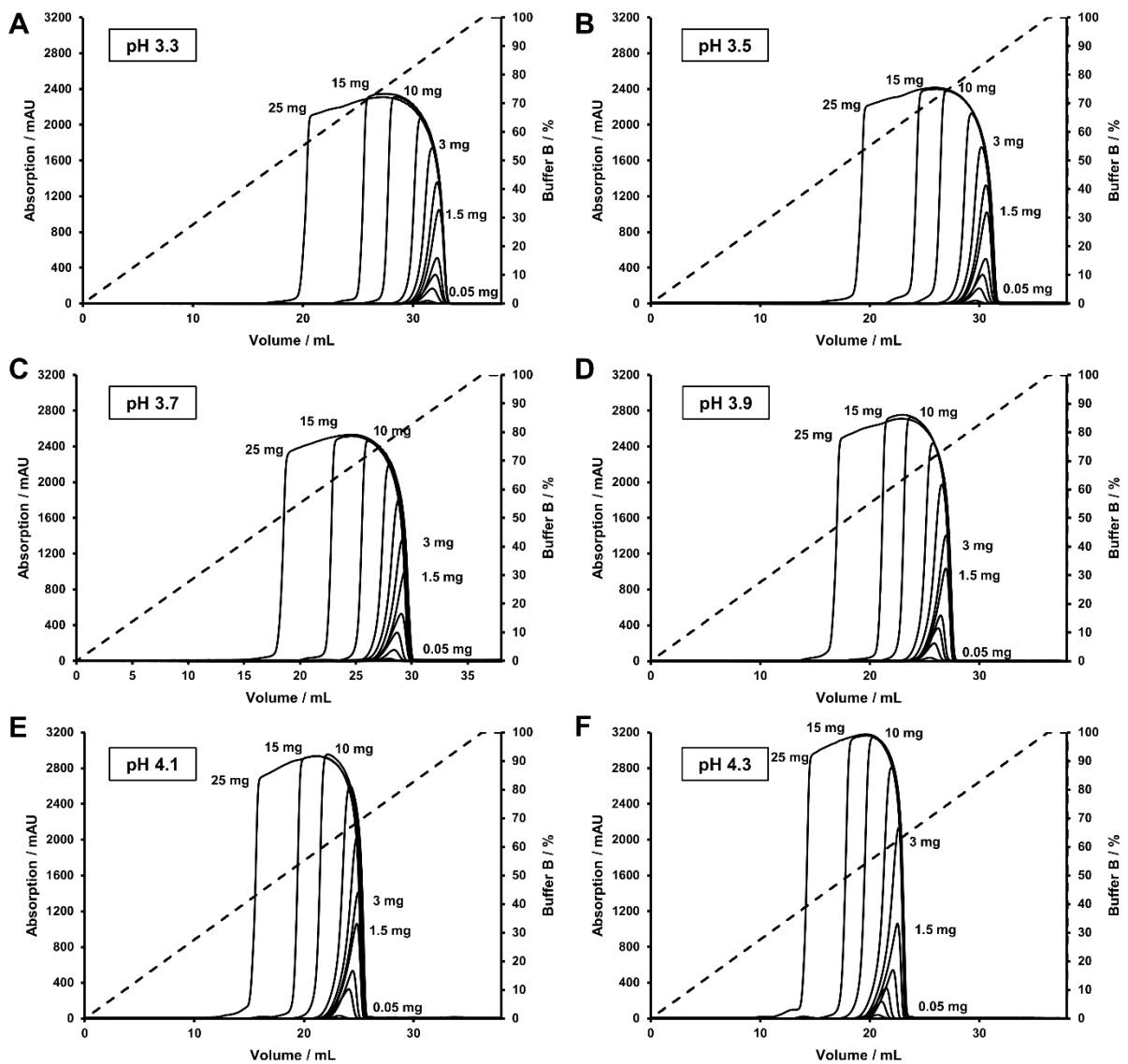


Figure 3: Gradient elution loading series with the polypeptide species at different pH values (A-F) from 0.05 mg/mL_{CV} up to 25 mg/mL_{CV} sample load with the CEX column 1. The smallest peak indicates the lowest load up to the highest load with the broadest and highest peak. The black dashed line represents the amount of Buffer B in the gradient. The UV signal from the high load runs was recorded with a 2 mm UV cell and then recalculated to a 10 mm UV cell.

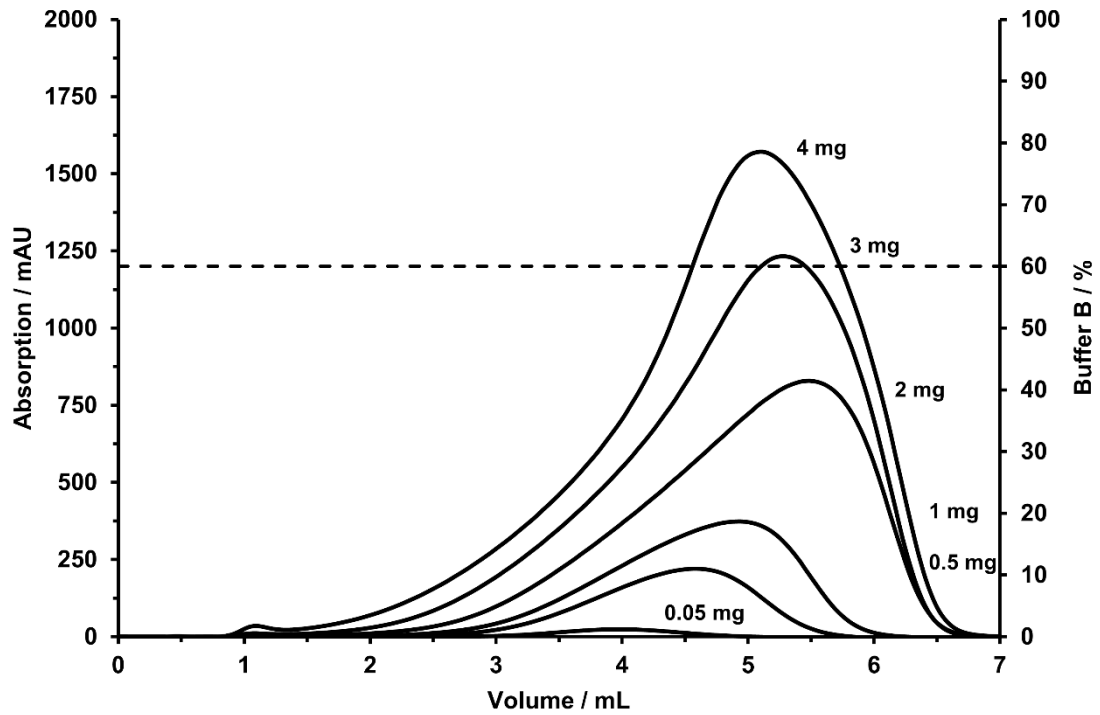


Figure 4: Isocratic elution loading series with the polypeptide at pH 3.7 from 0.05 mg/mL_{CV} up to 4 mg/mL_{CV} sample load with the CEX column 1. The polypeptide sample was diluted in 60 % Buffer B. The smallest peak indicates the lowest load up to the highest load with the broadest and highest peak. The black dashed line represents the amount of Buffer B during the isocratic run. The UV signal was recorded with a 10 mm UV cell.

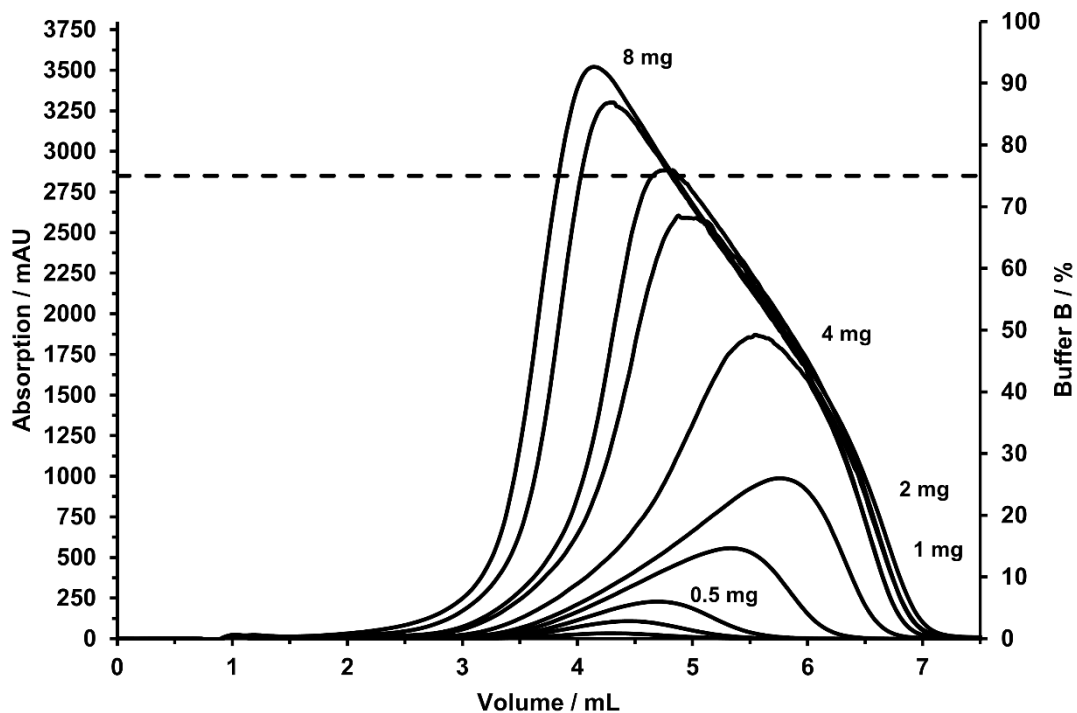


Figure 5: Isocratic elution loading series with the polypeptide at pH 3.7 from 0.05 mg/mL_{CV} up to 8 mg/mL_{CV} sample load with the CEX column 1. The polypeptide sample was diluted in Buffer A (without salt). The smallest peak indicates the lowest load up to the highest load with the broadest and highest peak. The black dashed line represents the amount of Buffer B during the isocratic run. The UV signal was recorded with a 10 mm UV cell.

3.2.1.2 Determination of model parameters in the linear range of the isotherm

The model parameters for the polypeptide, the two product-related, and the process-related impurities in the linear range of the isotherm ($\Delta G_i^0/RT$, $\Delta G_s^0/RT$, K_s^* , N_{acidic} , N_{basic} , $pK_{a_{acidic}}$, and $pK_{a_{basic}}$) were determined by applying LGE experiments at low loading conditions according to Yamamoto's approach^{108, 109}. Salt-induced LGE experiments at different gradient slopes were performed at various pH values (pH 3.3, pH 3.5, pH 3.7, pH 3.9, pH 4.1, and pH 4.3). By plotting the normalized slope of the salt gradients GH_{salt} against the eluting counterion concentration $c_{s,elu}$, the influence of the gradient slope and the pH on the eluting salt concentration is visualized. Figure 6 summarizes the evaluated LGE experiments for the polypeptide and the three impurity variants (variant A, variant B, and variant C) in the linear range with the CEX column 1. In general, the higher the pH, the earlier elutes the protein sample in the salt gradient. The steeper the gradient slope, the smaller becomes the peak retention

volume, whereas the eluting counterion concentration $c_{s,elu}$ increases. Variant A elutes in general later in the salt gradient than the polypeptide species, and variant B elutes later than variant A. Variant C elutes in a relatively small salt concentration range compared to the other proteins. The fitted parameters for the three protein samples from the final step of the fitting procedure in the linear range of the isotherm are summarized in Table 10. The parameters are in a similar range and will be discussed in section 4.2.1.2.

Table 10: Overview of the fitted model parameters in the linear range of the isotherm for the polypeptide, variant A, variant B, and variant C determined with the CEX column 1.

Parameter	Polypeptide	Variant A	Variant B	Variant C
$\Delta G_s^0/RT$	2.25	2.45	2.08	2.20
$\Delta G_i^0/RT$	4.87	5.84	3.22	-3.66
N_{acidic}	0.80	0.61	0.56	0.32
N_{basic}	3.70	3.81	3.80	2.45
pKa_{acidic}	4.10	3.95	3.85	4.34
pKa_{basic}	11*	11*	11*	11*
$K_s^* (L/mol)$	-295.15	-310.38	-286.66	-290.72

* Fixed parameter.

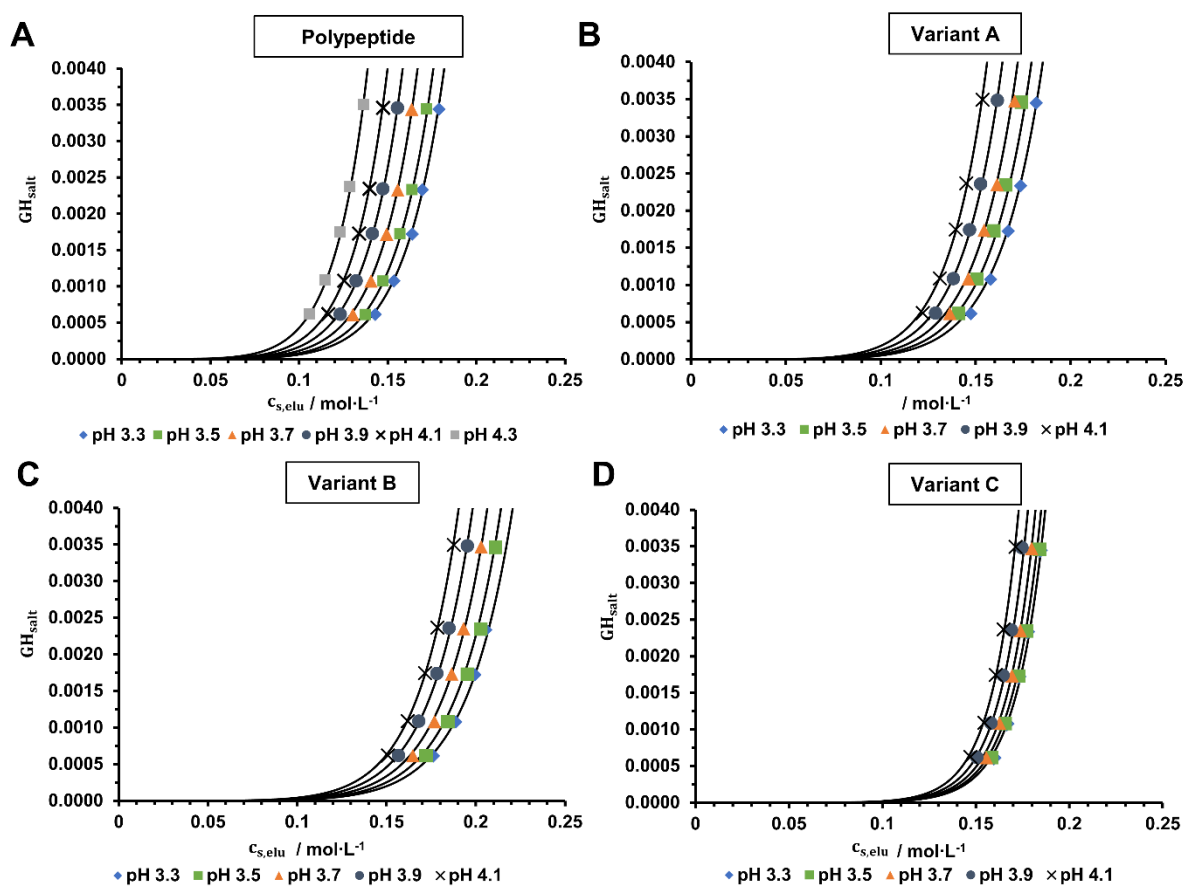


Figure 6: Normalized salt gradient slope as a function of the eluting counterion concentration ($c_{s,elu}$) at different pH values for the polypeptide (A), variant A (B), variant B (C), and variant C (D) determined with the CEX column 1. The solid lines indicate the respective $GH_{salt}/c_{s,elu}$ -function, generated with the final fit of the parameters in the linear range of the isotherm (see Table 10).

The pH dependency of the salt-dependent parameter K_S from the activity coefficient model from the protein species in solution and the characteristic charge v_i for the four sample proteins determined with the CEX column 1 are visualized in Figure 7. Due to the introduced empirical pH dependency, K_S decreases from -23.05 L/mol at pH 3.3 to -28.22 L/mol at pH 4.3 for the polypeptide species. The K_S curve for variant A has a similar starting point at pH 3.1 and a slightly higher value at pH 4.3 than the polypeptide. Variant B has, in general, a slightly higher curve. However, K_S from variant C starts from -50 L/mol and is, therefore, significantly lower compared to the other proteins. The pH dependency of the characteristic charge v_i for all three proteins is similar for the polypeptide and the two product-related impurities (variant A and variant B), whereas the shape of the curve for the polypeptide species is slightly under the curve of the two variant species. However, variant C has a significantly lower characteristic charge. In general, v_i decreases with increasing pH almost linearly.

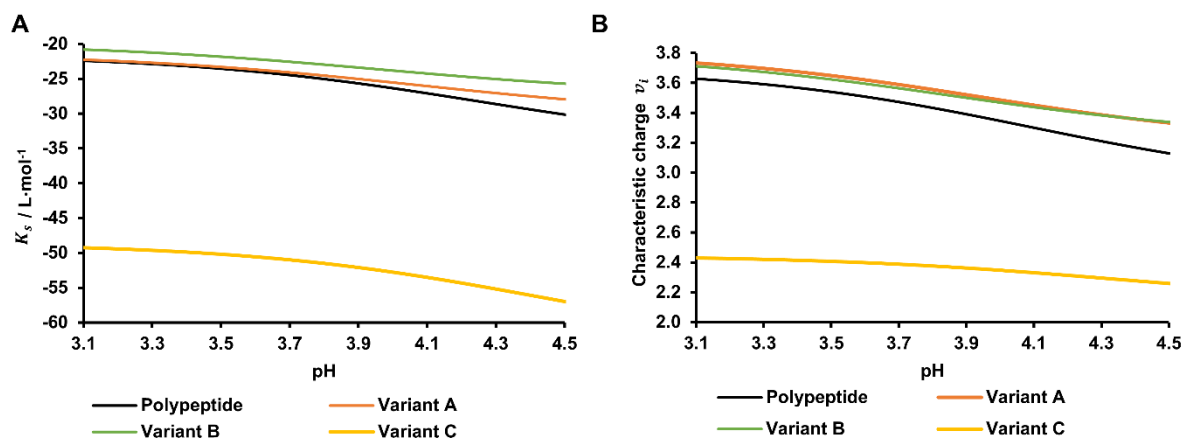


Figure 7: Parameter K_s from the asymmetric activity coefficient model (A) and the characteristic charge (B) as a function of the pH for the four sample proteins determined in the linear range with the CEX column 1.

The importance of the introduced activity coefficients used in the CEX model is investigated by comparing the modeling fits in the linear range with and without the influence of the activity coefficients. In this work, the fits from the first step of the fitting procedure without the influence of the asymmetric activity coefficient of the protein species in solution $\tilde{\gamma}_i$ were compared to the fits from the last step of the fitting procedure with the influence of the asymmetric activity coefficients for the protein species in solution $\tilde{\gamma}_i$. The difference and the correlation between the parameters for the polypeptide species are shown in Figure 8. The activity coefficient has a significant influence on the equilibrium constant K_i and the characteristic charge v_i . Notably, the equilibrium constant and the characteristic charge significantly differ based on the dependence of the asymmetric activity coefficient. The equilibrium constant changes with the influence of the asymmetric activity coefficient from about 5×10^{-3} to 2×10^{-2} to about 10 to 28 (Figure 8A). The characteristic charge changes from about 6.5 to 7.5 to about 3.2 to 3.5 (Figure 8B). However, the correlation between the parameters is linear.

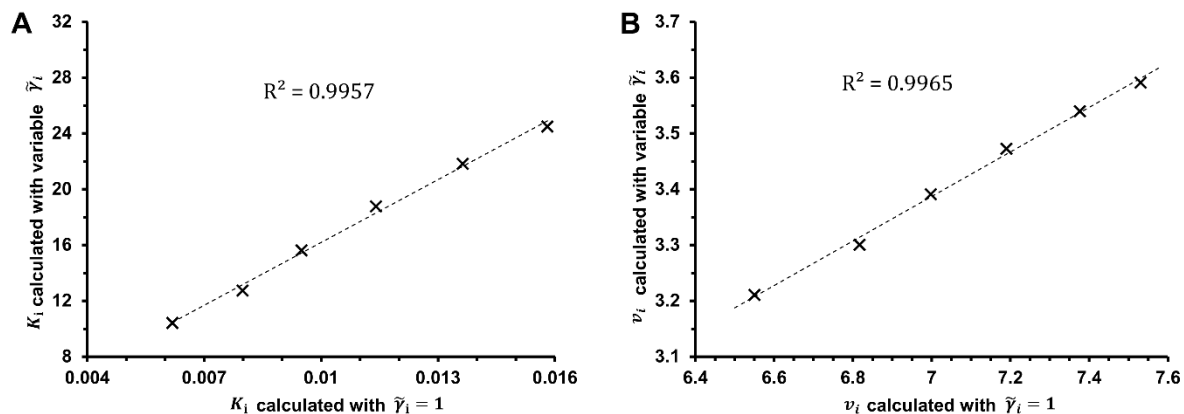


Figure 8: Influence of the asymmetric activity coefficient on (A) the equilibrium constant and (B) the characteristic charge for the polypeptide species. The asymmetric activity coefficient is calculated according to Equation (18). For ideal behavior, the activity coefficient $\tilde{\gamma}_i$ is set to one.

The influence of the two activity coefficients and the ratio of these two activity coefficients on the equilibrium constant for the four protein species used for the model calibration is shown in Figure 9. In general, the asymmetric activity coefficient of the protein species $\tilde{\gamma}_i$ is smaller than the asymmetric activity coefficient of the counterion species $\tilde{\gamma}_s$ in the experimental Na^+ concentration range for all four modeled proteins. Furthermore, it is shown that the ratio of the activity coefficients weakly depends on the pH and is dominated by $\tilde{\gamma}_i$, especially seen for variant C. Since the elution counterion concentration is between 0.12 and 0.20 mol/L, the impact on the equilibrium constant and indirectly on the characteristic charge, therefore, varies between a factor of 0.2 and 0.05 for all four protein species.

Since the model for the chromatographic CEX step should be representative in high loading situations, experiments with the polypeptide species in the non-linear range of the isotherm were performed. The parameter estimation results are shown in the next section (3.2.1.3).

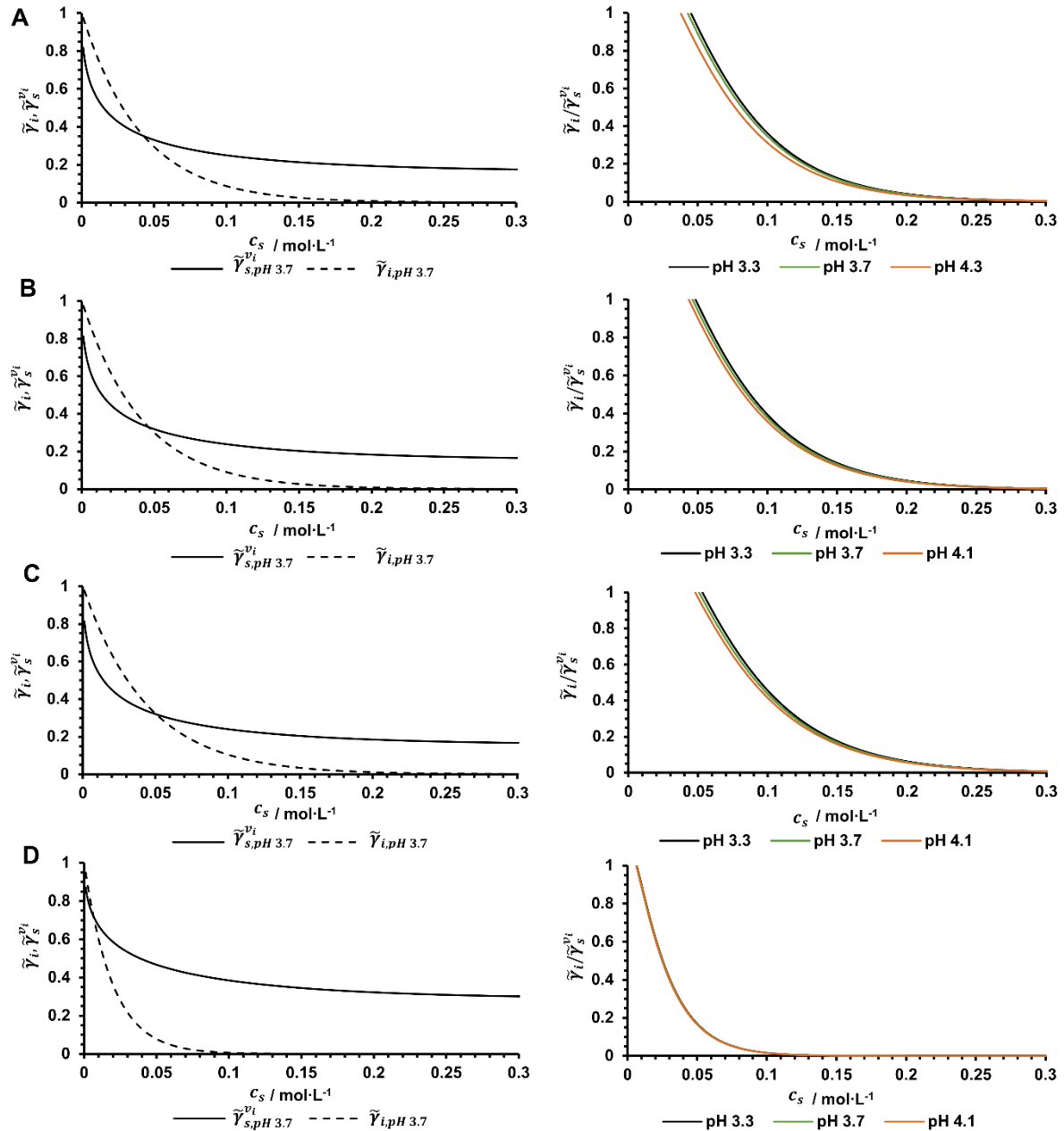


Figure 9: The asymmetric activity coefficients of the counterion and the protein in solution as a function of the counterion concentration (c_s) at pH 3.7 in the linear range and the ratio of the activity coefficients as a function of the counterion concentration at different pH values in the linear range for the polypeptide (A), variant A (B), variant B (C), and variant C (D).

3.2.1.3 Determination of model parameters in the non-linear range of the isotherm

After the parameter estimation in the linear range of the isotherm, the parameters in the non-linear range using high load experiments were determined. The applied

method for the parameter estimation was the inverse fitting method. In the inverse fitting method, the simulated UV signal, using the previously determined linear range parameter set, is fitted to the experimentally determined UV signal by varying the non-linear range model parameters. The model parameters influencing the peak shape in high loading situations are in the SAS isotherm, the equilibrium constant of the dimer layer $K_{D,i}/c$, the shielding factor of the protein species ξ_i , and the protein dependent parameter in the activity coefficient model of the protein species in solution K_p^* .

Table 11: Overview of the model parameters fitted in the non-linear range of the isotherm for the polypeptide species with the CEX column 1.

pH	Shielding factor ξ_i	$K_{D,i}/c$ (L/mol)	K_p^* (L/mol)
pH 3.3	8.30	2.2×10^5	1.046×10^6
pH 3.5	8.30	2.2×10^5	1.046×10^6
pH 3.7	8.40	2.2×10^5	1.046×10^6
pH 3.9	8.20	2.2×10^5	1.046×10^6
pH 4.1	8.10	2.2×10^5	1.046×10^6
pH 4.3	8.20	2.2×10^5	1.046×10^6

The estimated parameters for the polypeptide species in the non-linear range are given in Table 11. Since the model should be representative at different pH values from pH 3.3 to pH 4.3, pH-dependent high loading model parameters were determined. However, as shown in Table 11, $K_{D,i}/c$ was fitted to be constant and pH-independent since no thermodynamic pH dependency was introduced in the mechanistic model due to the assumption that equilibrium dimer formation in the relatively small applied pH range can be seen as constant. The values for the shielding factor of the polypeptide species show a slight variation at different pH values around 8.3. Furthermore, K_p^* was also fitted to be constant and pH-independent since this parameter, combined with the characteristic charge, describes the pH dependency for K_p according to Equation (20).

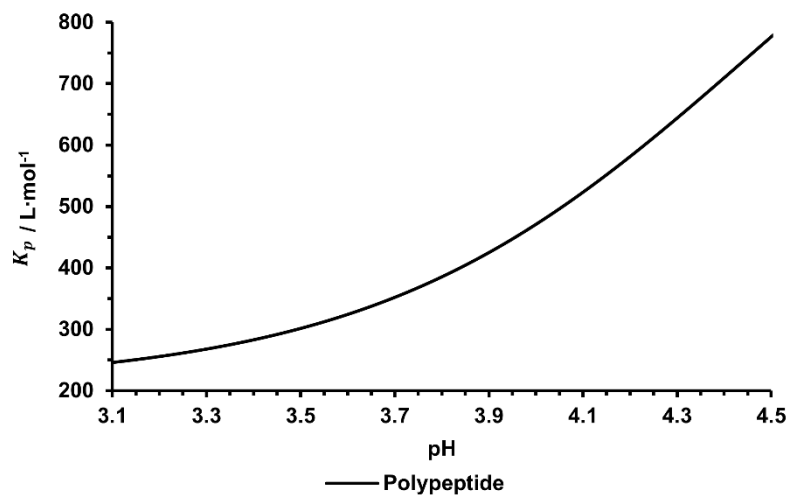


Figure 10: Parameter K_p from the asymmetric activity coefficient model as a function of the pH for the polypeptide species determined in the non-linear range with the CEX column 1.

The overall pH dependency for the protein dependent parameter K_p of the polypeptide species is shown in Figure 10. K_p ranges from about 250 L/mol to about 600 L/mol in the calibrated pH range.

There was not enough sample load available for variant A, variant B, and variant C to do experiments in high loading conditions. The non-linear range parameters for the product-related (variant A and variant B) and the process-related (variant C) impurities were first assumed to be the same as determined for the polypeptide species. In later multicomponent experiments and simulations, new non-linear range parameters were fitted for the impurity species (see section 3.2.1.5).

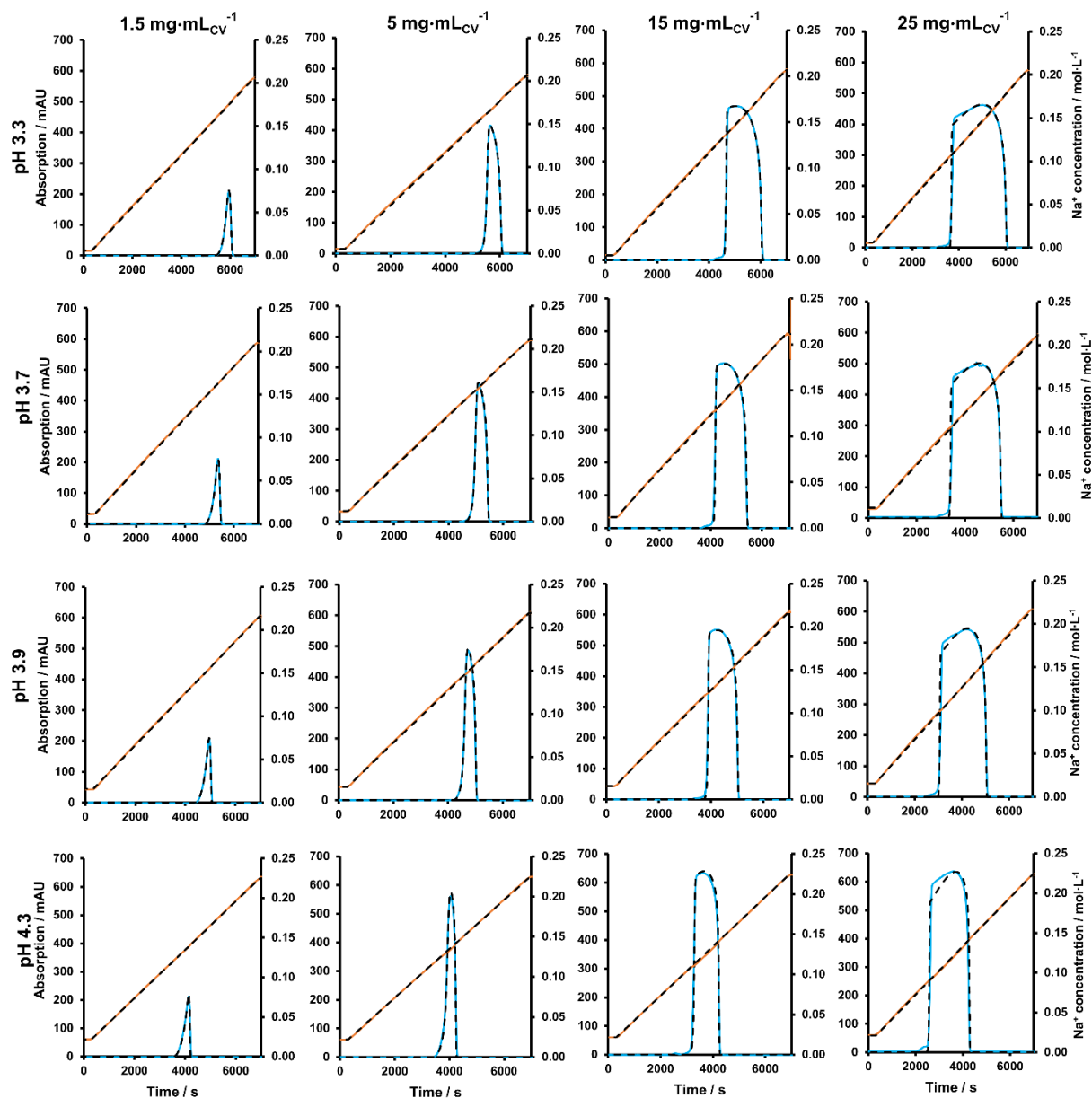


Figure 11: Simulated chromatograms compared to the experimentally determined chromatograms with the polypeptide species at different pH values and column loads for the CEX column 1. The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. The experimental UV signal was determined with a 2 mm UV cell.

Figure 11 shows the calibration runs for the polypeptide species at high loading situations and the corresponding simulated signals from the overall CEX model calculated with the determined model parameters from the linear and non-linear range. As shown in Figure 11, the developed model is able to describe the high loading situations with high accuracy using the determined parameter set. Note that in Figure 11, not all calibration experiments done in this work for the polypeptide species are shown. The

experiments at pH 3.5 and 4.1 and the loads: 0.05, 0.25, 0.5, and 0.75 mg/mL_{CV} are additionally displayed in Appendix 8.3 in Figure 41 and Figure 42. The change of the peak shape from the fronting peaks in the lower loads (anti-Langmuirian behavior) to the steep rise in the front and tailing peaks in higher loads (Langmuirian behavior) is described with high accuracy. Even the atypical peak shape in high loading situations, the peak broadening, and the flattened top with increasing load could be simulated correctly. The empirical pH dependency introduced in the asymmetric activity coefficient model describes the influence of the pH with high precision, at least in the tested pH range. However, the simulations with a column load of 25 mg/mL_{CV} slightly differ from the experimental UV elution signal in front of the peak.

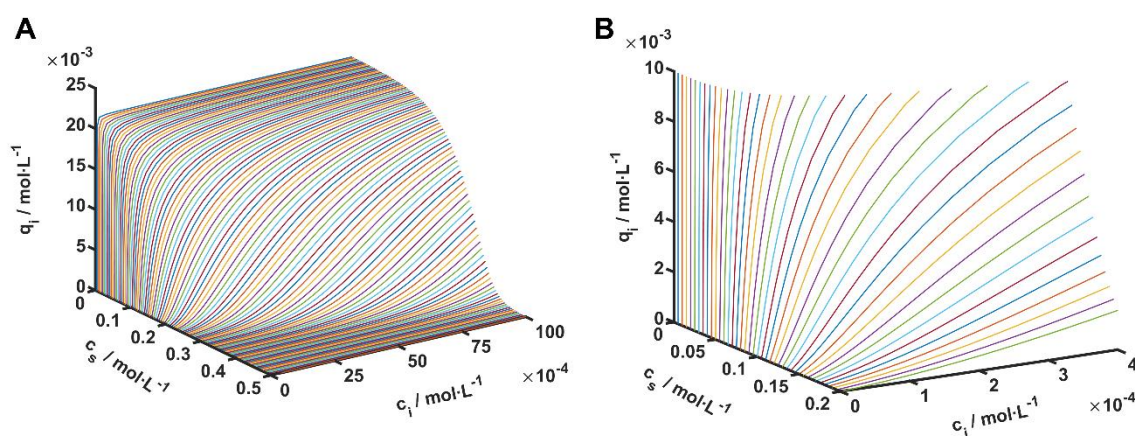


Figure 12: (A) 3D-Plot of the SAS isotherm used in this work. The bound polypeptide concentration is plotted against the polypeptide concentration in solution (c_i) at different counterion concentrations (c_s). The values of the isotherm were calculated at pH 3.7 for the polypeptide species. (B) 3D-Plot of the isotherm in the experimental concentration range.

Figure 12 shows a 3D-Plot of the SAS isotherm at different counterion concentrations using the determined isotherm parameters for the polypeptide species in Table 10 and Table 11. The bound protein species increases with increasing solute protein species with a sigmoid-shaped curvature till a maximum bounded concentration. As shown in Figure 12, the isotherm shape strongly depends on the salt concentration. With increasing salt concentration, the polypeptide binds less to the adsorbent. Furthermore, the distribution of calculated bound monomeric and dimeric polypeptide species for different counterion concentrations ranging from strong binding conditions ($c_s = 0.01 \text{ mol/L}$) to elution conditions ($c_s = 0.1$ to 0.2 mol/L) is shown in Figure 13. In general, at linear range conditions the monomer-dimer equilibrium is shifted to the

monomer species and shifts towards the dimer species with increasing polypeptide concentration.

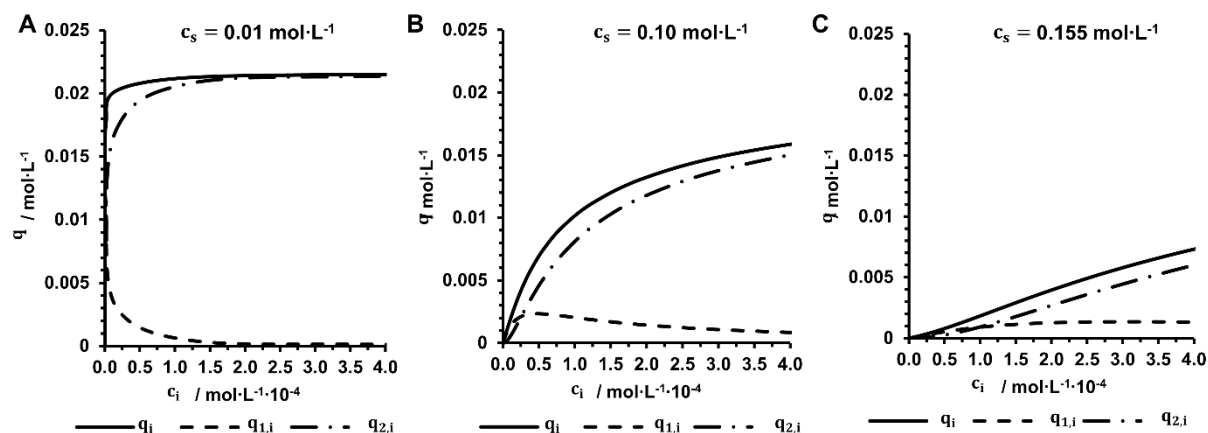


Figure 13: Isotherm plot with values for the total protein (q_i), the monomer layer ($q_{1,i}$), and the dimer layer concentration ($q_{2,i}$) at different counterion concentrations (c_s): (A) 0.01 mol/L, (B) 0.10 mol/L, and (C) 0.155 mol/L NaCl. The values of the isotherm were calculated for the polypeptide species at pH 3.7.

3.2.1.4 Model validation

In order to validate the developed CEX model calibrated with the CEX column 1 for the polypeptide species in the previous sections, two gradient elution validation experiments, shown in Figure 14, and several isocratic elution experiments, shown in Figure 15, were done. Both gradient elution validation experiments were done at non-calibrating conditions. The first gradient elution experiment was done at pH 4.0 and 7.5 mg/mL_{CV} column load (Figure 14A). The second gradient elution experiment was done as a dual gradient elution with a pH gradient from pH 4.3 to pH 3.4 and an additional salt gradient from 0 mol/L to 0.2 mol/L sodium chloride. As shown in Figure 14, both experiments could be successfully predicted with model simulations. The simulated UV signal describes the experimentally determined UV signal with high accuracy. The validation experiments were done with a replacement column for the CEX column 1 with the same column dimension packed with the same batch of the stationary phase since the original CEX column 1 has broken. Nevertheless, the column-specific parameters for the replacement column were determined and applied for the model simulations. The experimentally determined total porosity for the replacement column was ε_t : 0.736, consequently the particle porosity was calculated to be ε_p : 0.594, and

the ionic capacity based on the packed bed was determined to be Λ_{pack} : 0.105 mol/L. The interstitial porosity was not again determined experimentally with Blue Dextran for the replacement column due to the observed and discussed inaccuracies (see section 4.2.3.3). Therefore, the interstitial porosity was assumed to have the same value as the CEX column 1 (ϵ_0 : 0.349). The total porosity from the replacement column has a slightly lower value than the CEX column 1. However, as already shown in the column characterization for the CEX columns (section 3.1.1), this value is in the range of the determined uncertainty range for the column volume. The ionic capacity is also in a similar range. For the dual gradient validation experiment, the mean value of the shielding factors in the pH range (Table 11) of ξ_i : 8.3 was chosen in the simulation.

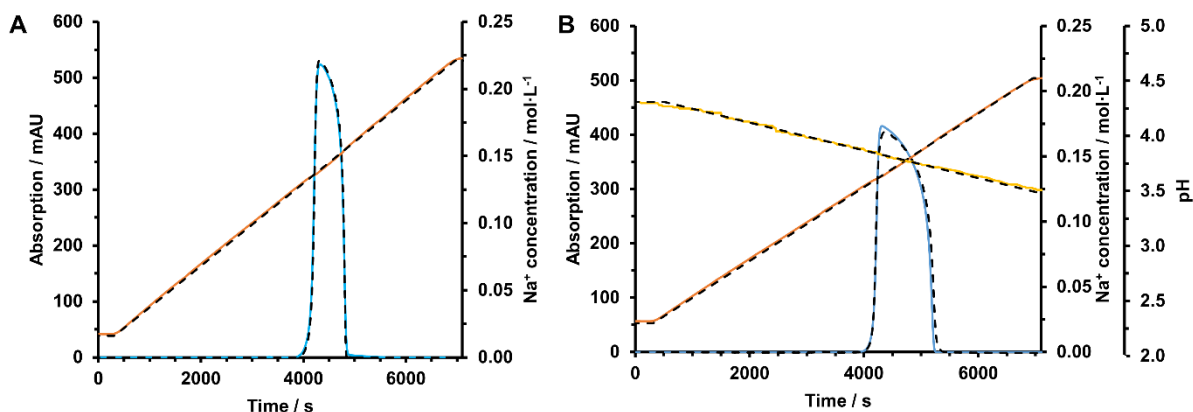


Figure 14: (A) Salt gradient validation experiment at pH 4.0. (B) Dual gradient validation experiment. pH gradient from pH 4.3 to pH 3.4 combined with a salt gradient from 0 to 0.2 mol/L NaCl. The column load in both experiments was 7.5 mg/mL_{CV} pure polypeptide. The orange, yellow, and blue lines indicate the experimental sodium concentration, pH, and UV signal. The dashed lines indicate the corresponding simulated signals.

Since the CEX model, developed in the previous section, was calibrated using LGE experiments, the ability of a transfer to isocratic elutions was tested and simultaneously considered as additional model validation. Figure 15 shows the comparison of the experimentally determined and the simulated isocratic validation experiments from the elution behavior investigation shown in Figure 4 (section 3.2.1.1) with a column load of 0.05 mg/mL_{CV} to 4 mg/mL_{CV}. As shown, the model is able to describe the whole isocratic elution data set with high accuracy. In Figure 15, the simulated signals were corrected with a marginal change in the elution salt concentration range of ± 0.0035 mol/L Na⁺. An almost linear dependency of the salt correction with the column load could be observed and is discussed in section 4.2.1.3.

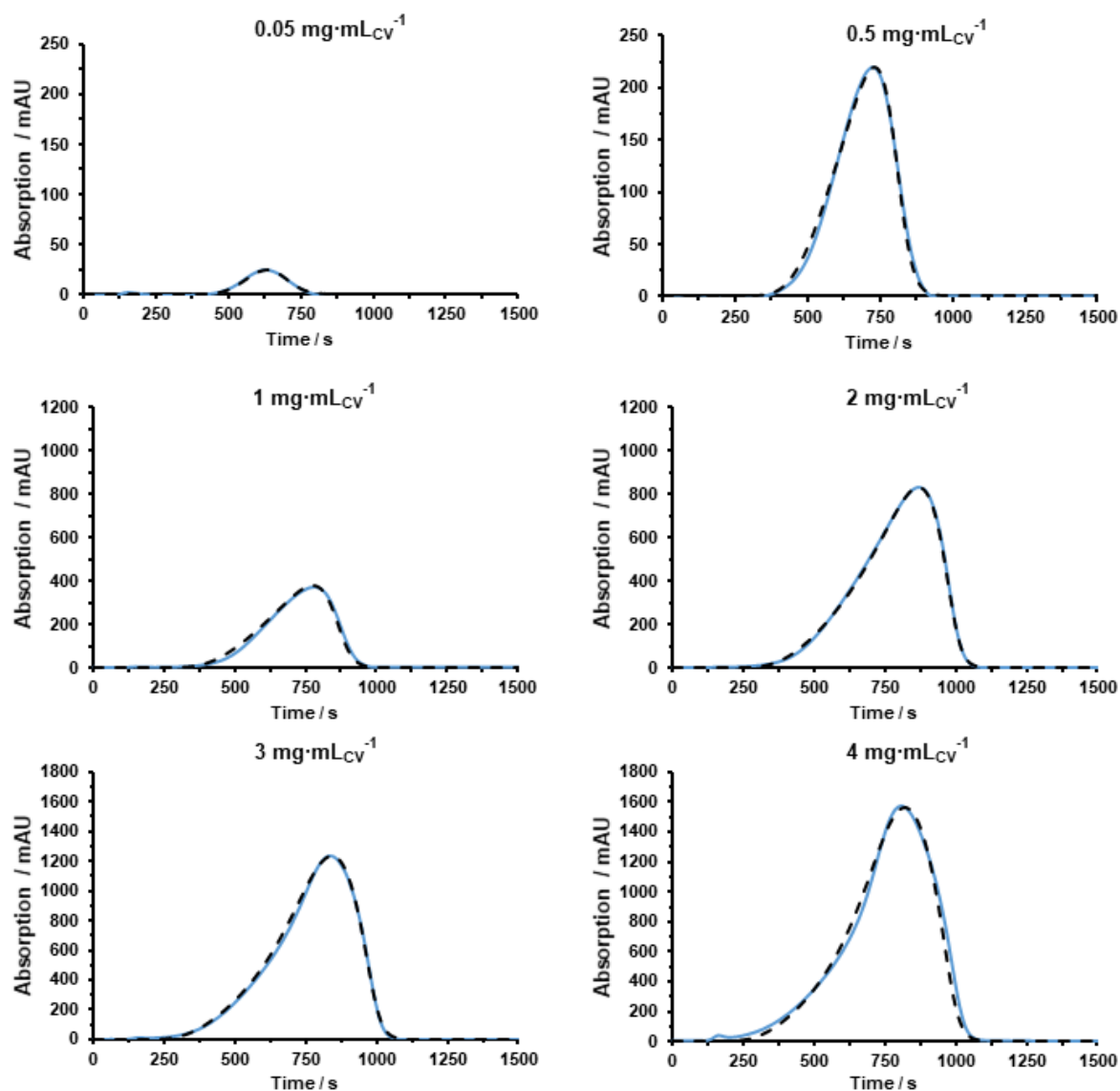


Figure 15: Simulated isocratic elution chromatograms compared to the experimentally determined chromatograms with the polypeptide species at pH 3.7. The solid blue and dashed black lines indicate the experimental and simulated UV signals, respectively. The experimental UV signal was determined with a 10 mm UV cell.

3.2.1.5 Multicomponent experiments and simulation with CEX column 1

The previous section showed a successful model development with the CEX column 1 for the four sample proteins in the analytical loading range and for the polypeptide species at high loading conditions. Since this model should represent a multicomponent chromatographic step at linear and non-linear range conditions, different high load experiments with multicomponent samples were done.

Figure 16, Figure 17, and Figure 18 show salt gradients performed on the CEX column 1 with a multicomponent sample solution including the polypeptide species, the product-related impurity variant B, and the process-related impurity variant C (Figure 16) and a sample with the polypeptide species and the two product-related impurities variant A and variant B (Figure 17 and Figure 18). The peak of the polypeptide species shown in Figure 16 shows an anti-Langmuirian peak shape as the load was at moderate non-linear range conditions. The peaks of the impurities are relatively symmetrical since the column load of the impurity species was in the linear range. The simulated signals calculated with the developed model for the three protein species fit with high accuracy to the experimentally determined chromatogram. Figure 17 and Figure 18 show experiments with the impurity variants variant A and variant B. As shown in the figures, the variant A elutes in the front of the polypeptide peak, and the determined parameters for variant A are able to describe the elution behavior with high precision in a multicomponent sample. Variant B partially coelutes with the polypeptide species, and most of variant B elutes behind the polypeptide peak. However, the calibrated model is able to describe the elution behavior with high accuracy. At the beginning of the modeling procedure, the non-linear range parameters for the product-related (variant A and variant B) and process-related (variant C) impurities were assumed to be the same as determined for the polypeptide species. Nevertheless, in Figure 16, Figure 17, and Figure 18, the high load parameters for the impurity variant B was adapted in order to describe the experimentally determined signal of the multicomponent experiments. The new non-linear range parameter sets for the four protein species are given in Table 12. The adjustment was possible by only varying the equilibrium constant for the dimer layer $K_{D,i}/c$ for one impurity species (variant B).

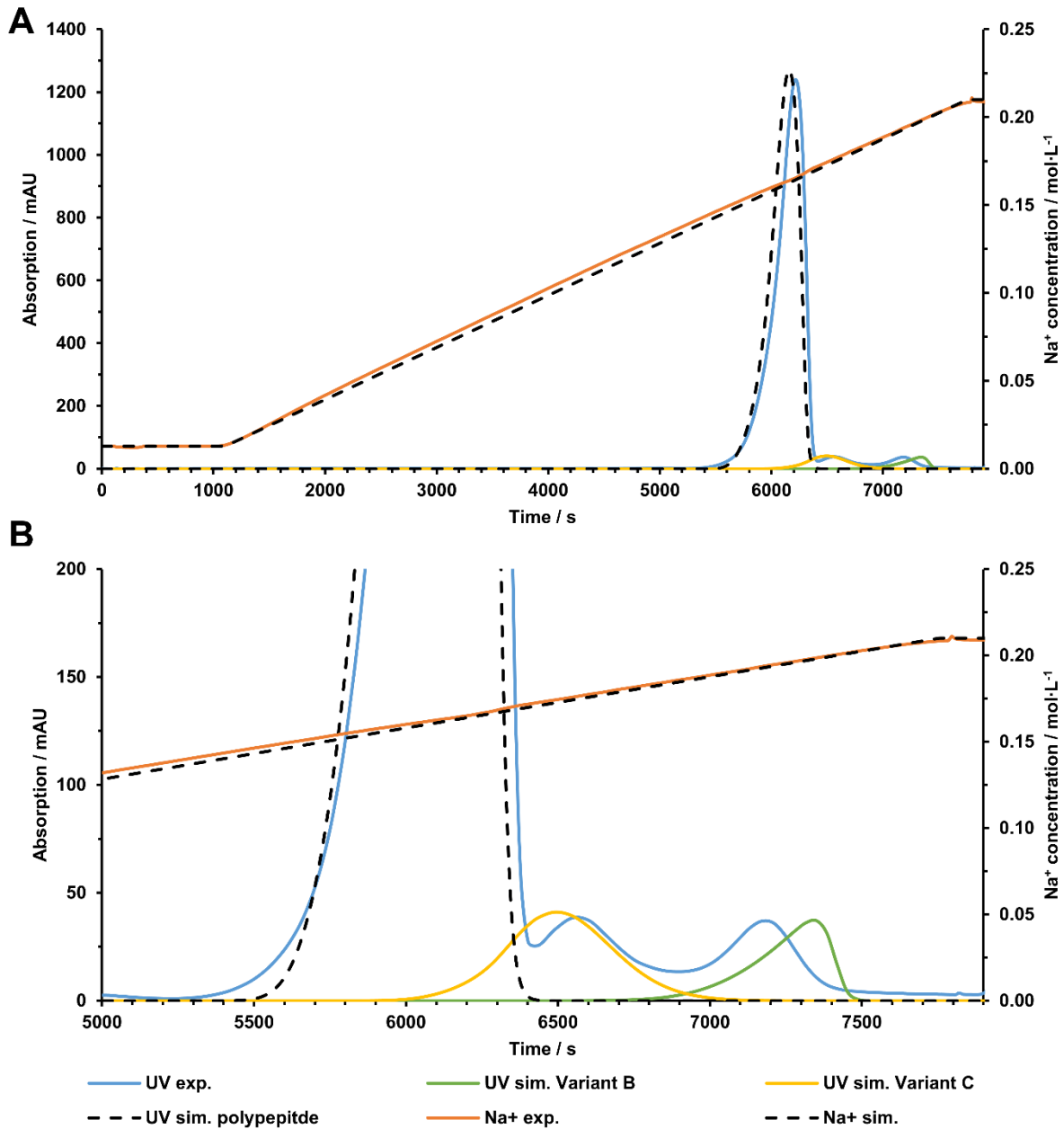


Figure 16: Salt gradient elution with the CEX column 1 and a three-component sample solution, which included the polypeptide species, variant B, and variant C. The column load was 1.5 mg/mL_{CV} polypeptide, 0.05 mg/mL_{CV} variant B, and 0.08 mg/mL_{CV} variant C. The solid blue line indicates the experimentally determined UV signal, and the dashed black and the colored solid lines are the simulated signals.

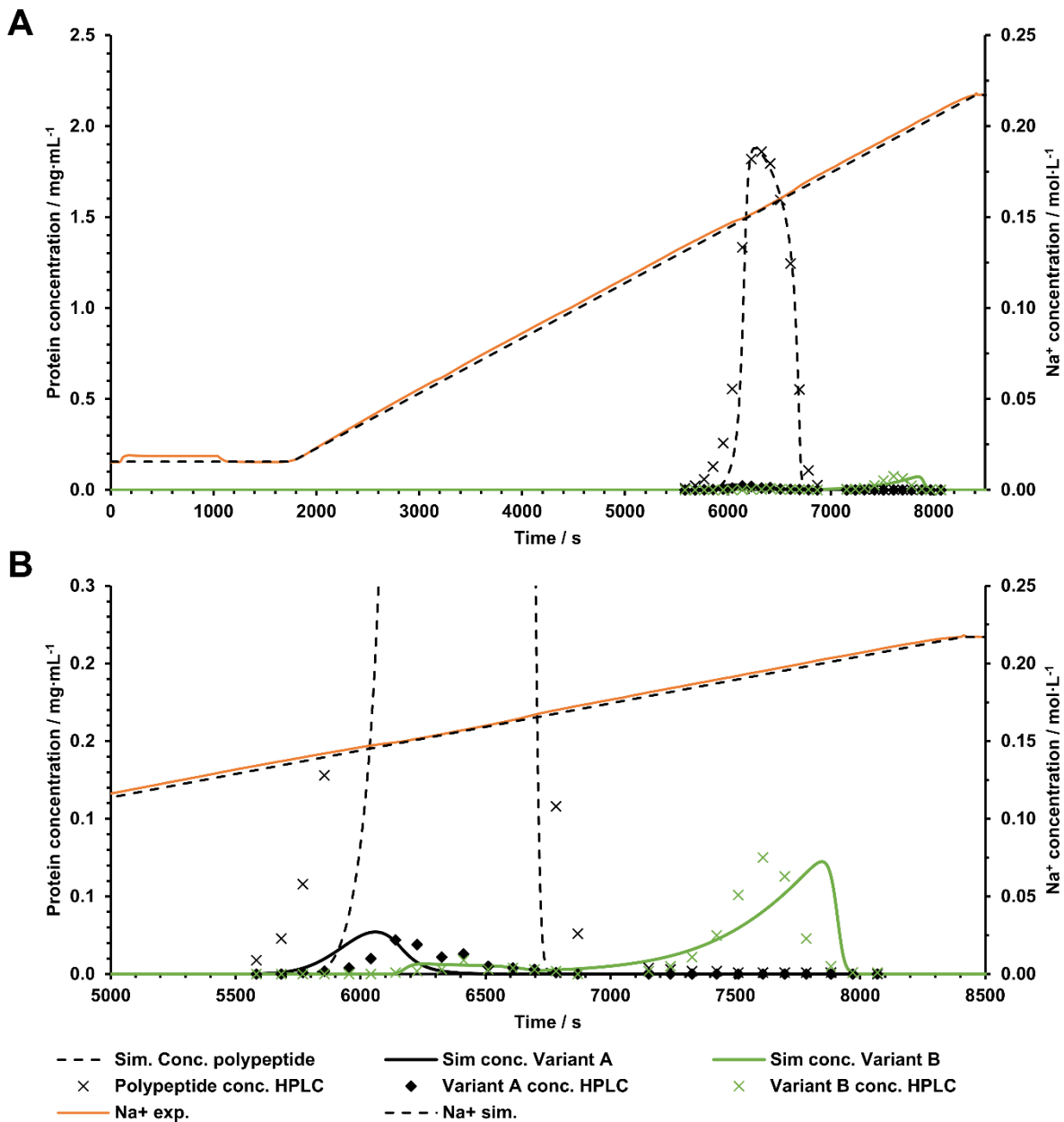


Figure 17: Salt gradient elution with the CEX column 1 and a three-component sample solution, which included the polypeptide species, variant A, and variant B. The column load was $6 \text{ mg}/\text{mL}_{\text{CV}}$ polypeptide, $0.05 \text{ mg}/\text{mL}_{\text{CV}}$ variant A, and $0.18 \text{ mg}/\text{mL}_{\text{CV}}$ variant B. The data points represent the protein concentration experimentally determined by HPLC. The corresponding simulated signals are indicated with the colored solid line, and the solid orange line indicates the experimental salt gradient.

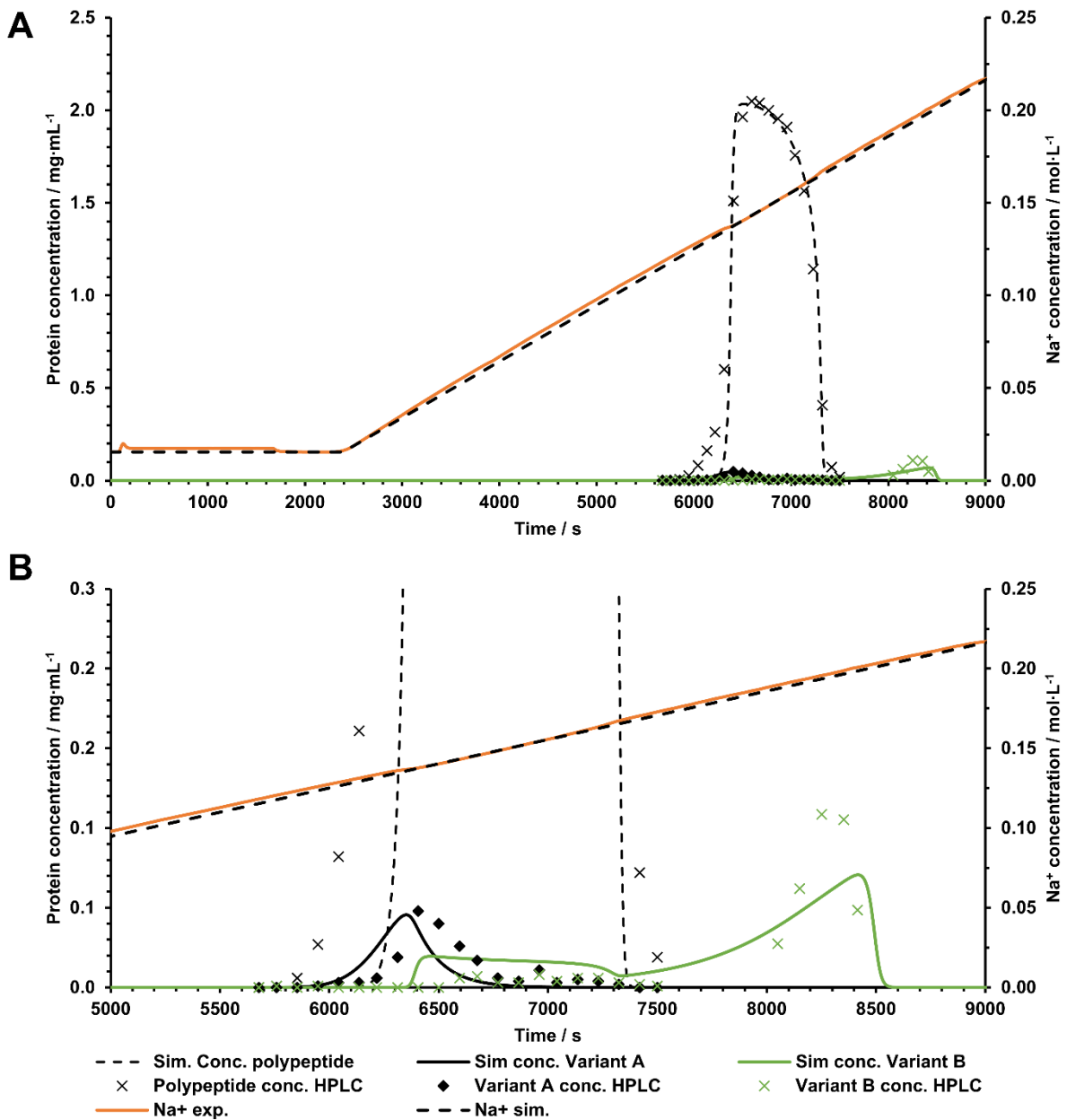


Figure 18: Salt gradient elution with the CEX column 1 and a three-component sample solution, which included the polypeptide species, variant A, and variant B. The column load was $10 \text{ mg}/\text{mL}_{\text{CV}}$ polypeptide, $0.08 \text{ mg}/\text{mL}_{\text{CV}}$ variant A, and $0.3 \text{ mg}/\text{mL}_{\text{CV}}$ variant B. The data points represent the protein concentration experimentally determined by HPLC. The corresponding simulated signals are indicated with the colored solid line, and the solid orange line indicates the experimental salt gradient.

Table 12: Overview of the model parameters fitted in the non-linear range of the isotherm for the four protein species with the CEX column 1.

Parameter	Polypeptide	Variant A	Variant B	Variant C
Shielding factor $\xi_{i,pH\ 3.7}$	8.40	8.40	8.40	8.40
$K_{D,i}/c$ (L/mol)	2.2×10^5	2.2×10^5	2.0×10^6	2.2×10^5
K_p^* (L/mol)	1.046×10^6	1.046×10^6	1.046×10^6	1.046×10^6

For the investigation, if the developed model is also able to describe a multicomponent sample solution, including the product and partially unknown impurity species, an experiment with such a sample solution was performed and compared to simulated signals. Figure 19 shows a salt gradient performed on the CEX column 1 with a multicomponent sample solution, including the polypeptide species as the product of interest and an unknown amount of impurities, which include the product-related impurities variant A and variant B. It is shown that the polypeptide species in the multicomponent solution exhibits a similar peak shape in the gradient as observed with a single component elution. Generally, the major part of the impurities elutes in the front and rear of the polypeptide peak and smear into the middle of the main peak. Variant A elutes in the front, and variant B in the middle and the rear of the main peak. It is also shown that the product-related impurities (variant A and variant B) only represent a small part of the overall impurities in the multicomponent sample. Furthermore, Figure 19 shows that the offline HPLC data fit the experimentally determined UV signal, which was recalculated from UV cell calibration done with the polypeptide species.

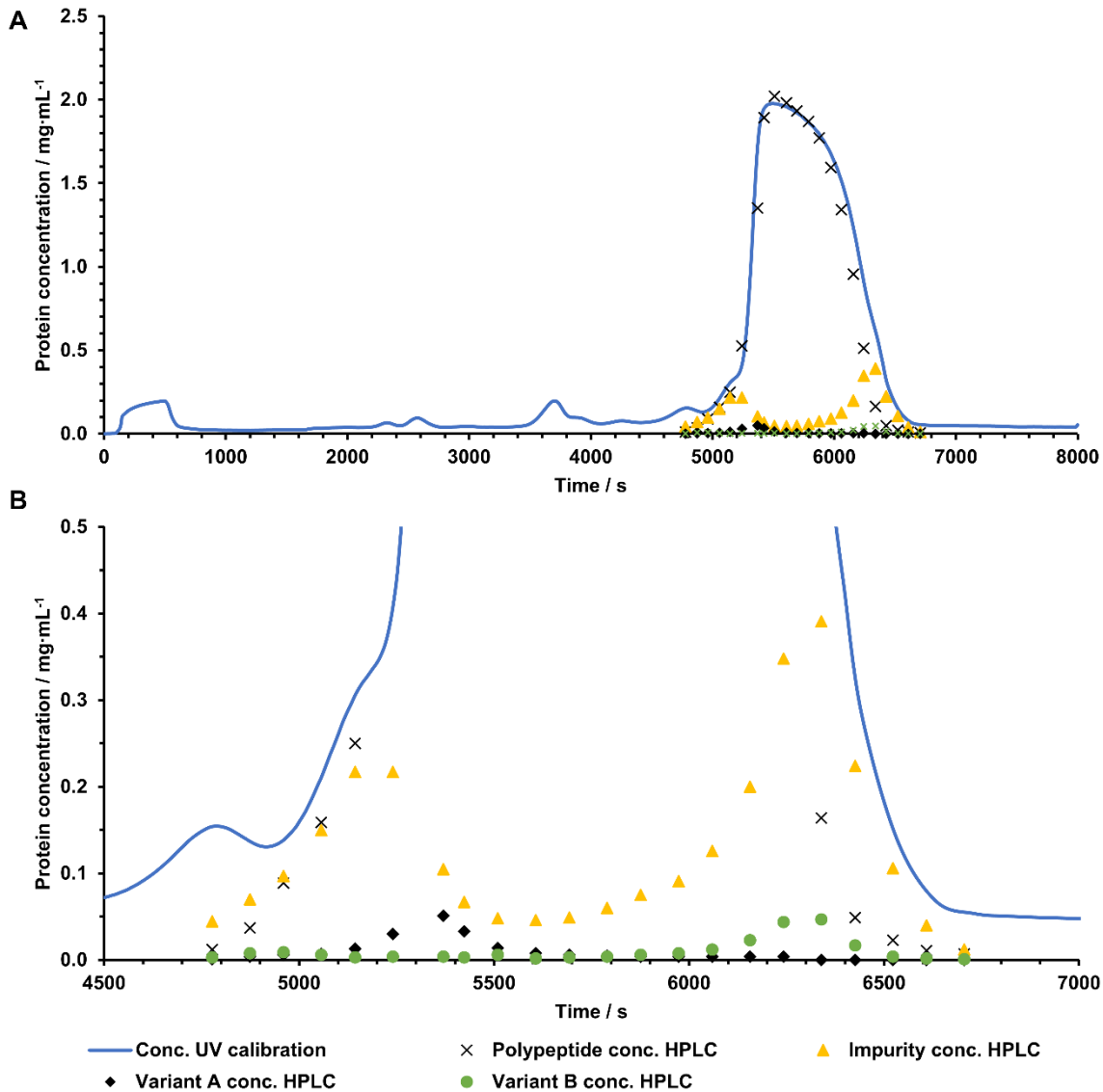


Figure 19: Salt gradient elution with the CEX column 1 and a multicomponent sample solution, which included the polypeptide species, variant A, variant B, and further non-specified impurities. The column load was 10 mg/mL_{CV} polypeptide and an unknown amount of impurities. The solid blue line indicates the protein concentration profile recalculated from the UV signal using a UV cell calibration (data not shown). The colored points represent the offline HPLC protein concentration data points.

In order to investigate if the developed model for the four sample proteins is able to simulate the high load multicomponent experiment, the experimental data from Figure 19 was compared with the simulated signal from the model by using the experimental conditions for the polypeptide species, variant A and variant B in the model. The result of the three-component simulation and the comparison with the experimental multicomponent solution run is displayed in Figure 20. The polypeptide peak has the same shape as the experimental peak, determined with the offline HPLC data. Nevertheless, the simulated polypeptide peak does not fit precisely the offline data since the peak is

slightly shifted to the right-hand side. The simulated signal from variant A describes the offline HPLC signal with high accuracy. However, the simulated signal from variant B represents not the experimental signal.

To further investigate the influence of the impurities on the simulated signal, another simulation, with a higher loaded amount of variant B to mimic the experimental conditions, was conducted. The comparison of the simulated and the experimental signal is shown in Figure 21. The higher amount of variant B, which elutes in the middle and the rear of the main peak, significantly influences the overall elution behavior of the simulated signals. The simulation of the main peak is shifted to the left, and now the simulated polypeptide signal fits the experimentally determined signal. The Simulated signal from variant A is also shifted to the left through the higher amount of variant B. However, the simulated signal still fits the experimental signal.

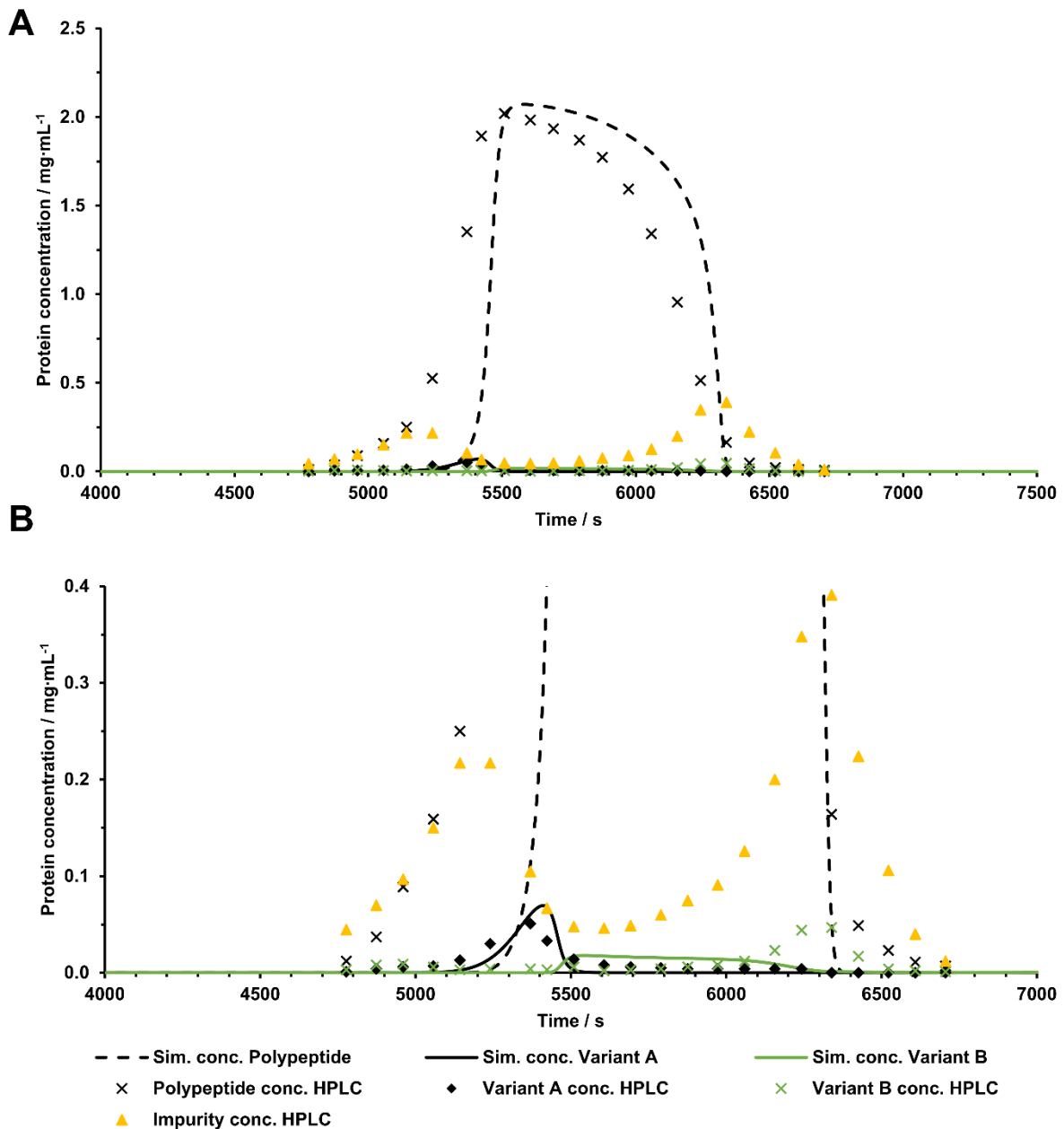


Figure 20: Simulated salt gradient elution with the CEX column 1 and a multicomponent sample solution, which included the polypeptide species, variant A, variant B, and further non-specified impurities. The column load was $10 \text{ mg}/\text{mL}_{\text{CV}}$ polypeptide and an unknown amount of impurities. The colored solid and dashed lines indicate the simulated protein concentration, and the colored points represent experimental offline HPLC protein concentration data points. The simulated column load for variant A and variant B was $0.075 \text{ mg}/\text{mL}_{\text{CV}}$.

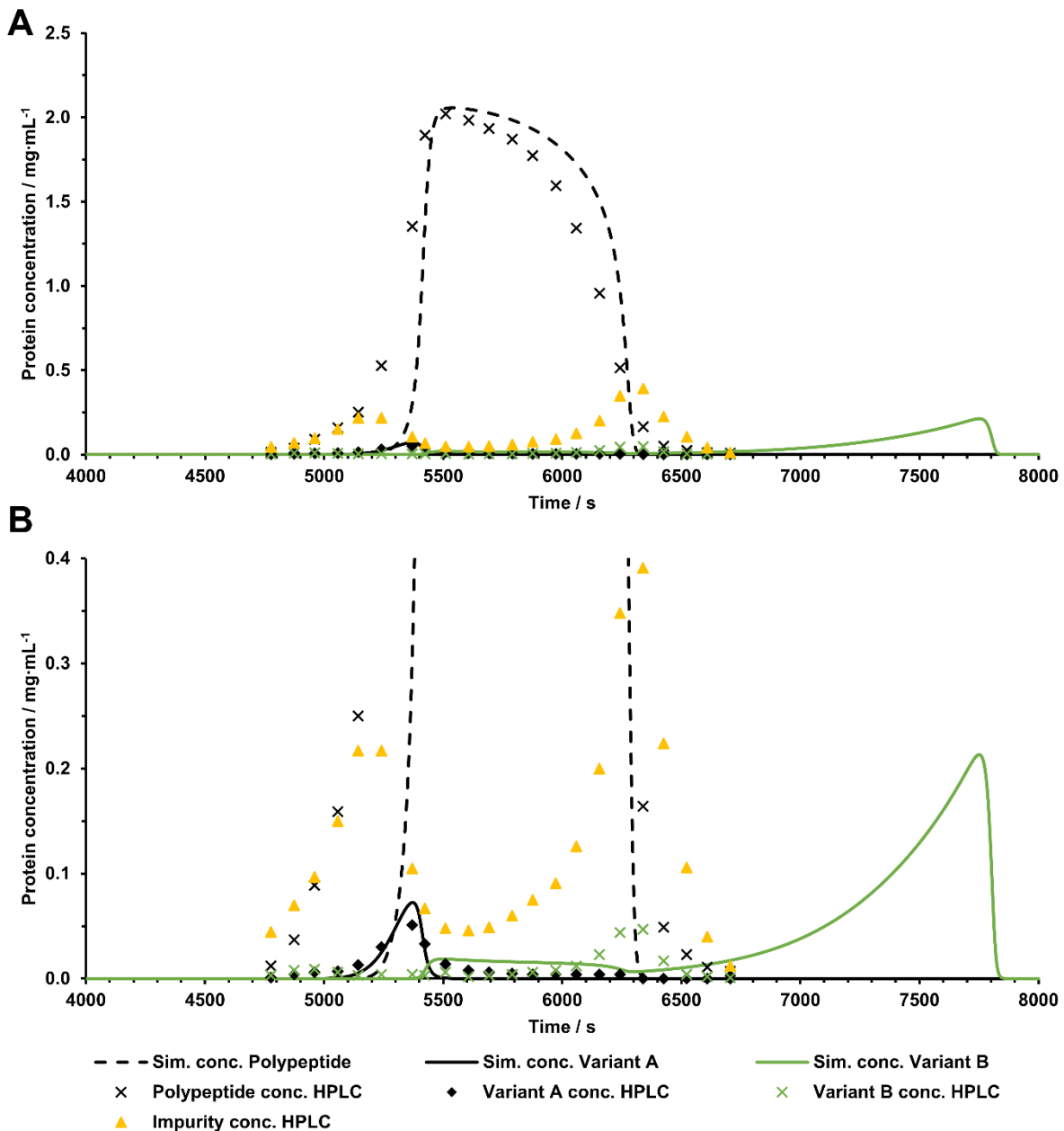


Figure 21: Simulated salt gradient elution with the CEX column 1 and a multicomponent sample solution, which included the polypeptide species, variant A, variant B, and further non-specified impurities. The column load was 10 mg/mL_{CV} polypeptide and an unknown amount of impurities. The colored solid and dashed lines indicate the simulated protein concentration, and the colored points represent experimental offline HPLC protein concentration data points. The simulated column load for variant A was 0.075 mg/mL_{CV} , and for variant B 0.75 mg/mL_{CV} .

3.2.2 CEX model calibration at small scale (TOYOPEARL S-650S)

To investigate the application of the developed CEX model to other CEX resin types and to further validate the model, experiments with the polypeptide species on another CEX column, including another type of CEX resin (TOYOPEARL GigaCap S-650S),

were performed. The elution behavior of the polypeptide on the TOYOPEARL S-650S column was investigated, and model parameters applying the same CEX model for the column were calibrated in the linear and non-linear range of the isotherm.

3.2.2.1 Investigation of the polypeptide elution behavior

As mentioned in the previous section, a loading series performed with a relatively pure sample solution is an appropriate method to make a statement about the general elution behavior of the sample protein and the curvature of the adsorption isotherm.

Figure 22 shows a comparison of the loading series with the polypeptide species conducted on the TOYOPEARL S-650S column (Figure 22A) and the CEX column 1 (Figure 22B). Note that all experimentally determined UV signals are in the linear range of the UV cell since the UV signals were recorded with a 2 mm UV cell and recalculated to a 10 mm UV cell. The polypeptide exhibits the same anti-Langmuirian to Langmuirian elution behavior as observed with the CEX column 1. The anti-Langmuirian elution behavior is still present at a load of 10 mg/mL_{cv} on the TOYOPEARL S-650S column, whereas the anti-Langmuirian elution behavior on the CEX column only is present up to a load of about 1.5 mg/mL_{cv}. Furthermore, the polypeptide elution peak on the TOYOPEARL S-650S column is significantly higher than on the CEX column 1. Additionally, the loading capacity is markedly higher. The eluting salt concentration in the gradient cannot be compared with each other since Buffer B in the experiment from Figure 22A contained more sodium chloride. Therefore, the gradient slope in both loading series is slightly different.

Since the polypeptide has the same elution behavior on the TOYOPEARL S-650S column as on the CEX column 1, the same model used in the previous section for the CEX column 1 should be usable for the description of the elution behavior of the polypeptide on the new column.

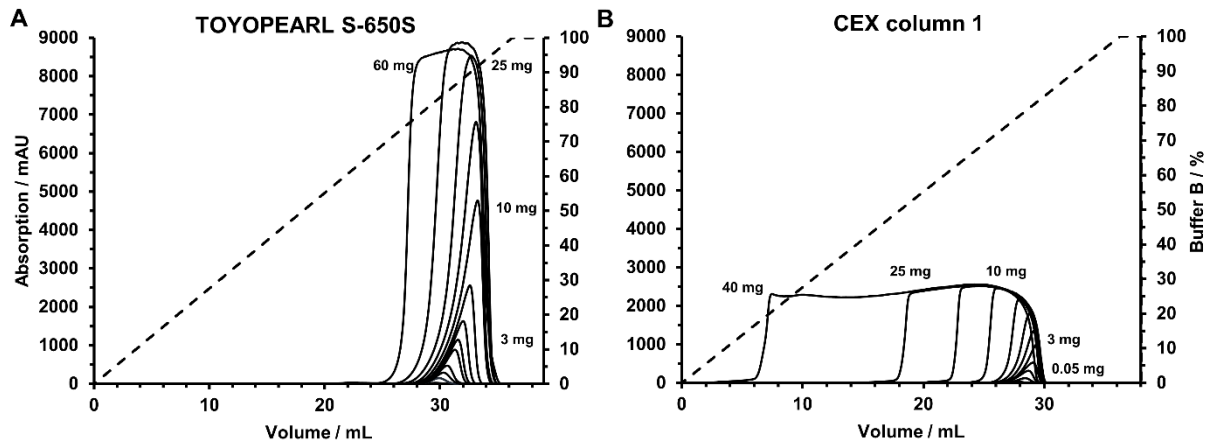


Figure 22: Loading series with the polypeptide at pH 3.7 for the TOYOPEARL S-650S column (A) and the CEX column 1 (B). The load ranges from 0.05 mg/mL_{CV} to 60 mg/mL_{CV}. The smallest peak indicates the lowest load up to the highest load with the broadest and highest peak. The black dashed line represents the amount of Buffer B in the gradient. Note that the high loading experiments were determined with a 2 mm UV cell and recalculated to 10 mm UV cell signals. All experiments were performed in the linear range of the UV cell.

3.2.2.2 Determination of model parameters in the linear and non-linear range of the isotherm

For the determination of the model parameters in the linear range, the same approach as described in section 3.2.1.2 was used. The only difference was that LGE experiments were only performed at pH 3.7, and no pH dependence was investigated since these experiments were done for a quick applicability test for other CEX columns and were intended to serve as additional model validation.

Figure 23 compares the plotted normalized gradient slope of the salt gradients GH_{salt} against the eluting counterion concentration $c_{s,elu}$ for the TOYOPEARL S-650S column and the CEX column 1 at pH 3.7. The eluting counterion concentrations for the TOYOPEARL S-650S column are generally higher than the concentrations of the CEX column 1. Table 13 shows the, in section 4.2.2.2 discussed, corresponding fitted model parameters in the linear range of the adsorption isotherm for the TOYOPEARL S-650S column. The solid line in Figure 23 represents the curve calculated with Equation (22) using the fitted parameters in Table 13 and fits well with the experimentally determined data points.

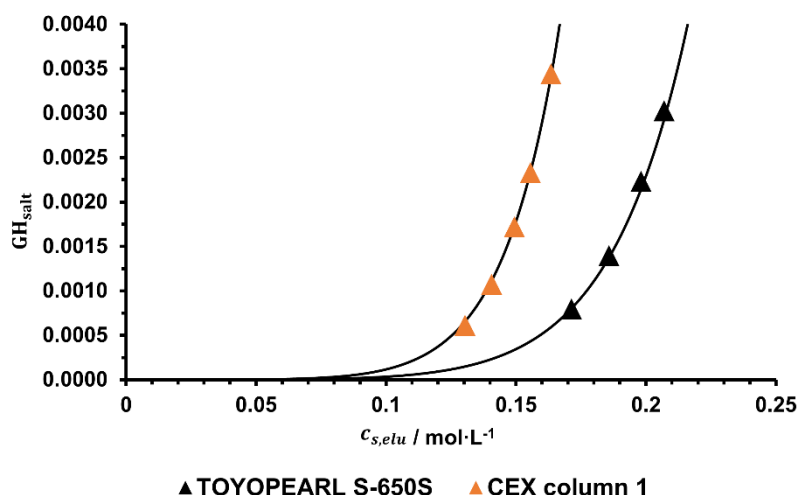


Figure 23: Comparison of the normalized salt gradient slope as a function of the eluting counterion concentration ($c_{s,elu}$) at pH 3.7 for the polypeptide species determined with the TOYOPEARL S-650S column and the CEX column 1. The solid lines indicate the respective GH_{salt} -function, generated with the final fit of the parameters in the linear range of the isotherm (see Table 13).

Table 13: Comparison of the fitted model parameters in the linear range of the isotherm for the polypeptide species determined with the TOYOPEARL S-650S column and the CEX column 1.

Parameter	TOYOPEARL S-650S	CEX column 1
$\Delta G_s^0/RT$	0.21	2.25
$\Delta G_i^0/RT$	0.77	4.87
N_{acidic}	0.42	0.80
N_{basic}	3.48	3.70
$pK_{a_{acidic}}$	3.83	4.10
$pK_{a_{basic}}$	11*	11*
K_s^* (L/mol)	-190.00	-295.15

* Fixed parameter.

The calculated pH dependency of K_s and the characteristic charge v_i for the polypeptide species determined with both CEX columns is visualized and compared in Figure 24. The salt-dependent parameter K_s starts for the TOYOPEARL S-650S column at about -16 L/mol and decreases with increasing pH, as observed for the CEX column 1,

whereas the CEX column 1 has a slightly lower K_s value. The characteristic charge v_i is for the CEX column 1 slightly higher at pH 3.1, but both models for the CEX columns end in a v_i value of about 3.2 and in both models, the characteristic charge decreases almost linearly with increasing pH in the calculated pH range. Only experiments at pH 3.7 were done for the TOYOPEARL S-650S column. Therefore, the calculated pH dependency of this model on that column is only of limited significance. However, since the same fitting strategy for the CEX column 1 was used, the values are not entirely fitted empirically.

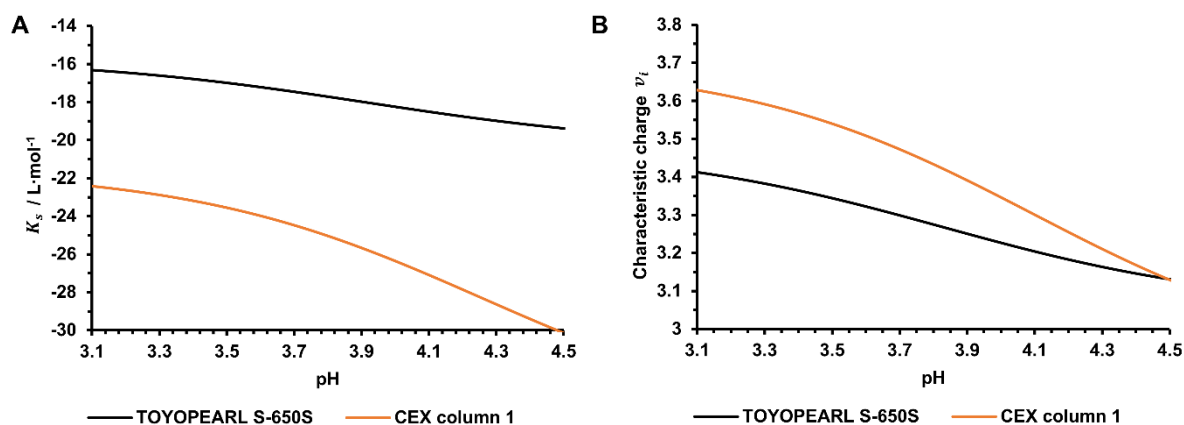


Figure 24: Comparison of the parameter K_s from the asymmetric activity coefficient model (A) and the characteristic charge (B) as a function of the pH for the polypeptide species determined in the linear range with the TOYOPEARL S-650S column and the CEX column 1.

The influence of the two activity coefficients and the influence of the ratio of these two activity coefficients on the equilibrium constant for the polypeptide species for both CEX columns is shown in Figure 25. In general, the asymmetric activity coefficients $\tilde{\gamma}_i$ and $\tilde{\gamma}_s$ in the model for the TOYOPEARL S-650S column are higher than the asymmetric activity coefficients in the model for the CEX column 1. The ratio of both activity coefficients is also shifted to higher salt concentrations compared to the model for the CEX column 1. However, all activity coefficients are in a similar range and do not differ much.

Overall, the model parameters in the linear range of the adsorption isotherm for the TOYOPEARL S-650S column are comparable with the parameters for the model from the CEX column 1 and are in a similar range.

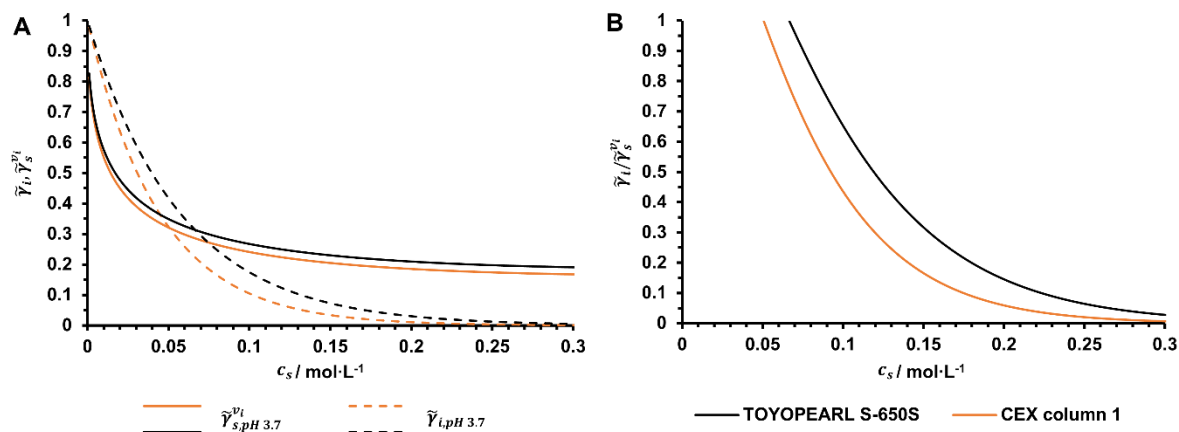


Figure 25: (A) Comparison of the asymmetric activity coefficients of the counterion and the protein in solution as a function of the counterion concentration (c_s) at pH 3.7 in the linear range for the polypeptide species determined with the TOYOPEARL S-650S column (black) and the CEX column 1 (orange). (B) Comparison of the ratio of the activity coefficients as a function of the counterion concentration at different pH 3.7 in the linear range for the polypeptide species.

After the parameter estimation in the linear range of the isotherm, the parameters in the non-linear range using high load experiments were determined. Model parameter estimation in the non-linear range of the adsorption isotherm was done by applying the inverse fitting method, as performed in section 3.2.1.3. The finished results from the inverse fitting method for the TOYOPEARL S-650S column are shown in Figure 26. As shown in the figure, the simulated UV signal fits with high accuracy to the experimentally determined UV signal from low loading conditions up to the highest tested column load.

The estimated non-linear range parameters species with the TOYOPEARL S-650S column and the CEX column 1 for the polypeptide are given in Table 14. Both model parameter sets are in a similar range and do not differ much. A detailed discussion of these values can be found in section 4.2.2.2.

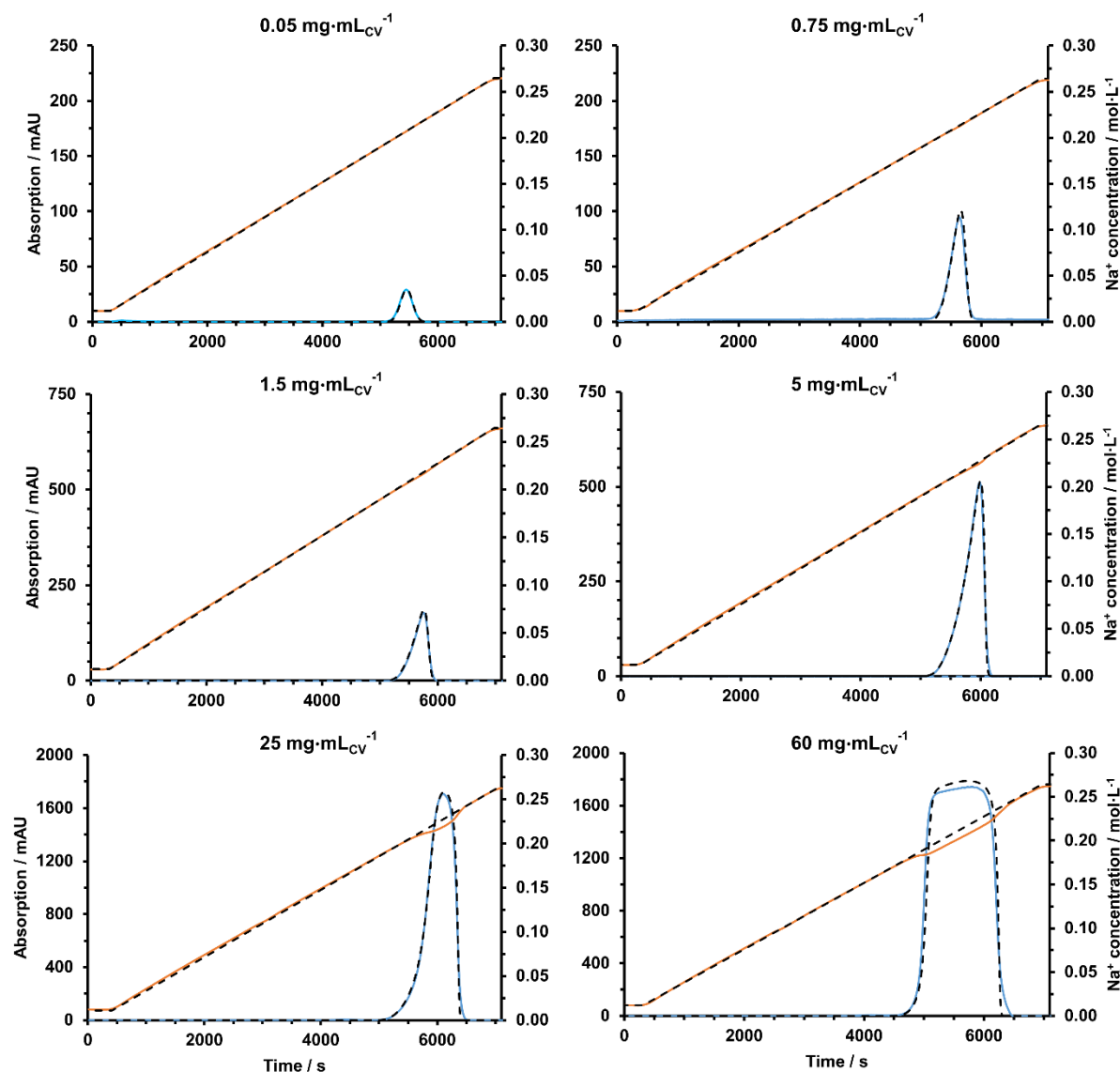


Figure 26: Simulated chromatograms compared to the experimentally determined with the polypeptide species at pH 3.7 and different column loads for the TOYOPEARL S-650S column. The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. The experimental UV signal for the 0.05 mg/mL_{CV} was determined with a 10 mm UV cell. The other experiments were determined with a 10 mm UV cell.

Table 14: Comparison of the model parameters fitted in the non-linear range of the isotherm for the polypeptide species with the TOYOPEARL S-650S column and the CEX column 1.

pH	TOYOPEARL S-650S	CEX column 1
Shielding factor ξ_i	7.2	8.4
$K_{D,i}/c$ (L/mol)	7.5×10^4	2.2×10^5
K_p^* (L/mol)	8.65×10^5	1.046×10^6

3.2.3 Scale-up investigation of the CEX column 1 and the CEX model

To investigate whether the small-scale calibration column (CEX column 1) and the developed mechanistic model are scalable to larger column volumes, several experiments with the polypeptide species were performed on different column scales with the resin from the CEX column 1. Therefore, a 5 mL (50 x 11.3) column (CEX column 2), a 10 mL (100 x 11.3) column (CEX column 3), and a 28.2 mL (359 x 10) column (CEX column 4) were tested.

The following sections show the results of the scalability investigation of the CEX column 1 and the mechanistic model to larger column volumes. The investigation involves analysis of the influence of the flow rate and residence time and the influence of the column-specific parameters on the elution behavior of the polypeptide. Scale-up simulations were then done by considering the column-specific parameters and by considering packing inhomogeneities.

3.2.3.1 Influence of the flow rate and residence time in the linear and non-linear range

The experiments for the scale-up simulations with the CEX column 1-4 and the polypeptide species in this work were performed using the same linear flow rates and, therefore, different residence times for each column scale. The influence of the linear flow rate and the residence time (see Table 15) on the elution behavior of the polypeptide species in the linear and non-linear loading range was investigated and compared (Figure 27). At low loading conditions (load: 0.05 mg/mL_{CV}), the flow rate influences the peak shape but not the elution volume. The lower the residence time at one column scale, the broader and smaller elutes the peak with the same elution volume. In the non-linear range (load: 10 mg/mL_{CV}), the residence time shows the same effect on the peak shape at one column scale, but it is not as pronounced as in the linear range.

Table 15: Linear flow rate and corresponding residence time values for the different CEX column scales.

Column	Flow rate (cm/h)	Residence time (min)
CEX column 1 (1 mL)	100	3.00
	50	6.00
	13.93	21.54
CEX column 2 (5 mL)	100	3.00

Column	Flow rate (cm/h)	Residence time (min)
	50	6.00
	13.93	21.54
CEX column 3 (10 mL)	100	6.00
	27.85	21.54
CEX column 4 (28.2 mL)	100	21.54

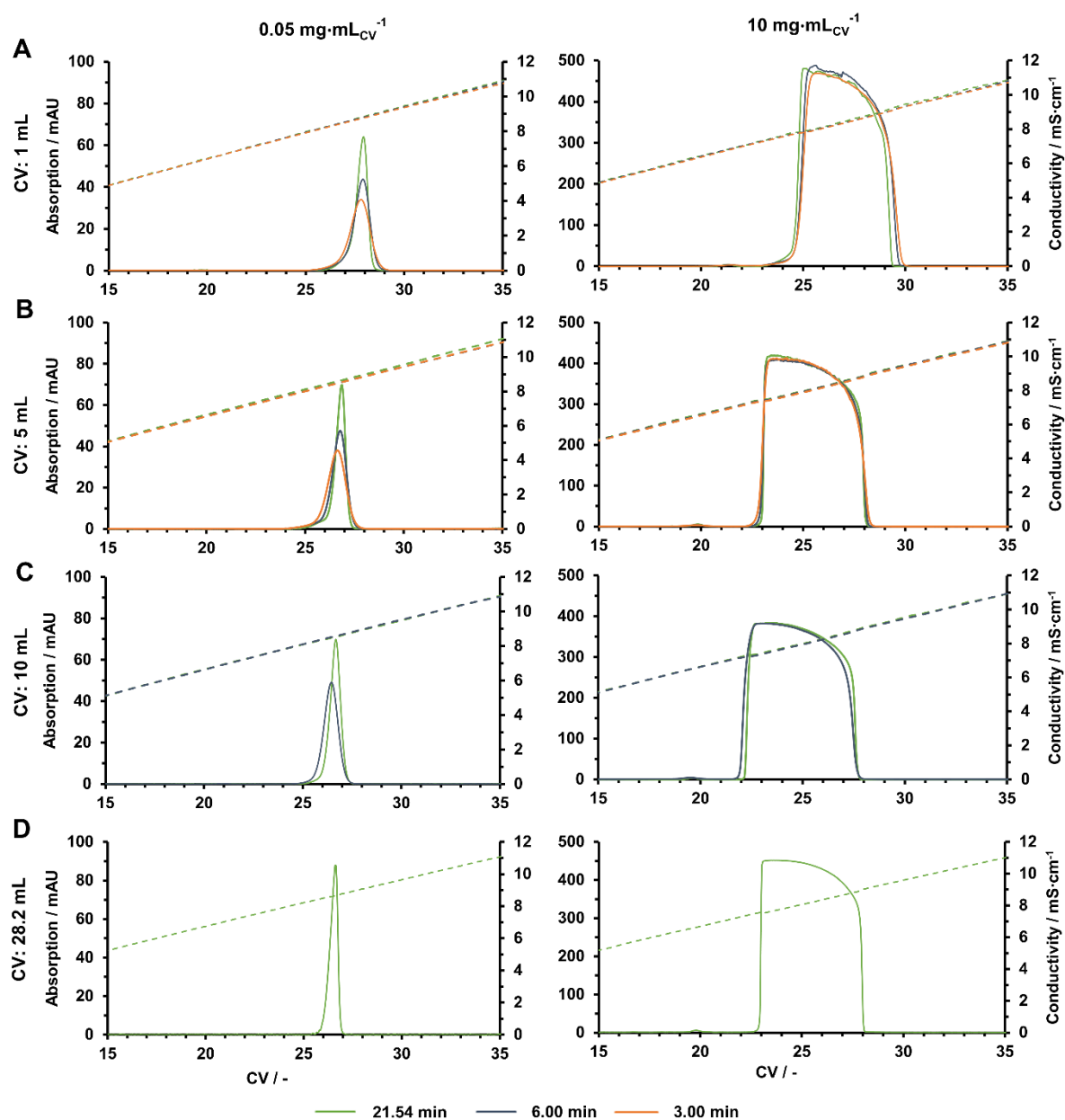


Figure 27: Residence time comparison with the polypeptide species for the CEX column 1 (1 mL) (A), the CEX column 2 (5 mL) (B), the CEX column 3 (10 mL) (C), and the CEX column 4 (28.2 mL) (D) in

the linear (0.05 mg/mL_{CV} with 10 mm UV cell) and non-linear range (10 mg/mL_{CV} with 2 mm UV cell) at pH 3.7.

Figure 28A shows a polypeptide load comparison at 0.05 mg/mL_{CV} and 10 mg/mL_{CV} with the same linear flow rate and gradient slope across the different CEX column scales. As shown above, peak widths differ in the linear range due to different residence times at different column scales. Furthermore, the peaks in the linear range have different salt elution concentrations. The peak retention order in the non-linear range is similar to the retention order at low loading conditions. In Figure 28B, a comparison with the same residence time and the same gradient slope is shown. Now, the peaks in the linear range show comparable peak widths. The peaks' retention order and eluting salt concentration are similar to the retention order with the same linear flow rate (Figure 28A). As observed above, in the non-linear range, a neglectable effect in the peak shape of the residence time is shown. However, there are slight differences in the peak height for peaks with the same residence time at low loading conditions.

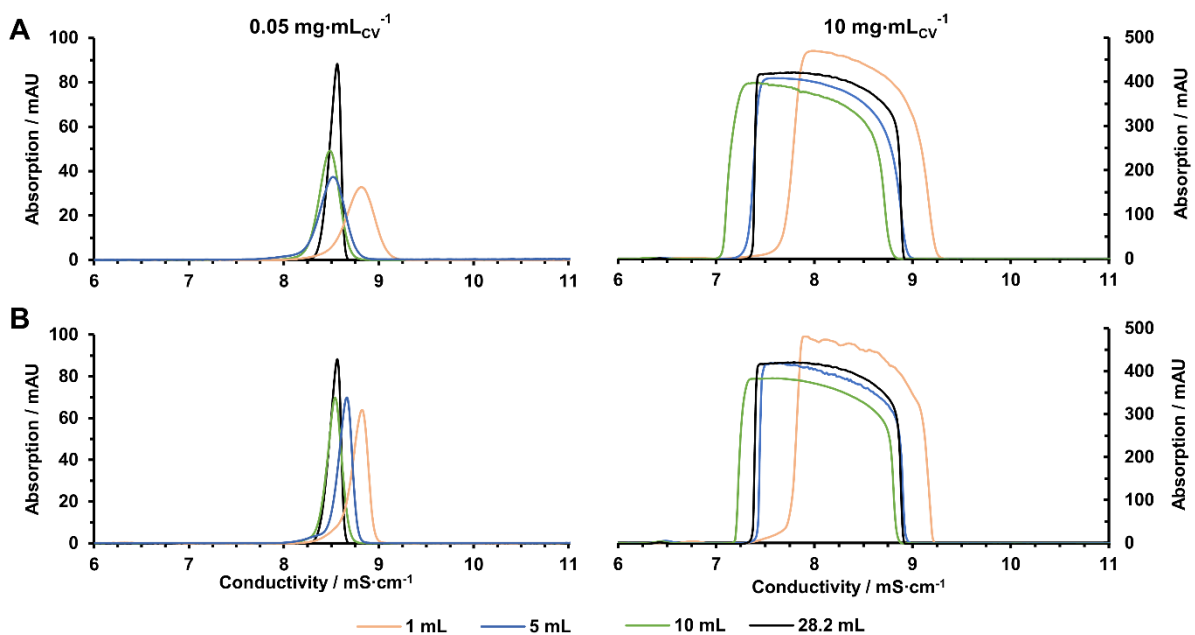


Figure 28: Load comparison for the polypeptide species with the different CEX column scales (CEX column 1-4) at 0.05 mg/mL_{CV} (10 mm UV cell) and 10 mg/mL_{CV} (2 mm UV cell). (A) shows the comparison with the same linear flow rate (100 cm/h) and the same gradient slope at pH 3.7. (B) shows the comparison with the same residence time (21.54 min) and the same gradient slope.

3.2.3.2 Load comparison at different column scales and scaling based on the normalized gradient slope

The different eluting salt concentrations in the linear and non-linear range in Figure 28 can be explained by the different normalized salt gradient slope $GH_{salt,j}$ values for each column scale j caused by the column scale specific parameters ($\varepsilon_{0,j}$, $\varepsilon_{t,j}$, $\varepsilon_{p,j}$, Λ_j , and $k_{d,i,j}$). Therefore, based on the relationship between the normalized gradient slope and the eluting salt concentration described by Equations (21) and (22), a new normalized gradient slope was calculated considering the unique column parameters for each CEX scale-up column. By varying the gradient slope for each CEX column, new experiments were performed to ensure that the peak shape and the eluting salt concentration coincide. The resulting experiments are shown in Figure 29. Figure 29A shows the load comparison with individually calculated normalized gradient slopes and the same linear flow rate for each column scale. The eluting salt concentrations for the peaks at low and high loading conditions are now in a similar range, and the peak heights and shapes at 10 mg/mL_{CV} column load coincide with each other. Figure 29B shows the load comparison with the individually calculated normalized gradient slopes and the same residence time for each CEX column scale. The peak heights at low loading conditions are now in a similar range.

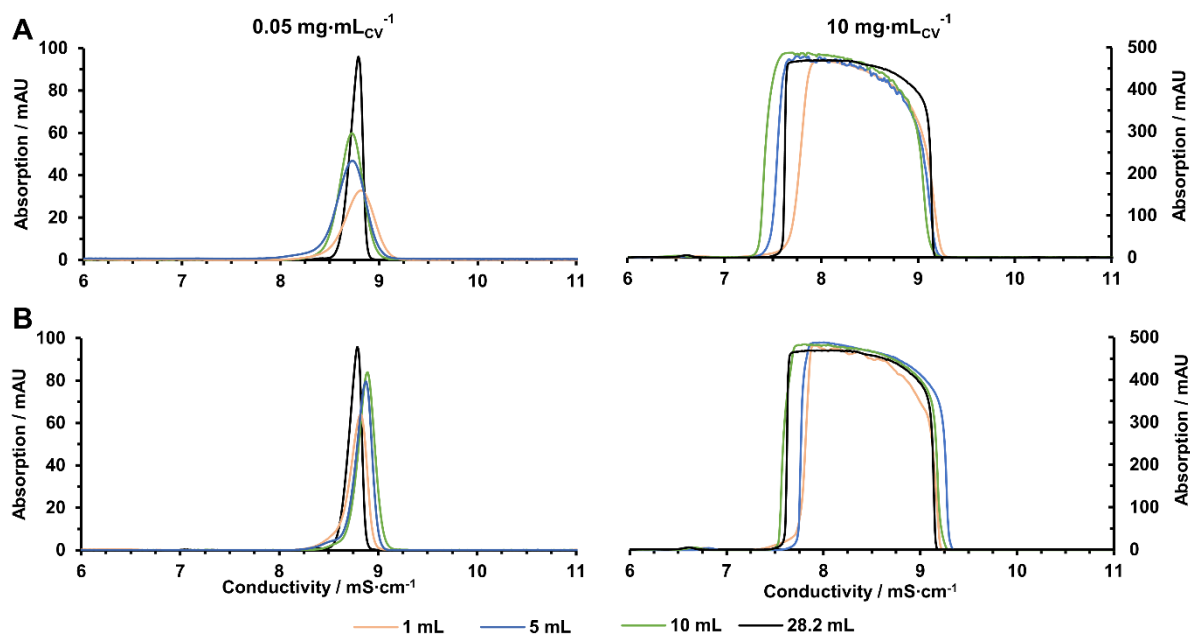


Figure 29: Load comparison for the polypeptide species with the different column scales CEX column scales (CEX column 1-4) at 0.05 mg/mL_{CV} (10 mm UV cell) and 10 mg/mL_{CV} (2 mm UV cell). (A) shows the comparison with adjusted normalized gradient slopes according to Equations (21) and (22) with the

same linear flow rate (100 cm/h). (B) shows the comparison with adjusted normalized gradient slopes and the same residence time (21.54 min).

3.2.3.3 Scale-up simulations

After the scalability investigation of the CEX column 1 to larger columns, the, with the calibration column (CEX column 1), developed mechanistic CEX model was investigated for the scalability to larger column volumes up to 28.2 mL (CEX column 4). Therefore, with the experimentally determined column-specific parameters shown in Table 7 and the model parameters for the CEX model in Table 10 and Table 11, various experiments with different column loads for each CEX column scale were simulated and compared to experimental data. In this work, scaling up to larger column volumes was done by keeping the linear flow rate constant due to the observed neglectable influence of the residence time on the peak shape in the non-linear range (see previous sections). Figure 30A compares the simulated and experimental data at a column load of 10 mg/mL_{CV} for each CEX column scale with the same linear flow rate and gradient slope at pH 3.7. For the simulation of the chromatograms, the shielding factor, determined with the 1 mL calibration column (CEX column 1) at pH 3.7 ($\xi_{i,1} = 8.4$), was used. The simulated UV signals for the CEX column 2 (5 mL), the CEX column 3 (10 mL), and the CEX column 4 (28.2 mL) fit not as precisely as the UV signal from the CEX column 1 (1 mL). To adapt the peak shape, the effective mass transfer coefficient $k_{eff,i,j}$ was first fitted to the experimental peak shape. However, the parameter had no significant effect on the peak shape, so a good fit could not be achieved. In a first empirical approach, the shielding factor for each column scale was additionally fitted to the experimental data by applying the inverse fitting method. Only after the shielding factor was adapted, the simulated signals correlate with high accuracy to the experimentally determined signals (Figure 30B). The fitted shielding factor for the CEX column 2 (5 mL) was $\xi_{i,2} = 9.5$, for the CEX column 3 (10 mL) $\xi_{i,3} = 10.3$, and for the CEX column 4 (28.2 mL) $\xi_{i,4} = 8$.

After re-fitting the shielding factor for each CEX scale-up column, a linear correlation can be found by plotting the fitted shielding factors against the asymmetry factor of each CEX column scale (Figure 31). This correlation indicates an influence of the packing quality and the resulting packing inhomogeneities on the peak shape.

In a second approach, scale-up simulations were done by considering the packing inhomogeneities of the column. In this work, the inhomogeneous packing quality of the CEX columns was described by dividing the scale-up columns into $p = 50$ concentric annular columns. For the sake of simplicity, all fifty annular columns had equal face areas and, therefore, equal column volumes. Furthermore, the concentric annular column interstitial porosity was changed to describe a radial interstitial velocity profile. Subsequently, the factors describing how many annular columns with different interstitial porosities were fitted to the experimental elution profile. Table 16 shows the fitted values for the factor $b_{\varepsilon_{0,j,k}}$ how many of the concentric annular columns are needed with the corresponding interstitial porosity $\varepsilon_{0,j,k}$ for the description of the experimental elution profiles with a constant shielding factor across column scale. Figure 32 visualizes the interstitial porosity and interstitial velocity profile as a function of radial distance inside the corresponding column. The column combinations with the resulting radial interstitial porosity profile lead to the final simulation results for the scale-up columns. The comparison between the experimental and simulated signals is shown in Figure 33. As shown in Figure 33, the model describes the whole tested loading range from linear (0.05 mg/mL_{CV}) to the non-linear range (up to 10 mg/mL_{CV}) with high accuracy under consideration of a fitted radial porosity profile. An interpretation of the scale-up simulation results and a detailed discussion can be found in section 4.2.3.3.

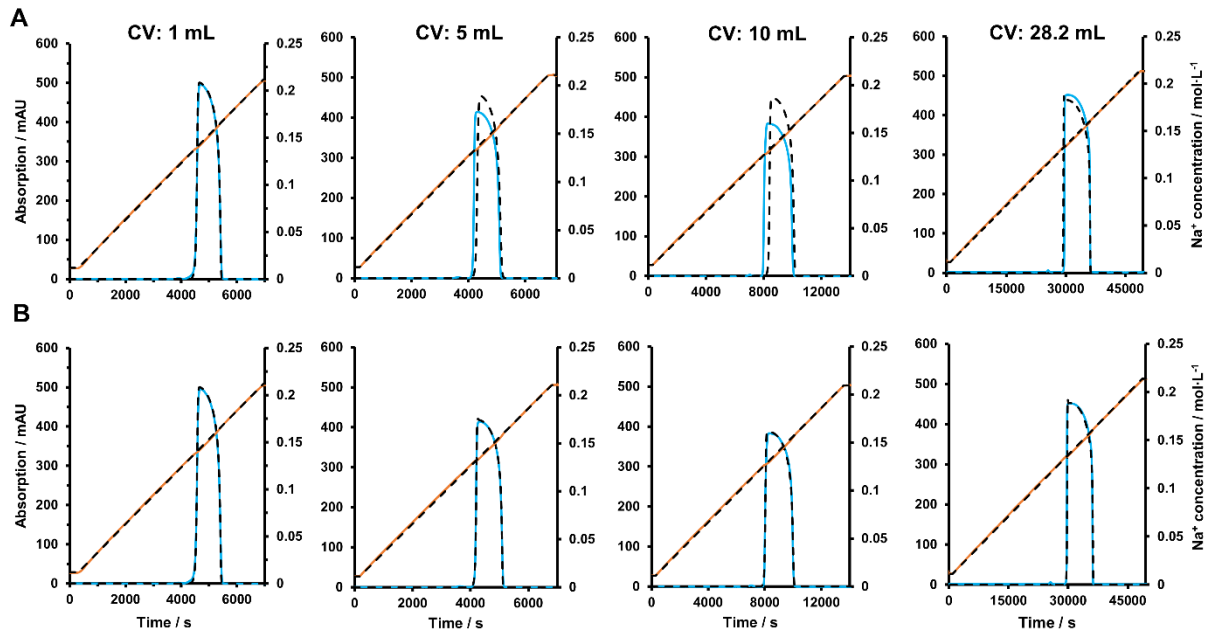


Figure 30: Simulated chromatograms for the CEX column 1-4 compared to the experimentally determined chromatograms with the same gradient slope, linear flow rate (100 cm/h), and polypeptide column load (10 mg/mL_{CV}). The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. (A) shows the simulations with a shielding factor of 8.4 (determined with the 1 mL column scale at pH 3.7), and (B) the simulations with the fitted shielding factors from the first empirical approach.

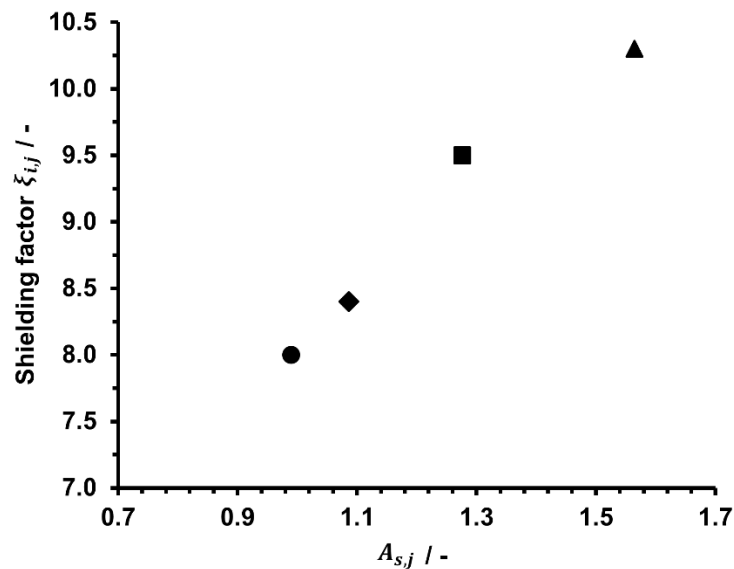


Figure 31: The Shielding factor determined for the polypeptide species in the first empirical approach for each CEX column scale at pH 3.7 as a function of the asymmetry factor. The symbol \blacklozenge represents

the data points for the CEX column 1 (1 mL), ■ for the CEX column 2 (5 mL), ▲ for the CEX column 3 (10 mL), and ● for the CEX column 4 (28.2 mL).

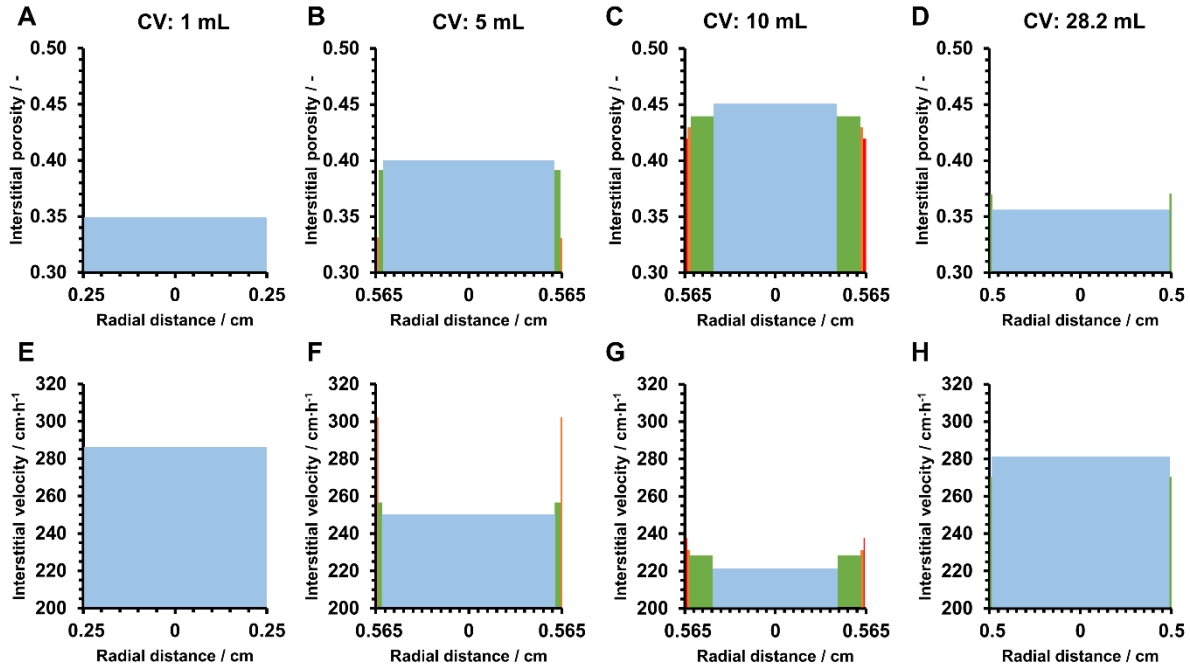


Figure 32: (A-D) Fitted interstitial porosity and interstitial velocity (E-H) profile as a function of radial distance for the different CEX columns 1-4.

Table 16: Values for the concentric annular column interstitial porosity with the corresponding fitted factor for each scale-up CEX column 2-4.

Column	Interstitial porosity in con. ann. column $\varepsilon_{0,j,k}$	Factor $b_{\varepsilon_{0,j,k}}$
	0.330	1
CEX column 2 (5 mL)	0.390	8
	0.400	41
	0.420	3
CEX column 3 (10 mL)	0.430	4
	0.440	19
	0.450	24
	0.370	2
CEX column 4 (28.2 mL)	0.355	48

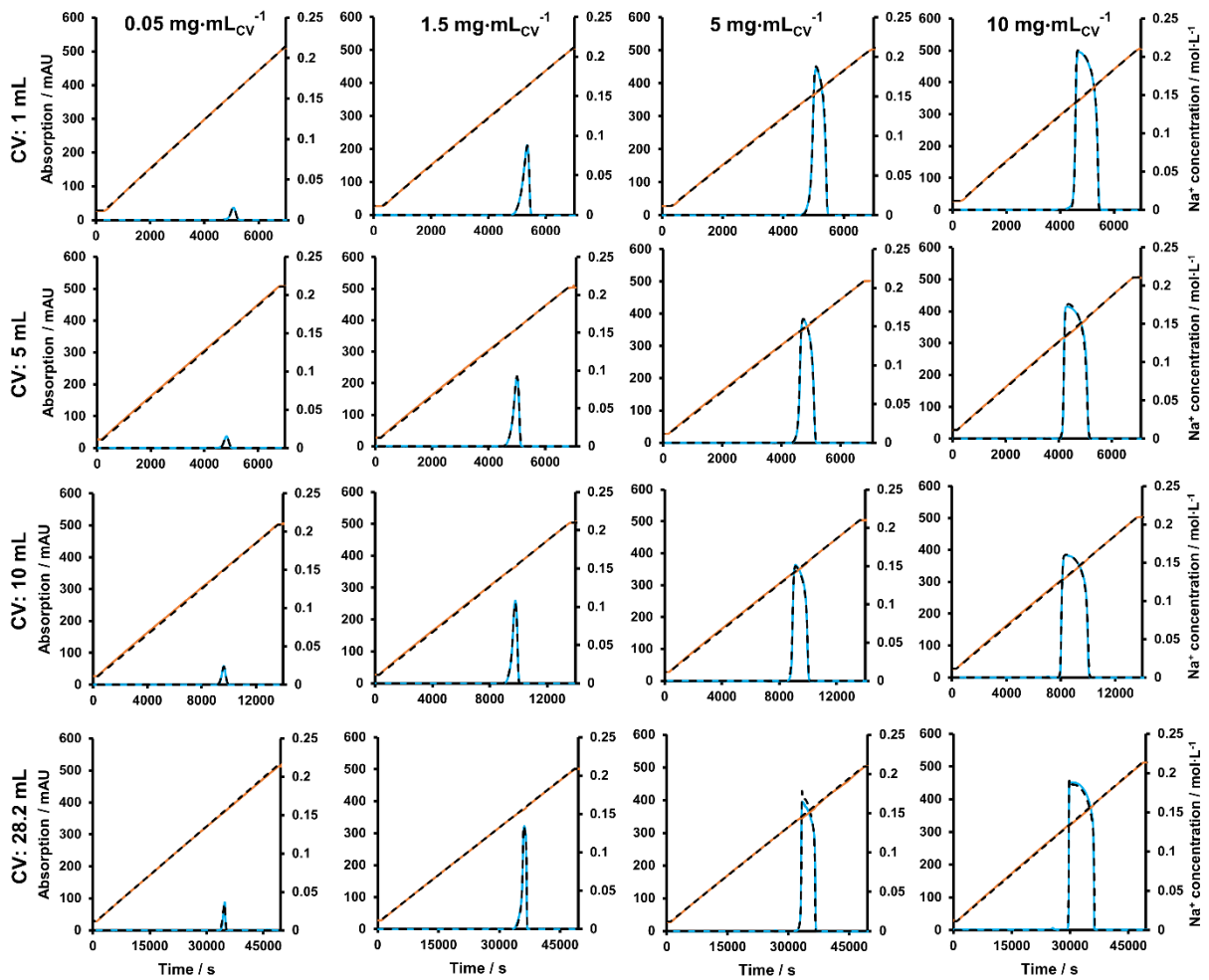


Figure 33: Simulated chromatograms for the different CEX columns 1-4 with the fitted radial interstitial porosity profile compared to the experimentally determined chromatograms at pH 3.7 with different polypeptide column loads, the same gradient slope, the same linear flow rate (100 cm/h), and a constant shielding factor ($\xi_i = 8.4$). The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. For the 0.05 mg/mL_{CV} load experiments, a 10 mm UV cell was used. All other experiments were determined with a 2 mm UV cell.

3.3 RP model development

This section presents the model development results for the chromatographic RP column 1.

3.3.1 RP model calibration at small scale (RP column 1)

3.3.1.1 Investigation of the polypeptide elution behavior

As mentioned in the previous sections, a loading series performed with a relatively pure sample solution is an appropriate method to make a statement about the general elution behavior of the protein and the curvature of the isotherm. Therefore, several column loads from 0.05 mg/mL_{CV} up to 25 mg/mL_{CV} of the polypeptide species were eluted in LGE experiments at different pH values between pH 3.3 and pH 4.1. The gradient elution with the organic modifier was conducted with constant salt modulator concentrations. The results of the different loading series are shown in Figure 34. It can be observed that between pH 3.3 and pH 3.7, the polypeptide exhibits a typical Langmuirian elution behavior with a steep rise at the peak front and diffuse and coincide rears¹³⁴. At pH 3.9 and especially at pH 4.1, the polypeptide exhibits an anti-Langmuirian elution behavior at higher loadings. So, in general, at pH values above pH 3.9, the polypeptide shows an atypical Langmuirian to anti-Langmuirian elution behavior. Possible reasons for that unusual pH-dependent elution behavior are discussed in section 4.3.1.1. Furthermore, it can be seen that the polypeptide always elutes at the same organic modifier concentration when it exhibits a Langmuirian elution behavior and is, therefore, pH-independent under those conditions. The shape of the adsorption isotherm is discussed in section 4.3.1.1. Based on that elution behavior, an appropriate adsorption isotherm was chosen.

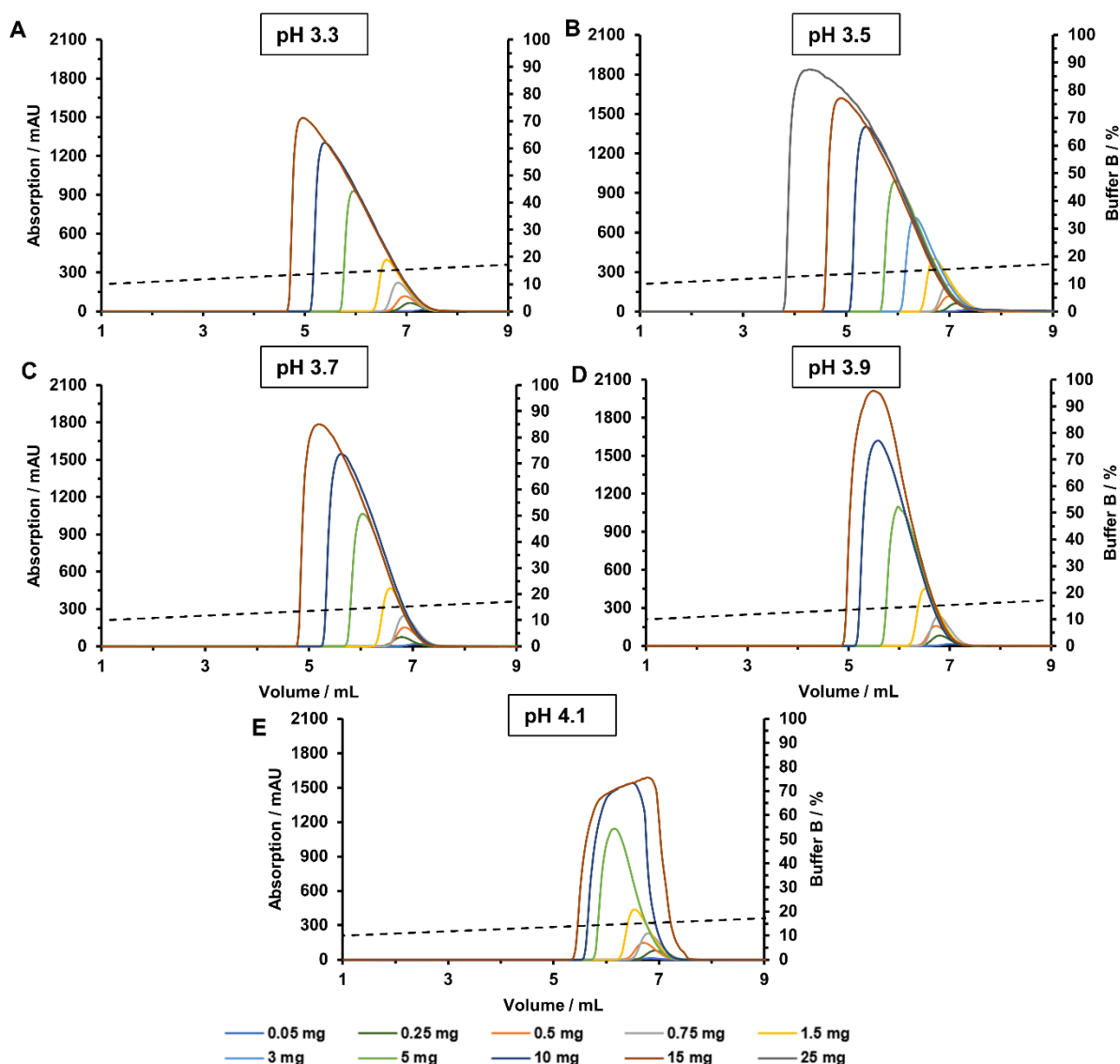


Figure 34: Loading series with the polypeptide species at different pH values (A-E) from 0.05 mg/mL_{CV} up to 25 mg/mL_{CV} sample load with the RP column 1. The smallest peak indicates the lowest load up to the highest load with the broadest and highest peak. The black dashed line represents the amount of Buffer B in the gradient.

3.3.1.2 Determination of model parameters in the linear range of the isotherm

For the determination of the linear range isotherm parameters, the LGE approach, according to Yamamoto^{108, 109}, as already done for the CEX model development, was applied. Several LGE experiments at linear loading conditions and different pH values were performed with the polypeptide species. According to the results in the previous section, a model-based description of the elution behavior at pH 4.1 was neglected. Therefore, LGE experiments from pH 3.1 to pH 3.9 were performed. Furthermore, due

to sample availability, the LGE experiments for impurities were only performed with variant A and variant B at pH 3.1, pH 3.5, and pH 3.9. The resulting evaluated LGE experiments according to Equation (57) are displayed with the normalized gradient slope of the organic modifier LGE run GH_M value as a function of the eluting concentration of the organic modifier in the gradient $x_{M,elu}$ (Figure 35). The plots show that there is a negligible pH dependency in the linear range, as observed in the previous section during the elution behavior investigation (Figure 34). The solid lines in Figure 35 describe the fitted relationship between GH_M and $x_{M,elu}$ and fit with high accuracy to the experimentally determined data points.

The fitted linear range isotherm parameters for each protein species are summarized in Table 17, and a detailed discussion of these parameters is shown in section 4.3.1.2.

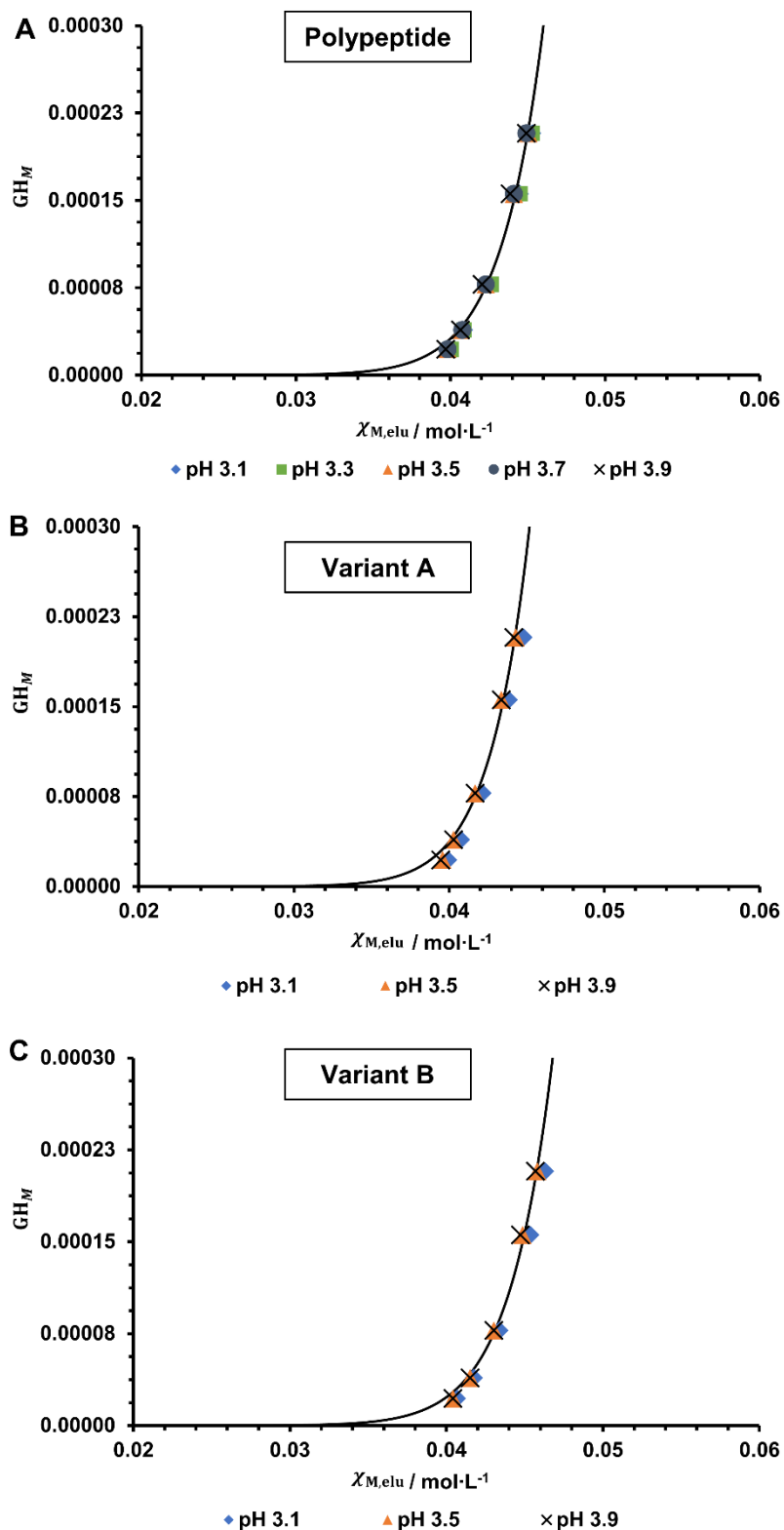


Figure 35: Normalized salt gradient slope as a function of the eluting organic modifier concentration ($\chi_{M,elu}$) at different pH values for the polypeptide (A), variant A (B), and variant B (C) determined with the RP column 1. The solid line indicates the respective $GH_M/\chi_{M,elu}$ - function, generated with the final fit of the parameters in the linear range of the isotherm (see Table 17).

Table 17: Overview of the fitted model parameters in the linear range of the isotherm for the polypeptide, variant A, and variant B determined with the RP column 1.

Parameter	Polypeptide	Variant A	Variant B
$(\eta\tau^2)_i$	1.82×10^{-45}	1.49×10^{-44}	1.26×10^{-46}
$A'_{0,i}$	1.71×10^{-7}	1.50×10^{-10}	1.88×10^{-7}
α_H	34.43	24.21	35.12
$\vartheta_i\zeta_i$	11.29	13.79	11.40

3.3.1.3 Determination of model parameters in the non-linear range of the isotherm

After the parameter estimation in the linear range of the isotherm, the parameters in the non-linear range for the polypeptide species using high load experiments were determined by applying the inverse fitting method. Experiments from 0.05 mg/mL_{CV} to 25 mg/mL_{CV} were performed. The result of the inverse fitting, the comparison of the simulated signals to the experimentally determined signals, is shown in Figure 36. According to the results in the previous section and for simplicity, a pH dependency for the RP model was neglected. Therefore, only experiments at pH 3.5, representing the average of the tested pH range, were performed and fitted for the model development. The resulting non-linear range parameter set is shown in Table 18. Figure 36 shows that the applied model is able to describe the elution behavior of the polypeptide with high accuracy. At higher loads, the simulated UV signals are slightly different compared to the experimentally determined UV signal at the top of the peak, where the polypeptide concentration is the highest (approximately at 1600 mAU). However, the general elution behavior of the polypeptide on the RP column 1 could be described with high precision. A detailed discussion of the fitted non-linear range parameter set is done in section 4.3.1.2.

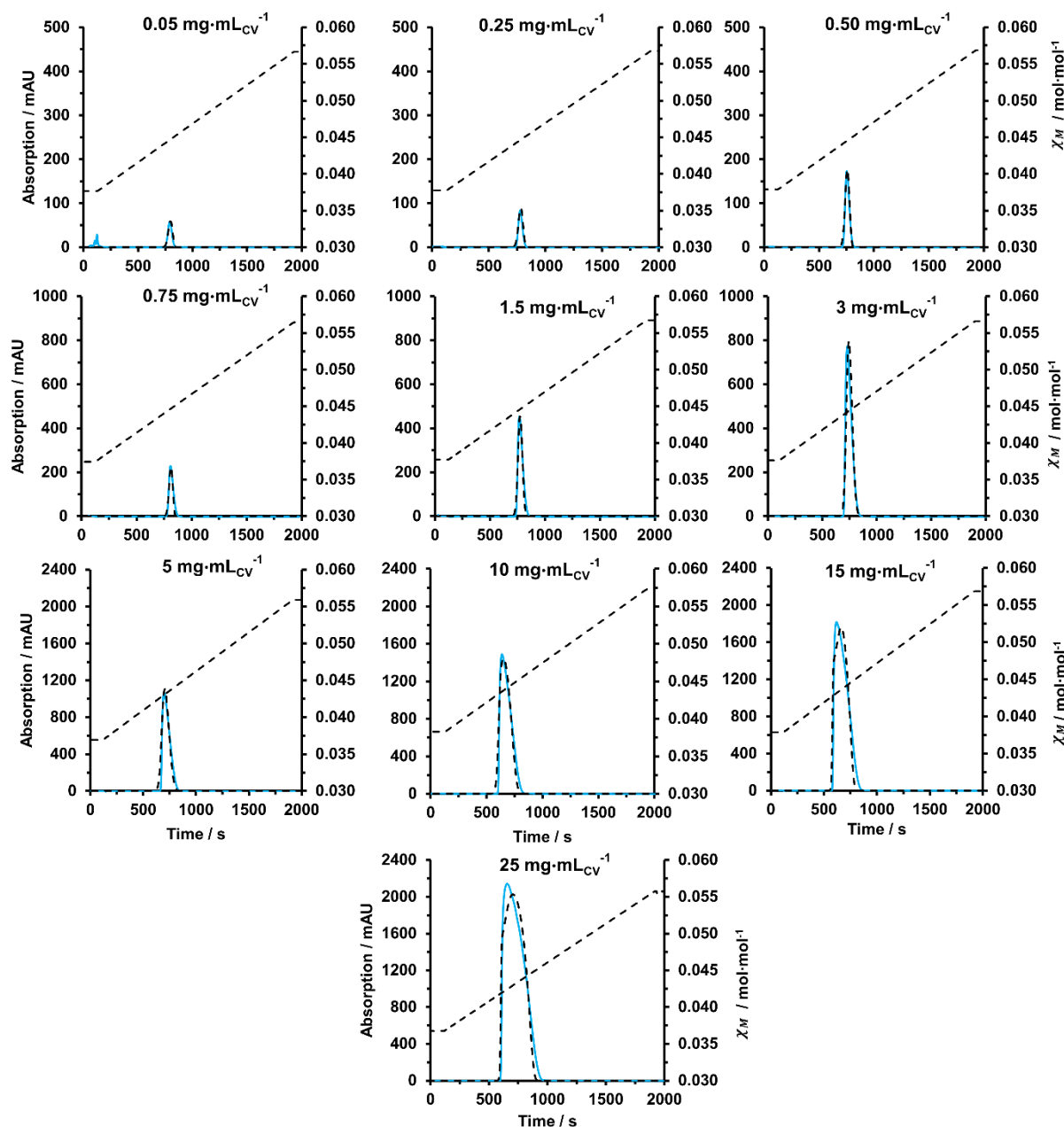


Figure 36: Simulated chromatograms compared to the experimentally determined chromatograms with the polypeptide species at different column loads for the RP column 1. The blue line indicates the experimental UV signal. The black dashed lines indicate the simulated signals. The experimental signal for the 0.05 $\text{mg}\cdot\text{mL}^{-1}$ was determined with a 10 mm UV cell. The other experiments were determined with a 2 mm UV cell.

Table 18: Overview of the fitted model parameters in the non-linear range of the isotherm for the polypeptide determined with the RP column 1.

Parameter	Polypeptide
ξ_i	38
$A'_{0,i,n}$	7.07×10^{-8}
ϑ_i	0.25

3.3.1.4 Multicomponent experiment and simulation with the RP column 1

After successful model development for the polypeptide species in the linear and the non-linear range of the adsorption isotherm and since the developed model should describe multicomponent elutions, a multicomponent experiment was done. This experiment was used as a calibration experiment at high loading conditions for the product-related impurity variants variant A and variant B. The non-linear range parameters of the impurities were fitted to the experimentally determined offline HPLC analysis data. The comparison of the simulated to the experimental signal for the multicomponent run is shown in Figure 37. It can be seen that the applied model and the determined model parameters for the polypeptide species accurately represent the elution behavior of the species. Furthermore, non-linear range model parameters could be fitted, representing the product-related impurities with high accuracy. The resulting non-linear range parameter set for all protein species is summarized in Table 19.

Table 19: Overview of the fitted model parameters in the non-linear range of the isotherm for the polypeptide, variant A, and variant B determined with the RP column 1.

Parameter	Polypeptide	Variant A	Variant B
ξ_i	38	38	38
$A'_{0,i,n}$	7.07×10^{-8}	6.32×10^{-11}	4.15×10^{-8}
ϑ_i	0.25	0.20	0.45

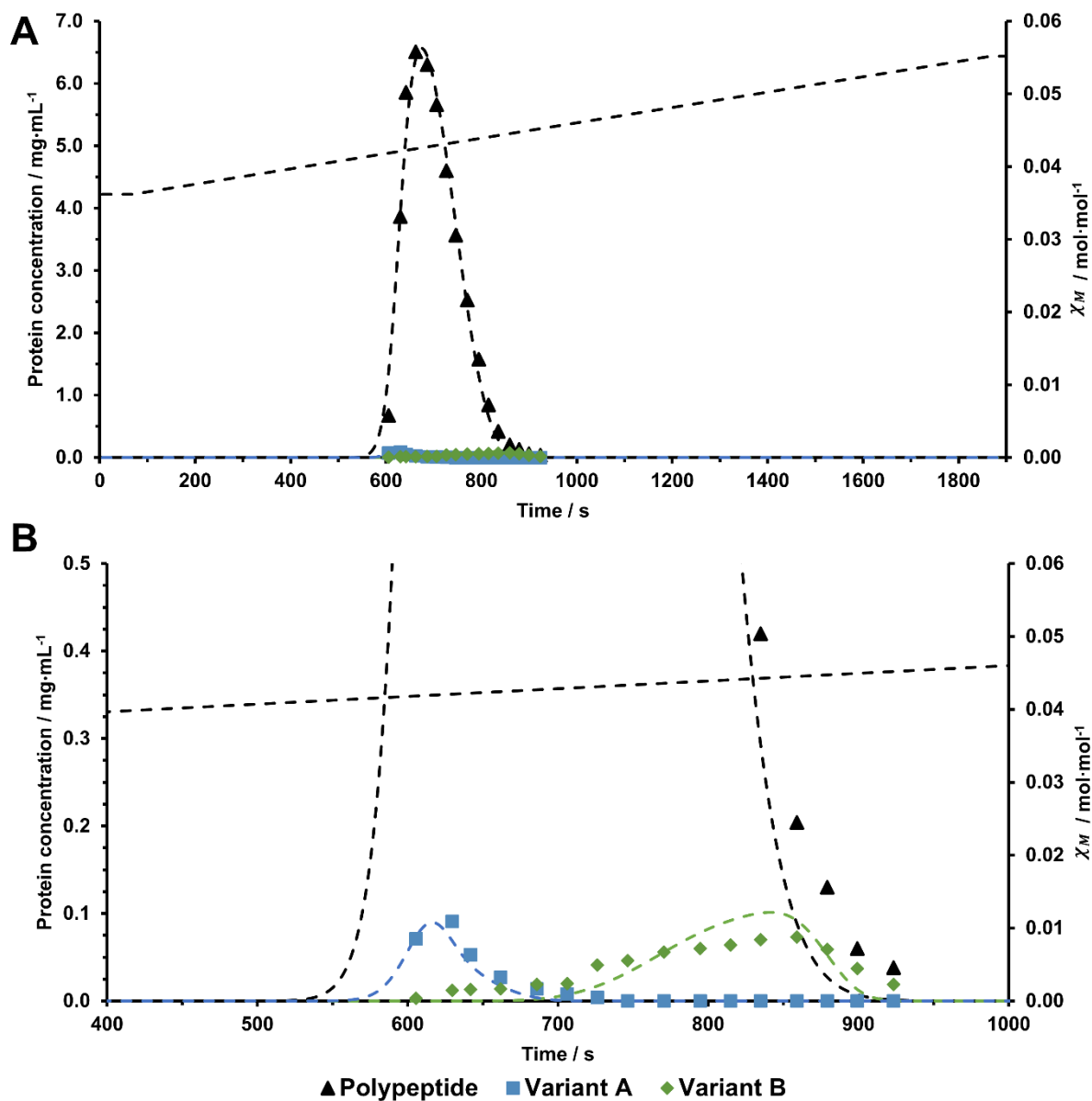


Figure 37: Simulated chromatogram compared to the experimentally determined chromatogram with a multi component sample including the polypeptide species, variant A, and variant B for the RP column 1. The column load was $10 \text{ mg}/\text{mL}_{\text{CV}}$ polypeptide, $0.07 \text{ mg}/\text{mL}_{\text{CV}}$ variant A, and $0.20 \text{ mg}/\text{mL}_{\text{CV}}$ variant B. The colored points represent experimental offline HPLC protein concentration data points, and the dashed lines indicate the corresponding simulated signals.

3.4 CEX model and RP model coupled simulations sampling study

Since this work aimed to develop and calibrate a representative model for a chromatographic two-column purification step and should investigate the influence of loading and pooling parameters on both chromatographic steps, a sampling study using the developed model was done. For the sampling study, the CEX model was coupled with the RPC model (small column scale), and after each step, theoretical pools with a specific pooling criterion were built. The pools after each chromatographic step were visualized using multivariate analysis. The parameters by which the pools were analyzed were pool volume, pool concentration of the polypeptide, polypeptide step yield, and polypeptide step recovery. The simulated sampling study and the multivariate analysis done for the first step, the CEX model, are visualized in Figure 38. The corresponding data set is displayed in Appendix 8.5 in Table 21. Several simulations were carried out for the sampling study for the first chromatographic step. In general, the loads were varied for each simulation to mimic a pool variation from the purification step before the CEX chromatography step. The load was varied in order to mimic different pool volumes with equal polypeptide pool purities and different polypeptide pool purities with equal load volumes. The pooling criterion for the resulting pool of the CEX step was $> 92\%$ polypeptide purity. The reference load volume for the CEX step was 2.75 mL with a polypeptide concentration of 3.8 mg/mL, a concentration of 0.254 mg/mL for variant A, and 0.254 mg/mL for variant B resulting in the polypeptide load purity of 88.2%. Based on that reference load, the load was varied around $\pm 20\%$ load volume, which mimics a varied pool volume with equal purity for the previous purification step, and around $\pm 5\%$ polypeptide load purity, which mimics a varied pool purity with equal pool volume in the previous purification step.

The theoretical pool built in the first step was consequently loaded on the second chromatographic model, the RP model. In the second step, the pooling criterion for the theoretical pool was $> 96\%$ polypeptide purity. As done in the first chromatographic step, the second step was analyzed using multivariate analyses for the pool volume, the pool concentration of the polypeptide, the polypeptide step yield, and the polypeptide step recovery. The visualized analysis and the corresponding data are shown in Figure 39 and in Appendix 8.5 in Table 22. An interpretation of the multivariate analysis and a discussion of the results was made in section 4.4.

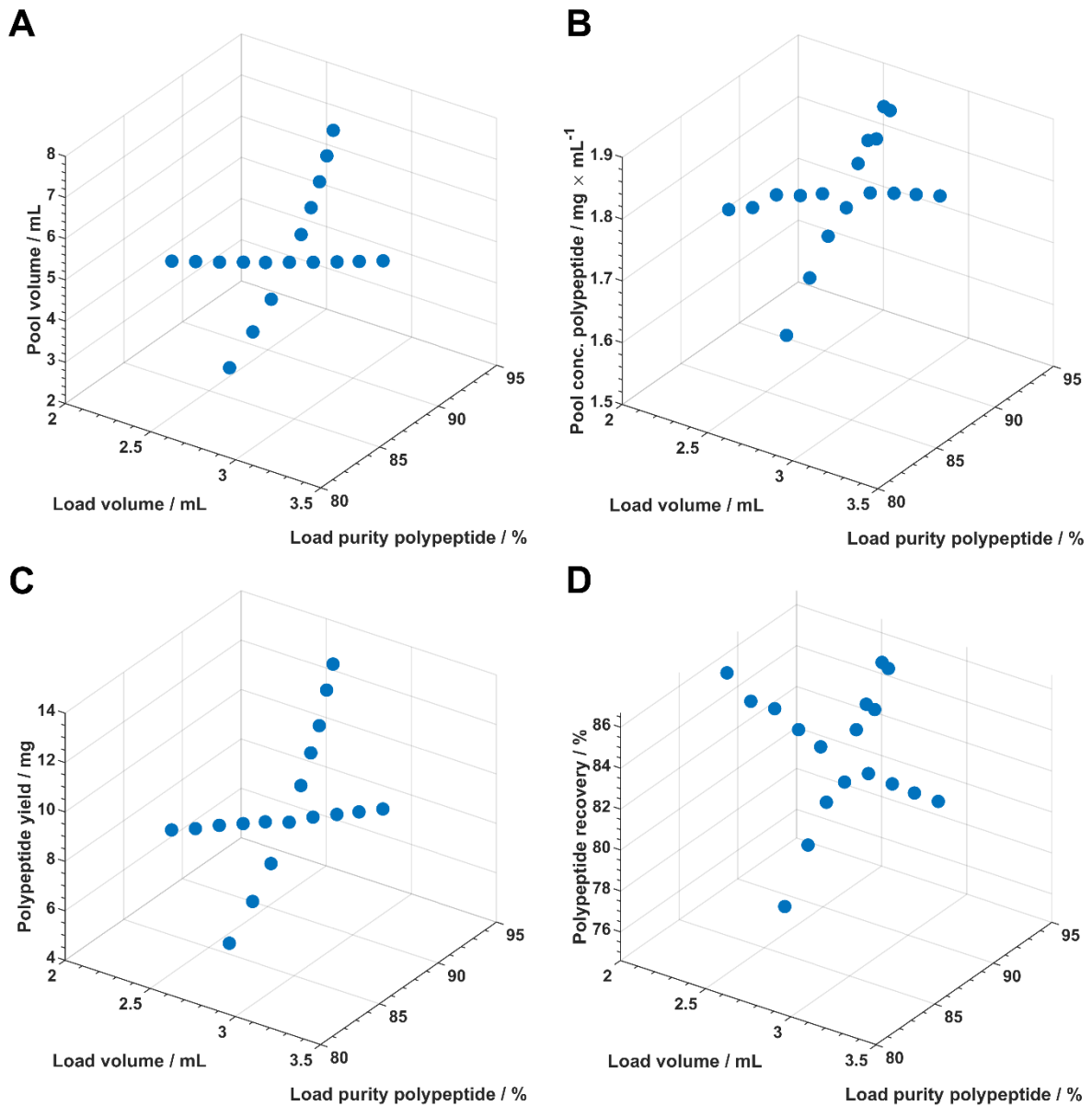


Figure 38: Calculated values from the sampling study for the different pools simulated with the model for the CEX column 1. The pool volume (A), the pool concentration of the polypeptide (B), the polypeptide yield (C), and the polypeptide recovery of the chromatographic step (D) are displayed as a function of the load volume and the load purity.

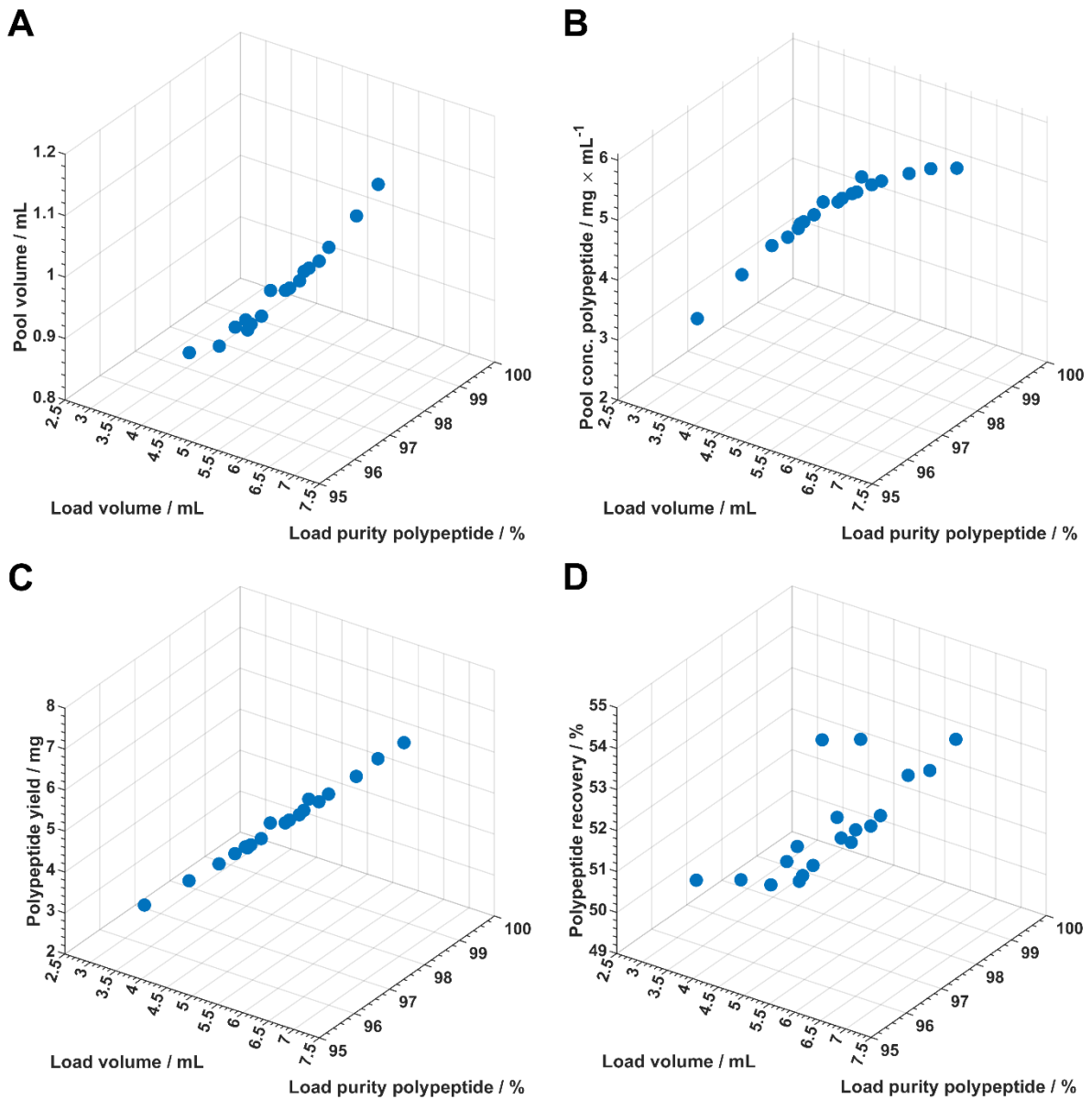


Figure 39: Calculated values from the sampling study for the different pools simulated with the model for the RP column 1. The pool volume (A), the pool concentration of the polypeptide (B), the polypeptide yield (C), and the polypeptide recovery of the chromatographic step (D) are displayed as a function of the load volume and the load purity.

4 DISCUSSION

Parts of this work have been summarized in three different publications. Two of these papers have already been published during the doctoral project in two peer-reviewed scientific publications, and one paper is currently in draft mode. The results and the corresponding discussion can be found in the references from Koch et al. ^{73, 131, 132}.

4.1 Column characterization

This section discusses the column characterization done in section 3.1. In this work, two CEX resins and one RP resin were used. One CEX resin was purchased in different column dimensions to investigate the scalability of one resin to larger column volumes. The second CEX resin was purchased in one dimension to compare the two different CEX resins with each other.

4.1.1 CEX resin columns

Figure 1 shows a comparison of the porosities, ligand densities, and exclusion factors for each CEX column 1-4 and represents the scale-up investigation within one CEX resin. The interstitial porosities for the CEX column 1-4 are within a theoretically plausible range. Literature shows that the interstitial porosity can vary significantly between different column scales packed with the same resin ¹³⁵⁻¹³⁷. This effect could be explained by different packing densities and the arrangement of the particles within the column ¹³⁵ and the technical conditions of column packing of different sizes. When modeling and simulating scale-up processes, identical porosities are often assumed for different column sizes ^{45, 47, 56-60}. An interstitial porosity variation, as observed in this work, significantly influences the resulting simulated peak shape (data not shown). However, these results indicate that each column's interstitial porosity should be determined, particularly when transferring mechanistic models to different column scales. It should be noted that there is considerable uncertainty in the porosity determination of the 1 mL column (CEX column 1) due to the relatively large system and column dead volumes compared to the overall column volume. The ionic capacity measurements show all similar values for all columns and resin batches. Three different resin batches for the CEX columns applied in this work were used (batch 1 for the 1 mL scale (CEX column 1), batch 2 for the 5 mL (CEX column 2) and 10 mL scale (CEX

column 3), and batch 3 for the 28.2 mL scale (CEX column 4)). The ionic capacity measurements show that there might be lot-to-lot variations of the resin and minor measurement uncertainty within one resin lot. However, the experimentally determined ionic capacities agree with the manufacturer information of 0.08 mol/L to 0.12 mol/L based on resin sediment. The exclusion factors show a decreasing dependency with increasing column scale. Nevertheless, the exclusion factor for the 1 mL column (CEX column 1) is inaccurate due to the significant experimental uncertainties. The axial dispersion coefficients in Table 7 show different values for each column scale from the CEX column 1-4. Theoretically, D_{ax} depends only on the packing properties and the flow conditions in lumped rate models¹¹². Since different packing characteristics were determined for each column scale, the axial dispersion coefficient, determined with the same linear flow rates for each scale, varies as expected in a small range¹³⁸. No explicit dependency of the axial dispersion coefficient with the packing characteristics of the columns was observed.

Table 8 shows an overview of the determined packing quality parameters for each CEX resin column. Figure 2 compares the parameters for the CEX column 1-4 for the visualization of the scalability of columns with the same resins and different dimensions. The first moments (peak center) of the peaks from the CEX columns 1-4 increase linearly with increasing column volume, indicating a comparable total porosity for the different column scales and a high reproducibility of the column packing for the same resins with different column dimensions^{136, 139}. The standard deviations in Figure 2A tend to increase with increasing column volume. However, it is worth mentioning that $\sigma_{EMG,salt}$ of the salt peak from the CEX column 3 (10 mL) appears to be higher than expected from the trend of the other columns. The empty column of the CEX column 4 involves filters, frits, and adapters that increase the ECV broadening and significantly affect the overall peak broadening of the column¹²². Therefore, the standard deviation for the CEX column 4 (28.2 mL) decreases significantly when the empty column is also considered, although the first moment is not influenced. The determined *HETP* values and plate numbers for each column are shown in Figure 2B. The *HETP* values for the CEX column 1 (1 mL), the CEX column 2 (5 mL), and the CEX column 4 (28.2 mL) are comparable and are within the theoretically desirable range of 2 - 6 particle diameters¹⁴⁰. The value for the CEX column 3 (10 mL) is outside this range, indicating poor packing quality of this column. The plate numbers of the CEX column 1 and the CEX

column 2 are around 500 and, as expected, are comparable because they have the same column length. The CEX column 3 has the same value as the CEX column 1 and the CEX column 2, although the CEX column 3 is twice as long, again indicating the poor packing quality of this column. As expected, the CEX column 4 (28.2 mL) has the largest plate number with a value of about 4000 or 7500 (blue data point) since it is the longest column. The asymmetry factors for the CEX columns 1-4 and the ratio between the ECV and the column volume are shown in Figure 2C. The highest asymmetry factor, with a value of almost 1.6, belongs to CEX column 3. An asymmetry factor of 0.8 – 1.2 is within a desirable range and indicates adequate packing¹⁴⁰. The ECV is more than 20 % for the 1 mL column, about 6 % for CEX column 2, and less than 4 % for larger column volumes. Nevertheless, ECV broadening has a significant influence¹³⁹ and can be more than 50 % for large column volumes considering column filters, frits, and adapters¹²², as shown above for the CEX column 4. ECV effects only significantly influence the retention volume (first moment) for the CEX column 1. Nevertheless, ECV effects on the retention volume were considered for the later column characterization at each column scale. These results show that scalability to larger columns is comparable. However, the packing quality has to be considered, which was still sufficient for the CEX column 3, for example, but not comparable with the other column scales. Such packing quality differences may influence the elution behavior of the protein and the peak shape on that column.

Comparing the TOYOPEARL S-650S column with the CEX column 1, having the same dimensions, shows that the packing quality is similar since the asymmetry factor and the plate number for both columns are comparable. Furthermore, this shows that the purchased pre-packed columns in that dimensions are comparable based on the packing quality.

4.1.2 RP resin column

The interstitial porosity of the RP column 1 was assumed to be 0.35, which seems to be a plausible value for RP columns^{116, 141}. The exclusion factor was estimated to be 0.75 for the polypeptide and the product-related impurities. The total porosity was determined with a salt pulse and calculated to be 0.49. Nevertheless, it is known that the protein's local porosity and the exclusion factor change with the organic modifier concentration^{126, 141}. Therefore, the experimentally determined and the estimated column

parameters may not reflect the actual values in the gradient elution. However, this assumption is reasonable and commonly made in modeling RPC ^{116, 118, 142-144}.

The axial dispersion coefficient was fitted to be 0.001 cm²/s at analytical conditions. This value seems plausible since Jakobsson et al. estimated a similar value using an empirical correlation and a similar RP column with a silica-based resin ¹⁴⁵.

The column packing quality parameters were determined with a salt pulse at a 10 % organic modifier concentration. The asymmetry factor for the RP column 1 is 1.7, and according to general classifications ¹⁴⁰, outside a desirable range. One reason could be that the salt pulse determination may not be accurate enough since the salt concentration in RPC changes the ligand pattern and can influence particle porosity ¹²⁵. The plate number for the 5 cm height column was determined to be 1741, which is according to the product catalog of the resin under the minimum of 2000, indicating in combination with the asymmetric factor for poor packing quality. One reason for the poor values could be that the salt pulse determination may not be accurate enough since the salt concentration in RPC changes the ligand pattern and can influence particle porosity ¹²⁵.

4.2 CEX model development

This section discusses the results of the CEX model development done with the different used CEX columns and the results from the investigation of the scale-up ability from the CEX column 1 to larger column volumes (CEX column 2-4).

4.2.1 CEX model calibration at small scale (CEX column 1)

4.2.1.1 Investigation of the polypeptide elution behavior

The general polypeptide elution behavior on the CEX column 1 in a salt gradient elution and with an isocratic elution is shown in Figure 3, Figure 4, and Figure 5. As described in Section 3.2.1.1, the peak exhibits a complex anti-Langmuirian to Langmuirian elution behavior with increasing load. Complex elution behaviors in liquid chromatography, especially in high loading situations, are common since additional mechanisms often occur at high protein concentrations. For example, intraparticle diffusion effects and protein conformational changes cause ^{76, 146} multiple peak formations and peak shoulders. However, no multiple peak structures or shoulders are observed in this work. The results show that a convex isotherm describing Langmuir's elution behavior, such as the SMA isotherm commonly used for IEC modeling, does not apply to the elution behavior observed for the polypeptide species on the CEX column 1. An anti-Langmuirian elution behavior reveals itself through a fronting peak with a steep rear and coinciding peak fronts with increasing load. Langmuirian elution behavior is shown by steep fronts and diffuse and coinciding rears with increasing load ¹³⁴. The fronting peak at lower loadings, which is part of the anti-Langmuirian behavior, is not related to the column characteristics because the asymmetry factor of the CEX column 1 is 1.09 (see Table 8). These findings show that the adsorption isotherm must be concave at low protein concentrations (anti-Langmuirian elution behavior) and convex-shaped at higher concentrations (Langmuirian elution behavior), resulting in a sigmoid-shaped adsorption isotherm. The anti-Langmuirian behavior indicates that the adsorption process observed in this work could include dimer or multimer formation on the resin ¹³⁴. Similar elution behavior was observed with β -lactoglobulin A by Jen and Pinto ⁷⁵ and in the work related to the SAS isotherm by Mollerup ^{74, 88, 147, 148}. The elution behavior found in the literature only coincides with the elution behavior observed in this work until a column load of about 10 mg/mL_{CV}. In this work, higher loadings were tested in gradient elution experiments, where only peak broadening and peak flattening are observed as

the loading continues to increase, as shown in Figure 3. Increasing the pH leads, as expected, to an earlier elution and a narrower and taller peak since the mobile phase pH approximates the protein's isoelectric point by increasing the pH.

According to the results discussed above, the sigmoidal SAS isotherm^{88, 102, 149}, which describes dimer formation on the resin, was chosen for modeling the polypeptide species on the CEX columns in this work. For the description of non-idealities, the isotherm was extended with the activity coefficients for the counterion and protein species in solution. A detailed deviation of the isotherm is shown in Section 1.4.1.

4.2.1.2 Determination of model parameters in the linear range of the isotherm

The model parameter for the polypeptide species, the product-related (variant A and variant B), and the process-related (variant C) impurities in the linear range of the isotherm were determined by applying salt-induced LGE experiments at low loading conditions and by applying Yamamoto's approach^{108, 109}. The evaluated LGE experiments are summarized in Figure 6. As expected, increasing the pH leads to an earlier elution of the polypeptide in the salt gradient. In general, the steeper the gradient slope, the smaller becomes the peak retention volume, whereas the eluting counterion concentration $c_{e,elu}$ increases as described elsewhere^{150, 151}. The solid lines in Figure 6 show a good correlation for each pH between the experimental and calculated elution points using the fitted parameters. The different elution behavior induced by varying the pH value and the gradient slope can be described with high accuracy using the relationship in Equation (22) for all four applied protein species. The parameters for the polypeptide and the three variants (variant A, variant B, and variant C) fitted from the final step of the fitting procedure in the linear range of the isotherm are summarized in Table 10. The parameter $\Delta G_s^0/RT$ values for all species protein species range from 2.08 to 2.45 and are thus in a similar range. Since $\Delta G_s^0/RT$ describes the differences in the reference potential when the counterion changes its state from bound to solute and since the same counterion was used for all three protein species, the values coincide in a similar range with each other as expected⁸³. $\Delta G_i^0/RT$ shows, as expected, different values in a range from -3.66 to 5.84 since different protein species were used. The values for the polypeptide species and the product-related impurities (variant A and variant B) are in a similar range, whereas the process-related impurity (variant C) is out of this range and shows a negative value for $\Delta G_i^0/RT$, explainable with the

significant different structure for this impurity compared to the structure difference of the polypeptide and the product-related impurities. The parameters describing the characteristic charge v_i (N_{acidic} , N_{basic} , $pK_{a_{acidic}}$, and $pK_{a_{basic}}$) for the polypeptide, variant A and variant B are all in a similar range and differ slightly from each other, since they result in similar v_i curves. However, variant C shows a significantly lower characteristic charge due to its different structure. Note that the pka value of the basic amino acid-like group $pK_{a_{basic}}$ was set to be 11 since the calibration experiments were conducted in a relatively small pH range from pH 3.3 to pH 4.3. The parameter K_s^* ranges from -286.66 L/mol to -310.38 L/mol. However, the comparison of K_s^* is only of limited significance, since it is part of the overall pH dependency from K_s , and is only comparable in combination with the characteristic charge v_i according to equations (18) and (19).

Since there is no information about the parameter range of the asymmetric activity coefficient of the protein $\tilde{\gamma}_i$, the fitting procedure in the linear range for the polypeptide data set in this work starts without the influence of this coefficient. Fits without knowing the parameter range would lead to less robust fits and multiple solutions for K_s^* . The parameter K_s shows a significant influence in the linear and non-linear range of the isotherm, previously shown by Huuk et al. ⁷⁸. However, in this work, the parameter range of K_s^* for the polypeptide was fitted from high load experiments. Therefore, the fitting procedure starts without the influence of the asymmetric activity coefficient leading to a parameter set, which is used in high loading simulations for determining the parameter range of K_s^* . With the knowledge of the parameter range K_s^* , a new fit in the linear range of the isotherm with the influence of the asymmetric activity coefficient was done, resulting in a robust fit and a set of reproducible modeling parameters (Table 10). Note that for the K_s^* parameter determination for the three impurities, the same previously determined range with the polypeptide data set was used. The parameter describes the interaction between the solute salt and protein and thus depends on the individual species. Therefore, the influence should slightly differ in a similar range for the polypeptide and the product-related impurities since the same salt species was used, and the variants only had minor differences in protein structure. However, variant C is also in the same range, although the significant differences in the structure compared to the other proteins. But as already mentioned, this value is only comparable by considering the overall pH dependency of K_s . The interacting parameter K_s for

the salt species depends on the charge of the protein and is, thus, pH-dependent⁸³. The pH dependency of K_s and the characteristic charge v_i for the four sample proteins determined with the CEX column 1 is visualized in Figure 7. In general, K_s decreases with increasing pH for all four protein species. Variant A has a similar starting point at pH 3.1 and a slightly higher value at pH 4.3 than the polypeptide. Variant B has, in general, a slightly higher curve. All three K_s curves are, as expected, marginally different. However, variant C has a significantly lower K_s values compared to the other proteins, explainable with the different structures. A negative K_s value indicates that the protein-salt interaction is stronger than the protein-water interaction⁸³, which leads to an activity coefficient of the protein in solution less than 1 (see Equation (18)). According to Equations (6) and (10), an activity coefficient $\tilde{\gamma}_i < 1$ disfavors the interaction of the monomeric protein with the ligand, that is, binding of the protein and dimer formation. This effect depends on the sodium chloride concentration and is more pronounced at higher pH. As mentioned above, K_s depends on the salt type and the protein. Negative values for K_s for different mAbs and NaCl as the mobile phase modulator were also determined by Huuk et al.⁷⁸ and Lee et al.¹¹⁰. The characteristic charge v_i decreases with increasing pH for all three protein species almost linearly. Since the pH approximates the isoelectric point of the protein with increasing the pH v_i decreases as expected. Furthermore, the polypeptide species has a lower v_i curve than the product-related impurity variants (variant A and variant B), which coincides with the elution order of the three protein samples since the polypeptide species elutes the earliest in the salt gradient. However, variant A has a slightly higher v_i curve than variant B, although variant A elutes earlier in the gradient. Furthermore, the v_i value of variant C is significantly lower and elutes in a similar range. This elution behavior can be explained by the slightly different K_s curve of the four species, since the parameter K_s has a strong influence on the elution behavior in the linear range, as observed during the fitting procedure in this work and by Huuk et al.⁷⁸. The slightly different values for the thermodynamic parameters $\Delta G_i^0/RT$ and $\Delta G_s^0/RT$ also have an influence on the elution behavior in the linear range and may be a reason for the different elution order compared to the values of the characteristic charge v_i .

The importance of the activity coefficients was investigated by comparing the modeling fits in the linear range with and without the influence of the activity coefficients. The difference and the correlation between the parameters for the polypeptide species are

shown in Figure 8. The correlation between the parameters with and without the influence of $\tilde{\gamma}_i$ is linear. These results coincide with observations made in the literature¹⁵². Pinto et al. investigated the influence of the activity coefficient of the counterion species in solution on the equilibrium constant K_i and the characteristic charge ν_i . They found the same influence which was observed in this work. Since both asymmetric activity coefficients $\tilde{\gamma}_i$ and $\tilde{\gamma}_s$ depend on the counterion concentration, the influence should be similar. Note that in this work, only the influence of the asymmetric activity coefficient of the polypeptide species in solution $\tilde{\gamma}_i$ on the equilibrium constant and the characteristic charge during the fitting procedure was investigated. The influence of the activity coefficient on the parameters for variant A, variant B, and variant C was not investigated since no experiments in the non-linear range of the species were done. The asymmetric activity coefficient of the counterion species in solution $\tilde{\gamma}_s$ is always calculated according to equation (16). The further the value of the activity coefficient is from one, the stronger the influence on the equilibrium constant and the characteristic charge¹⁵².

4.2.1.3 Determination of model parameters in the non-linear range of the isotherm

The estimated parameters for the polypeptide species for the non-linear range are given in Table 11. Within the small pH range used in this investigation, dimer formation is assumed to be pH-independent and identical $K_{D,i}/c$ values for each pH were fitted. To consider pH and salt effects for $K_{D,i}/c$, a detailed understanding of the dimerization process and appropriate models for the activity coefficients for the protein species in the bound state $\hat{\gamma}_{1,i}$ and $\hat{\gamma}_{2,i}$ are needed. Westerberg et al. determined $K_{D,i}$ values for a therapeutic protein and two product-related impurities from about 1×10^4 L/mol to 1×10^5 L/mol¹⁵³. Mollerup determined a $K_{D,i}$ value of 1.8×10^5 L/mol for the Glucagon-like Peptide 1 analogue and a $K_{D,i}$ value of 5×10^5 L/mol for the β -lactoglobulin A on an anion exchange resin¹⁰². The values determined by Westerberg et al. and Mollerup are lower but are in the same range compared to the $K_{D,i}$ value of 9.8×10^6 L/mol determined in this work. Since the equilibrium constant for the dimer formation $K_{D,i}$ strongly depends on the dimerization propensity of the protein in the investigated environment, the comparison of values of different protein species is only possible to a limited extent.

As already mentioned above, the interacting parameter K_s and K_p for the salt and the protein species depend on the charge of the protein and, thus, are pH-dependent⁸³. The introduced empirical pH dependency of K_p determined for the polypeptide species in the experimental pH range is shown in Figure 10. The protein-interacting parameter K_p only influences the non-linear range of the isotherm since $K_p \times c_i$ approaches zero in the linear range of the isotherm (low loading conditions). The overall influence of K_p on the peak shape is small compared to the influence of K_s ⁷⁸. In our case, however, the simulations would not be as precise as shown in Figure 11 with a K_p of zero. As shown in Figure 10, K_p is positive and increases with increasing pH. A positive value for K_p indicates that the protein-protein interaction is weaker than the water-protein interaction¹⁵⁴, which supports the assumption that dimerization of the polypeptide takes place on the stationary phase surface and not in solution.

The marginal shielding factor variation at different pH values shown in Table 11 is relatively small (~1 % around the mean value) and shows no pH dependency. Nevertheless, unique shielding factors were fitted for each pH value since this adaption improved the simulation results.

As shown in Figure 11, the developed model is able to describe the whole tested loading range up to 25 mg/mL_{CV} and pH range from pH 3.3 to pH 4.3 with high accuracy. However, the simulations with a column load of 25 mg/mL_{CV} slightly differ from the experimental UV elution signal in the front region of the peak. This slight deviation could be explained by the fact that, at higher loads, not only dimers but also multimers might be formed on the resin, which is not considered in the isotherm model. Additionally, the deviation might be the result of the assumption that the activity coefficients of the bound protein species $\hat{\gamma}_{1,i}$ and $\hat{\gamma}_{2,i}$ were set to one. This assumption could be inappropriate for loads ≥ 25 mg/mL_{CV} when the binding density approaches the monomer layer binding capacity of the resin. An extrapolation outside the tested pH range must be investigated. Since the pH dependency of the parameters K_s and K_p were not described mechanistically, it might be that simulations outside the calibrated pH range might not fit the experimental data. However, the rest of the developed model is based on mechanistic relationships, which should be extrapolatable within a certain range.

As expected, the SAS isotherm calculated for the polypeptide species shows a sigmoidal type V shape¹⁵⁵, as visualized in the 3D-Plot in Figure 12. The inflection point

of the isotherm is shifted to lower analyte mobile phase concentrations c_i and the slope at the inflection is increasing when the counterion concentration decreases. At binding conditions ($c_s < 0.05$ mol/L), the isotherm is appearing favorable and of rectangular shape. The concave shape at lower protein concentrations leads to coincident eluting peak fronts. A convex shape at higher protein loads lead to coincident rear edges of the eluting peaks^{134, 148}, as observed and described in the investigation of the elution behavior of the polypeptide (Figure 3, Figure 4, and Figure 5). As displayed in Figure 13, at protein concentrations of $c_i > 10^{-4}$ mol/L, the monomer-dimer equilibrium is shifted towards the dimer species and almost all the binding sites are occupied by this species. At analytical load conditions, the dominating bound state is the monomer showing anti-Langmuirian binding behavior with peak fronting during elution. At increased loads, the dimer is formed with classical Langmuirian saturation behavior. During elution under these conditions, an asymptotic triangular peak with a shock layer at the front is formed.

The gradient elution validation experiments for the developed CEX column 1 with the polypeptide species shown in Figure 14 demonstrate that the model is able to predict the elution of the polypeptide correctly in the calibrated pH and loading range. Even though no pH or dual gradients were performed in the model calibration, the partially empiric pH dependency of the model seems to be able to describe the dependency correctly. Furthermore, the validation experiment was done with a replacement column for the CEX column 1, which shows that this model is robust enough to cover this process variation as long as the packing-specific column parameters (ε_t , ε_p , Λ_{pack}) are used.

Furthermore, the simulated isocratic elution experiments in Figure 15 show that the model is able to describe the general elution behavior of the polypeptide, although only LGE experiments were applied for model calibration. As mentioned in section 3.2.1.4, the simulated signals were corrected with a marginal change in the elution salt concentration in a range of ± 0.0035 mol/L Na⁺, and an almost linear dependency of the salt correction with the column load could be observed. This linear correlation could be due to the fact that the introduced empirical pH dependence for the parameters K_s and K_p may not reflect the actual pH dependency. However, the salt correction is in a marginal range and below the error range of the system.

4.2.1.4 Multicomponent experiments and simulation with CEX column 1

After a successful model development for the four sample proteins in the linear range of the isotherm and model validation for the polypeptide species at loading conditions, multicomponent experiments and simulations in the non-linear range of the isotherm were conducted. These experiments should serve as a validation for the overall developed CEX model and should investigate the limitations of different multicomponent sample solutions.

Figure 16 shows an experiment at moderate non-linear loading conditions for the polypeptide species and linear loading conditions for variant B and variant C. The simulated signals for each species fit precisely to the experimentally determined signals, which shows that the developed model, calibrated for each species separately, can describe experimental multicomponent conditions. Figure 17 and Figure 18 show experiments with the impurities variant A and variant B and the polypeptide species at non-linear loading conditions. Variant A elutes in the front of the main peak. However, according to the data in the linear range, it should elute after the polypeptide species. This elution behavior could be explained by the displacement effect^{156, 157}, whereby a competitive behavior of the two species leads to an out crowding of the stationary phase of one species and different velocities of the species through the column. Variant B mainly elutes after the polypeptide peak, as investigated at low loading concentrations. However, it partially coelutes with the polypeptide. This effect is also called the tag-along effect, resulting from the competitive behavior of the different species^{156, 157}. After adjusting the equilibrium constant for the dimer layer for one impurity species (variant B), the model can describe the elution behavior of each species, including the displacement and the tag-along effect, with high accuracy. An adjustment of the parameter $K_{D,i}/c$ for the impurity species (variant B) is plausible, since this parameter depends on the dimerization capability of the protein and each species may have a different dimerization capability. However, this could not be investigated since not enough pure impurity sample was available for representative non-linear range experiments.

In order to investigate if the model is also able to predict multicomponent experiments with unknown impurities in it, an experiment with such a multicomponent sample was performed. Figure 19, Figure 20, and Figure 21 show the multicomponent experiment

and the corresponding simulated signal performed with a sample solution which includes the polypeptide, variant A, and variant B as known and model calibrated species and additionally with an unknown amount of undefined impurities. Considering only the three calibrated species with the experimentally applied load, the simulated polypeptide species is shifted to the right-hand side compared to the offline HPLC data. Variant A elutes in the front of the main peak as determined experimentally, and variant B coelutes with the polypeptide species. This simulated elution behavior of the three components can be explained by comparing it with the elution behavior in Figure 21, in which the amount of variant B was increased. It is shown that increasing the amount of variant B has a significant influence on the general elution behavior of the three components, whereby the whole peak is shifted to the left. However, by increasing the amount, variant B exhibits a strong tag-along effect in the polypeptide peak and therefore influences the overall eluting behavior of the polypeptide. These results show that the general elution behavior of the polypeptide species strongly depends on the concentration amount of the impurities. Simulation of a multicomponent sample with unspecified and uncalibrated impurities in the model is only partially meaningful in terms of representative peak intersections and pooling studies. However, the general elution behavior of the polypeptide could be predicted with high accuracy by considering the overall impurity amount.

Overall, the results in this section showed that each component species should be considered and modeled separately for an accurate and representative multicomponent model. With a multicomponent sample including unknown impurities, no accurate model could be developed with the calibrated species. Nevertheless, it is often impossible to have each species in a pure and isolated sample available. Therefore, for model development of such samples, a solution could be to create a hybrid model, whereas the empiric addition of several sample species, describing the impurity profile of the calibration experiments with additional parameter fitting. However, this method is, as described, rather empiric than mechanistic, resulting in an overall hybrid model. In summary, if modeling of a defined multicomponent with fewer than five impurity species is desired and sufficient samples of each species are available for model calibration in the linear range, separate consideration of each species, as at the beginning of this section, is recommended. Note that calibration in the linear range does not necessarily require a pure sample. Linear range calibration applying LGE experiments is also possible with a multicomponent sample. If modeling an undefined multicomponent

sample with an unknown amount of impurities is required, the hybrid modeling approach might be the better choice.

4.2.2 CEX model calibration at small scale (TOYOPEARL S-650S)

4.2.2.1 Investigation of the polypeptide elution behavior

To further validate the, in section 3.2.1, developed CEX model for the polypeptide species and investigate the applicability of the model to other CEX resins, experiments on an alternative CEX column with a different CEX resin (TOYOPEARL GigaCap S-650S) with the polypeptide species were performed. Consequently, the model parameters in the linear and the non-linear range according to the same scheme as for the CEX column 1 were determined.

As an initial investigation of the general elution behavior of the polypeptide on the column, a loading series up to 60 mg/mL_{CV} column load was performed (Figure 22). As observed for the CEX column 1, the polypeptide exhibits the same uncommon anti-Langmuirian to Langmuirian elution behavior with increasing load. These findings show that the adsorption isotherm must be sigmoid-shaped (type V¹⁵⁵), and dimer or multimer formation on the resin might occur¹³⁴. According to these results, the SAS isotherm used for the CEX column 1 should also be able to describe the elution behavior of the polypeptide on the TOYOPEARL S-650S column.

4.2.2.2 Determination of model parameters in the linear and non-linear range of the isotherm

Comparing the $GH_{salt}/c_{s,elu}$ - curve of the polypeptide species determined with the TOYOPEARL S-650S column and the CEX column 1 (Figure 23) shows that higher salt concentrations are necessary to elute the polypeptide in the TOYOPEARL S-650S column due to the determined significant higher ligand density. The resulting parameters for the linear range of the isotherm of the fit are summarized in Table 13. The parameter $\Delta G_s^0/RT$ for the TOYOPEARL S-650S column is with 0.21 lower than the determined value of 2.25 with the CEX column 1. As already mentioned in section 4.2.1.2, the parameter $\Delta G_s^0/RT$ describes the differences in the reference potential when the counterion changes its state from bound to solute state. Therefore, the values of both columns differ from each other as expected since the different ligand structures may lead to different $\Delta G_s^0/RT$ values although the same counterion was applied. The

same differences are observed in the $\Delta G_i^0/RT$ values for both columns. Although the same polypeptide species was used, the different ligand structures may lead to different $\Delta G_i^0/RT$ values. The parameters describing the pH-dependent characteristic charge v_i (N_{acidic} , N_{basic} , pka_{acidic} , and pka_{basic}) are all in a similar range except the number of acidic amino acid-like groups N_{acidic} for the TOYOPEARL S-650S column which is 0.42 and about half the amount from the model of the CEX column 1. According to the model, this value means that half the amount of acidic groups with a comparable pka value are present. The salt-dependent parameter K_s^* from the activity coefficient model of the protein species in solution was fitted to be -190 L/mol for the model of the TOYOPEARL S-650S column and is higher than the determined value for the polypeptide with the CEX column 1. However, as already mentioned above, the comparison of the value is only of limited significance since K_s depends on the characteristic charge v_i . Therefore, the pH dependency of K_s and the characteristic charge v_i for the polypeptide species determined with both CEX columns is compared in Figure 24. As can be seen, the K_s in both columns decrease with increasing pH and the value for the TOYOPEARL S-650S column is higher than the value for the CEX column 1. However, the values are in a similar range, and both are negative, which means that the protein-salt interaction is stronger than the protein-water interaction⁸³, leading to an activity coefficient of the protein in solution less than 1. The pH dependency for the characteristic charge for the polypeptide species fitted on both CEX columns is similar and varies in a range of 3.65 and 3.20 in the tested pH range. Since the characteristic charge is a resin-dependent parameter, the values vary as expected.

The influence of the two activity coefficients and the ratio of these two activity coefficients on the equilibrium constant for the polypeptide species were compared for both CEX column models in Figure 25. In general, both activity coefficients from the TOYOPEARL S-650S column model are slightly higher than the activity coefficients of the CEX column model. The difference of both activity coefficients is relatively small and only occurs due to the slightly different fitted v_i and K_s^* values (see Equations (16) and (18)). Note that both activity coefficients are referred to the solute state of the species^{83, 106} and, therefore, only slightly differ as expected in a small range since the same protein and salt species were used for the modeling procedure.

As done for the CEX column 1, the inverse fitting method was applied for the determination of the non-linear range isotherm parameters for the polypeptide species on the

TOYOPEARL S-650S column. The parameters for both CEX column models are compared in Table 14. The resulting comparison of the simulated signals with the experimentally determined signal of the calibration experiments is shown in Figure 26. As one can see, the applied SAS isotherm is able to describe the general elution behavior and peak shape with high accuracy up to the maximum tested load of 60 mg/mL_{CV}, which validates the developed model for the polypeptide species on CEX resins and indicates that the polypeptide shows the same general anti-Langmuirian to Langmuirian elution behavior on these two CEX resins. Comparing the non-linear range parameters shows that the protein-dependent parameter K_p^* from the activity coefficient model for the solute protein species is, as expected, in a similar range for both CEX column models since the same protein species in solution was used. The fitted parameter for the equilibrium constant for the dimer layer $K_{D,i}/c$ is also in both CEX column models in a similar range. Since the $K_{D,i}$ describes the dimer formation of a monomer to an already bound monomer^{88, 102, 149}, it is plausible that $K_{D,i}$ is for both CEX column models in a similar range, because the dimer equilibrium of the dimer formation to an already bound monomer should be ligand-independent. The fitted shielding factor of the polypeptide species at pH 3.7 is for the TOYOPEARL S-650S column model and the CEX column 1 model 7.2 and 8.4, respectively. Since the shielding factor describes the ligand shielding effects of the protein species⁸², one could expect that with a significantly higher ligand density, as observed with the TOYOPEARL S-650S column compared to the CEX column 1, the shielding factor should increase, which was not observed in this work. One effect could be that the TOYOPEARL GigaCap S-650S resin has a higher solid phase surface area than the CEX column 1 resin. However, this effect is probably not significant enough to lead to a smaller shielding factor for the TOYOPEARL S-650S column model for the polypeptide. Another reason could be the difference in the salt-dependent parameter K_s (Figure 24), which is although in the same range. However, a change in this range leads to a significant difference in the simulated peak shape, which might lead to the theoretically not plausible comparable shielding factors for both CEX column models.

Overall, the application of the in section 3.2.1 developed model to a different CEX resin shows that the model is generally able to describe the elution behavior of the polypeptide species on the CEX resins mechanistically.

4.2.3 Scale-up investigation of the CEX column 1 and the CEX model

After successful model development at small column scale with the CEX column 1, the column itself and the developed mechanistic model were investigated for the scalability to larger column volumes with different column dimensions. Therefore, the influence of the flow rate, the residence time, and the column-specific parameters on the elution behavior of the polypeptide on the different column scales was investigated. Scale-up simulations were then done by considering the column-specific parameters and by considering packing inhomogeneities.

4.2.3.1 Influence of the flow rate and residence time in the linear and non-linear range

The influence of the linear flow rate and the residence time (see Table 15) on the elution behavior in the linear and non-linear loading range was investigated and compared (Figure 27). As expected, the flow rate influences the peak shape but not the elution volume at low loading conditions (load: 0.05 mg/mL_{CV}). The lower the residence time at one column scale, the broader and smaller elutes the peak with the same elution volume as reported^{108, 150, 158}. The peak broadening is a consequence of flow rate-dependent mass transfer and axial dispersion effects in the column. Since, in this work, a lumped rate model was used, the mass transfer effects were considered by fitting different values for the lumped effective mass transfer coefficient $k_{eff,i,j}$ for each column scale j and column load. In the non-linear range (load: 10 mg/mL_{CV}), the residence time shows the same effect on the peak shape at one column scale, but as expected, it is not as pronounced as in the linear range. Therefore, mass transfer and axial dispersion effects on the peak shape induced by different flow rates in the non-linear range can be neglected and scaling up to larger column scales could be done without keeping the residence time constant for the CEX column 1-4 with the polypeptide species.

Figure 28A shows a polypeptide load comparison at 0.05 mg/mL_{CV} and 10 mg/mL_{CV} with the same linear flow rate and gradient slope across the different CEX column scales. The same effects as discussed above are observed. In Figure 28B, a comparison with the same residence time and the same gradient slope is shown and also shows similar effects as already discussed. The peaks in the linear and non-linear range have different salt elution concentrations. Furthermore, there are slight

differences in the peak height for peaks with the same residence time at low loading conditions. Al-Jibbouri compared the elution profile of β -lactoglobulin A and B in the linear range at different column scales with the same gradient slope and scaled them by applying the concept of time scales. The same variation as observed in this work in the linear range for the peak heights could be found¹³⁷. If the columns had identical column parameters, the peaks should be eluting at the same salt concentration and with a comparable peak height and shape with the same residence time and gradient slope.

4.2.3.2 Load comparison at different column scales and scaling based on the normalized gradient slope

The different eluting salt concentrations in the linear and non-linear range in Figure 28 can be explained by the different normalized salt gradient slope $GH_{salt,j}$ values for each column scale j caused by the column scale specific parameters ($\varepsilon_{0,j}$, $\varepsilon_{t,j}$, $\varepsilon_{p,j}$, Λ_j , and $k_{d,i,j}$). Different resin batches were considered in the model by considering resin-specific parameters. To achieve similar eluting salt concentrations $c_{s,elu}$ across different scales, the normalized salt gradient slope $GH_{salt,j}$ should be adjusted by varying the gradient volume $V_{g,j}$ and considering the column-specific parameters according to Equations (21) and (22). Individual $GH_{salt,j}$ values were calculated to obtain the same eluting salt concentration. The experimental data is shown in Figure 29. The eluting salt concentrations for the peaks at low and high loading conditions are now in a similar range, and the peak heights and shapes at 10 mg/mL_{CV} column load coincide with each other. The remaining slight difference in the position of the peaks could be caused by the imprecise experimentally determined interstitial porosities. The peak widths in the linear range differ due to the already mentioned mass transfer and axial dispersion effects with the same linear flow rates. Figure 29B shows the load comparison with individually calculated normalized gradient slopes and the same residence time for each column scale. The peak heights at low loading conditions are now in a similar range due to the same residence time. However, the residence time could be considered based on the interstitial or total column volume to obtain even more similar peak shapes. Additionally, a flow rate dependent *HETP* value could be considered as described by Yamamoto et al.¹⁵¹. The elution profiles in the non-linear range coincide with each other due to the neglectable effect of the residence time. It should be

mentioned that the peak areas in Figure 28 and Figure 29 are not identical for each experiment at different column scales because different batches of samples were used that did not always have the same concentration. Furthermore, the elution order in Figure 28A and Figure 29A is different from the elution order in Figure 28B and Figure 29B. However, the elution order varied only in a small conductivity range (equivalent to < 0.005 mol/L NaCl) which could be due to the use of different buffer batches and system inaccuracies.

These results of the scalability investigation of the pre-packed small-scale column show that the packing quality of each column scale could vary. Therefore, individual column parameters for each column scale should be considered to achieve identical peak shapes and retention volumes when scaling a chromatographic step.

4.2.3.3 Scale-up simulations

After the scalability investigation of the small-scale pre-packed calibration column (CEX column 1) to larger column volumes, the scalability of the developed mechanistic model was investigated. Therefore, various experiments with different column loads for each column scale were simulated and compared to experimental data (Figure 30). The simulated UV signals for the 5 mL, 10 mL, and 28.2 mL column scale fit not as precisely as the UV signal from the 1 mL column. To adapt the peak shape, the effective mass transfer coefficient was first fitted to the experimental peak shape. However, the parameter had no significant effect on the peak shape, so a good fit could not be achieved. In the first empirical approach, the shielding factor for each column scale was additionally fitted to the experimental data by applying the inverse fitting method. Only after the shielding factor was adapted did the simulated signals correlate with high accuracy to the experimentally determined signals. However, the shielding factor describes ligand shielding effects of the protein and should be constant at similar ligand densities as observed in this work for the different scale-up columns. A linear correlation can be found by plotting the fitted shielding factors against the asymmetry factor of each column scale (Figure 31). These findings indicate that different packing qualities, which lead to different asymmetry factors, might significantly influence the peak shape in the non-linear range and, therefore, the shielding factor. The worse the quality of the packed column, the more the shielding factor changes. This effect could explain the varying peak shape, which was compensated in the first empirical approach with a

shielding factor variation at different column scales. Various publications can be found in the literature investigating the influence of an inhomogeneous column packing in the linear and non-linear range ⁶¹⁻⁷². It is shown that packing inhomogeneities in the radial direction of the column lead to different radial velocity profiles, which can significantly influence the peak shape in low and high loading conditions ⁶³. For the description and the simulation of the radial velocity profile in the column, it is possible to divide the column into several concentric annular columns with different flow rates. The elution profiles from each column can then be summed up ⁶¹. The results in Figure 32 and Figure 33 show that by considering a radial porosity profile for each CEX scale-up column, the model is able to describe the entire data set with constant model parameters. In addition, the salt gradient could also be accurately described, indicating that a more detailed model for the elution of the salt species is not needed. One should note that this work focuses on the simulation of high loading situations. The effective mass transfer coefficient and the axial dispersion coefficient at these loading conditions showed a negligible effect on the simulated peak. This observation suggests that the change in radial porosity is largely responsible for the elution behavior of the polypeptide and that the neglect of radial dispersion is not important. If the elution behavior is considered only under analytical loading conditions, the assumption that there is no radial dispersion may be oversimplified.

As shown in Table 16, for the CEX column 2 (5 mL), the main interstitial porosity was fitted between 0.390 and 0.400, and only a minor part of the column had a fitted interstitial porosity of 0.330. The small part of the column with lower interstitial porosity can be seen as the outer rim of the column. The CEX column 3 (10 mL) shows a radial porosity profile from 0.450 to 0.420. The CEX column 4 (28.2 mL) column has a main fitted porosity of 0.355 and an outer rim with a higher fitted interstitial porosity. It should be noted that it cannot be said in which radial position the concentric annular columns are located due to the assumption that each column has the same face area and column volume. However, remembering the wall effect in packed columns ^{159, 160}, we assume that the concentric annular columns with the most significant deviation are located in the outer part of the column near the column wall. Therefore, the CEX column 2 (5 mL) and the CEX column 3 (10 mL) have lower fitted interstitial porosities, and the CEX column 4 (28.2 mL) has a higher fitted interstitial porosity near the column wall than at the column center. This radial interstitial porosity deviation correlates with the asymmetry factor for each column. The CEX column 3 (10 mL) has the most significant

deviation and the highest asymmetry factor (> 1), indicating the poorest packing quality, followed by the CEX column 2 (5 mL) with the second largest deviation. The CEX column 4 (28.2 mL) has the smallest deviation, a mirrored interstitial porosity profile, and an asymmetry factor smaller than 1. In the literature, values can be found for the thickness at which the wall effect mainly occurs. Values for analytical columns are given with about 30-50 d_p ^{161, 162}. For small preparative columns with a diameter of 50 mm, the affected wall region was found to be thicker ¹⁶³. Comparing these values, assuming that the thickness for the scale-up column is similar to small preparative columns, with the radial distance of the concentric annular column with the most significant deviation near the column wall in Figure 32, we find that they are in a similar range.

If one compares the experimental interstitial porosity determined with Blue Dextran to the fitted main porosity values in Table 16, it becomes clear that these values differ. The CEX column 3 (10 mL) shows the most significant difference, and the CEX column 4 (28.2 mL) has the smallest difference, which also correlates with the asymmetry factor. It should be noted that the porosity determination with a high molecular weight tracer (Blue Dextran) often causes difficulties since an ideal high molecular weight tracer should be spherical and inert to interactions with the resin. In addition, the tracer should be large enough to prevent pore penetration and still be highly mobile ¹²³. It is challenging to meet all the requirements with the resins commonly used in liquid chromatography. The blue dye from the Blue Dextran molecule (Cibacron blue F3GA), covalently bonded to the large dextran molecule, is hydrophobic ¹⁶⁴ and could form non-specific hydrophobic interactions with the resin matrix. Therefore, it is likely that incorrect porosity values were determined experimentally. Comparing the fitted interstitial porosity values with the experimentally determined interstitial porosities is, therefore, only of limited significance. It is worth mentioning that the scaling based on the relationship between the normalized gradient slope and the eluting salt concentration, as done in Figure 29, might be more accurate with proper experimentally determined interstitial porosity values. Furthermore, no porosity inhomogeneity was considered in the model calibration process with the CEX column 1 (1 mL) ⁷³. Therefore, possible inappropriate peak shape describing parameters (mainly shielding factor) could be determined for the CEX column 1 (1 mL). It is not necessarily clear that the real shielding factor for all column scales is 8.4 at pH 3.7. However, we decided to take the calibration

column (CEX column 1) as a reference column for model calibration, which should be scaled up.

Overall, the investigation of the scale-up ability of the devolved model with a pre-packed small-scale column showed that a column's packing quality significantly influences the peak shape of an eluting protein sample. Therefore, this may lead to inappropriate model parameters determined at one column scale if the packing quality and the radial porosity profile of a scale-up column are not considered in the mechanistic model.

4.3 RP model development

This section discusses the results of the RP model development done with the RP column 1.

4.3.1 RP model calibration at small scale (RP column 1)

4.3.1.1 Investigation of the polypeptide elution behavior

For the model development of the RP chromatography step, the same method used for the CEX model was applied. At the beginning of model development, the general polypeptide elution behavior on the RP column 1 in an organic modifier gradient elution at different pH values, shown in Figure 34, was investigated. As motioned in section 3.3.1.1, the polypeptide exhibits in the range from pH 3.3 to pH 3.7 a Langmuirian elution behavior, which should be describable by a convex isotherm. At higher pH values and loads, the elution behavior changes from Langmuirian to anti-Langmuirian. In general, at pH values above pH 3.9, the polypeptide exhibits the opposite elution behavior on the RP column 1 as observed with the CEX column 1. Remembering that the anti-Langmuirian elution behavior on the CEX column 1 was described considering dimer formation, the anti-Langmuirian elution behavior observed on the RP column 1 might also result due to pH-dependent multilayer formation on the resin. A similar pH-dependent elution behavior was observed by Gritti et al. with acids and bases at non-linear range conditions on RPC¹⁶⁵. They observed a pH-dependent shift from a Langmuirian to an anti-Langmuirian elution behavior and explained it with a competitive adsorption behavior of couples of acido-basic conjugated compounds. This mechanism could be transferred to the results observed in this work. As the polypeptide approaches its pka by increasing pH, it could be possible that the polypeptide then forms multimers at higher concentrations leading to an anti-Langmuirian elution behavior. Furthermore, an anti-Langmuirian elution behavior of a peptide could also be caused in RPC by increasing peptide porosity with increasing the local negative charges, as observed by Gétaz et al.¹²⁵. Since the net negative charge increases by increasing the pH, this mechanism could also be an explanation for the, in this work, observed uncommon pH-dependent Langmuirian to anti-Langmuirian elution behavior in RPC.

However, since the focus of the model should lie in the range of pH 3.1 to pH 3.7, the elution behavior on the RP column 1 was described using a Langmuirian adsorption

isotherm. The isotherm used in this work is the convex-shaped RP isotherm model developed by Arkell et al. ¹¹⁶⁻¹¹⁸. The model is based on the principle of the SMA formalism extended with various activity coefficient models for each modulator and protein species. A detailed derivation of the model is presented in section 1.4.2.

4.3.1.2 Determination of model parameters in the linear and non-linear range of the isotherm

Comparing the experimental $GH_M/x_{M,elu}$ data points determined for several pH values show that there is, as expected, a neglectable influence of the pH on the retention behavior in the linear loading range for all three calibrated protein species. In RPC, the retention of a protein at analytical conditions mainly depends on hydrophobic interactions. As long as a pH change does not markedly change the hydrophobicity conditions, the retention should be pH-independent. In this work, a pH variation from pH 3.1 to pH 3.9 was tested, which is relatively small and should therefore be considered pH-independent in the linear range of the adsorption isotherm. Nevertheless, as mentioned in the previous section, the pH influences the elution behavior at high loading conditions. Due to the neglectable influence of the pH in the linear range, Equation (57) was fitted only to the data points determined at pH 3.5. The fitted linear range parameter set for all three protein species is shown in Table 17. The lumped parameter $(\eta\tau^2)_i$ is for all three protein species in the same range and relatively small with a factor in the range of the power of -45. However, the assumption zero for that parameter leads to a marginally different fit. Therefore, this assumption was not made. The parameter $(\eta\tau^2)_i$ is a lumped parameter in the RPC model consisting of the protein- and salt-specific parameters η_i , τ_{p_i} , and θ_{p_i} ¹¹⁶. The lumped parameter $A'_{0,i}$, which initially contains the equilibrium constant of the adsorption mechanism and further lumped constants and which was estimated in the linear range (see section 1.4.2 or the reference from Arkell et al.), is also in the same range for all protein species and appears to be in a plausible range, as Arkell et al. determined similar values for three insulin variants on a C₁₈ reversed-phase resin ¹¹⁶. The parameter α_H , from the activity coefficient describing the effect of the organic modifier on the protein activity, was also fitted in the linear range and is for the three calibrated polypeptide species between 24 and 35. These values seem to be in a reasonable range, since Arkell et al. determined similar values for the insulin variants on the a C₁₈ reversed-phase resin ¹¹⁶. Furthermore, since the structure of the three protein species is similar, the effect of the organic modifier

should be similar too, which coincides with similar α_H values. However, as already stated in the work from Arkell et al. the parameter α_H is not representative for any physiochemical property, and therefore difficult to compare to literature values. Furthermore, similar values for the parameters $\vartheta_i \zeta_i$ were fitted. Note that ϑ_i and ζ_i have the same meaning as the characteristic charge in the IEC model, except that ϑ_i describes displacement of the protein to the ligand and ζ_i the displacement of the modulator molecule (organic modifier). In contrast, in the IEC model the modulator molecule is salt and it is assumed that only one salt molecule binds to one ligand. Similar $\vartheta_i \zeta_i$ values for the different protein species were fitted, which seem to be plausible since the structure of the three protein species does not differ much. However, for a detailed comparison of $\vartheta_i \zeta_i$ the parameter ϑ_i , which is determined in the non-linear range, is needed.

After parameter estimation in the linear range according to the Yamamoto approach, the parameters in the non-linear range were estimated by applying the inverse fitting method to all chromatographic calibration experiments. The final results and the comparison between the simulated and the experimentally determined signals are shown in Figure 36. The corresponding fitted non-linear range parameter set for the polypeptide species is shown in Table 18. The ligand density, which could also have been fitted in the non-linear range, was estimated from literature data from the resin vendor in order to reduce the fitted parameter set and was calculated to be 2.66×10^{-3} mol/L based on pore volume. The shielding factor fitted for the polypeptide species was 38 and is significantly higher than the shielding factor assumed by Arkell et al. for the insulin variants having a similar molecular size as the polypeptide species used in this work. However, the ligand density used by Arkell et al. was significantly higher, which could explain the difference between both estimated values. The binding parameter ϑ_i was fitted to be 0.25 for the polypeptide. The physiochemical meaning of that value is that the protein binds with a net binding site number of 0.25 to the hydroponic ligands and is, therefore, unexpectedly small. Consequently, this value indicates that the organic modifier displaces around 50 counter molecules with one molecule, which is an unexpectedly large number. Remembering that the silica-based resin might have ligands that are long tentacles, the values, however, become reasonable. During the non-linear range parameter fitting, it became apparent that no satisfactory fit could be achieved without refitting the lumped linear range parameter $A'_{0,i}$. Therefore, new fits

were started by varying $A'_{0,i}$, which is then named $A'_{0,i,n}$. The parameter $A'_{0,i,n}$ was fitted to be 7.07×10^{-8} and is marginal lower than the value fitted in the linear range. Nevertheless, without an adaption of the lumped parameter $A'_{0,i}$ to $A'_{0,i,n}$, no good fit could be achieved for the polypeptide species. Since the whole model for the simulation of the RPC was adopted from the work of Arkell et al. and since the model was developed for a specific case with different general experimental conditions (sample protein, buffer substances), the adjustment of lumped model parameters is theoretically plausible. An inappropriate assumption might be the linear dependency of $\ln(A_{H,i})$ with the mole fraction of the organic modifier, consequently leading to further assumptions in the activity coefficient describing the effect of the organic modifier on the protein activity. In this work, the dependency of $\ln(A_{H,i})$ with the mole fraction of the organic modifier is not straight linear but instead shows a quadratic relationship (data not shown), which may lead to inappropriate values for the parameter $A'_{0,i}$. Furthermore, the assumption made in this work that the exclusion factor and the local porosity are constant during different organic modifier concentrations may be inappropriate since it is known that these parameters change with the organic modifier concentration ¹²⁶. Nevertheless, in general, the simulated UV signals calculated from the model describe the experimentally determined UV signal in the calibration experiments with high accuracy. However, there are minor deviations at the top of the peak at higher loads than 15 mg/mL_{CV}. These deviations could also be a consequence of the simplified mass transfer kinetic used in this work.

After successful model calibration in the linear and non-linear range with the polypeptide species, a multicomponent experiment was conducted to calibrate the product-related impurities variant A and variant B. The fitted parameter set and the simulated signals compared to the experimental signals are shown in Table 19 and Figure 37. In order to reduce the fitted parameters, it was assumed that the impurities have the same shielding factor as the polypeptide due to their similar structure. Similar values for the ϑ_i parameter were fitted and furthermore, the linear range parameter $A'_{0,i}$ was also adjusted, which resulted in marginal different values for $A'_{0,i,n}$. Nevertheless, the experiment validated the calibrated parameters for the polypeptide species. Furthermore, it shows that the developed model is able to describe the elution behavior of the product-related impurities. Variant A shows a typical displacement behavior in the front of the product peak. Variant B shows a tag-along effect and extends from the end of the

main peak into the middle of the main peak with a broad elution peak. The physical mechanism behind the displacement and the tag-along effect is a consequence of different velocities of the protein species in the chromatographic column and competitive out crowding effects ¹⁵⁶. Overall, the multicomponent system could be calibrated with high accuracy. Furthermore, it shows that, as already discussed during the multicomponent simulations with the CEX model (section 4.2.1.4), it is more meaningful to consider each component separately for a representative model.

4.4 CEX model and RP model coupled simulations sampling study

After successful model development for both chromatographic steps, a sampling study was performed, and theoretical pools were built to investigate the influence of the loading volume and the polypeptide load purity on the resulting pool volume, the pool concentration of the polypeptide, the polypeptide step yield, and the polypeptide step recovery. The pooling criterion for the first chromatographic step (CEX model) was > 92 % polypeptide purity, and for the second chromatographic step (RP model), > 96 % polypeptide purity. The visualized multivariate analysis and the corresponding data are shown in Figure 38 and Figure 39 and Table 21 and Table 22 for the first and second chromatographic steps, respectively.

In order to mimic different pool sizes and polypeptide purities from the previous purification step, the sampling study for the first chromatographic step was performed with different load volumes and polypeptide purities. In general, in the first chromatographic step, a higher load volume with the same polypeptide purity as the reference load (2.75 mL with a polypeptide purity of 88.2 %) resulted in a higher theoretical pool volume (Figure 38A), an increased polypeptide pool concentration (Figure 38B), an increased polypeptide step yield (Figure 38C), and a decreased polypeptide step recovery (Figure 38D). In contrast, a decreasing load volume compared to the reference load results in decreased values. Keeping the load volume constant with increasing polypeptide load purity compared to the reference load results in increased pool volumes, polypeptide pool concentrations, polypeptide step yields, and polypeptide step recoveries. A decreased polypeptide load purity with the same load volume as the reference load results in decreased values. In general, analyzing the overall sampling study of the first chromatographic step shows that the best step yields and the best step recoveries could be achieved with runs 18 and 19, as one may expect since these runs have the highest specific polypeptide load (about 15 mg/mL_{CV} to 16 mg/mL_{CV}) with the highest load purity (91.4% to 92%). However, since a higher polypeptide pool concentration and purity may not have been achieved in the previous step, run 6 results in a similar step recovery with half of the polypeptide yield. The specific load of run 6 is half the load from runs 18 and 19, meaning that two times run 6 equals the same resulting polypeptide recovery and purity as one times run 18 or 19 without changing the pool concentration from the reference pool. This yield with two times run 6 could

be achieved only by lowering the load volume from 2.75 mL to 2.06 mL with the same polypeptide purity as the reference load.

All built theoretical pools from the first chromatographic step were consequently applied in the second chromatographic step (RP model). Further theoretical pools were built to investigate the influence of the different load conditions. As visible in Figure 39, applying the theoretical pools, which resulted in the previous step to the highest polypeptide step yield and step recovery (runs 18 and 19), also results in the best step yield and step recoveries for the second chromatographic step. However, run 6 results in that step in a slightly lower step recovery than runs 18 and 19. Run 2 also provides a good step yield with a higher polypeptide yield than run 6. Returning to the first step shows that run 2 delivers similar values to run 6. Summarizing, as expected, the runs with the highest specific polypeptide load in the first step also result in the highest recovery and polypeptide yield. However, as already mentioned, if a higher polypeptide pool concentration and purity before the first step are not possible, two times run 2 or 6 delivers the same results without changing the pool concentration and the purity (reference load).

One should note that this sampling study was performed by hand in which each condition for each chromatographic step was simulated and evaluated by hand, because the custom simulations file written with the ODE solver software Berkely Madonna does not include such automatic studies. Commercial software is available for the simulation of liquid chromatography, for example, the software DSPX from GoSilico. This software includes such sampling studies, which can be done entirely automatized. However, the included binding models in DSPX cannot describe the polypeptide's uncommon elution behavior in the first step. Moreover, implementing custom adsorption isotherms is only possible with commercial customer support.

Nevertheless, this sampling study successfully showed that the two representative models developed at small column scale could be coupled for a sampling study. Furthermore, the model could be used to examine the most appropriate conditions for the purification process, which may not be apparent when defining the different pooling criteria and load volumes for each chromatographic step.

5 CONCLUSION

In a preparative chromatographic two-column purification process, a complex elution behavior of a polypeptide in a counterion-induced CEX chromatography step was observed. Isotherm research showed that anti-Langmuirian and Langmuirian behavior occurred at low and high sample concentrations, resulting in a sigmoid-shaped isotherm with a concave shape at low and a convex shape at high protein concentrations. The uncommon elution behavior of the polypeptide on the CEX resin in the linear and non-linear range could be successfully described using an extended version of the SAS isotherm^{74, 88, 102} under consideration of the activity coefficients for the salt and the protein species in solution. On an RPC resin, the polypeptide exhibits a Langmuirian elution behavior which could be successfully described using a mechanistic RP chromatography model developed by Arkell et al.¹¹⁶. Additional to the polypeptide, two product-related and one process-related impurity variants were successfully modeled on both chromatographic steps. Multicomponent experiments showed that the developed mechanistic model is able to describe the competitive behavior of the protein species with high accuracy. The modeling study with the polypeptide and the impurity variants on both chromatographic steps demonstrated that in mechanistic modeling of liquid chromatography, it is crucial to be aware of the sample's overall elution behavior and the shape of the isotherm. In addition, activity coefficients should not generally be neglected in preparative chromatography due to non-ideal thermodynamic conditions.

Accurate elution prediction of the sample at high loading situations is mandatory under commercial and industrial conditions. Simulations performed with a mechanistic model lead to improved process understanding, which regulatory agencies desire. The mechanistic model developed in this work could be used for industrial high load multicomponent process conditions. A drawback of this method is that sufficient amounts of pure product and pure impurities are needed to determine the model parameters of each species. Parameter determination in the non-linear range of the isotherm with a multicomponent sample applying fraction analysis as done in references^{166, 167} avoids the need for a pure sample. However, it does not clarify the influence of the impurities in high load situations on the elution behavior of the product molecule. Therefore, no representative model parameters, for example, shielding effects in the non-linear range

of the isotherm, might be determined for the species in the multicomponent sample. Nevertheless, this work provides a working scheme for the mechanistic modeling of a chromatographic step in the linear and non-linear range of the isotherm.

Model calibration for both chromatographic steps was done with pre-packed mini-columns. Detailed column characterization with the CEX columns showed that the scalability of the pre-packed mini-columns to larger column volumes with different column dimensions strongly depends on each column's specific parameters and their unique packing quality. Although the quality parameters of the column packing are within a reasonable range, the columns are not comparable in terms of performance and elution behavior with a sample. Scaling based on the relationship between the normalized gradient slope and the eluting salt concentration by considering the unique column parameters leads to similar elution salt concentrations and peak shapes. By considering a radial interstitial porosity profile, the mechanistic model is able to describe the whole scale-up data set up to a large benchtop column scale with constant model parameters¹³¹. The scalability investigation of the pre-packed columns and the developed model showed that the assumption of the same packing quality and column characteristics, often made in the literature, might be inappropriate when scaling a chromatographic step because column properties and packing quality can significantly affect the elution profile. The scale-up procedure used in this work could be applied for faster scouting runs at small column scales with different residence times compared to the process scale. If peak widths and heights are held constant by considering the normalized gradient slope across different column scales, the pooling criteria based on peak widths can also be scaled up to process conditions. Furthermore, it was shown that too small benchtop columns (column volume less than 1 mL) might be inappropriate for mechanistic model calibration and scale-up due to the uncertainties in the column characterization and the fact that the packing quality can significantly influence the peak shape in the non-linear range. A more appropriate approach might be to use these mini-columns for the initial selection of a suitable isothermal model. Consequently, the detailed calibration of the model on an appropriate SDM column, considering possible packing inhomogeneities, could be done. A transfer of the model from SDM to the manufacturing scale might be possible if one is aware of the packing quality of the column used in the manufacturing scale and by considering it in the model.

Finally, with the coupling of the two developed models for each chromatographic type, it was shown that a *in silico* study could be used for evaluating the best purification conditions in the two-column purification step.

6 ABSTRACT

Biopharmaceuticals represent some of the best accomplishments of modern medicine. Over the last decades, more and more new biopharmaceutics have been developed to treat a wide range of diseases. Nevertheless, the purification process of a biopharmaceutical is still a challenging task. Currently, the method of choice for high-resolution product purification mainly focuses on chromatographic techniques.

In this work, a preparative multicomponent chromatographic two-column purification process was modeled using mini-columns for model calibration. A polypeptide showed an uncommon anti-Langmuirian to Langmuirian elution behavior on the cation-exchange chromatography resin at low and high loading conditions. On the reversed-phase resin, however, the polypeptide exhibits a Langmuirian elution behavior. The elution behavior of the polypeptide on the cation-exchange resin could be successfully described using an extended version of the self-association isotherm under consideration of the activity coefficients for the solute salt and the protein species. The elution behavior on the reversed-phase resin was successfully modeled using a mechanistic model describing multiple effects between the organic modifier, salt modulator, and protein species. Multicomponent experiments showed that the mechanistic model can describe the competitive behavior between the polypeptide, two product-, and one process-related impurity variant with high accuracy on both chromatographic steps.

Since model calibration was done with pre-packed mini-columns, the scalability of these columns and the developed mechanistic model to larger column volumes was investigated and showed a strong dependency on each column's specific parameters and their unique packing quality. Under consideration of column-specific parameters, the pre-packed mini-column is scalable to larger column dimensions. Furthermore, under consideration of the packing inhomogeneities of each column, successful scale-up simulations with the developed mechanistic model could be done.

With a coupled multicomponent two-column simulation sampling study, the influence of product purity and load was investigated for each chromatographic step and evaluated in a multivariate analysis. This work showed the full potential of the developed model for the two-column purification process, which could be used in further process development and optimization studies.

7 REFERENCES

1. Jozala, AF, Geraldés, DC, Tundisi, LL, Feitosa, VdA, Breyer, CA, Cardoso, SL, Mazzola, PG, Oliveira-Nascimento, Ld, Rangel-Yagui, CdO, Magalhães, PdO, Oliveira, MAd, Pessoa, A: Biopharmaceuticals from microorganisms: from production to purification. *Brazilian Journal of Microbiology*, 47: 51-63, 2016. <https://doi.org/10.1016/j.bjm.2016.10.007>
2. Kesik-Brodacka, M: Progress in biopharmaceutical development. *Biotechnology and applied biochemistry*, 65: 306-322, 2018. <https://doi.org/10.1002/bab.1617>
3. Biotechnology inspection guide reference materials and training aids. In: FISH, R. C., SZE, T. T., RICE, K. A. (Hrsg.) *Biotechnology Inspection Guide (11/91)*. U.S. Department of Health and Human Services, 1991.
4. Johnson, IS: Human Insulin from Recombinant DNA Technology. 219: 632-637, 1983. <https://doi.org/10.1126/science.6337396>
5. McElwain, L, Phair, K, Kealey, C, Brady, D: Current trends in biopharmaceuticals production in *Escherichia coli*. *Biotechnology Letters*, 44: 917-931, 2022. <https://doi.org/10.1007/s10529-022-03276-5>
6. Donini, M, Marusic, C: Current state-of-the-art in plant-based antibody production systems. *Biotechnology Letters*, 41: 335-346, 2019. <https://doi.org/10.1007/s10529-019-02651-z>
7. Walsh, G: Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, 36: 1136-1145, 2018. <https://doi.org/10.1038/nbt.4305>
8. Kaplon, H, Chenoweth, A, Crescioli, S, Reichert, JM: Antibodies to watch in 2022. *mAbs*, 14: 2014296, 2022. <https://doi.org/10.1080/19420862.2021.2014296>
9. Carter, PJ, Rajpal, A: Designing antibodies as therapeutics. *Cell*, 185: 2789-2805, 2022. <https://doi.org/10.1016/j.cell.2022.05.029>
10. Madhavan, A, Arun, KB, Sindhu, R, Krishnamoorthy, J, Reshmy, R, Sirohi, R, Pugazhendi, A, Awasthi, MK, Szakacs, G, Binod, P: Customized yeast cell factories for biopharmaceuticals: from cell engineering to process scale up. *Microbial Cell Factories*, 20: 124, 2021. <https://doi.org/10.1186/s12934-021-01617-z>
11. Butler, M, Meneses-Acosta, A: Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Applied Microbiology and Biotechnology*, 96: 885-894, 2012. <https://doi.org/10.1007/s00253-012-4451-z>
12. Weber, W, Fussenegger, M: Insect Cell-Based Recombinant Protein Production. In: *Cell and Tissue Reaction Engineering: With a Contribution by Martin Fussenegger and Wilfried Weber*. Herausgegeben von EIBL, R., EIBL, D., PÖRTNER, R., CATAPANO, G., CZERMAK, P., Berlin, Heidelberg, Springer Berlin Heidelberg, 2009, S. 263-277.
13. Rosales-Mendoza, S, Solís-Andrade, KI, Márquez-Escobar, VA, González-Ortega, O, Bañuelos-Hernández, B: Current advances in the algae-made biopharmaceuticals field. *Expert Opinion on Biological Therapy*, 20: 751-766, 2020. <https://doi.org/10.1080/14712598.2020.1739643>
14. Müller, D, Bayer, K, Mattanovich, D: Potentials and limitations of prokaryotic and eukaryotic expression systems for recombinant protein production – a comparative view. *Microbial Cell Factories*, 5: P61, 2006. <https://doi.org/10.1186/1475-2859-5-S1-P61>
15. Muyldermans, S: Applications of Nanobodies. 9: 401-421, 2021. <https://doi.org/10.1146/annurev-animal-021419-083831>

16. Reuters: Why Sanofi Is Buying Belgian Biotech Company Ablynx for Nearly \$5 Billion, 2018. <https://fortune.com/2018/01/29/sanofi-ablynx-deal-caplacizumab/>. Abgerufen 25.10.2022.
17. Rosenmayr-Templeton, L: An industry update: December 2019, what is new in the field of therapeutic delivery this month? 11: 289-296, 2020. <https://doi.org/10.4155/tde-2020-0032>
18. Sanofi: Sanofi : Sanofi to acquire Synthorx to bolster its immuno-oncology pipeline for \$2.5 Billion, 2019. <https://www.sanofi.com/en/media-room/press-releases/2019/2019-12-09-06-00-00-1957603>. Abgerufen 25.10.2022.
19. Janku, F, Abdul-Karim, R, Azad, A, Bendell, J, Falchook, G, Gan, HK, Tan, T, Wang, JS, CHEE, CE, Ma, L, Mooney, J, Marina, N, Abbadessa, G, Milla, M, Meniawy, T: Abstract LB041: THOR-707 (SAR444245), a novel not-alpha IL-2 as monotherapy and in combination with pembrolizumab in advanced/metastatic solid tumors: Interim results from HAMMER, an open-label, multicenter phase 1/2 Study. *Cancer Research*, 81: LB041-LB041, 2021. <https://doi.org/10.1158/1538-7445.Am2021-lb041>
20. Sahin, U, Karikó, K, Türeci, Ö: mRNA-based therapeutics — developing a new class of drugs. *Nature Reviews Drug Discovery*, 13: 759-780, 2014. <https://doi.org/10.1038/nrd4278>
21. Qin, S, Tang, X, Chen, Y, Chen, K, Fan, N, Xiao, W, Zheng, Q, Li, G, Teng, Y, Wu, M, Song, X: mRNA-based therapeutics: powerful and versatile tools to combat diseases. *Signal Transduction and Targeted Therapy*, 7: 166, 2022. <https://doi.org/10.1038/s41392-022-01007-w>
22. Lin, Y-X, Wang, Y, Blake, S, Yu, M, Mei, L, Wang, H, Shi, J: RNA Nanotechnology-Mediated Cancer Immunotherapy. *Theranostics*, 10: 281-299, 2020. <https://doi.org/10.7150/thno.35568>
23. Watson, OJ, Barnsley, G, Toor, J, Hogan, AB, Winskill, P, Ghani, AC: Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*, 22: 1293-1302, 2022. [https://doi.org/10.1016/S1473-3099\(22\)00320-6](https://doi.org/10.1016/S1473-3099(22)00320-6)
24. Valeur, E, Guéret, SM, Adihou, H, Gopalakrishnan, R, Lemurell, M, Waldmann, H, Grossmann, TN, Plowright, AT: New Modalities for Challenging Targets in Drug Discovery. 56: 10294-10323, 2017. <https://doi.org/10.1002/anie.201611914>
25. Flickinger, MC: *Downstream industrial biotechnology recovery and purification*, Hoboken, N.J. : John Wiley & Sons Inc., 2013., 2013.
26. Strube, J, Grote, F, Josch, JP, Ditz, R: Process Development and Design of Downstream Processes. 83: 1044-1065, 2011. <https://doi.org/10.1002/cite.201100017>
27. Baumann, P, Hubbuch, J: Downstream process development strategies for effective bioprocesses: Trends, progress, and combinatorial approaches. *Engineering in Life Sciences*, 17: 1142-1158, 2017. <https://doi.org/10.1002/elsc.201600033>
28. Destro, F, Barolo, M: A review on the modernization of pharmaceutical development and manufacturing – Trends, perspectives, and the role of mathematical modeling. *International Journal of Pharmaceutics*, 620: 121715, 2022. <https://doi.org/10.1016/j.ijpharm.2022.121715>
29. Babi, DK, Griesbach, J, Hunt, S, Insaïdo, F, Roush, D, Todd, R, Staby, A, Welsh, J, Wittkopp, F: Opportunities and challenges for model utilization in the biopharmaceutical industry: current versus future state. *Current Opinion in Chemical Engineering*, 36: 100813, 2022. <https://doi.org/10.1016/j.coche.2022.100813>

30. Hanke, AT, Ottens, M: Purifying biopharmaceuticals: knowledge-based chromatographic process development. *Trends in Biotechnology*, 32: 210-220, 2014. <https://doi.org/10.1016/j.tibtech.2014.02.001>
31. Fields, C, Li, P, O'Mahony, JJ, Lee, GU: Advances in affinity ligand-functionalized nanomaterials for biomagnetic separation. 113: 11-25, 2016. <https://doi.org/10.1002/bit.25665>
32. Rathore, AS, Kumar, D, Kateja, N: Recent developments in chromatographic purification of biopharmaceuticals. *Biotechnology Letters*, 40: 895-905, 2018. <https://doi.org/10.1007/s10529-018-2552-1>
33. Shukla, AA, Hubbard, B, Tressel, T, Guhan, S, Low, D: Downstream processing of monoclonal antibodies—Application of platform approaches. *Journal of Chromatography B*, 848: 28-39, 2007. <https://doi.org/10.1016/j.jchromb.2006.09.026>
34. Reck, JM, Pabst, TM, Hunter, AK, Carta, G: Separation of antibody monomer-dimer mixtures by frontal analysis. *Journal of Chromatography A*, 1500: 96-104, 2017. <https://doi.org/10.1016/j.chroma.2017.04.014>
35. Lee, YF, Jöhnck, M, Frech, C: Evaluation of differences between dual salt-pH gradient elution and mono gradient elution using a thermodynamic model: Simultaneous separation of six monoclonal antibody charge and size variants on preparative-scale ion exchange chromatographic resin. *Biotechnology Progress*, 34: 973-986, 2018. <https://doi.org/10.1002/btpr.2626>
36. Boschetti, E: Antibody separation by hydrophobic charge induction chromatography. *Trends in Biotechnology*, 20: 333-337, 2002. [https://doi.org/10.1016/S0167-7799\(02\)01980-7](https://doi.org/10.1016/S0167-7799(02)01980-7)
37. Maria, S, Joucla, G, Garbay, B, Dieryck, W, Lomenech, A-M, Santarelli, X, Cabanne, C: Purification process of recombinant monoclonal antibodies with mixed mode chromatography. *Journal of Chromatography A*, 1393: 57-64, 2015. <https://doi.org/10.1016/j.chroma.2015.03.018>
38. Gagnon, P: IgG Aggregate Removal by Charged-Hydrophobic Mixed Mode Chromatography. *Current Pharmaceutical Biotechnology*, 10: 434-439, 2009. <https://doi.org/10.2174/138920109788488888>
39. Gao, D, Wang, L-L, Lin, D-Q, Yao, S-J: Evaluating antibody monomer separation from associated aggregates using mixed-mode chromatography. *Journal of Chromatography A*, 1294: 70-75, 2013. <https://doi.org/10.1016/j.chroma.2013.04.018>
40. Manzke, O, Tesch, H, Diehl, V, Bohlen, H: Single-step purification of bispecific monoclonal antibodies for immunotherapeutic use by hydrophobic interaction chromatography. *Journal of Immunological Methods*, 208: 65-73, 1997. [https://doi.org/10.1016/S0022-1759\(97\)00129-4](https://doi.org/10.1016/S0022-1759(97)00129-4)
41. Diogo, MM, Queiroz, JA, Monteiro, GA, Martins, SAM, Ferreira, GNM, Prazeres, DMF: Purification of a cystic fibrosis plasmid vector for gene therapy using hydrophobic interaction chromatography. 68: 576-583, 2000. [https://doi.org/10.1002/\(SICI\)1097-0290\(20000605\)68:5<576::AID-BIT13>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1097-0290(20000605)68:5<576::AID-BIT13>3.0.CO;2-5)
42. Johansson, K, Frederiksen, SS, Degerman, M, Breil, MP, Mollerup, JM, Nilsson, B: Combined effects of potassium chloride and ethanol as mobile phase modulators on hydrophobic interaction and reversed-phase chromatography of three insulin variants. *Journal of Chromatography A*, 1381: 64-73, 2015. <https://doi.org/10.1016/j.chroma.2014.12.081>

43. Łącki, KM: High-throughput process development of chromatography steps: Advantages and limitations of different formats used. 7: 1192-1202, 2012. <https://doi.org/10.1002/biot.201100475>
44. Treier, K, Lester, P, Hubbuch, J: Application of genetic algorithms and response surface analysis for the optimization of batch chromatographic systems. *Biochemical Engineering Journal*, 63: 66-75, 2012. <https://doi.org/10.1016/j.bej.2012.02.003>
45. Keller, WR, Evans, ST, Ferreira, G, Robbins, D, Cramer, SM: Understanding operational system differences for transfer of miniaturized chromatography column data using simulations. *Journal of Chromatography A*, 1515: 154-163, 2017. <https://doi.org/10.1016/j.chroma.2017.07.091>
46. Keller, WR, Evans, ST, Ferreira, G, Robbins, D, Cramer, SM: Understanding the effects of system differences for parameter estimation and scale-up of high throughput chromatographic data. *Journal of Chromatography A*, 1661: 462696, 2022. <https://doi.org/10.1016/j.chroma.2021.462696>
47. Keller, WR, Evans, ST, Ferreira, G, Robbins, D, Cramer, SM: Use of MiniColumns for linear isotherm parameter estimation and prediction of benchtop column performance. *Journal of Chromatography A*, 1418: 94-102, 2015. <https://doi.org/10.1016/j.chroma.2015.09.038>
48. FDA: ICH Guidance for Industry: Q11 Development and Manufacture of Drug Substances. 2012.
49. European Medicines Agency (EMA), ICH guideline Q11 on development and manufacture of drug substances. London, 2012.
50. Mollerup, JM, Hansen, TB, Kidal, S, Sejergaard, L, Staby, A: Development, modelling, optimisation and scale-up of chromatographic purification of a therapeutic protein. *Fluid Phase Equilibria*, 261: 133-139, 2007. <https://doi.org/10.1016/j.fluid.2007.07.047>
51. Rathore, AS, Velayudhan, A: *Scale-Up and Optimization in Preparative Chromatography: Principles and Biopharmaceutical Applications*, CRC PRESS, 2003.
52. Hansen, EB: Chromatographic Scale-Up on a Volume Basis. In: *Preparative Chromatography for Separation of Proteins*. Herausgegeben von STABY, A., RATHORE, A. S., AHUJA, S., Hoboken, New Jersey, 2017, S. 227-246.
53. Kidal, S, Jensen, OE: Using volumetric flow to scaleup chromatographic processes. *BioPharm International*, 19: 34-44, 2006.
54. Prentice, J, Evans, ST, Robbins, D, Ferreira, G: Pressure-Flow experiments, packing, and modeling for scale-up of a mixed mode chromatography column for biopharmaceutical manufacturing. *Journal of Chromatography A*, 1625: 461117, 2020. <https://doi.org/10.1016/j.chroma.2020.461117>
55. Antoniou, C, McCue, J, Natarajan, V, Thömmes, J, Yuan, QS: Scaling Up Industrial Protein Chromatography. In: *Preparative Chromatography for Separation of Proteins*. Herausgegeben von STABY, A., RATHORE, A. S., AHUJA, S., Hoboken, New Jersey, 2017, S. 247-268.
56. Saleh, D, Wang, G, Mueller, B, Rischawy, F, Kluters, S, Studts, J, Hubbuch, J: Cross-scale quality assessment of a mechanistic cation exchange chromatography model. *Biotechnology Progress*, 37: e3081, 2021. <https://doi.org/10.1002/btpr.3081>
57. Benner, SW, Welsh, JP, Rauscher, MA, Pollard, JM: Prediction of lab and manufacturing scale chromatography performance using mini-columns and mechanistic modeling. *Journal of Chromatography A*, 1593: 54-62, 2019. <https://doi.org/10.1016/j.chroma.2019.01.063>

58. Kołodziej, M, Sauer, DG, Beck, J, Marek, WK, Hahn, R, Jungbauer, A, Dürauer, A, Piątkowski, W, Antos, D: Scale up of a chromatographic capture step for a clarified bacterial homogenate – Influence of mass transport limitation and competitive adsorption of impurities. *Journal of Chromatography A*, 1618: 460856, 2020. <https://doi.org/10.1016/j.chroma.2020.460856>
59. Gu, T, Hsu, K-H, Syu, M-J: Scale-up of affinity chromatography for purification of enzymes and other proteins. *Enzyme and Microbial Technology*, 33: 430-437, 2003. [https://doi.org/10.1016/S0141-0229\(03\)00141-8](https://doi.org/10.1016/S0141-0229(03)00141-8)
60. Li, Z, Gu, Y, Gu, T: Mathematical modeling and scale-up of size-exclusion chromatography. *Biochemical Engineering Journal*, 2: 145-155, 1998. [https://doi.org/10.1016/S1369-703X\(98\)00027-8](https://doi.org/10.1016/S1369-703X(98)00027-8)
61. Nicoud, RM, Perrut, M: Hydrodynamics of Preparative Chromatography Columns. In: *Preparative and Production Scale Chromatography*. Herausgegeben von GANETSOS, G., BARKER, P. E., Boca Raton, CRC Press, 1992, S. 47-78.
62. Smith, MS, Guiochon, G: Theoretical analysis of the behavior of an heterogeneously packed chromatographic column. *Journal of Chromatography A*, 827: 241-257, 1998. [https://doi.org/10.1016/S0021-9673\(98\)00810-3](https://doi.org/10.1016/S0021-9673(98)00810-3)
63. Yun, T, Guiochon, G: Modeling of radial heterogeneity in chromatographic columns Columns with cylindrical symmetry and ideal model. *Journal of Chromatography A*, 672: 1-10, 1994. [https://doi.org/10.1016/0021-9673\(94\)80589-X](https://doi.org/10.1016/0021-9673(94)80589-X)
64. Huygens, B, Song, H, Cabooter, D, Desmet, G: Detailed numerical analysis of the effect of radial column heterogeneities on peak parking experiments with slowly diffusing analytes. *Journal of Chromatography A*, 1656: 462557, 2021. <https://doi.org/10.1016/j.chroma.2021.462557>
65. Miyabe, K, Guiochon, G: Numerical Method for the Estimation of Column Radial Heterogeneity and of the Actual Column Efficiency from Tailing Peak Profiles. *Analytical Chemistry*, 83: 182-192, 2011. <https://doi.org/10.1021/ac102195x>
66. Roberts, JA, Carta, G: Relationship between HETP measurements and breakthrough curves in short chromatography columns. *Biotechnology Progress*, 37: e3065, 2021. <https://doi.org/10.1002/btpr.3065>
67. Miyabe, K: Numerical correction for asymmetrical peak profiles for moment analysis of chromatographic behavior. *Journal of Chromatography A*, 1567: 155-163, 2018. <https://doi.org/10.1016/j.chroma.2018.06.067>
68. Miyabe, K, Matsumoto, Y, Niwa, Y, Ando, N, Guiochon, G: An estimation of the column efficiency made by analyzing tailing peak profiles. *Journal of Chromatography A*, 1216: 8319-8330, 2009. <https://doi.org/10.1016/j.chroma.2009.09.006>
69. Gonzalez, MJ, Jaulmes, A, Valentin, P, Vidal-Madjar, C: Modelling of peak profiles: Application to the preparative liquid chromatography of steroids. *Journal of Chromatography A*, 386: 333-344, 1987. [https://doi.org/10.1016/S0021-9673\(01\)94610-2](https://doi.org/10.1016/S0021-9673(01)94610-2)
70. Villiermaux, J: Deformation of Chromatographic Peaks Under the Influence of Mass Transfer Phenomena. *Journal of Chromatographic Science*, 12: 822-831, 1974. <https://doi.org/10.1093/chromsci/12.12.822>
71. Jaulmes, A, Ignatiadis, I, Cardot, P, Vidal-Madjar, C: Characterization of peak asymmetry with overloaded capillary columns. *Journal of Chromatography A*, 395: 291-306, 1987. [https://doi.org/10.1016/S0021-9673\(01\)94118-4](https://doi.org/10.1016/S0021-9673(01)94118-4)
72. McCue, JT, Engel, P, Thömmes, J: Modeling the effects of column packing quality and residence time changes on protein monomer/aggregate separation. *Journal of Chromatography A*, 1216: 4895-4901, 2009. <https://doi.org/10.1016/j.chroma.2009.04.041>

73. Koch, J, Scheps, D, Gunne, M, Boscheinen, O, Hafner, M, Frech, C: Mechanistic modeling and simulation of a complex low and high loading elution behavior of a polypeptide in cation exchange chromatography. *Journal of separation science*, 45: 2008-2023, 2022. <https://doi.org/10.1002/jssc.202200098>
74. Mollerup, JM, Hansen, TB, Frederiksen, SS, Staby, A: *Thermodynamic Modeling of Chromatographic Separation*, Bosa Roca, CRC Press Taylor & Francis Group, 2010.
75. Jen, S-CD, Pinto, NG: Nonlinear Chromatography of beta-Lactoglobulins A and B Non-Langmurian Behaviour. *Industrial and Engineering Chemistry Research*, 34: 2685-2691, 1995. <https://doi.org/10.1021/ie00047a018>
76. Seelinger, F, Wittkopp, F, von Hirschheydt, T, Frech, C: Application of the Steric Mass Action formalism for modeling under high loading conditions: Part 2. Investigation of high loading and column overloading effects. *Journal of Chromatography A*, 1676: 463266, 2022. <https://doi.org/10.1016/j.chroma.2022.463266>
77. Guo, J, Creasy, AD, Barker, G, Carta, G: Surface induced three-peak elution behavior of a monoclonal antibody during cation exchange chromatography. *Journal of Chromatography A*, 1474: 85-94, 2016. <https://doi.org/10.1016/j.chroma.2016.10.061>
78. Huuk, TC, Hahn, T, Doninger, K, Griesbach, J, Hepbildikler, S, Hubbuch, J: Modeling of complex antibody elution behavior under high protein load densities in ion exchange chromatography using an asymmetric activity coefficient. *Biotechnology Journal*, 12: 1600336, 2017. <https://doi.org/10.1002/biot.201600336>
79. Creasy, A, Barker, G, Carta, G: Systematic interpolation method predicts protein chromatographic elution with salt gradients, pH gradients and combined salt/pH gradients. *Biotechnology Journal*, 12: 1600636, 2017. <https://doi.org/10.1002/biot.201600636>
80. Creasy, A, Barker, G, Yao, Y, Carta, G: Systematic interpolation method predicts protein chromatographic elution from batch isotherm data without a detailed mechanistic isotherm model. *Biotechnology Journal*, 10: 1400-1411, 2015. <https://doi.org/10.1002/biot.201500089>
81. Creasy, A, Reck, J, Pabst, T, Hunter, A, Barker, G, Carta, G: Systematic Interpolation Method Predicts Antibody Monomer-Dimer Separation by Gradient Elution Chromatography at High Protein Loads. *Biotechnology Journal*, 14: 1800132, 2019. <https://doi.org/10.1002/biot.201800132>
82. Brooks, CA, Cramer, SM: Steric Mass-Action Ion Exchange: Displacement Profiles and Induced Salt Gradients. *AIChE Journal*, 38: 1969-1979, 1992. <https://doi.org/10.1002/aic.690381212>
83. Mollerup, JM: Applied thermodynamics: A new frontier for biotechnology. *Fluid Phase Equilibria*, 241: 205-215, 2006. <https://doi.org/10.1016/j.fluid.2005.12.037>
84. Diedrich, J, Heymann, W, Leweke, S, Hunt, S, Todd, R, Kunert, C, Johnson, W, von Lieres, E: Multi-state steric mass action model and case study on complex high loading behavior of mAb on ion exchange tentacle resin. *Journal of Chromatography A*, 1525: 60-70, 2017. <https://doi.org/10.1016/j.chroma.2017.09.039>
85. Briskot, T, Hahn, T, Huuk, T, Hubbuch, J: Adsorption of colloidal proteins in ion-exchange chromatography under consideration of charge regulation. *Journal of Chromatography A*, 1611: 460608, 2020. <https://doi.org/10.1016/j.chroma.2019.460608>

86. Briskot, T, Hahn, T, Huuk, T, Hubbuch, J: Protein adsorption on ion exchange adsorbers: A comparison of a stoichiometric and non-stoichiometric modeling approach. *Journal of Chromatography A*, 1653: 462397, 2021. <https://doi.org/10.1016/j.chroma.2021.462397>
87. Briskot, T, Hahn, T, Huuk, T, Wang, G, Kluters, S, Studts, J, Wittkopp, F, Winderl, J, Schwan, P, Hagemann, I, Kaiser, K, Trapp, A, Stamm, SM, Koehn, J, Malmquist, G, Hubbuch, J: Analysis of complex protein elution behavior in preparative ion exchange processes using a colloidal particle adsorption model. *Journal of Chromatography A*, 1654: 462439, 2021. <https://doi.org/10.1016/j.chroma.2021.462439>
88. Mollerup, JM: A Review of the Thermodynamics of Protein Association to Ligands, Protein Adsorption, and Adsorption Isotherms. *Chemical Engineering and Technology*, 31: 864-874, 2008. <https://doi.org/10.1002/ceat.200800082>
89. Khalaf, R, Coquebert de Neuville, B, Morbidelli, M: Protein adsorption in polyelectrolyte brush type cation-exchangers. *Journal of Chromatography A*, 1471: 126-137, 2016. <https://doi.org/10.1016/j.chroma.2016.10.024>
90. Stoddard, BL: *Computational Design of Ligand Binding Proteins*, New York, NY, Humana New York, NY, 2016.
91. Daviter, T, Johnson, CM, McLaughlin, SH, Williams, MA: *Protein-Ligand Interactions*, New York, NY, Humana New York, NY, 2021.
92. Shirai, H, Prades, C, Vita, R, Marcatili, P, Popovic, B, Xu, J, Overington, JP, Hirayama, K, Soga, S, Tsunoyama, K, Clark, D, Lefranc, M-P, Ikeda, K: Antibody informatics for drug discovery. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844: 2002-2015, 2014. <https://doi.org/10.1016/j.bbapap.2014.07.006>
93. Macalino, SJY, Gosu, V, Hong, S, Choi, S: Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, 38: 1686-1701, 2015. <https://doi.org/10.1007/s12272-015-0640-5>
94. Leach, AR, Shoichet, BK, Peishoff, CE: Prediction of Protein–Ligand Interactions. Docking and Scoring: Successes and Gaps. *Journal of Medicinal Chemistry*, 49: 5851-5855, 2006. <https://doi.org/10.1021/jm060999m>
95. Bhattacharjee, A, Wallin, S: Exploring Protein-Peptide Binding Specificity through Computational Peptide Screening. *PLoS Comput Biol*, 9: e1003277, 2013. <https://doi.org/10.1371/journal.pcbi.1003277>
96. Damborsky, J, Brezovsky, J: Computational tools for designing and engineering enzymes. *Curr Opin Chem Biol*, 19: 8-16, 2014. <https://doi.org/10.1016/j.cbpa.2013.12.003>
97. Heinzelman, P, Komor, R, Kanaan, A, Romero, P, Yu, X, Mohler, S, Snow, C, Arnold, F: Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination. *Protein Eng Des Sel*, 23: 871-880, 2010. <https://doi.org/10.1093/protein/gzq063>
98. Chen, JL, Stewart-Jones, G, Bossi, G, Lissin, NM, Wooldridge, L, Choi, EM, Held, G, Dunbar, PR, Esnouf, RM, Sami, M, Boulter, JM, Rizkallah, P, Renner, C, Sewell, A, van der Merwe, PA, Jakobsen, BK, Griffiths, G, Jones, EY, Cerundolo, V: Structural and kinetic basis for heightened immunogenicity of T cell vaccines. *J Exp Med*, 201: 1243-1255, 2005. <https://doi.org/10.1084/jem.20042323>
99. Fisher, HF, Singh, N: [9] Calorimetric methods for interpreting protein—Ligand interactions. In: *Methods in Enzymology*. Academic Press, 1995, S. 194-221.

100. Indyk, L, Fisher, HF: Theoretical aspects of isothermal titration calorimetry. *Methods Enzymol*, 295: 350-364, 1998. [https://doi.org/10.1016/s0076-6879\(98\)95048-0](https://doi.org/10.1016/s0076-6879(98)95048-0)
101. Johnson, CM: Isothermal Titration Calorimetry. *Methods in molecular biology (Clifton, NJ)*, 2263: 135-159, 2021. https://doi.org/10.1007/978-1-0716-1197-5_5
102. Mollerup, JM: Adsorption Isotherms: Fundamentals and Modeling Aspects. In: *Preparative Chromatography for Separation of Proteins*. Herausgegeben von STABY, A., RATHORE, A. S., AHUJA, S., Hoboken, 2017, S. 11-79.
103. Schmidt, M, Hafner, M, Frech, C: Modeling of salt and pH gradient elution in ion-exchange chromatography. *Journal of separation science*, 37: 5-13, 2014. <https://doi.org/10.1002/jssc.201301007>
104. Ribeiro, JM, Sillero, A: An algorithm for the computer calculation of the coefficients of a polynomial that allows determination of isoelectric points of proteins and other macromolecules. *Computers in Biology and Medicine*, 20: 235-242, 1990. [https://doi.org/10.1016/0010-4825\(90\)90048-T](https://doi.org/10.1016/0010-4825(90)90048-T)
105. Meireles Ribeiro, J, Sillero, A: A program to calculate the isoelectric point of macromolecules. *Computers in Biology and Medicine*, 21: 131-141, 1991. [https://doi.org/10.1016/0010-4825\(91\)90022-2](https://doi.org/10.1016/0010-4825(91)90022-2)
106. Davies, CW: *Ion Association*, London, Butterworths, 1962.
107. Debye, P, Hückel, E: Zur Theorie der Elektrolyte. *Physikalische Zeitschrift*, 9: 185-206, 1923.
108. Yamamoto, S, Nakanishi, K, Matsuno, R, Kamikubo, T: Ion exchange chromatography of proteins - prediction of elution curves and operating conditions. I. Theoretical considerations. *Biotechnology and bioengineering*, 25: 1465-1483, 1983. <https://doi.org/10.1002/bit.260250605>
109. Yamamoto, S, Nakanishi, K, Matsuno, R: *Ion-Exchange Chromatography of Proteins*, New York, Taylor & Francis Inc, 1988.
110. Lee, YF, Graalfs, H, Frech, C: Thermodynamic modeling of protein retention in mixed-mode chromatography: An extended model for isocratic and dual gradient elution chromatography. *Journal of Chromatography A*, 1464: 87-101, 2016. <https://doi.org/10.1016/j.chroma.2016.08.026>
111. Pedersen, L, Mollerup, J, Hansen, E, Jungbauer, A: Whey proteins as a model system for chromatographic separation of proteins. *Journal of Chromatography B*, 790: 161-173, 2003. [https://doi.org/10.1016/S1570-0232\(03\)00127-2](https://doi.org/10.1016/S1570-0232(03)00127-2)
112. Seidel-Morgenstern, A, Schmidt-Traub, H, Michel, M, Epping, A, Jupke, A: Modeling and Model Parameters. In: *Preparative Chromatography*. Herausgegeben von SCHMIDT-TRAUB, H., SCHULTE, M., SEIDEL-MORGENSTERN, A., Weinheim, 2012, S. 321-424.
113. Glueckauf, E, Coates, JI: Theory of chromatography. Part IV. The influence of incomplete equilibrium on the front boundary of chromatograms and on the effectiveness of separation. *Journal of the Chemical Society (Resumed)*: 1315-1321, 1947. <https://doi.org/10.1039/JR9470001315>
114. Ingham, J, Dunn, I, J., Heinzle, E, Přenosil, J, E., Snape, JB: Differential Flow and Reaction Applications. In: *Chemical Engineering Dynamics: An Introduction to Modelling and Computer Simulation*. Weinheim, Wiley VCH, 2008, S. 173-222.
115. Frederiksen, SS: *Computer aided development and optimisation of chromatographic separations*. Med. Dissertation. Medizinische Fakultät Mannheim, Universität Heidelberg, 2004.
116. Arkell, K, Breil, MP, Frederiksen, SS, Nilsson, B: Mechanistic Modeling of Reversed-Phase Chromatography of Insulins with Potassium Chloride and

- Ethanol as Mobile-Phase Modulators. *ACS Omega*, 2: 136-146, 2017. <https://doi.org/10.1021/acsomega.6b00248>
117. Arkell, K, Breil, MP, Frederiksen, SS, Nilsson, B: Mechanistic Modeling of Reversed-Phase Chromatography of Insulins within the Temperature Range 10–40 °C. *ACS Omega*, 3: 1946-1954, 2018. <https://doi.org/10.1021/acsomega.7b01527>
118. Arkell, K, Knutson, H-K, Frederiksen, SS, Breil, MP, Nilsson, B: Pareto-optimal reversed-phase chromatography separation of three insulin variants with a solubility constraint. *Journal of Chromatography A*, 1532: 98-104, 2018. <https://doi.org/10.1016/j.chroma.2017.11.065>
119. Arkell, K: *Modeling and Optimization of Reversed-Phase Chromatography - Effects of Modulators and Temperature*. Med. Dissertation. Medizinische Fakultät Mannheim, Universität Heidelberg, 2017.
120. Mollerup, JM, Breil, MP, Vogelpohl, C, Sadowski, G: Simultaneous correlation of hydrophobic interactions in HIC and protein solubility in aqueous salt solutions and mixed solvents. *Fluid Phase Equilibria*, 301: 163-170, 2011. <https://doi.org/10.1016/j.fluid.2010.11.028>
121. Wilson, GM: Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *Journal of the American Chemical Society*, 86: 127-130, 1964. <https://doi.org/10.1021/ja01056a002>
122. Schweiger, S, Jungbauer, A: Scalability of pre-packed preparative chromatography columns with different diameters and lengths taking into account extra column effects. *Journal of Chromatography A*, 1537: 66-74, 2018. <https://doi.org/10.1016/j.chroma.2018.01.022>
123. Seidel-Morgenstern, A, Schulte, M, Epping, A: Fundamentals and General Terminology. In: *Preparative Chromatography*. Herausgegeben von SCHMIDT-TRAUB, H., SCHULTE, M., SEIDEL-MORGENSTERN, A., Weinheim, 2012, S. 7-46.
124. Luo, H, Cheng, Y-K: A comparative study of void volume markers in immobilized-artificial-membrane and reversed-phase liquid chromatography. *Journal of Chromatography A*, 1103: 356-361, 2006. <https://doi.org/10.1016/j.chroma.2005.12.036>
125. Gétaz, D, Ströhlein, G, Morbidelli, M: Peptide pore accessibility in reversed-phase chromatography. *Journal of Chromatography A*, 1216: 933-940, 2009. <https://doi.org/10.1016/j.chroma.2008.12.019>
126. Gritti, F, Felinger, A, Guiochon, G: Overloaded gradient elution chromatography on heterogeneous adsorbents in reversed-phase liquid chromatography. *Journal of Chromatography A*, 1017: 45-61, 2003. [https://doi.org/10.1016/S0021-9673\(03\)01285-8](https://doi.org/10.1016/S0021-9673(03)01285-8)
127. Tugcu, N, Bae, SS, Moore, JA, Cramer, SM: Stationary phase effects on the dynamic affinity of low-molecular-mass displacers. *J Chromatogr A*, 954: 127-135, 2002. [https://doi.org/10.1016/s0021-9673\(02\)00164-4](https://doi.org/10.1016/s0021-9673(02)00164-4)
128. Osberghaus, A, Hepbildikler, S, Nath, S, Haindl, M, von Lieres, E, Hubbuch, J: Determination of parameters for the steric mass action model - a comparison between two approaches. *Journal of Chromatography A*, 1233: 54-65, 2012. <https://doi.org/10.1016/j.chroma.2012.02.004>
129. Jeansonne, MS, Foley, JP: Review of the Exponentially Modified Gaussian (EMG) Function Since 1983. *Journal of Chromatographic Science*, 29: 258-266, 1991. <https://doi.org/10.1093/chromsci/29.6.258>
130. Miyabe, K, Guiochon, G: Estimation of the column radial heterogeneity from an analysis of the characteristics of tailing peaks in linear chromatography. *Journal*

- of Chromatography A*, 830: 29-39, 1999. [https://doi.org/https://doi.org/10.1016/S0021-9673\(98\)00881-4](https://doi.org/https://doi.org/10.1016/S0021-9673(98)00881-4)
131. Koch, J, Scheps, D, Gunne, M, Boscheinen, O, Frech, C: Mechanistic modeling of cation exchange chromatography scale-up considering packing inhomogeneities. *Journal of separation science*, 46: 2300031, 2023. <https://doi.org/10.1002/jssc.202300031>
132. Koch, J, Scheps, D, Gunne, M, Boscheinen, O, Frech, C: DRAFT: Application of a mechanistic model describing an anti-Langmuirian to Langmuirian elution behavior for the separation of a polypeptide and product-related impurities. *currently under draft mode*, n/a.
133. Kaltenbrunner, O, Watler, P, Yamamoto, S: Column Qualification in Process Ion-Exchange Chromatography. In: *Progress in Biotechnology*. Herausgegeben von ENDO, I., NAGAMUNE, T., KATOH, S., YONEMOTO, T., Amsterdam, Elsevier, 2000, S. 201-206.
134. Guiochon, G, Felinger, A, Shirazi, DG, Katti, AM: The Mass Balance Equation of Chromatography and Its General Properties. In: *Fundamentals of Preparative and Nonlinear Chromatography*. 2 Aufl. San Diego, Elsevier Academic Press, 2006, S. 19-63.
135. Schweiger, S, Hinterberger, S, Jungbauer, A: Column-to-column packing variation of disposable pre-packed columns for protein chromatography. *Journal of Chromatography A*, 1527: 70-79, 2017. <https://doi.org/10.1016/j.chroma.2017.10.059>
136. Schweiger, S, Berger, E, Chan, A, Peyser, J, GebSKI, C, Jungbauer, A: Packing quality, protein binding capacity and separation efficiency of pre-packed columns ranging from 1 mL laboratory to 57 L industrial scale. *Journal of Chromatography A*, 1591: 79-86, 2019. <https://doi.org/10.1016/j.chroma.2019.01.014>
137. Al-Jibbouri, S: Scale-up of chromatographic ion-exchange processes in biotechnology. *Journal of Chromatography A*, 1116: 135-142, 2006. <https://doi.org/10.1016/j.chroma.2006.03.033>
138. Pfister, D, Morbidelli, M: Mass transfer coefficients determination from linear gradient elution experiments. *Journal of Chromatography A*, 1375: 42-48, 2015. <https://doi.org/10.1016/j.chroma.2014.11.068>
139. Kaltenbrunner, O, Jungbauer, A, Yamamoto, S: Prediction of the preparative chromatography performance with a very small column. *Journal of Chromatography A*, 760: 41-53, 1997. [https://doi.org/10.1016/S0021-9673\(96\)00689-9](https://doi.org/10.1016/S0021-9673(96)00689-9)
140. Carta, G, Jungbauer, A: Laboratory and Process Columns and Equipment. In: *Protein Chromatography: Process Development and Scale-Up*. Weinheim, Wiley-VCH, 2020, S. 139-158.
141. Cabooter, D, Lynen, F, Sandra, P, Desmet, G: Total pore blocking as an alternative method for the on-column determination of the external porosity of packed and monolithic reversed-phase columns. *Journal of Chromatography A*, 1157: 131-141, 2007. <https://doi.org/10.1016/j.chroma.2007.04.053>
142. Liu, X, Kaczmarski, K, Cavazzini, A, Szabelski, P, Zhou, D, Guiochon, G: Modeling of Preparative Reversed-Phase HPLC of Insulin. *Biotechnology Progress*, 18: 796-806, 2002. <https://doi.org/10.1021/bp020003a>
143. Liu, X, Szabelski, P, Kaczmarski, K, Zhou, D, Guiochon, G: Influence of pressure on the chromatographic behavior of insulin variants under nonlinear conditions. *Journal of Chromatography A*, 988: 205-218, 2003. [https://doi.org/10.1016/s0021-9673\(03\)00002-5](https://doi.org/10.1016/s0021-9673(03)00002-5)

144. Shekhawat, LK, Rathore, AS: Mechanistic modeling based process analytical technology implementation for pooling in hydrophobic interaction chromatography. 35: e2758, 2019. <https://doi.org/10.1002/btpr.2758>
145. Jakobsson, N, Karlsson, D, Axelsson, JP, Zacchi, G, Nilsson, B: Using computer simulation to assist in the robustness analysis of an ion-exchange chromatography step. *Journal of Chromatography A*, 1063: 99-109, 2005. <https://doi.org/10.1016/j.chroma.2004.11.067>
146. Seelinger, F, Wittkopp, F, von Hirschheydt, T, Hafner, M, Frech, C: Application of the Steric Mass Action formalism for modeling under high loading conditions: Part 1. Investigation of the influence of pH on the steric shielding factor. *Journal of Chromatography A*, 1676: 463265, 2022. <https://doi.org/10.1016/j.chroma.2022.463265>
147. Mollerup, JM: Chemical Engineering Thermodynamics and Protein Adsorption Chromatography. *Journal of Chemical & Engineering Data*, 59: 991-998, 2013. <https://doi.org/10.1021/je400739j>
148. Mollerup, JM, Hansen, TB, Kidal, S, Sejergaard, L, Hansen, E, Staby, A: Use of Quality by the Design for the Modelling of Chromatographic Separations. *Journal of Liquid Chromatography & Related Technologies*, 32: 1577-1597, 2009. <https://doi.org/10.1080/10826070902956345>
149. Mollerup, JM: Modelling oligomer formation in chromatographic separations. *Journal of Chromatography A*, 1218: 8869-8873, 2011. <https://doi.org/10.1016/j.chroma.2011.05.097>
150. Yamamoto, S, Nakanishi, K, Matsuno, R, Kamijubo, T: Ion exchange chromatography of proteins - predictions of elution curves and operating conditions. II. Experimental verification. *Biotechnology and bioengineering*, 25: 1373-1391, 1983. <https://doi.org/10.1002/bit.260250516>
151. Ishihara, T, Yamamoto, S: Optimization of monoclonal antibody purification by ion-exchange chromatography: Application of simple methods with linear gradient elution experimental data. *Journal of Chromatography A*, 1069: 99-106, 2005. <https://doi.org/10.1016/j.chroma.2004.10.040>
152. Li, Y, Pinto, NG: Model for ion-exchange equilibria of macromolecules in preparative chromatography. *Journal of Chromatography A*, 702: 113-123, 1995. [https://doi.org/10.1016/0021-9673\(94\)01281-1](https://doi.org/10.1016/0021-9673(94)01281-1)
153. Westerberg, K, Broberg Hansen, E, Degerman, M, Budde Hansen, T, Nilsson, B: Model-Based Process Challenge of an Industrial Ion-Exchange Chromatography Step. *Chemical Engineering and Technology*, 35: 183-190, 2012. <https://doi.org/10.1002/ceat.201000560>
154. Nfor, BK, Noverraz, M, Chilamkurthi, S, Verhaert, PDEM, van der Wielen, LAM, Ottens, M: High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *Journal of Chromatography A*, 1217: 6829-6850, 2010. <https://doi.org/10.1016/j.chroma.2010.07.069>
155. Sing, KSW: Reporting physisorption data for gas/solid systems with special reference to the determination of surface area and porosity. *Pure and Applied Chemistry*, 54: 2001-2018, 1982. <https://doi.org/10.1351/pac198254112201>
156. Golshan-Shirazi, S, Guiochon, G: Theoretical explanation of the displacement and tag-along effects. *Chromatographia*, 30: 613-617, 1990. <https://doi.org/10.1007/BF02269733>
157. Guiochon, G, Felinger, A, Shirazi, DG, Katti, AM: Band Profiles of Two Components with the Ideal Model. In: *Fundamentals of Preparative and*

- Nonlinear Chromatography*. 2 Aufl. San Diego, Elsevier Academic Press, 2006, S. 19-63.
158. Rüdts, M, Gillet, F, Heege, S, Hitzler, J, Kalbfuss, B, Guélat, B: Combined Yamamoto approach for simultaneous estimation of adsorption isotherm and kinetic parameters in ion-exchange chromatography. *Journal of Chromatography A*, 1413: 68-76, 2015. <https://doi.org/10.1016/j.chroma.2015.08.025>
159. Knox, JH, Parcher, JF: Effect of the column to particle diameter ratio on the dispersion of unadsorbed solutes in chromatography. *Analytical Chemistry*, 41: 1599-1606, 1969. <https://doi.org/10.1021/ac60281a009>
160. Guiochon, G, Farkas, T, Guan-Sajonz, H, Koh, J-H, Sarker, M, Stanley, BJ, Yun, T: Consolidation of particle beds and packing of chromatographic columns. *Journal of Chromatography A*, 762: 83-88, 1997. [https://doi.org/10.1016/S0021-9673\(96\)00642-5](https://doi.org/10.1016/S0021-9673(96)00642-5)
161. Knox, JH, Laird, GR, Raven, PA: Interaction of radial and axial dispersion in liquid chromatography in relation to the "infinite diameter effect". *Journal of Chromatography A*, 122: 129-145, 1976. [https://doi.org/10.1016/S0021-9673\(00\)82240-2](https://doi.org/10.1016/S0021-9673(00)82240-2)
162. Eon, CH: Comparison of broadening patterns in regular and radially compressed large-diameter columns. *Journal of Chromatography A*, 149: 29-42, 1978. [https://doi.org/10.1016/S0021-9673\(00\)80977-2](https://doi.org/10.1016/S0021-9673(00)80977-2)
163. Farkas, T, Sepaniak, MJ, Guiochon, G: Radial distribution of the flow velocity, efficiency and concentration in a wide HPLC column. *AIChE Journal*, 43: 1964-1974, 1997. <https://doi.org/10.1002/aic.690430806>
164. Viet, D, Beck-Candanedo, S, Gray, DG: Synthesis and characterization of blue dextrans. *Carbohydrate Polymers*, 74: 372-378, 2008. <https://doi.org/10.1016/j.carbpol.2008.03.003>
165. Gritti, F, Guiochon, G: Peak shapes of acids and bases under overloaded conditions in reversed-phase liquid chromatography, with weakly buffered mobile phases of various pH: A thermodynamic interpretation. *Journal of Chromatography A*, 1216: 63-78, 2009. <https://doi.org/10.1016/j.chroma.2008.11.020>
166. Rischawy, F, Saleh, D, Hahn, T, Oelmeier, S, Spitz, J, Kluters, S: Good modeling practice for industrial chromatography: Mechanistic modeling of ion exchange chromatography of a bispecific antibody. *Computers & Chemical Engineering*, 130: 106532, 2019. <https://doi.org/10.1016/j.compchemeng.2019.106532>
167. Saleh, D, Wang, G, Muller, B, Rischawy, F, Kluters, S, Studts, J, Hubbuch, J: Straightforward method for calibration of mechanistic cation exchange chromatography models for industrial applications. *Biotechnology Progress*: 2984, 2020. <https://doi.org/10.1002/btpr.2984>

8 APPENDIX

8.1 Calculation of the column model

The column model presented in Equations (24) and (25), consisting of partial differential equations (PDEs), was solved using the finite difference axial dispersion model, as shown by Ingham et al. ¹¹⁴. Therefore, the partial differential equation system is transformed into a set of difference-differential equations by dividing the volume of the column into several equally shaped discs. Figure 40 illustrates the principle of the finite difference model considering axial dispersion. The convective flow with the flow rate L_{in} and the diffusive flux in axial direction j_{ax} enter and leave the liquid phase of any section n . The concentration of species i in the liquid phase is denoted as c_i and the concentration in the solid phase is denoted as q_i^* . The rate of loss of species i due to transfer from the mobile to the solid phase is given by r_i . The concentrations of species i at the boundaries of each segment are averaged. Note that the solid phase available for binding is based on pore volume in the used model.

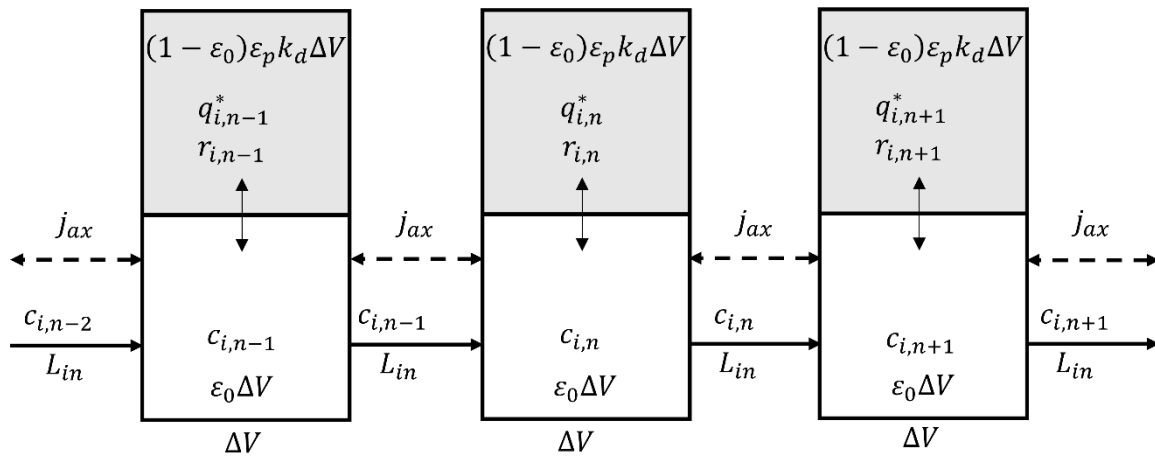


Figure 40: Finite difference axial dispersion model of a porous chromatographic column. The grey boxes indicate the solid phase based on pore volume, and the white boxes the liquid phase (adapted from Ingham et al. ¹¹⁴).

The partial differential equation column model presented in Equations (24) and (25) was transformed into a dynamic difference equation form, so the model is suitable for a solution in the ODE solver Berkeley Madonna, as shown by Ingham et al. ¹¹⁴.

The material balance in the liquid phase for each component i is given as ¹¹⁴:

$$\begin{aligned} \left(\begin{array}{c} \text{Accumulation} \\ \text{rate of } i \end{array} \right) &= \left(\begin{array}{c} \text{Convective} \\ \text{flow of } i \text{ in} \end{array} \right) - \left(\begin{array}{c} \text{Convective} \\ \text{flow of } i \text{ out} \end{array} \right) + \left(\begin{array}{c} \text{Diffusive} \\ \text{flow of } i \text{ in} \end{array} \right) \\ &- \left(\begin{array}{c} \text{Diffusive} \\ \text{flow of } i \text{ out} \end{array} \right) - \left(\begin{array}{c} \text{Rate of} \\ \text{transfer to} \\ \text{solid phase of } i \end{array} \right). \end{aligned} \quad (\text{A.1})$$

Consequently, the concentrations in the liquid phase of i accessing and leaving a section n are given by ¹¹⁴:

$$c_{i,in} = \frac{c_{i,n-1} + c_{i,n}}{2}, \quad (\text{A.2})$$

$$c_{i,out} = \frac{c_{i,n} + c_{i,n+1}}{2}, \quad (\text{A.3})$$

and the concentration gradients of i at the inlet and outlet for each n are given by ¹¹⁴:

$$\left(\frac{dc_i}{dx} \right)_{in} = \frac{c_{i,n-1} - c_{i,n}}{\Delta x}, \quad (\text{A.4})$$

$$\left(\frac{dc_i}{dx} \right)_{out} = \frac{c_{i,n} - c_{i,n+1}}{\Delta x}. \quad (\text{A.5})$$

The transfer rate of species i from the liquid to the solid phase is given by:

$$r_i(1 - \varepsilon_0)\varepsilon_p k_d A_c \Delta x = k_{eff} \frac{6}{d_p} (q_{eq,i}^* - q_i^*)(1 - \varepsilon_0)\varepsilon_p k_d A_c \Delta x. \quad (\text{A.6})$$

By multiplying the concentrations of i in Equations (A.2) and (A.3) with the volumetric flow rate L_{in} and by multiplying the concentration gradients of i in Equations (A.4) and (A.5) by the interstitial cross-sectional area $\varepsilon_0 A_c$, the material balance for i in the liquid phase in any section n is given by ¹¹⁴:

$$\begin{aligned}
\varepsilon_0 A_c \Delta x \frac{dc_{i,n}}{dt} &= L_{in} \left(\frac{c_{i,n-1} + c_{i,n}}{2} - \frac{c_{i,n} + c_{i,n+1}}{2} \right) \\
&+ \varepsilon_0 A_c D_{ax} \left(\frac{c_{i,n-1} - c_{i,n}}{\Delta x} - \frac{c_{i,n} - c_{i,n+1}}{\Delta x} \right) \\
&- (1 - \varepsilon_0) \varepsilon_p k_d k_{eff} \frac{6}{d_p} (q_{eq,i}^* - q_i^*) A_c \Delta x.
\end{aligned} \tag{A.7}$$

Rearranging Equation (A.7) results in the final dynamic difference equation form for any section n , which is suitable for the calculation in the software Berkely Madonna ¹¹⁴:

$$\begin{aligned}
\frac{dc_{i,n}}{dt} &= \frac{L_{in}}{\varepsilon_0 A_c \Delta x} \left(\frac{c_{i,n-1} - c_{i,n+1}}{2} \right) \\
&+ D_{ax} \left(\frac{c_{i,n-1} - 2c_{i,n} + c_{i,n+1}}{\Delta x^2} \right) - \frac{(1 - \varepsilon_0) \varepsilon_p k_d k_{eff}}{\varepsilon_0} \frac{6}{d_p} (q_{eq,i}^* - q_i^*).
\end{aligned} \tag{A.8}$$

At the column inlet and outlet, the model requires special considerations (boundary conditions) since an initial solute concentration of i must enter the column inlet, and no diffusive flow can enter and exit the closed inlet and outlet of the column. Consequently, the material balance for i in the liquid phase in the first element of the column is ¹¹⁴:

$$\begin{aligned}
\varepsilon_0 A_c \Delta x \frac{dc_{i,1}}{dt} &= L_{in} \left(c_{i,in} - \frac{c_{i,1} + c_{i,2}}{2} \right) - \varepsilon_0 A_c D_{ax} \left(\frac{c_{i,1} - c_{i,2}}{\Delta x} \right) \\
&- (1 - \varepsilon_0) \varepsilon_p k_d k_{eff} \frac{6}{d_p} (q_{eq,i}^* - q_i^*) A_c \Delta x.
\end{aligned} \tag{A.9}$$

Rearranging Equation (A.9) results in:

$$\begin{aligned}
\frac{dc_{i,1}}{dt} &= \frac{L_{in}}{\varepsilon_0 A_c \Delta x} \left(\frac{2c_{i,in} - c_{i,1} - c_{i,2}}{2} \right) - D_{ax} \left(\frac{c_{i,1} - c_{i,2}}{\Delta x^2} \right) \\
&- \frac{(1 - \varepsilon_0) \varepsilon_p k_d k_{eff}}{\varepsilon_0} \frac{6}{d_p} (q_{eq,i}^* - q_i^*).
\end{aligned} \tag{A.10}$$

Accordingly, the material balance for i in the last element of the column N is given by ¹¹⁴:

$$\begin{aligned} \varepsilon_0 A_c \Delta x \frac{dc_{i,N}}{dt} = & L_{in}(c_{i,N-1} - c_{i,N}) + \varepsilon_0 A_c D_{ax} \left(\frac{c_{i,N-1} - c_{i,N}}{\Delta x} \right) \\ & - (1 - \varepsilon_0) \varepsilon_p k_d k_{eff} \frac{6}{d_p} (q_{eq,i}^* - q_i^*) A_c \Delta x. \end{aligned} \quad (\text{A.11})$$

Rearranging Equation (A.11) results in:

$$\begin{aligned} \frac{dc_{i,N}}{dt} = & \frac{L_{in}}{\varepsilon_0 A_c \Delta x} (c_{i,N-1} - c_{i,N}) + D_{ax} \left(\frac{c_{i,N-1} - c_{i,N}}{\Delta x^2} \right) \\ & - \frac{(1 - \varepsilon_0) \varepsilon_p k_d}{\varepsilon_0} k_{eff} \frac{6}{d_p} (q_{eq,i}^* - q_i^*). \end{aligned} \quad (\text{A.12})$$

With Equations (A.8), (A.10), and (A.12), it is possible to calculate lumped rate column model (Equations (24) and (25)) consisting of several PDEs with the ODE solver software Berkeley Madonna. Note that the solid phase protein concentration is calculated from the isotherm model. How the isotherm model is calculated in the Berkeley Madonna file is shown in Appendix 8.2.

8.2 Calculation of the SAS isotherm

The SAS isotherm derived in Section 1.4.1 was calculated according to the SMA formalism, as shown by Frederiksen ¹¹⁵. According to the scheme, Equation (15) has to be rearranged into the following form:

$$\frac{q_i}{c_i} = K_i \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} c_s^{-v_i} \left[\Lambda - \sum_{j=1}^{\alpha} (v_j + \xi_j) q_j \right]^{v_i} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{\gamma}_i \right); i, j = 1, 2 \dots \alpha. \quad (\text{A.13})$$

With the assumption $q_i = q_{1,i} + 2q_{2,i}$ (Equation (11)) follows:

$$\frac{q_i}{c_i} = K_i \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} c_s^{-v_i} \left[\Lambda - \sum_{j=1}^{\alpha} [(v_j + \xi_j) q_{1,j} + 2q_{2,ij}(v_j + \xi_j)] \right]^{v_i} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{\gamma}_i \right). \quad (\text{A.14})$$

With the following defined relationship for the ionic capacity:

$$\Lambda = q_s + \sum_{j=1}^{\alpha} [(v_j + \xi_j) q_{1,j} + 2q_{2,ij}(v_j + \xi_j)], \quad (\text{A.15})$$

and the relationship for $q_{1,i}$:

$$q_{1,i} = K_i \left(\frac{q_s}{c_s} \right)^{v_i} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} c_i, \quad (\text{A.16})$$

and for $2q_{2,i}$:

$$q_{2,i} = q_{1,i} \frac{K_{D,i}}{c} \tilde{\gamma}_i c_i = K_i \left(\frac{q_s}{c_s} \right)^{v_i} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} \frac{K_{D,i}}{c} \tilde{\gamma}_i c_i^2, \quad (\text{A.17})$$

the SAS isotherm can be extended with the ionic capacity as described with the following relationship:

$$\frac{q_i}{c_i} = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} c_s^{-v_i} [\Lambda - \sum_{j=1}^{\alpha} [(v_j + \xi_j) q_{1,j} + 2q_{2,j}(v_j + \xi_j)]]^{v_i} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{Y}_i\right)}{q_s + \sum_{j=1}^{\alpha} [(v_j + \xi_j) q_{1,j} + 2q_{2,j}(v_j + \xi_j)]} \quad (\text{A.18})$$

$$\frac{q_i}{c_i} = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} c_s^{-v_i} [\Lambda - \sum_{j=1}^{\alpha} [(v_j + \xi_j) q_{1,j} + 2q_{2,j}(v_j + \xi_j)]]^{v_j} \left(1 + 2 \frac{K_{D,j}}{c} c_j \tilde{Y}_j\right)}{q_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{q_s}{c_s}\right)^{v_j} \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} c_j (v_j + \xi_j) + 2K_j \left(\frac{q_s}{c_s}\right)^{v_j} \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} \frac{K_{D,j}}{c} \tilde{Y}_j c_j^2 (v_j + \xi_j) \right]} \quad (\text{A.19})$$

By inserting the adsorbed concentration of the counterion q_s available for binding, the relationship simplifies to:

$$\frac{q_i}{c_i} = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} \left(\frac{q_s}{c_s}\right)^{v_i} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{Y}_i\right)}{q_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{q_s}{c_s}\right)^{v_j} \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} c_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} c_j \tilde{Y}_j\right) \right]} \quad (\text{A.20})$$

By extending with c_s/q_s , the SAS isotherm can be written as:

$$\frac{q_i}{c_i} = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} \left(\frac{q_s}{c_s}\right)^{v_i} \left(\frac{c_s}{q_s}\right) \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{Y}_i\right)}{q_s \left(\frac{c_s}{q_s}\right) + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{q_s}{c_s}\right)^{v_j} \left(\frac{c_s}{q_s}\right) \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} c_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} c_j \tilde{Y}_j\right) \right]} \quad (\text{A.21})$$

$$\frac{q_i}{c_i} = \frac{\Lambda K_i \frac{\tilde{Y}_i}{\tilde{Y}_s^{v_i}} \left(\frac{q_s}{c_s}\right)^{v_i-1} \left(1 + 2 \frac{K_{D,i}}{c} c_i \tilde{Y}_i\right)}{c_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{q_s}{c_s}\right)^{v_j-1} \frac{\tilde{Y}_j}{\tilde{Y}_s^{v_j}} c_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} c_j \tilde{Y}_j\right) \right]} \quad (\text{A.22})$$

For the calculation of the isotherm with the ODE solver Berkely Madonna, Equation (A.22) is rearranged with the dimensionless variables for the liquid and solid phase protein and counterion concentration:

$$\ddot{y}_i = \frac{c_i z_s}{\Lambda} = \frac{c_i}{\Lambda} \rightarrow c_i = \ddot{y}_i \Lambda, \quad (\text{A.23})$$

$$\ddot{y}_s = \frac{c_s z_s}{\Lambda} = \frac{c_s}{\Lambda} \rightarrow c_s = \ddot{y}_s \Lambda, \quad (\text{A.24})$$

$$Q_i = \frac{q_i z_s}{\Lambda} = \frac{q_i}{\Lambda} \rightarrow q_i = Q_i \Lambda. \quad (\text{A.25})$$

$$Q_s = \frac{q_s z_s}{\Lambda} = \frac{q_s}{\Lambda} \rightarrow q_s = Q_s \Lambda. \quad (\text{A.26})$$

Introduction into Equation (A.22) results in the following relationship:

$$Q_i \Lambda = \frac{\Lambda K_i \frac{\tilde{y}_i}{\tilde{y}_s^{v_i}} \left(\frac{Q_s \Lambda}{\ddot{y}_s \Lambda} \right)^{v_i-1} \ddot{y}_i \Lambda \left(1 + 2 \frac{K_{D,i}}{c} \tilde{y}_i \ddot{y}_i \Lambda \right)}{\ddot{y}_s \Lambda + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{Q_s \Lambda}{\ddot{y}_s \Lambda} \right)^{v_j-1} \frac{\tilde{y}_j}{\tilde{y}_s^{v_j}} \ddot{y}_j \Lambda (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} \tilde{y}_j \ddot{y}_j \Lambda \right) \right]}. \quad (\text{A.27})$$

Rearranging results in:

$$Q_i = \frac{K_i \frac{\tilde{y}_i}{\tilde{y}_s^{v_i}} \left(\frac{Q_s}{\ddot{y}_s} \right)^{v_i-1} \ddot{y}_i \left(1 + 2 \frac{K_{D,i}}{c} \tilde{y}_i \ddot{y}_i \Lambda \right)}{\ddot{y}_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{Q_s}{\ddot{y}_s} \right)^{v_j-1} \frac{\tilde{y}_j}{\tilde{y}_s^{v_j}} \ddot{y}_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} \tilde{y}_j \ddot{y}_j \Lambda \right) \right]}. \quad (\text{A.28})$$

With the equilibrium concentration of the adsorbed protein $q_{eq,i}^* = q_i = Q_i \Lambda$, the relationship, which is implemented in the ODE solver Berkely Madonna, can be described with:

$$q_{eq,i}^* = \frac{\Lambda K_i \frac{\tilde{y}_i}{\tilde{y}_s^{v_i}} \left(\frac{Q_s}{\ddot{y}_s} \right)^{v_i-1} \ddot{y}_i \left(1 + 2 \frac{K_{D,i}}{c} \tilde{y}_i \ddot{y}_i \Lambda \right)}{\ddot{y}_s + \sum_{j=1}^{\alpha} \left[K_j \left(\frac{Q_s}{\ddot{y}_s} \right)^{v_j-1} \frac{\tilde{y}_j}{\tilde{y}_s^{v_j}} \ddot{y}_j (v_j + \xi_j) \left(1 + 2 \frac{K_{D,j}}{c} \tilde{y}_j \ddot{y}_j \Lambda \right) \right]}. \quad (\text{A.29})$$

As Frederiksen stated in his work, the simplest way to calculate the SMA formalism is to determine the free salt Q_s bonded to the ligands available for binding first since it is given as one implicit equation in one variable ¹¹⁵. Q_s can be calculated from ¹¹⁵:

$$Q_s = \frac{\dot{y}_s}{G_{(y)}}. \quad (\text{A.30})$$

Whereas $G_{(y)}$ is given by:

$$G_{(y)} = \dot{y}_s + \sum_{i=1}^{\alpha} \left[K_i \left(\frac{Q_s}{\dot{y}_s} \right)^{v_i-1} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} \dot{y}_i (v_i + \xi_i) \left(1 + 2 \frac{K_{D,i}}{c} \tilde{\gamma}_i \dot{y}_i \Lambda \right) \right], \quad (\text{A.31})$$

$$G_{(y)} = \dot{y}_s + \sum_{i=1}^{\alpha} (L_i \dot{y}_i), \quad (\text{A.32})$$

with:

$$L_i = K_i \left(\frac{Q_s}{\dot{y}_s} \right)^{v_i-1} \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} (v_i + \xi_i) \left(1 + 2 \frac{K_{D,i}}{c} \tilde{\gamma}_i \dot{y}_i \Lambda \right). \quad (\text{A.33})$$

By applying a Newton iteration and defining the functions:

$$F_1 = -\dot{y}_s + Q_s G_{(y)} = 0, \quad (\text{A.34})$$

$$F_2 = \frac{dF_1}{dQ_s} = \dot{y}_s + \dot{y}_s + \sum_{i=1}^{\alpha} (L_i v_i \dot{y}_i). \quad (\text{A.35})$$

the dimensionless variable Q_s can be calculated. In this work, a Newton iteration with 20 steps was coded in the Berkely Madonna simulation file. Consequently, the equilibrium concentration bound to the ligands can be calculated from the following:

$$q_{eq,i}^* = \frac{\Lambda K_i \frac{\tilde{\gamma}_i}{\tilde{\gamma}_s^{v_i}} \left(\frac{Q_s}{\dot{y}_s} \right)^{v_i-1} \dot{y}_i \left(1 + 2 \frac{K_{D,i}}{c} \tilde{\gamma}_i \dot{y}_i \Lambda \right)}{G_{(y)}}. \quad (\text{A.36})$$

Note that the isotherm model for the RPC step derived in section 1.4.2 was calculated according to the same scheme shown above since the model applied for the RPC step is also based on the SMA formalism.

8.3 CEX column 1 modeling and simulation

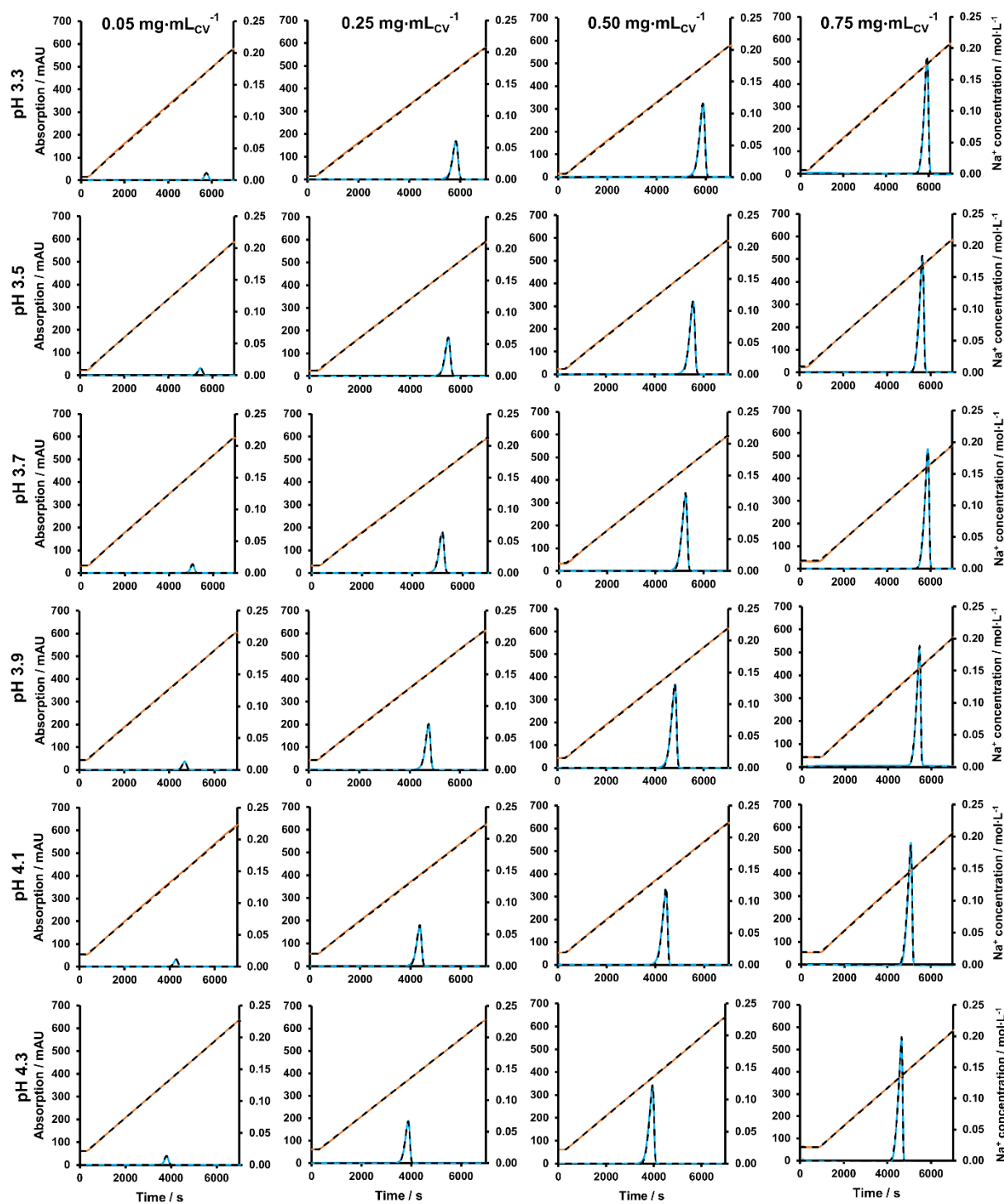


Figure 41: Simulated chromatograms compared to the experimentally determined chromatograms with the polypeptide species at different pH values and column loads. The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. The experimental UV signal was determined with a 10 mm UV cell.

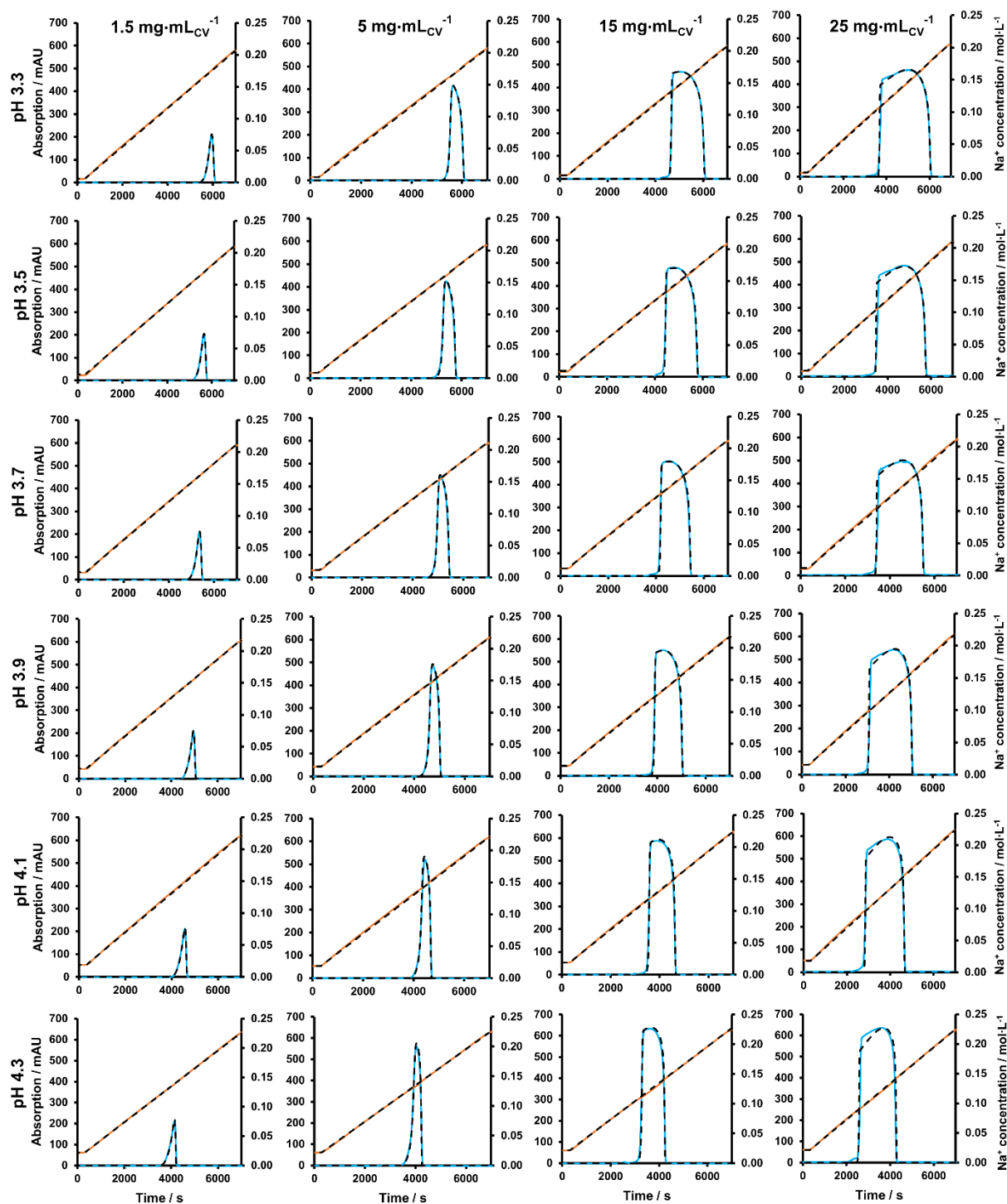


Figure 42: Simulated chromatograms compared to the experimentally determined chromatograms with the polypeptide species at different pH values and column loads. The orange and blue lines indicate the experimental sodium concentration and UV signal. The black dashed lines indicate the corresponding simulated signals. The experimental UV signal was determined with a 2 mm UV cell.

8.4 CEX column scale-up modeling

Table 20: Fitted empiric loading dependency for the effective mass transfer coefficient of the polypeptide species for each CEX column scale.

	CEX column 1	CEX column 2	CEX column 3	CEX column 4
Load (mg·mL _{CV} ⁻¹)	$k_{eff,i}$ (cm ² ·s ⁻¹)	$k_{eff,i}$ (cm ² ·s ⁻¹)	$k_{eff,i}$ (cm ² ·s ⁻¹)	$k_{eff,i}$ (cm ² ·s ⁻¹)
0.05	3.5×10^{-5}	5.0×10^{-5}	3.8×10^{-5}	3.2×10^{-5}
1.5	4.8×10^{-5}	6.5×10^{-5}	6.0×10^{-5}	7.0×10^{-5}
5	4.4×10^{-5}	5.8×10^{-5}	4.2×10^{-5}	6.0×10^{-5}
10	3.5×10^{-5}	4.5×10^{-5}	3.7×10^{-5}	4.0×10^{-5}

8.5 CEX model and RP model coupled simulations sampling study

Table 21: Overview of the parameters for the sampling study with the CEX column 1. The pooling criteria for the theoretical pool was > 92 % polypeptide purity.

Run	Load (mL)	Conc. Poly. (mg/mL)	Variant A conc. (mg/mL)	Variant B conc. (mg/mL)	Load Poly. (mg/mL _{cv})	Load variant A (mg/mL _{cv})	Load variant B (mg/mL _{cv})	Purity Poly. Load (%)	Pool volume (mL)	Pool conc. Poly. (mg/mL)	Yield Poly. (mg)	Recovery Poly. (%)
1	2.75	3.8	0.254	0.254	10.450	0.699	0.699	88.21	4.825	1.778	8.580	82.11
2	2.61	3.8	0.254	0.254	9.918	0.663	0.663	88.21	4.628	1.788	8.276	83.44
3	2.48	3.8	0.254	0.254	9.424	0.630	0.630	88.21	4.460	1.773	7.909	83.92
4	2.34	3.8	0.254	0.254	8.892	0.594	0.594	88.21	4.268	1.762	7.520	84.57
5	2.2	3.8	0.254	0.254	8.360	0.559	0.559	88.21	4.089	1.728	7.068	84.54
6	2.06	3.8	0.254	0.254	7.828	0.523	0.523	88.21	3.910	1.712	6.696	85.54
7	2.89	3.8	0.254	0.254	10.982	0.734	0.734	88.21	5.016	1.815	9.105	82.90
8	3.03	3.8	0.254	0.254	11.514	0.770	0.770	88.21	5.217	1.827	9.532	82.79
9	3.16	3.8	0.254	0.254	12.008	0.803	0.803	88.21	5.406	1.837	9.931	82.70
10	3.3	3.8	0.254	0.254	12.540	0.838	0.838	88.21	5.612	1.847	10.367	82.67
11	2.75	3.3	0.254	0.254	9.075	0.699	0.699	86.66	4.232	1.753	7.418	81.74
12	2.75	2.9	0.254	0.254	7.975	0.699	0.699	85.09	3.753	1.706	6.402	80.27
13	2.75	2.5	0.254	0.254	6.875	0.699	0.699	83.11	3.274	1.639	5.367	78.06
14	2.75	2	0.254	0.254	5.500	0.699	0.699	79.74	2.652	1.547	4.103	74.60
15	2.75	4.2	0.254	0.254	11.550	0.699	0.699	89.21	5.300	1.836	9.732	84.26
16	2.75	4.6	0.254	0.254	12.650	0.699	0.699	90.05	5.784	1.863	10.774	85.17
17	2.75	5	0.254	0.254	13.750	0.699	0.699	90.78	6.270	1.856	11.636	84.62
18	2.75	5.4	0.254	0.254	14.850	0.699	0.699	91.40	6.775	1.900	12.871	86.68
19	2.75	5.8	0.254	0.254	15.950	0.699	0.699	91.95	7.286	1.886	13.743	86.16

APPENDIX

Table 22: Overview of the parameters for the sampling study with the RP column 1. The pooling criteria for the theoretical pool was > 96 % polypeptide purity.

Run	Load (mL)	Conc. Poly. (mg/mL)	Variant A conc. (mg/mL)	Variant B conc. (mg/mL)	Load Poly. (mg/mL _{cv})	Load variant A (mg/mL _{cv})	Load variant B (mg/mL _{cv})	Purity Poly. Load (%)	Pool volume (mL)	Pool conc. Poly. (mg/mL)	Yield Poly. (mg)	Recovery Poly. (%)
1	4.825	1.778	0.0189	0.0190	8.580	0.0914	0.092	97.91	0.927	4.773	4.423	51.54
2	4.628	1.788	0.0229	0.0179	8.276	0.1059	0.083	97.77	0.927	4.773	4.423	53.44
3	4.460	1.773	0.0240	0.0167	7.909	0.1072	0.074	97.76	0.881	4.518	3.979	50.31
4	4.268	1.762	0.0254	0.0154	7.520	0.1084	0.066	97.74	0.863	4.355	3.760	50.00
5	4.089	1.728	0.0242	0.0139	7.068	0.0988	0.057	97.84	0.861	4.151	3.574	50.57
6	3.910	1.712	0.0258	0.0126	6.696	0.1009	0.049	97.80	0.846	3.971	3.359	50.15
7	5.016	1.815	0.0213	0.0205	9.105	0.1066	0.103	97.75	0.942	4.951	4.663	51.21
8	5.217	1.827	0.0204	0.0218	9.532	0.1063	0.114	97.74	0.959	5.087	4.879	51.19
9	5.406	1.837	0.0195	0.0230	9.931	0.1055	0.124	97.74	0.985	5.421	5.342	53.79
10	5.612	1.847	0.0187	0.0242	10.367	0.1048	0.136	97.73	1.003	5.351	5.366	51.76
11	4.232	1.753	0.0256	0.0158	7.418	0.1083	0.067	97.69	0.855	4.329	3.699	49.87
12	3.753	1.706	0.0293	0.0130	6.402	0.1100	0.049	97.58	0.820	3.878	3.179	49.65
13	3.274	1.639	0.0333	0.0101	5.367	0.1090	0.033	97.42	0.802	3.323	2.666	49.68
14	2.652	1.547	0.0406	0.0064	4.103	0.1077	0.017	97.05	0.794	2.566	2.036	49.63
15	5.300	1.836	0.0204	0.0219	9.732	0.1080	0.116	97.75	0.977	5.135	5.015	51.53
16	5.784	1.863	0.0183	0.0244	10.774	0.1060	0.141	97.76	1.029	5.453	5.610	52.07
17	6.270	1.856	0.0141	0.0269	11.636	0.0884	0.169	97.84	1.090	5.680	6.191	53.20
18	6.775	1.900	0.0147	0.0295	12.871	0.0997	0.200	97.72	1.160	5.949	6.899	53.60
19	7.286	1.886	0.0118	0.0321	13.743	0.0857	0.234	97.73	1.230	6.099	7.499	54.57

9 CURRICULUM VITAE

PERSONAL DATA

Name: Jonas Martin Maria Koch
Date of birth: 17.04.1993
Place of birth: Achern, Ortenaukreis, Germany

SCHOOL CAREER

09/2010 – 08/2013 **University Entrance Qualification** | Biotechnologisches
Gymnasium Offenburg, Germany

- Final grade: 2.0 (good)

ACADEMIC CAREER

10/2020 – today **Doctoral Student** | Ruprecht-Karls-Universität Heidelberg,
Germany

- Research worker at Steinbeis Transferzentrum an der Hochschule Mannheim (Medical Faculty Mannheim – Heidelberg University)
- Focus on modeling and simulation of liquid chromatography

08/2017 – 07/2019 **Bioprocess Development M.Sc.** | Hochschule Mannheim –
University of Applied Sciences, Germany

- Thesis: Mechanistic modeling of CEX- and RP-Chromatography for peptide purification (Sanofi-Aventis Deutschland GmbH, Frankfurt)
- Final grade: 1.2 (with distinction)

03/2014 – 07/2017 **Biotechnologie B.Sc.** | Hochschule Mannheim – University
of Applied Sciences, Germany

- Thesis: Modellierung des Elutionsverhaltens monoklonaler Antikörper auf mixed-mode Kationenaustauscherprototypen (Steinbeis Transferzentrum an der Hochschule Mannheim)
- Final grade: 1.4

WORK EXPERIENCE

- 12/2022 – today **Senior Data Science Expert** | Sanofi-Aventis Deutschland GmbH, Frankfurt
- Data Science Department (MSAT)
- 10/2020 – 11/2022 **Research Worker** | Steinbeis Transferzentrum an der Hochschule Mannheim, Germany
- Doctoral student at Medical Faculty Mannheim – Heidelberg University, Germany
 - Focus on modeling and simulation of liquid chromatography
- 08/2019 – 09/2020 **Project Engineer** | Steinbeis Transferzentrum an der Hochschule Mannheim, Germany
- Working on protein purification projects with external industry partners
- 03/2015 – 08/2015 **Internship** | InstrAction GmbH, Mannheim, Germany
- Product development – Chromatography resins

PUBLICATIONS

- 03/2022 **Jonas Koch**, Daniel Scheps, Matthias Gunne, Oliver Boscheinen, Mathias Hafner and Christian Frech
- Mechanistic modeling and simulation of a complex low and high loading elution behavior of a polypeptide in cation exchange chromatography, Journal of Separation Science, <https://doi.org/10.1002/jssc.202200098>
- 02/2023 **Jonas Koch**, Daniel Scheps, Matthias Gunne, Oliver Boscheinen and Christian Frech
- Mechanistic modeling of cation exchange chromatography scale-up considering packing inhomogeneities, Journal of Separation Science, <https://doi.org/10.1002/jssc.202300031>

- 05/2023 **Jonas Koch**, Daniel Scheps, Matthias Gunne, Oliver Boscheinen and Christian Frech
Effect of salt modulators on the elution behavior of insulin and the separation performance of product-related impurities in reversed-phase chromatography, Journal of Chromatography B, <https://doi.org/10.1016/j.jchromb.2023.123735>
- n/a **Jonas Koch**, Daniel Scheps, Matthias Gunne, Oliver Boscheinen and Christian Frech
DRAFT: Application of a mechanistic model describing an anti-Langmuirian to Langmuirian elution behavior for the separation of a polypeptide and product-related impurities, Currently under draft mode

SCIENTIFIC TALKS

- 03/2022 Jonas Koch, Daniel Scheps, Matthias Gunne, Oliver Boscheinen, and Christian Frech
ACS Spring 2022, San Diego (March 20-24),
“Mechanistic modeling and simulation of an uncommon elution behavior of a recombinant polypeptide in preparative cation exchange chromatography”
- 05/2022 Jonas Koch, Daniel Scheps, Matthias Gunne, Oliver Boscheinen and Christian Frech
PREP Symposium 2022, Baltimore (May 15-18),
“Mechanistic modeling, simulation and scalability investigation of an uncommon elution behavior of a recombinant polypeptide in preparative cation exchange chromatography”

POSTER PRESENTATIONS

12/2021

Jonas Koch, Daniel Scheps, Matthias Gunne, Oliver Boscheinen, and Christian Frech

Global Sanofi R&D Week 2021, Frankfurt (December 15),

“Mechanistic modeling and simulation of a complex low- and high loading elution behavior of in cation exchange chromatography”

10 ACKNOWLEDGEMENT

First, I would like to sincerely thank Prof. Dr. rer. nat. Mathias Hafner for the supervision and support of my doctoral thesis.

Special thanks to Prof. Dr. Christian Frech from the Institute for Biochemistry at the University of Applied Sciences in Mannheim for the excellent education I was allowed to enjoy in the field of liquid chromatography and for the support of my doctoral thesis.

Further, I would like to thank Dr. Daniel Scheps, Dr. Matthias Gunne, and Dr. Oliver Boscheinen from Sanofi-Aventis Deutschland GmbH for organizing the resources for this doctoral project and for their support during the entire time.

My sincere appreciation goes to my colleagues from the Institute for Biochemistry at the University of Applied Sciences in Mannheim for the pleasant working atmosphere and the tremendous mutual support. Special thanks to Dr. Felix Seelinger for the intensive support at the beginning of my doctoral project.

In addition, I would like to thank my parents and my entire family for their immense support during my studies. Without you, this would not have been possible. Finally, I thank my girlfriend for her great support in good and hard times.