Journal of Chromatography A 1718 (2024) 464706

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Predicting multimodal chromatography of therapeutic antibodies using multiscale modeling

Rudger Hess [a,b], Jan Faessler [b], Doil Yun [b], Ahmed Mama [b], David Saleh [b], Jan-Hendrik Grosch [b], Gang Wang [b], Thomas Schwab [b], Jürgen Hubbuch [a,*]

[a] Karlsruhe Institute of Technology (KIT), Institute of Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe, Germany
[b] DSP Development, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany

ARTICLE INFO

Keywords:
Multispecific monoclonal antibody (mAb) formats
Mechanistic chromatography modeling
Mixed-mode antibody descriptors
Quantitative structure-activity/property relationship (QSAR/QSPR)
In silico / a priori process development

ABSTRACT

Multimodal chromatography has emerged as a powerful method for the purification of therapeutic antibodies. However, process development of this separation technique remains challenging because of an intricate and molecule-specific interaction towards multimodal ligands, leading to time-consuming and costly experimental optimization.

This study presents a multiscale modeling approach to predict the multimodal chromatographic behavior of therapeutic antibodies based on their sequence information. Linear gradient elution (LGE) experiments were performed on an anionic multimodal resin for 59 full-length antibodies, including five different antibody formats at pH 5.0, 6.0, and 7.0 that were used for parameter determination of a linear adsorption model at low loading density conditions. Quantitative structure-property relationship (QSPR) modeling was utilized to correlate the adsorption parameters with up to 1374 global and local physicochemical descriptors calculated from antibody homology models. The final QSPR models employed less than eight descriptors per model and demonstrated high training accuracy ($R^2 > 0.93$) and reasonable test set prediction accuracy ($Q^2 > 0.83$) for the adsorption parameters. Model evaluation revealed the significance of electrostatic interaction and hydrophobicity in determining the chromatographic behavior of antibodies, as well as the importance of the HFR3 region in antibody binding to the multimodal resin. Chromatographic simulations using the predicted adsorption parameters showed good agreement with the experimental data for the vast majority of antibodies not employed during the model training.

The results of this study demonstrate the potential of sequence-based prediction for determining chromatographic behavior in therapeutic antibody purification. This approach leads to more efficient and cost-effective process development, providing a valuable tool for the biopharmaceutical industry.

1. Introduction

Liquid column chromatography has established itself as a prominent technique for the purification of biotherapeutics and is virtually indispensable for the production of monoclonal antibodies (mAb) [1]. Current research targets to improve this separation method by combining multiple forms of physicochemical interactions into a single chromatographic ligand, referred to as multimodal chromatography [2]. The improved ligand structure enables increased selectivity and salt tolerant adsorption compared to unimodal interaction resins [3].

The increased selectivity of multimodal chromatography originates from orthogonal interactions towards the molecule surface, such as complex electrostatic and hydrophobic effects and leads to increased sensitivity towards molecule characteristics and process conditions [4]. However, this increased sensitivity often necessitates a molecule-specific process development, which involves extensive experimentation for process optimization and characterization to ensure adequate product quality and process yield [5]. To expedite this time and resource-intensive process development, high-throughput experimentation and *in silico* tools have been deployed. Mechanistic and statistical modeling serve as valuable tools for the acceleration, optimization, and characterization of chromatography process development by either simulating the mass-transport and accumulation within the chromatographic column or by capturing convoluted

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

https://doi.org/10.1016/j.chroma.2024.464706
Received 4 November 2023; Received in revised form 30 January 2024; Accepted 31 January 2024
Available online 5 February 2024

interrelations where mechanistic understanding is scarce [6,7]. While reliable mass transfer models for chromatographic simulation exist, current isotherm equations struggle to capture the complex adsorption and desorption kinetics of multimodal chromatography [8]. Deviations of chromatographic simulation from experimental data are frequently accounted to the anisotropic shape, charge, and hydrophobicity distribution of the molecule surface, leading to a conformation- and orientation-dependency of the adsorption process that is especially pronounced for complex multidomain proteins such as mAbs [9,10]. Nevertheless, electrostatic, and hydrophobic effects are combined into a complex but lumped parameter model to reflect stoichiometric displacement [11] or colloidal particle interaction [12] of the molecule-adsorber system in a continuum representation, leading to semi-mechanistic isotherm equations. Additional factors like dewetting phenomena or pH dependence can be incorporated into the mechanistic description, further increasing model complexity [13]. Elevated complexity and semi-mechanistic model structure increase the difficulty in calibrating isotherm parameters as they cannot be directly measured but must be fitted to experimental observations [14]. In contrast, model applicability is enhanced through simplification, by focusing on either electrostatic [15] or hydrophobic effects [16] as the governing form of physicochemical interaction.

To gain a deeper understanding of the multimodal antibody binding, various biophysical methods have recently been employed and include: thermodynamic modeling of macroscopic observations [17], investigation of single amino acid substitutions in homologous fragment antigen binding (Fab) libraries [18], assessment of full-length antibody domain contributions [19–21], and identification of preferred ligand-binding regions of the fragment crystallizable (Fc) by using solution-phase nuclear magnetic resonance spectroscopy in concert with molecular dynamics (MD) simulations [22,23], as well as covalent labeling and mass spectrometry [24]. This experimental data can then be correlated with physicochemical antibody descriptors to infer quantitative structure-property relationships and improve mechanistic interpretability [25,26].

Despite recent advancements in isotherm discrimination and biophysical investigation, current models lack mechanistic insight into the multimodal adsorption process and are impeded by an elaborate and resource-intensive model calibration. A significant gap remains in linking macroscopic adsorption parameters with antibody structure to debottleneck the expensive model calibration and enhance mechanistic understanding, to facilitate an *a priori* process development of multimodal chromatography [8,27].

In this study, we aim to bridge this gap by introducing a multiscale modeling approach to predict multimodal chromatographic behavior of therapeutic antibodies based on their sequence information, as illustrated in Fig. 1. Previous studies have demonstrated the potential of multiscale modeling in predicting unimodal interactions within cation exchange chromatography for model proteins, both with [28] and without [29] pH dependency, as well as mAbs [30]. To the best of our knowledge, this is the first approach of predicting adsorption parameters of therapeutic antibodies in multimodal chromatography at various pH conditions. Our approach involves conducting LGE experiments with a multimodal anion exchange resin for a diverse set of therapeutic antibodies, determining linear adsorption parameters, and employing QSPR modeling to correlate these parameters with physicochemical descriptors calculated from antibody homology models. The accuracy of our multiscale predictions, provide valuable insights into the physicochemical aspects of multimodal chromatography by segregating contributions of the adsorption parameters. Ultimately, our multiscale approach paves the way for a more efficient and cost-effective development process in the purification of therapeutic antibodies.

2. Material and methods

2.1. Chromatography resin, buffers, and molecules

In this study, we employed the multimodal strong anion exchanger Capto adhere (Cytiva, Marlborough, USA) for all chromatographic experiments. Model calibration and validation were conducted using a prepacked Capto adhere HiScreen column (7.7 × 100 mm, Cytiva) with a column volume (CV) of 4.7 mL [31].

All chemicals utilized in this study were purchased from Sigma-Aldrich Co LLC (Saint Louis, USA). Highly purified water was used for buffer preparation, while 1 M hydrochloric acid and 1 M sodium hydroxide were employed for pH adjustment. A multicomponent buffer system at pH 5.0, 6.0, and 7.0 was used for equilibration, elution, and buffer exchange during all chromatographic experiments. This buffer system, adapted from Kröner and Hubbuch [32], consisted of 9.1 mM 1, 2-ethanediamine, 6.4 mM 1-methylpiperazine, 13.7 mM 1,4-dimethylpiperazine, 5.8 mM bis-tris, and 7.7 mM hydroxylamine, with the addition of 60 mM hydrochloric acid [16]. Furthermore, 1.5 M sodium chloride was added to the multicomponent buffer used for equilibration (at inlet B), resulting in a total of 1.56 M chloride counterions. Other buffers

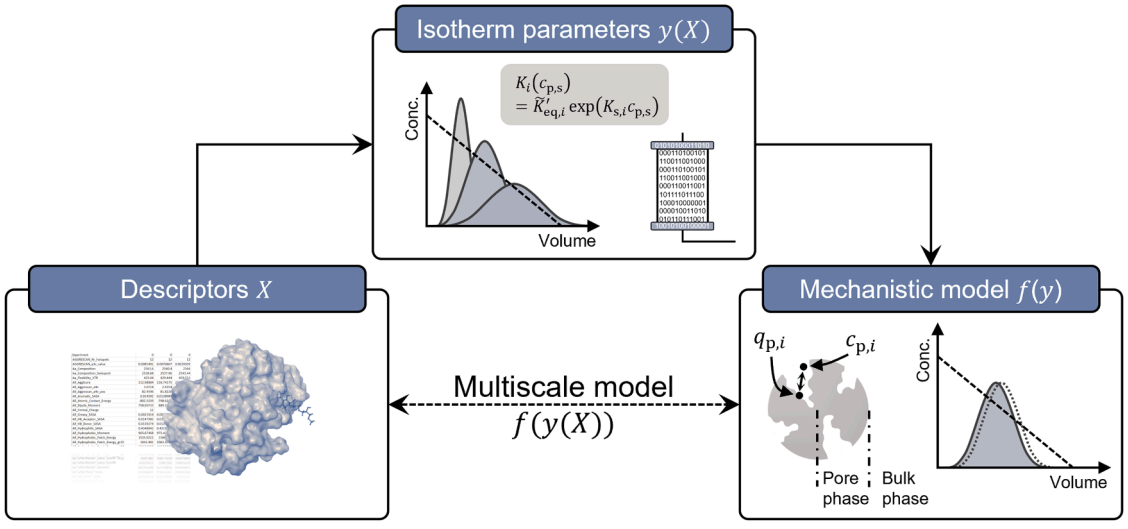

Fig. 1. Multiscale modeling workflow. This figure illustrates how QSPR models predict mechanistic adsorption parameters y from protein descriptors X. These parameters are then used in a priori process development through chromatographic simulation $f(y)$. Deviations in these simulations from standard operating conditions can guide molecule design to mitigate manufacturing risks. The depicted antibody structure is a modification of PDB entry 1HZH [35].

employed in the chromatographic experiments included 1 M acetic acid for column regeneration, 1 M sodium hydroxide for column cleaning, and 20 % ethanol for column storage.

This study utilized a diverse set of 59 full-length immunoglobulin G (IgG) derivatives, provided by Boehringer Ingelheim Pharma GmbH & Co. KG (Biberach, Germany) as introduced in our previous study [26]. This set included 57 human-origin and 2 humanized murine-origin antibodies, exhibiting a broad range of physicochemical characteristics. The antibody collection comprised 30 IgG1s, 19 IgG bispecifics with two single-chain fragment variables (scFv) appended to each heavy chain C-terminus (IgG(H)-scFv), 6 IgG4s, 3 Knob-in-Hole bispecifics (KiH), and 1 KiH trispecific with a single scFv attached to the C-terminus of the Hole chain (KiH-scFv). Antibody expression was performed using Chinese hamster ovary cell lines. Subsequently, the antibodies were captured through protein A affinity chromatography. The protein solutions were then neutralized to pH 5.5, sterile-filtered using a 0.2 μm filter, and frozen at −70 °C. Upon thawing, the final load material was adjusted to a concentration between 1.25 and 5 g/L, as determined by a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, USA). Prior to sample application, the load material underwent buffer exchange into the equilibration buffer using a 5 mL HiTrap Desalting column (Cytiva), following the manufacturer's instructions.

2.2. Homology modeling and descriptor calculation

The antibody structure prediction was conducted using homology modeling in Maestro Bioluminate 4.9 (Schrödinger Inc., New York, USA). The homology modeling protocol involved a five-stage process adapted from Zhu et al. [26], followed by further structure refinement based on the protocol of Sastry et al. [33]. The Enhanced Chothia numbering scheme [34] was employed, utilizing the human IgG1 1HZH [35] as full-length template for all molecules, except for IgG4 subtypes, which used 5DK3 [36] instead. Complex bi- and trispecific formats were modeled by grafting intra- and intermolecular linkers to append independently modeled scFv domains to the full-length mAb structures [37].

Subsequently, physicochemical descriptors were calculated using Bioluminate at the experimental pH value. A set of 165 unique features was calculated for each antibody structure, divided into sequence-based ($n = 69$), structural ($n = 59$), and patch-specific ($n = 37$) categories [38]. Additionally, a region-specific subset (n*=31) of the initial descriptor set was calculated for 37 antibody subdomains. These subdomains are specific to the antibody structure and encompass the light and heavy chain variable regions (VL, VL_Fv, VH, VH_Fv) focusing on the complementarity-determining regions (CDR, CDRL, CDRH) and framework regions (FR, FRL, FRH). These regions consist of individual loops (L1, L2, L3, H1, H2, H3) and frameworks (LFR1, LFR2, LFR3, LFR4, HFR1, HFR2, HFR3, HFR4). In addition to the variable region, the antibody constant regions (CL, CH1, CH2, CH3) and the hinge region (Hinge) are considered. The regional descriptors were further expanded by incorporating seven custom regions to accommodate the fragment variable (Fv), Fab, Fc, and the sum of the constant regions (CR). Moreover, the single-chain fragment variable regions (scFv, VLscFv, VLscFv) of the bi- and trispecific formats were taken into account. A more comprehensive explanation of the structure prediction and descriptor calculation can be found in our previous study [26].

In addition to the descriptors provided by Bioluminate, 62 supplementary global descriptors were calculated to complement the final feature set. These custom descriptors combine electrostatic, hydrophobic, and topological characteristics of the molecule solvent-accessible surface, calculated by MSMS [39] as implemented into the open-source tool SURFMAP [40]. SURFMAP provides the electrostatic potential calculated by the Adaptive Poisson–Boltzmann Solver [41], surface hydrophobicity using the residue-based Wimley-White hydrophobicity scale [42], and surface exposure, quantified through circular variance calculation [43] to consider surface cavity formation. To further characterize the surface topography, the normalized distance to the molecule center of mass was calculated for each surface grid point. Hereafter, these protein surface characteristics (electrostatic potential, hydrophobicity, circular variance, and normalized protein radius), were combined and summated, averaged, or binned into positive and negative contributions. An overview of the custom descriptors is detailed in Appendix B, Table B1. The final descriptor set consists of 1374 features per molecule, including 227 global descriptors and 1147 local descriptors.

2.3. Chromatographic experimentation

The chromatographic experiments involved a system and column characterization, as well as salt LGE experiments for the standardized calibration of the semi-mechanistic chromatography model as performed in our previous study [16]. These experiments were conducted using an ÄKTA Avant 25 (Cytiva) preparative chromatography system, controlled by Unicorn 7.5 (Cytiva) software, and maintained at a residence time of 5 min.

During the system and column characterization, extra-column effects and column-specific parameters were assessed by evaluating sensor dead volumes, column porosities, and ionic capacity. The characterization protocols were based on published protocols of Hunt et al. [44] and Huuk et al. [45]. Hereto, pulse injections of tracer solutions were applied to the system and the column at set flow rate. Subsequently, salt LGEs were performed to elute a total of 59 antibodies. Linear pH gradients performed in a previous study [46] identified a molecule-specific pH range in which salt elution was viable, reducing excessive experimentation. Three salt LGEs were then performed per molecule with a gradient length of 10, 20, and 30 CV at up to three distinct pH values (5.0, 6.0, 7.0). The column was initially equilibrated for 3 CV with equilibration buffer. After equilibration, the antibody solution was loaded onto the column up to a loading density of 1.0 g/L. Thereafter, the antibodies were eluted by linearly decreasing the sodium chloride concentration from 1560 to 60 mM within 10, 20, and 30 CV using the elution buffer. Retention times of the molecules were determined by measuring the first moments of the elution peaks through the UV trace at a wavelength of 280 nm, while recording the corresponding conductivity corrected by sensor dead volume. The column was then subjected to a 4 CV regeneration step, a 5 CV cleaning in place procedure, and a 4 CV storage step.

2.4. Chromatographic simulation

Following wet-lab experimentation, the chromatographic behavior was numerically simulated using the software ChromX (Cytiva) [47]. The column mass transfer was simulated by a transport dispersive model [48,49] to account for dispersion and diffusion effects, while the adsorption-desorption mechanism was described with a linear adsorption model introduced by Hess et al. [16], shown in Eq. (1).

$$K_i(c_{\mathrm{p,s}}) = \widetilde{K}'_{\mathrm{eq},i}\exp\left(K_{\mathrm{s},i}c_{\mathrm{p,s}}\right) \tag{1}$$

The utilized adsorption model follows an exponential form comparable to the hydrophobic interaction chromatography (HIC) isotherm introduced by Mollerup [50] and effectively neglects electrostatic attractive contributions that would be observed in ion-exchange operation. The equilibrium ratio $K_i(c_{\mathrm{p,s}})$ $[-]$ represents the ratio between the adsorbed protein $i \in [1,n]$ to the solute protein and is dependent from the modulator $s \in [1,m]$ concentration within the pore volume of the resin bead $c_{\mathrm{p,s}}$ $[molm^{-3}]$. The modified equilibrium constant $\widetilde{K}'_{\mathrm{eq},i} \approx k_{\mathrm{ads},i}/k_{\mathrm{des},i}$ $[-]$ is derived from the law of mass action at thermodynamic equilibrium and equals the ratio of the adsorption rate $k_{\mathrm{ads},i}$ to the desorption rate $k_{\mathrm{des},i}$ of the protein species from the ligand binding site and should be dependent on the ligand density of the resin. The salt-protein interaction parameter $K_{\mathrm{s},i}$ $[\mathrm{mol}^{-1}\mathrm{m}^3]$ introduced by

Mollerup [50] describes the thermodynamic activity of the solute protein and expresses the excess Gibbs free energy of the thermodynamic association during salt-protein interaction. Conceptually, this parameter should represent the intermolecular force of hydrophobic interaction caused by the salt-protein association.

The kinetic material balance of the multimodal adsorption process at low load density conditions can be expressed by Eq. (2), following the kinetic notation often employed in chromatographic simulation [14,51].

$$k_{\mathrm{kin},i}\frac{\partial q_{p,i}}{\partial t}(x,t) = \widetilde{K}_{\mathrm{eq},i}\exp\left(K_{\mathrm{s},i}c_{\mathrm{p,s}}\right)\, c_{\mathrm{p},i}(x,t) - q_{p,i}(x,t) \tag{2}$$

The kinetic adsorption model describes the time t [s] and location $x \in [0, L_c]$ [m] dependent binding of the protein to the chromatographic resin. Within the pore phase of the resin, the exchange of the unbound protein concentration $c_{\mathrm{p},i}$ [molm^{-3}] with adsorbed protein concentration $q_{\mathrm{p},i}$[molm^{-3}] is expressed, as depicted in Fig. 1. Consistent with Hahn et al. [52], the kinetic coefficient is defined as $k_{\mathrm{kin},i} = k_{\mathrm{des},i}^{-1}$ [s molm^{-3}].

Linear adsorption parameters were determined according to the protocol recently published by Hess et al. [16], based on the theoretical foundation of Yamamoto et al. [53]. The estimation method relies on linear modulator gradients and requires the determination of the normalized gradient slope GH [molm^{-3}], defined in Eq. (3).

$$GH = \frac{c_{\mathrm{s}}^{\mathrm{F}} - c_{\mathrm{s}}^{\mathrm{I}}}{V_{\mathrm{G}}}(1 - \varepsilon_{\mathrm{b}})V_{\mathrm{col}} \tag{3}$$

The slope of the gradient is given by the initial $c_{\mathrm{s}}^{\mathrm{I}}$ [molm^{-3}] and the final $c_{\mathrm{s}}^{\mathrm{F}}$ [molm^{-3}] salt concentrations at the start and the end of the gradient, divided by the linear gradient volume V_{G} [m^3]. The slope is than normalized by considering the interstitial porosity of the column bed ε_{b} [−] and the volume of the column V_{col} [m^3]. According to Yamamoto [54], the salt concentration at the peak maximum $c_{\mathrm{s},i}^{\mathrm{R}}$ [molm^{-3}] and GH are employed in an equilibrium chromatography model, which considers the zone movement of the components through the eluting column, shown in Eq. (4).

$$GH = \int_{c_{\mathrm{s}}^{\mathrm{I}}}^{c_{\mathrm{s},i}^{\mathrm{R}}} \frac{dc_{\mathrm{s}}}{\widetilde{K}_{\mathrm{eq},i}\exp\left(K_{\mathrm{s},i}c_{\mathrm{s}}\right) + K_{sec,i} - K_{sec,\mathrm{s}}} \tag{4}$$

Eq.(4), includes Eq. (1) and is integrated between the peak and the initial salt concentration. Furthermore, the distribution coefficients $K_{sec,i} = \varepsilon_{\mathrm{p,i}}$ [−] and $K_{sec,\mathrm{s}} = \varepsilon_{\mathrm{p,s}}$ [−] are considered, denoting the fraction of the pore volume accessible to the protein and the salt component at non-interacting conditions. The adsorption parameters $\widetilde{K}_{\mathrm{eq},i}$ and $K_{\mathrm{s},i}$ can be numerically estimated when inserting the first moments of the elution peaks from the salt LGEs into Eq. (4). In this study, the first moments were set equal to the peak maxima due to peak symmetry and were determined by using the python package SciPy [55]. Lastly, the kinetic parameter $k_{\mathrm{kin},i}$ is estimated by simultaneously curve fitting the three LGE experiments for each molecule and pH value, as described by Hahn et al. [51].

2.5. Multiscale model development and evaluation

In this study, we developed a multiscale model for predicting adsorption parameters from protein sequences using quantitative structure-property relationship modeling. Due to high dimensionality of the antibody-specific descriptor set ($n = 1374$) and the small dataset size ($m = 59$), descriptor preprocessing, dimensionality reduction, and model evaluation were required. The QSPR workflow was implemented using Python 3.9.12 and scikit-learn 1.0.2 [56], and modified from our previous study [26]. Gaussian process regression (GPR) was employed for non-linear predictions and heteroscedastic uncertainty estimation [57,58].

GPR, is rooted in Bayesian inference, and operates on the assumption that similar input variables (X) will yield similar target values (y). The primary objective of GPR is to identify a Gaussian distribution of mapping functions, commonly referred to as a kernel. This kernel is used to predict the target variable from the input features with minimal uncertainty, leveraging prior knowledge of their interrelationship. Once the joint distribution of mapping functions is updated or conditioned based on the training information, the model generates predictions derived from the mean of this updated distribution. The uncertainty, on the other hand, is represented by the distribution's standard deviation. This process allows GPR to make accurate predictions while providing an individual measure of confidence in each prediction.

According to the Bayesian framework, the adsorption parameters were predicted by conditioning a prior distribution of functions $P(y(x))$ with the training data $D = \{y,\ X\}$ to derive the posterior distribution $P(y(x)|\mathscr{D})$ using the Bayesian update rule, shown in Eq. (5).

$$P(y(x)|\mathscr{D}) \propto P(y|y(x), X)P(y(x)) \tag{5}$$

A mixed covariance function was chosen as the prior by multiplying a linear kernel with a Matérn class kernel and subsequently adding a white noise kernel [56,58,59]. Posterior conditioning was achieved by maximizing the likelihood $P(y|y(x), X)$ through minimization of log (marginal likelihood) using the l-BFGS-B algorithm [59].

The QSPR workflow begins with data preprocessing, where irrelevant descriptors were discarded, and several operations are performed to account for structural diversity of the IgG-like molecules. Furthermore, inaccurately determined adsorption parameters with a p-value above 0.1 are removed from the data set according to a two-sided Wald test proving the linearity of the initial parameter determination [16]. Thereafter, the dataset is randomly divided into 80 % training and 20 % test data, scaling the descriptors by their standard deviation (SD), and centering them based on the training data, while the adsorption parameters are transformed with the natural logarithm to assure normality of the target distribution. Two QSPR models are then created for both adsorption parameters using an identical workflow, training, and test set. Due to the strong correlation between the two adsorption parameters $\widetilde{K}_{\mathrm{eq},i}$ and $K_{\mathrm{s},i}$, a chained regression is performed, with $\widetilde{K}_{\mathrm{eq},i}$ predictions serving as possible feature for the $K_{\mathrm{s},i}$ parameter model, which is referred to as multioutput model. Dimensionality reduction is conducted by removing invariant descriptors, discarding low variance features, sorting remaining descriptors based on F-test results, removing collinear features, and selecting the 20 highest scoring features. Recursive feature elimination is then conducted using feature permutation importance [60]. Permutation importance is calculated as the average increase in model deviation when assessing model performance after shuffling a single feature one hundred times while keeping the remaining features constant. Model evaluation involved assessing overall model reliability and performance, feature interdependence, sensitivity, and contribution. Goodness of fit and prediction are inspected, and a 95 % confidence interval estimated. Fivefold cross-validation with ten repetitions is used for internal validation, while model reliability is analyzed through y-scrambling and MAE calculation [61]. Lastly, feature interdependence, sensitivity, and contribution are assessed using pairwise relationships, feature permutation importance, and partial dependence [60,62,63].

3. Results and discussion

3.1. Elution behavior of antibodies and antibody formats

LGE experiments were conducted using a multimodal anion exchange resin to evaluate 59 full-length antibodies across various formats and three pH values (5.0, 6.0, 7.0). The antibodies were eluted by linearly decreasing the sodium chloride concentration. An inverse relationship between salt concentration and molecule retention was observed, indicating a hydrophobically driven adsorption process. Fig. 2

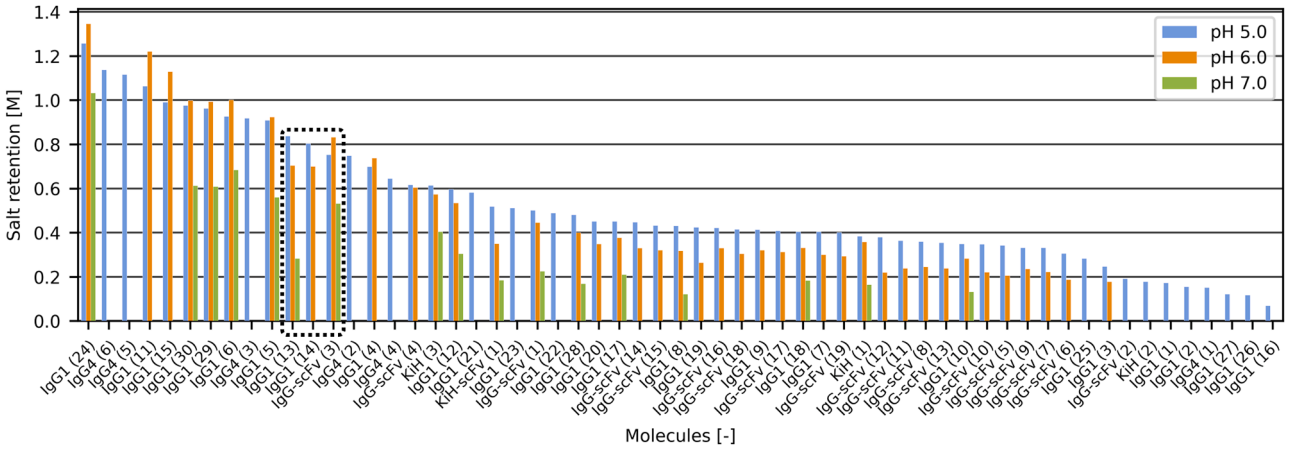

Fig. 2. Peak salt concentration of molecules in LGE elution experiments. The ionic strength at peak elution is shown for all molecules during 30 CV long negative sodium chloride gradients at pH 5.0, 6.0, and 7.0, arranged in ascending order based on pH 5.0 retention. A limited number of pH conditions were evaluated for some mAbs, as determined by linear pH gradients from a previous study [26]. The dashed black box highlights the pH-dependent selectivity reversal, exemplified by IgG1 (13) and IgG-scFv (3).

presents the ionic strength [M] at peak elution during a 30 CV gradient, arranged in ascending order based on pH 5.0 retention (decreasing ionic strength indicates increased multimodal retention during negative salt LGEs). Despite the structural similarities among the examined formats, antibody elution varied significantly, ranging from 1.347 [M] for IgG1 (24) at pH 6.0 to 0.069 [M] for IgG1 (16) at pH 5.0. This variability within the structurally conserved IgG1 antibodies indicates a weak format-dependent elution behavior. In contrast to the weak format-dependent elution behavior, a strong relationship was observed between antibody retention and elution pH, with retention increasing under elevated pH conditions. While most antibodies exhibited minimal variation in retention between pH 5.0 and 6.0, a substantial increase in retention was observed from pH 6.0 to 7.0. Similar observations have been reported for model proteins on a cation exchange resin [29] and a mAb on multimodal-cation exchange resins [20]. These findings were attributed to multiple, pH-dependent binding domains on the protein surface, leading to an altered protein binding orientation [64]. The pronounced pH dependency of antibody elution resulted in selectivity reversals, exemplified by IgG1 (13) and IgG-scFv (3), enclosed in a dashed box shown in Fig. 2. Upon further examination of pH trends, all antibodies eluting at a salt concentration above 0.7 [M] at pH 5.0 displayed a retention decrease from pH 5.0 to 6.0, followed by a larger increase from pH 6.0 to 7.0. Conversely, antibodies eluting below 0.7 [M] demonstrated a monotonous increase in retention with rising pH. The observed phenomena suggest a pH dependency in the multimodal chromatographic retention that is sensitive to salt concentration. To further elucidate these observations, additional experimentation will be required.

Following the gradient elution experiments, the equilibrium constant $\widetilde{K}'_{\mathrm{eq},i}$ and salt-protein interaction parameter $K_{\mathrm{s},i}$ of the multimodal adsorption model were calculated using Eq. (4). Initially, 118 parameter pairs were determined from 59 molecules at various pH levels. Out of the initial parameters, 103 parameter pairs from 55 molecules met the accuracy requirements as defined by the p-value in the parameter determination workflow, detailed in SubSection 2.5 and listed in the Appendix B, Table B4. Fig. 3. displays the natural logarithm of the scattered parameter pairs, providing format and pH information (a), as well as the distribution of training and test sets (b). Unimodal histograms of the adsorption parameters suggest an approximately normal distribution of the transformed parameters, framing the joint data as marginal axes. The parameter pairs appear to be broadly distributed and format independent. The logarithmic parameter $\log(\widetilde{K}'_{\mathrm{eq},i})$ ranges from −3.848 [-] for IgG1 (1) at pH 5.0 to 1.814 [-] for IgG1 (24) at pH 7.0 ($\widetilde{K}'_{\mathrm{eq},i}$: 0.021 - 6.132 [-]). Conversely, the parameter $\log(K_{\mathrm{s},i})$ ranges from −0.251 [-] for IgG1 (11) at pH 6.0 to 3.229 [-] for IgG1 (1) at pH 5.0 ($K_{\mathrm{s},i}$: 3.229 - 25.243 [-]). A strong negative correlation between the two parameters is evident, with a Pearson correlation coefficient p of −0.837. Parameter correlation has been observed in similar adsorption equations and is an inherent feature of the model structure used, which can impede parameter estimation [16,65]. Upon closer examination of the scattered adsorption parameters, a cluster of IgG4 antibodies is noticeable at higher $\log(\widetilde{K}'_{\mathrm{eq},i})$ and lower $\log(K_{\mathrm{s},i})$ values. This observation may indicate a format-specific relationship between the IgG4 structure and the multimodal elution behavior, as previously reported [26]. Furthermore, a strong and positive pH dependency of the $\log(\widetilde{K}'_{\mathrm{eq},i})$ values is indicated by perpendicular bands of parameter pairs at constant pH on the $\log(\widetilde{K}'_{\mathrm{eq},i})$ plane. This pattern could imply electrostatic contributions to the equilibrium constant within the multimodal adsorption model. This hypothesis is supported by the strong and negative correlation of $\log(\widetilde{K}'_{\mathrm{eq},i})$ values with the calculated net charge of the protein surface, as shown in Appendix A, Fig. A3, Subfigure (a8), with a Pearson correlation coefficient of −0.767. Consequently, a reduced positive surface charge at higher pH levels could decrease electrostatic repulsion towards the partially cationic ligand, thereby increasing the adsorption-to-desorption ratio, as expressed by the equilibrium constant $\widetilde{K}'_{\mathrm{eq},i}$. However, a similar pH dependency was not observed for the salt-protein interaction parameter $K_{\mathrm{s},i}$.

3.2. QSPR modeling of thermodynamic adsorption parameters

To realize the multiscale model, pH sensitive and molecule-specific parameters of the multimodal adsorption model had to be predicted from antibody structure. QSPR modeling was employed to correlate 103 parameter pairs from 55 molecules at three distinct pH values with 1374 global and local physicochemical descriptors, calculated from antibody homology models.

Initially, the data set was randomly divided into 80 % training (n = 44) and 20 % testing (n = 11) molecules. To prevent sharing structural information between the training and testing sets, different pH conditions of the molecules were grouped accordingly. Fig. 3(b) displays the scattered distribution of the logarithmic equilibrium constant $\log(\widetilde{K}'_{\mathrm{eq},i})$ and salt-protein interaction parameter $\log(K_{\mathrm{s},i})$ for the training and testing data. The parameter pairs of the test set molecules, including

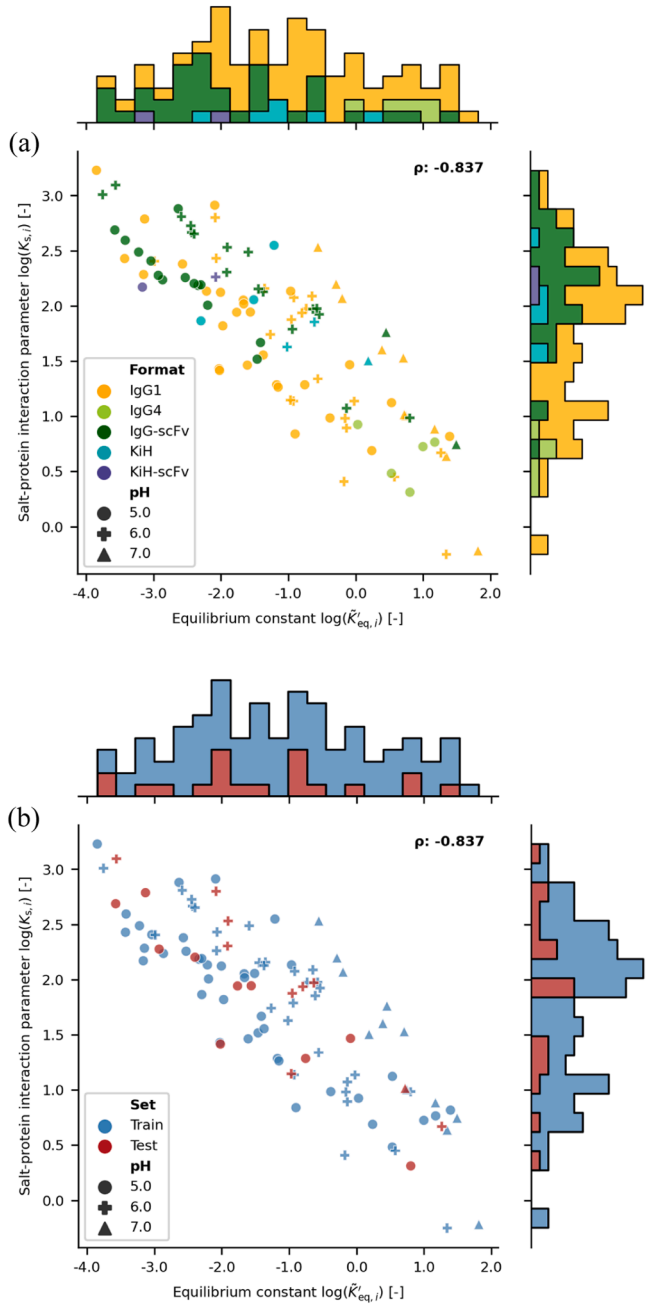

Fig. 3. Adsorption parameter-pair distributions from salt LGE experiments. a) Scatterplot showing antibody format- and pH-specific adsorption parameters, and b) Scatterplot presenting the distribution of model training and testing data, including pH information. The marginal axes of each subplot exhibit the unimodal histogram for each logarithmic adsorption parameter. The Pearson correlation coefficient (ρ) between both adsorption parameters is displayed in the upper right corner of the joint axes.

IgG1, IgG4, and IgG-scFv formats at pH 5.0, 6.0, and 7.0, are evenly distributed within the training data, as listed in Table 1. Two consecutive QSPR models were built from the training data for the two adsorption parameters. The prediction of $\log(\widetilde{K}'_{\mathrm{eq},i})$ served as a potential input feature for the subsequent $\log(K_{\mathrm{s},i})$ model. This multioutput model, also known as chained regression, aimed to enhance the accuracy of the $\log(K_{\mathrm{s},i})$ predictions, given their strong correlation displayed in Fig. 3. During the two-staged feature selection, informative descriptors were identified by reducing the initial 1374 descriptors to seven for the $\log(\widetilde{K}'_{\mathrm{eq},i})$ model and five for the $\log(K_{\mathrm{s},i})$ model, as can be assessed by the results of the recursive feature elimination shown in Appendix A, Fig. A1.

After model training, the two QSPR models were externally validated by predicting 22 parameter pairs of the test set molecules at multiple pH values, as listed in Table 1. The goodness of fit to the training data and the prediction accuracy for the test data is shown for $\log(\widetilde{K}'_{\mathrm{eq},i})$ (a) and $\log(K_{\mathrm{s},i})$ (b) in Fig. 4. High training accuracy was achieved for both logarithmic adsorption parameters, with R^2 values of 0.933 and 0.941 for $\log(\widetilde{K}'_{\mathrm{eq},i})$ and $\log(K_{\mathrm{s},i})$, respectively. The test set predictions exhibited reasonable accuracy, with Q^2 values of 0.834 and 0.835, and mean absolute errors (MAE) of 0.453 and 0.237 for the logarithmic parameters. Minor deviations in model prediction could originate from inaccuracies in the underlying descriptor models or adsorption parameter determination. Experimental parameter uncertainties from the standardized calibration workflow of up to 91 % for $\widetilde{K}'_{\mathrm{eq},i}$ determination and 20 % for $K_{\mathrm{s},i}$ determination had to be accounted for, as observed for IgG-scFv (6) at pH 6.0, listed in Appendix B, Table B4. Improving the standardized calibration workflow to increase data quality could significantly enhance the model predictions. However, the 95 % confidence interval (CI_{95}) of the training and test set predictions captured the model deviations, demonstrating the reliability of GPR to capture model uncertainty based on its training data. The reliability of the two QSPR models was further emphasized by the results of the internal model validation with a cross-validated Q^2 of 0.815 and 0.835 for $\log(\widetilde{K}'_{\mathrm{eq},i})$ and $\log(K_{\mathrm{s},i})$, as well as the results of y-scrambling, shown in Appendix A, Fig. A2. The absence of systematic model deviations was indicated by homogeneously dispersed residuals, as shown in the lower sections of Fig. 4, (a) and (b).

After validation of the QSPR models, an inspection of the descriptor contributions to the model predictions was conducted. The descriptor contributions were expressed through their partial dependencies, a method for examining non-linear feature interactions within a multivariate model. Fig. 5 displays the partial dependencies for $\log(\widetilde{K}'_{\mathrm{eq},i})$ (a) and $\log(K_{\mathrm{s},i})$ (b), which capture the target relationships depicted in Appendix A, Fig. A3 and Fig. A4. In brief, a single factor perturbation is performed by marginalizing all features except one and permutating it within the full feature range. Recording the output predictions yields the partial dependence of the descriptor. Additionally, feature permutation importance was calculated for each descriptor, as displayed on the upper right of the subfigures. Permutation importance measures model deterioration by shuffling individual features, enabling an assessment of the feature significance within a multivariate model.

In general, the partial dependencies of the test set molecules shown in Fig. 5 were comparable to the training data for both QSPR models, covering dense regions within the feature space as indicated by the decile lines on the bottom of the subfigures. The descriptors of the $\log(\widetilde{K}'_{\mathrm{eq},i})$ model primarily exhibited linear relationships towards the target variable and consisted of global (a1, a2, a7) and local (a5, a6) charge characteristics that emphasize the antibody variable region (a3, a4). In contrast, the descriptors of the $\log(K_{\mathrm{s},i})$ model displayed non-linear contributions and considered the molecule's global hydrophobicity, expressed by a supplemented descriptor (b4). The descriptors of the $\log(K_{\mathrm{s},i})$ model displayed similarities to the $\log(\widetilde{K}'_{\mathrm{eq},i})$ descriptors (b2, b3, b5) and even included the $\log(\widetilde{K}'_{\mathrm{eq},i})$ predictions (b1), which is unsurprising given the strong mutual correlation between the adsorption parameters discussed earlier. Moreover, the significance of the HFR3 region, an area in-between the H2 and H3 loop of the antibody variable region, was identified for both adsorption parameters (a3, b2). A detailed derivation of the utilized descriptors is given by Sankar et al. [38] and in Appendix B, Table B1.

Table 1
Observed and predicted thermodynamic model parameters of test set molecules after back transformation.

Molecule	pH	Observed $\tilde{K}'_{eq,i}$ [−]	Predicted $\tilde{K}'_{eq,i}$ [−]	Residual $\tilde{K}'_{eq,i}$ [%]	Observed $K_{s,i}$ [M]	Predicted $K_{s,i}$ [M]	Residual $K_{s,i}$ [%]
IgG1 (3)	5.0	0.044	0.073	67.434	16.239	12.782	−21.287
	6.0	0.124	0.157	26.094	16.452	12.188	−25.918
IgG1 (14)	5.0	0.469	0.406	−13.375	3.617	3.558	−1.649
	6.0	3.538	1.346	−61.954	1.950	2.597	33.196
IgG1 (20)	5.0	0.209	0.278	32.769	6.987	6.489	−7.130
	6.0	0.449	0.818	82.156	6.923	6.004	−13.278
IgG1 (22)	5.0	0.914	1.199	31.247	4.337	7.558	74.274
IgG1 (28)	5.0	0.171	0.273	59.693	6.976	5.496	−21.213
	6.0	0.385	0.921	139.177	6.536	3.374	−48.377
IgG1 (30)	5.0	0.132	0.152	14.642	4.118	4.200	2.008
	6.0	0.380	0.457	20.211	3.143	3.112	−0.992
	7.0	2.054	2.222	8.200	2.760	2.312	−16.243
IgG4 (6)	5.0	2.233	1.492	−33.199	1.366	1.809	32.455
IgG-scFv (6)	5.0	0.028	0.036	30.259	14.698	11.399	−22.443
	6.0	0.028	0.087	206.069	22.055	13.005	−41.032
IgG-scFv (11)	6.0	0.149	0.092	−37.920	12.564	15.669	24.712
IgG-scFv (15)	5.0	0.053	0.070	31.117	9.742	9.608	−1.382
	6.0	0.147	0.226	53.548	10.008	7.507	−24.993
IgG-scFv (18)	5.0	0.091	0.074	−18.133	9.050	10.110	11.716
	6.0	0.531	0.168	−68.399	7.183	9.867	37.356

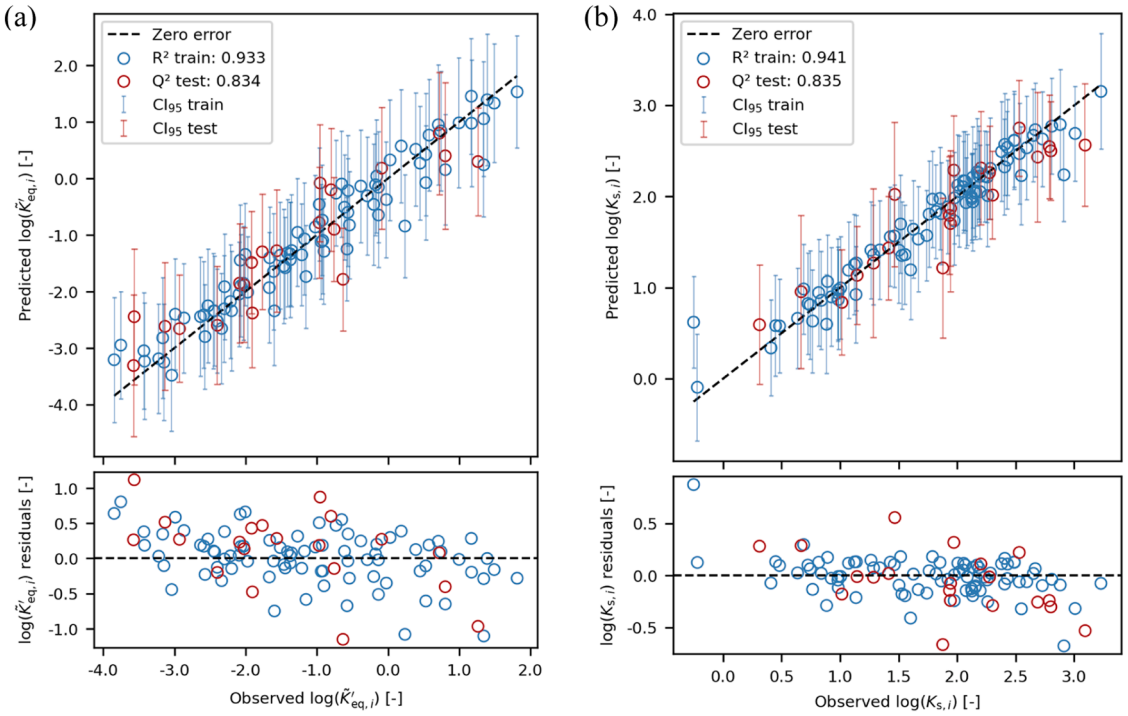

Fig. 4. Evaluation of adsorption parameter fit and prediction (R^2, Q^2). a) Logarithmic equilibrium constant $\log(\tilde{K}'_{eq,i})$ model, and b) Logarithmic salt-protein interaction parameter $\log(K_{s,i})$ model. The top section of each subfigure compares predicted and experimentally derived adsorption parameters for the molecules in the random training and testing data post-logarithmic transformation. An ideal model is depicted by a straight line, indicating zero error between predicted and experimental observations. The 95 % confidence interval (CI_{95}) of prediction is calculated for each molecule. The figure's lower section presents a residual plot, with the y-axis normalized to the absolute deviation of model prediction.

Interpreting the descriptor contributions of the adsorption parameters, the back-transformed equilibrium constant $\tilde{K}'_{eq,i}$ may capture an exponential electrostatic effect of positive repulsion or reduced negative attraction of the entire protein surface towards the anionic multimodal resin with emphasis on the antibody variable region. This effect is indicated by strong negative electrostatic descriptor correlations and importance in Fig. 5 (a1, a2, a5, a6, a7) and was already speculated about in SubSection 3.1. Interestingly, the results of our previous QSPR study [30] conducted for cation exchange chromatography, identified similar features for predicting the logarithmic equilibrium constant of the stoichiometric displacement model [66,67]. This finding could indicate the generalizability of the identified antibody characteristics to be relevant for other chromatographic resins utilizing electrostatic interactions. Therefore, improvements to the multimodal adsorption model should incorporate a pH extension of the $\tilde{K}'_{eq,i}$ parameter, already introduced for ion exchange adsorption models [44,68]. Conversely, the salt-protein interaction parameter $K_{s,i}$ suggests a more complex mechanism, potentially lumping multiple characteristics of significant binding domains into a single parameter. One such characteristic is the average surface hydrophobicity calculated after Wimley and White [42],

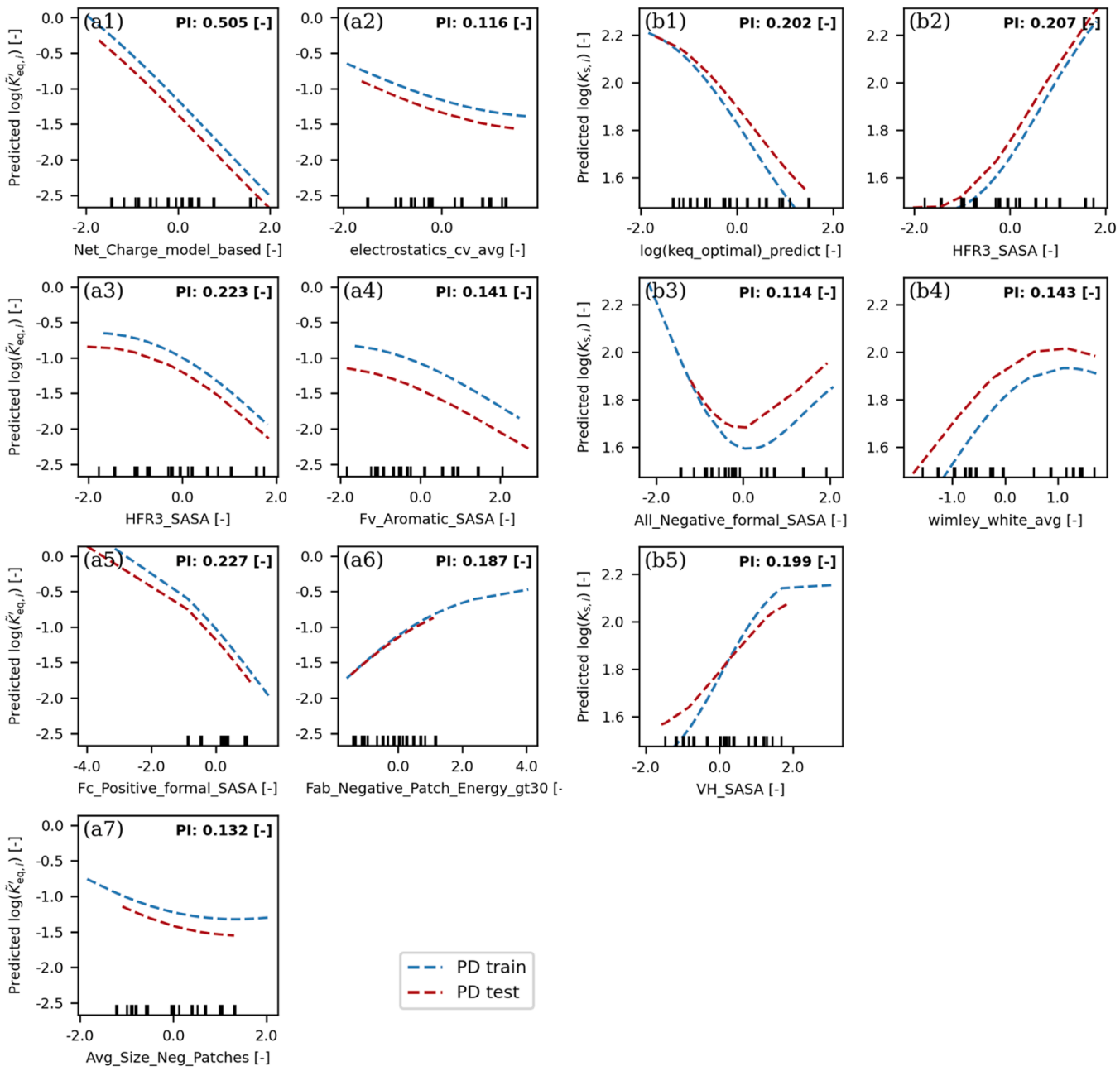

Fig. 5. Analysis of model features' partial dependence and permutation importance for training and testing data. (a) Logarithmic equilibrium constant $\log(\widetilde{K}_{\mathrm{eq},i})$ model, and (b) Logarithmic salt-protein interaction parameter $\log(K_{s,i})$ model. Each subfigure illustrates the model's response to feature permutations within their normalized range. The partial dependence (PD) displays the average target prediction for each feature. Permutation importance (PI) for target prediction is presented in the top-right corner of each subplot. Decile lines at each subplot's base indicate the frequency of feature values in the dataset.

which has recently proven to be a reliable predictor for HIC retention [69] and should be included as a standard descriptor in future QSPR studies. Lastly, the significance of the HFR3 region for antibody binding to multimodal resins has not yet been reported previously and warrants investigation in future studies.

3.3. Chromatographic process prediction

To validate the multiscale modeling method, the predicted adsorption parameters $\widetilde{K}_{\mathrm{eq},i}$ and $K_{s,i}$ of the test set molecules were employed in chromatographic simulations and compared against the measurement data. A macroscopic transport dispersive model was utilized for the chromatographic simulation, as implemented in our previous study [16], with $k_{\mathrm{kin},i}$ being determined by curve fitting to the experimental chromatograms. The results of the system and column characterization, as well as mass transfer parameters are provided in Appendix B, Table B2 and Table B3. Fig. 6 presents an overlay of the simulated and the measured elution of IgG1 (30) at pH 5.0, 6.0, and 7.0 during three different salt gradients with gradient lengths of 10, 20, and 30 CV, as shown in Subfigure a) to c). IgG1 (30) was selected for representation as the only test set molecule with experimental data available at all pH values. The chromatograms of the remaining test set molecules can be found in Appendix A, Fig. A5 to Fig. A14. A good agreement between the experimental and the simulated UV traces is evident in Fig. 6 across all operating conditions. Interestingly, the retention of IgG1 (30) was overpredicted at pH 5.0 and 6.0 and underpredicted at pH 7.0. Upon

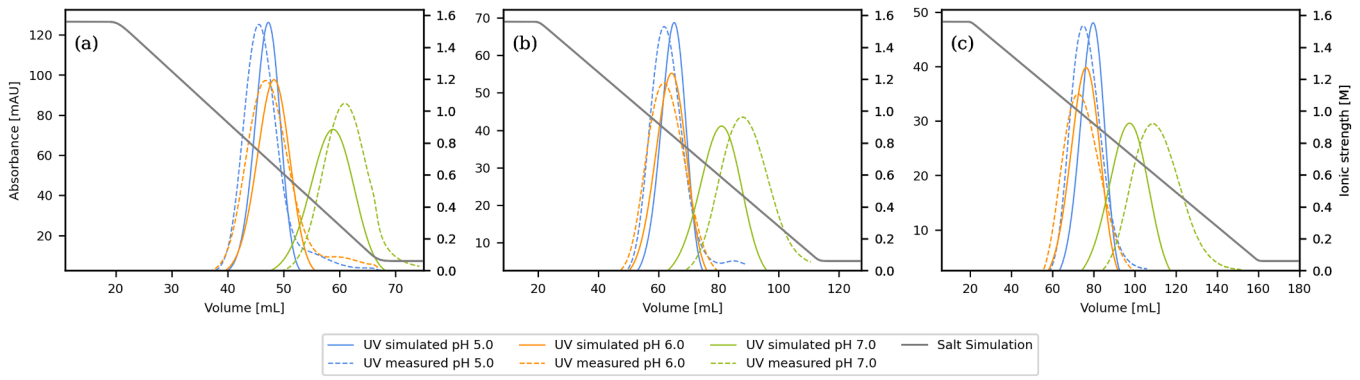

Fig. 6. Comparison of experimental and predicted chromatograms for IgG1 (30) under low loading density conditions. Subfigure (a) to (c) show the 10, 20, and 30 CV LGE experiments for the test set molecule at pH levels of 5.0, 6.0, and 7.0, respectively. These chromatograms are generated using multiscale modeling, which involves predicting thermodynamic parameters from antibody structure through QSPR models. These parameters are then incorporated into a transport dispersive model for chromatographic simulation.

evaluating the observed and predicted adsorption parameters in Table 1, an underprediction of the salt-protein interaction parameter $K_{s,i}$ at pH 7.0 of −16.2 % is evident in comparison to the other pH conditions, which explains the deviating model simulations. On the other hand, the chromatographic simulations appeared less sensitive to the equilibrium constant $\tilde{K}'_{eq,i}$, as previously observed by Altern et al. [14] employing a similar adsorption equation. The $K_{s,i}$ underprediction might be associated with the strong increase of antibody retention from pH 6.0 to 7.0, as discussed in SubSection 3.1. This increase could result from a transition of significant binding domains on the protein surface towards the multimodal ligands, which are challenging to be captured using static physicochemical descriptors and could translate to deteriorated model predictions. MD simulation of protein-ligand interaction could be one way of resolving this limitation as demonstrated by Banerjee et al. [70], but are currently not feasible for large multi-domain proteins. The validation of the remaining test set molecules, displayed in Appendix A, Fig. A5 to Fig. A14, revealed good agreement for all antibodies and process conditions, except for IgG1 (22) and IgG1 (28) at pH 6.0, as shown in Fig. A8 and Fig. A9. Notably, IgG1 (22) and IgG1 (28) also constituted the largest deviations in $K_{s,i}$ predictions, with residuals of 74.3 % and −48.4 % listed in Table 1, explaining the offset to the experimental chromatograms.

In conclusion, the presented multiscale model effectively predicts chromatographic behavior for a diverse set of therapeutic antibodies in a linear adsorption regime based solely on sequence information. Moreover, the incorporation of pH-sensitive descriptors into a mechanistic adsorption model enables the prediction of a broad range of chromatographic operating conditions, representing a significant step towards a priori process development for multimodal chromatography.

4. Conclusion

In this study, we successfully demonstrated the potential of a multiscale approach to predict the chromatographic behavior of a broad variety of therapeutic antibodies based solely on sequence information. LGE experiments were conducted on the multimodal resin Capto adhere for 59 full-length antibodies including five different formats and multispecific functionality. Multimodal adsorption parameters were determined across multiple pH values, consisting of the equilibrium constant $\tilde{K}'_{eq,i}$ and the salt-protein interaction parameter $K_{s,i}$ in a linear adsorption regime. QSPR modeling was employed to associate the adsorption parameters with physicochemical descriptors calculated from antibody homology models. The predicted adsorption parameters were incorporated into a semi-mechanistic chromatography model and validated experimentally.

The results revealed a strong correlation between antibody retention and elution pH as well as molecule specific interactions to the multimodal resin while the importance of the antibody format appeared less significant. The QSPR models exhibited high training accuracy and good test set prediction accuracy. The descriptor contributions of the adsorption parameters suggested the equilibrium constant to capture the electrostatic effects of the protein surface towards the anionic multimodal resin, while the salt-protein interaction parameter could represent a more intricate mechanism involving global surface hydrophobicity and significant binding domains. These finding could be generalizable to other chromatographic modes given the similarity of the selected antibody characteristics to previous QSPR studies. The chromatographic process predictions of the test set molecules showed good agreement with experimental data over a wide range of antibody formats and operating conditions.

This multiscale approach, incorporating pH-sensitive descriptors into an adsorption model, enables the prediction of multiple chromatographic operations and pH conditions (e.g., linear gradient, step gradient, or flow through chromatography). Thus, the presented model facilitates *a priori* process development for a variety of therapeutic antibodies in multimodal chromatography. As the underlining adsorption model is limited to low-loading density conditions, the multiscale model predictions are particularly useful for earlier rather than later stages of process development by specifying likely operating windows.

Increasing the amount of high-quality data and exploring the transition of significant binding domains on the protein surface could lead to potential model improvements. Future research should explore methods of quantifying the interaction of significant binding domains in multimodal chromatography like the impact of the HFR3 region or the role of adjacent surface patches in antibody binding. Next to these model improvements our results could facilitate the development of improved adsorption models for multimodal chromatography. Given the complexity of multimodal chromatography, our multiscale approach can serve as a template for unraveling complex physicochemical phenomena using structural information.

CRediT authorship contribution statement

Rudger Hess: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jan Faessler:** Formal analysis, Data curation. **Doil Yun:** Formal analysis, Data curation. **Ahmed Mama:** Software. **David Saleh:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Jan-Hendrik Grosch:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Gang Wang:**

Writing – review & editing, Conceptualization. **Thomas Schwab:** Resources, Project administration, Funding acquisition. **Jürgen Hubbuch:** Writing – review & editing, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank the department of Development Biologicals (Boehringer Ingelheim, Biberach, Germany) and the Biotherapeutics Discovery department Biologics Research CMC (Boehringer Ingelheim, Ridgefield, USA) and especially Michael Marlow, Helen Wu, and James Hu, for providing the therapeutic antibodies. Furthermore, we would like to thank the BI DSP modeling team for their constant support, as well as the long and fruitful discussions.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2024.464706.

References

[1] S.M. Cramer, M.A. Holstein, Downstream bioprocessing: recent advances and future promise, Curr. Opin. Chem. Eng. 1 (2011) 27–37, https://doi.org/10.1016/j.coche.2011.08.008.
[2] D. Sýkora, P. Řezanka, K. Záruba, V. Král, Recent advances in mixed-mode chromatographic stationary phases, J. Sep. Sci. 42 (2019) 89–129, https://doi.org/10.1002/jssc.201801048.
[3] V. Halan, S. Maity, R. Bhambure, A.S. Rathore, Multimodal chromatography for purification of biotherapeutics a review, Curr. Protein. Pept. Sc. 20 (2018) 4–13, https://doi.org/10.2174/1389203718666171020103559.
[4] L. Yu, L. Zhang, Y. Sun, Protein behavior at surfaces: orientation, conformational transitions and transport, J. Chromatogr. A. 1382 (2015) 118–134, https://doi.org/10.1016/j.chroma.2014.12.087.
[5] E. O'Connor, M. Aspelund, F. Bartnik, M. Berge, K. Coughlin, M. Kambarami, D. Spencer, H. Yan, W. Wang, Monoclonal antibody fragment removal mediated by mixed mode resins, J. Chromatogr. A 1499 (2017) 65–77, https://doi.org/10.1016/j.chroma.2017.03.063.
[6] A.T. Hanke, M. Ottens, Purifying biopharmaceuticals: knowledge-based chromatographic process development, Trends Biotechnol. 32 (2014) 210–220, https://doi.org/10.1016/j.tibtech.2014.02.001.
[7] M. Karlberg, M. von Stosch, J. Glassey, Exploiting mAb structure characteristics for a directed QbD implementation in early process development, Crit. Rev. Biotechnol. 38 (2018) 957–970, https://doi.org/10.1080/07388551.2017.1421899.
[8] V. Kumar, A.M. Lenhoff, Mechanistic modeling of preparative column chromatography for biotherapeutics, Annu. Rev. Chem. Biomol. Eng. 11 (2020) 235–255, https://doi.org/10.1146/annurev-chembioeng-102419-125430.
[9] J. Kittelmann, K.M.H. Lang, M. Ottens, J. Hubbuch, Orientation of monoclonal antibodies in ion-exchange chromatography: a predictive quantitative structure–activity relationship modeling approach, J. Chromatogr. A 1510 (2017) 33–39, https://doi.org/10.1016/j.chroma.2017.06.047.
[10] H.F. Tong, C. Cavallotti, S.J. Yao, D.Q. Lin, Molecular insight into protein binding orientations and interaction modes on hydrophobic charge-induction resin, J. Chromatogr. A 1512 (2017) 34–42, https://doi.org/10.1016/j.chroma.2017.06.071.
[11] B.K. Nfor, M. Noverraz, S. Chilamkurthi, P.D.E.M. Verhaert, L.A.M. van der Wielen, M. Ottens, High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents, J. Chromatogr. A 1217 (2010) 6829–6850, https://doi.org/10.1016/j.chroma.2010.07.069.
[12] T. Hahn, N. Geng, K. Petrushevska-Seebach, M.E. Dolan, M. Scheindel, P. Graf, K. Takenaka, K. Izumida, L. Li, Z. Ma, N. Schuelke, Mechanistic modeling, simulation, and optimization of mixed-mode chromatography for an antibody polishing step, Biotechnol. Progr. (2022) e3316, https://doi.org/10.1002/btpr.3316.
[13] Y.F. Lee, H. Graalfs, C. Frech, Thermodynamic modeling of protein retention in mixed-mode chromatography: an extended model for isocratic and dual gradient elution chromatography, J. Chromatogr. A 1464 (2016) 87–101, https://doi.org/10.1016/j.chroma.2016.08.026.
[14] S.H. Altern, J.P. Welsh, J.Y. Lyall, A.J. Kocot, S. Burgess, V. Kumar, C. Williams, A. M. Lenhoff, S.M. Cramer, Isotherm model discrimination for multimodal chromatography using mechanistic models derived from high-throughput batch isotherm data, J. Chromatogr. A 463878 (2023), https://doi.org/10.1016/j.chroma.2023.463878.
[15] H.S. Karkov, L. Sejergaard, S.M. Cramer, Methods development in multimodal chromatography with mobile phase modifiers using the steric mass action model, J. Chromatogr. A 1318 (2013) 149–155, https://doi.org/10.1016/j.chroma.2013.10.004.
[16] R. Hess, D. Yun, D. Saleh, T. Briskot, J.H. Grosch, G. Wang, T. Schwab, J. Hubbuch, Standardized method for mechanistic modeling of multimodal anion exchange chromatography in flow through operation, J. Chromatogr. A (2023) 463789, https://doi.org/10.1016/j.chroma.2023.463789.
[17] R.B. Gudhka, D.J. Roush, S.M. Cramer, A thermodynamic evaluation of antibody-surface interactions in multimodal cation exchange chromatography, J. Chromatogr. A 1628 (2020) 461479, https://doi.org/10.1016/j.chroma.2020.461479.
[18] H.S. Karkov, B.O. Krogh, J. Woo, S. Parimal, H. Ahmadian, S.M. Cramer, Investigation of protein selectivity in multimodal chromatography using in silico designed Fab fragment variants, Biotechnol. Bioeng. 112 (2015) 2305–2315, https://doi.org/10.1002/bit.25642.
[19] J. Robinson, D. Roush, S. Cramer, Domain contributions to antibody retention in multimodal chromatography systems, J. Chromatogr. A 1563 (2018) 89–98, https://doi.org/10.1016/j.chroma.2018.05.058.
[20] J. Robinson, D. Roush, S.M. Cramer, The effect of pH on antibody retention in multimodal cation exchange chromatographic systems, J. Chromatogr. A 1617 (2020) 460838, https://doi.org/10.1016/j.chroma.2019.460838.
[21] S.S. Parasnavis, B. Niu, M. Aspelund, W.K. Chung, M. Snyder, S.M. Cramer, Systematic workflow for studying domain contributions of bispecific antibodies to selectivity in multimodal chromatography, Biotechnol. Bioeng. 119 (2022) 211–225, https://doi.org/10.1002/bit.27967.
[22] R.B. Gudhka, C.L. Bilodeau, S.A. McCallum, M.A. McCoy, D.J. Roush, M.A. Snyder, S.M. Cramer, Identification of preferred multimodal ligand-binding regions on IgG1 FC using nuclear magnetic resonance and molecular dynamics simulations, Biotechnol. Bioeng. 118 (2021) 809–822, https://doi.org/10.1002/bit.27611.
[23] R.B. Gudhka, M. Vats, C.L. Bilodeau, S.A. McCallum, M.A. McCoy, D.J. Roush, M. A. Snyder, S.M. Cramer, Probing IgG1 FC–multimodal nanoparticle interactions: a combined nuclear magnetic resonance and molecular dynamics simulations approach, Langmuir 37 (2021) 12188–12203, https://doi.org/10.1021/acs.langmuir.1c02114.
[24] K. Dhingra, R.B. Gudhka, S.M. Cramer, Evaluation of preferred binding regions on ubiquitin and IgG1-FC for interacting with multimodal cation exchange resins using DEPC labeling/mass spectrometry, Biotechnol. Bioeng. (2023), https://doi.org/10.1002/bit.28361.
[25] J.R. Robinson, H.S. Karkov, J.A. Woo, B.O. Krogh, S.M. Cramer, QSAR models for prediction of chromatographic behavior of homologous fab variants, Biotechnol. Bioeng. 114 (2017) 1231–1240, https://doi.org/10.1002/bit.26236.
[26] R. Hess, J. Faessler, D. Yun, D. Saleh, J.H. Grosch, T. Schwab, J. Hubbuch, Antibody sequence-based prediction of pH gradient elution in multimodal chromatography, J. Chromatogr. A 1711 (2023) 464437, https://doi.org/10.1016/j.chroma.2023.464437.
[27] F. Insaidoo, S. Banerjee, D. Roush, S. Cramer, Preparative chromatography for separation of proteins, (2017) 177–225. https://doi.org/10.1002/9781119031116.ch6.
[28] A. Ladiwala, K. Rege, C.M. Breneman, S.M. Cramer, A priori prediction of adsorption isotherm parameters and chromatographic behavior in ion-exchange systems, P. Natl. Acad. Sci. U.S.A. 102 (2005) 11710–11715, https://doi.org/10.1073/pnas.0408769102.
[29] T. Yang, M.C. Sundling, A.S. Freed, C.M. Breneman, S.M. Cramer, Prediction of pH-dependent chromatographic behavior in ion-exchange systems, Anal. Chem. 79 (2007) 8927–8939, https://doi.org/10.1021/ac071101j.
[30] D. Saleh, R. Hess, M. Ahlers-Hesse, F. Rischawy, G. Wang, J. Grosch, T. Schwab, S. Kluters, J. Studts, J. Hubbuch, A multiscale modeling method for therapeutic antibodies in ion exchange chromatography, Biotechnol. Bioeng. 120 (2023) 125–138, https://doi.org/10.1002/bit.28258.
[31] Cytiva, Multimodal chromatography, (2021). https://www.cytivalifesciences.com/en/us/support/handbooks (accessed April 27, 2022).
[32] F. Kröner, J. Hubbuch, Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application, J. Chromatogr. A 1285 (2013) 78–87, https://doi.org/10.1016/j.chroma.2013.02.017.
[33] G.M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J. Comput. Aid Mol. Des. 27 (2013) 221–234, https://doi.org/10.1007/s10822-013-9644-8.
[34] K.R. Abhinandan, A.C.R. Martin, Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains, Mol. Immunol. 45 (2008) 3832–3839, https://doi.org/10.1016/j.molimm.2008.05.022.
[35] E.O. Saphire, P.W.H.I. Parren, R. Pantophlet, M.B. Zwick, G.M. Morris, P.M. Rudd, R.A. Dwek, R.L. Stanfield, D.R. Burton, I.A. Wilson, Crystal structure of a neutralizing human igg against HIV-1: a template for vaccine design, Science 293 (2001) 1155–1159, https://doi.org/10.1126/science.1061692.

[36] G. Scapin, X. Yang, W.W. Prosise, M. McCoy, P. Reichert, J.M. Johnston, R.S. Kashi, C. Strickland, Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab, Nat. Struct. Mol. Biol. 22 (2015) 953–958, https://doi.org/10.1038/nsmb.3129.
[37] K. Zhu, T. Day, Ab initio structure prediction of the antibody hypervariable H3 loop, Proteins Struct. Funct. Bioinform. 81 (2013) 1081–1089, https://doi.org/10.1002/prot.24240.
[38] K. Sankar, K. Trainor, L.L. Blazer, J.J. Adams, S.S. Sidhu, T. Day, E. Meiering, J.K. X. Maier, A descriptor set for quantitative structure-property relationship prediction in biologics, Mol. Inform. (2022) 2100240, https://doi.org/10.1002/minf.202100240.
[39] M.F. Sanner, A.J. Olson, J. Spehner, Reduced surface: an efficient way to compute molecular surfaces, Biopolymers 38 (1996) 305–320, https://doi.org/10.1002/(sici)1097-0282(199603)38:3<;305::aid-bip4>3.0.co;2-y.
[40] H. Schweke, M.H. Mucchielli, N. Chevrollier, S. Gosset, A. Lopes, SURFMAP: a software for mapping in two dimensions protein surface features, J. Chem. Inf. Model. 62 (2022) 1595–1601, https://doi.org/10.1021/acs.jcim.1c01269.
[41] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, D.H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D.W. Gohara, T. Dolinsky, R. Konecny, D.R. Koes, J.E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G. Wei, M.J. Holst, J. A. McCammon, N.A. Baker, Improvements to the APBS biomolecular solvation software suite, Protein Sci. 27 (2018) 112–128, https://doi.org/10.1002/pro.3280.
[42] W.C. Wimley, S.H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, Nat. Struct. Biol. 3 (1996) 842–848, https://doi.org/10.1038/nsb1096-842.
[43] M. Mezei, A new method for mapping macromolecular topography, J. Mol. Graph. Model. 21 (2003) 463–472, https://doi.org/10.1016/s1093-3263(02)00203-6.
[44] S. Hunt, T. Larsen, R.J. Todd, Preparative chromatography for separation of proteins, (2017) 399–427. https://doi.org/10.1002/9781119031116.ch13.
[45] T.C. Huuk, T. Briskot, T. Hahn, J. Hubbuch, A versatile noninvasive method for adsorber quantification in batch and column chromatography based on the ionic capacity, Biotechnol. Progr. 32 (2016) 666–677, https://doi.org/10.1002/btpr.2228.
[46] R.T. Hess, J. Faessler, D. Yun, D. Saleh, J.H. Grosch, T. Schwab, J. Hubbuch, Antibody sequence-based prediction of ph gradient elution in multimodal chromatography, SSRN [Preprint]. (2023). https://doi.org/10.2139/ssrn.4511168.
[47] T. Hahn, T. Huuk, V. Heuveline, J. Hubbuch, Simulating and optimizing preparative protein chromatography with chromx, J. Chem. Educ. 92 (2015) 1497–1502, https://doi.org/10.1021/ed500854a.
[48] H. Schmidt-Traub, M. Schulte, A. Seidel-Morgenstern, Preparative Chromatography, 2nd, completely revised and updated edition, Wiley-VCH Verlag GmbH &Co, KGaA, Weinheim, Germany, 2012.
[49] G. Guiochon, D.G. Shirazi, A. Felinger, A.M. Katti, Fundamentals of Preparative and Nonlinear Chromatography, 2nd ed., Academic Press, Boston, 2006.
[50] J.M. Mollerup, A Review of the thermodynamics of protein association to ligands, protein adsorption, and adsorption isotherms, Chem. Eng. Technol. 31 (2008) 864–874, https://doi.org/10.1002/ceat.200800082.
[51] T. Hahn, P. Baumann, T. Huuk, V. Heuveline, J. Hubbuch, UV absorption-based inverse modeling of protein chromatography, Eng. Life Sci. 16 (2016) 99–106, https://doi.org/10.1002/elsc.201400247.
[52] T. Hahn, A. Sommer, A. Osberghaus, V. Heuveline, J. Hubbuch, Adjoint-based estimation and optimization for column liquid chromatography models, Comput. Chem. Eng. 64 (2014) 41–54, https://doi.org/10.1016/j.compchemeng.2014.01.013.
[53] S. Yamamoto, K. Nakanishi, R. Matsuno, T. Kamikubo, Ion exchange chromatography of proteins prediction of elution curves and operating conditions. I. theoretical considerations, Biotechnol. Bioeng. 25 (1983) 1465–1483, https://doi.org/10.1002/bit.260250605.
[54] S. Yamamoto, Electrostatic interaction chromatography process for protein separations: impact of engineering analysis of biorecognition mechanism on process optimization, Chem. Eng. Technol. 28 (2005) 1387–1393, https://doi.org/10.1002/ceat.200500199.
[55] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, , S. 1 0 Contributors, A. Vijaykumar, A.P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C.N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D.A. Nicholson, D.R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G.A. Price, G.L. Ingold, G.E. Allen, G.R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J.T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J.V.de M. Cardoso, J. Reimer, J. Harrington, J.L.C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N.J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P.A. Brodtkorb, P. Lee, R.T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T.J. Pingel, T. P. Robitaille, T. Spura, T.R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y.O. Halchenko, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020) 261–272, https://doi.org/10.1038/s41592-019-0686-2.
[56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, G. Louppe, P. Prettenhofer, R. Weiss, R.J. Weiss, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[57] O. Obrezanova, G. Csányi, J.M.R. Gola, M.D. Segall, Gaussian processes: a method for automatic QSAR modeling of ADME properties, J. Chem. Inf. Model. 47 (2007) 1847–1857, https://doi.org/10.1021/ci7000633.
[58] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning), The MIT Press, 2005.
[59] P. Zhou, F. Tian, X. Chen, Z. Shang, Modeling and prediction of binding affinities between the human amphiphysin SH3 domain and its peptide ligands using genetic algorithm-Gaussian processes, Peptide Sci. 90 (2008) 792–802, https://doi.org/10.1002/bip.21091.
[60] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/a:1010933404324.
[61] M. Ojala, G.C. Garriga, Permutation tests for studying classifier performance, in: 2009 Ninth IEEE International Conference on Data Mining (ICDM), 2009, pp. 908–913, https://doi.org/10.1109/icdm.2009.108.
[62] T. Hastie, The Elements of Statistical learning : Data mining, inference, and Prediction, 2nd ed., Springer, New York, N.Y., 2009.
[63] A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin, Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation, J. Comput. Graph. Stat. 24 (2015) 44–65, https://doi.org/10.1080/10618600.2014.907095.
[64] R. Wälchli, M. Ressurreição, S. Vogg, F. Feidl, J. Angelo, X. Xu, S. Ghose, Z.J. Li, X. L. Saoût, J. Souquet, H. Broly, M. Morbidelli, Understanding mAb aggregation during low pH viral inactivation and subsequent neutralization, Biotechnol. Bioeng. 117 (2020) 687–700, https://doi.org/10.1002/bit.27237.
[65] F. Rischawy, D. Saleh, T. Hahn, S. Oelmeier, J. Spitz, S. Kluters, Good modeling practice for industrial chromatography: mechanistic modeling of ion exchange chromatography of a bispecific antibody, Comput. Chem. Eng. (2019) 106532, https://doi.org/10.1016/j.compchemeng.2019.106532.
[66] W. Kopaciewicz, M.A. Rounds, J. Fausnaugh, F.E. Regnier, Retention model for high-performance ion-exchange chromatography, J. Chromatogr. A 266 (1983) 3–21, https://doi.org/10.1016/s0021-9673(01)90875-1.
[67] A. Velayudhan, C. Horváth, Preparative chromatography of proteins: analysis of the multivalent ion-exchange formalism, J. Chromatogr. A 443 (1988) 13–29, https://doi.org/10.1016/s0021-9673(00)94779-4.
[68] M. Schmidt, M. Hafner, C. Frech, Modeling of salt and pH gradient elution in ion-exchange chromatography, J. Sep. Sci. 37 (2014) 5–13, https://doi.org/10.1002/jssc.201301007.
[69] F. Waibl, M.L. Fernández-Quintero, F.S. Wedl, H. Kettenberger, G. Georges, K. R. Liedl, Comparison of hydrophobicity scales for predicting biophysical properties of antibodies, Front. Mol. Biosci. 9 (2022) 960194, https://doi.org/10.3389/fmolb.2022.960194.
[70] S. Banerjee, S. Parimal, S.M. Cramer, A molecular modeling based method to predict elution behavior and binding patches of proteins in multimodal chromatography, J. Chromatogr. A 1511 (2017) 45–58, https://doi.org/10.1016/j.chroma.2017.06.059.