Journal of Chromatography A 1680 (2022) 463418

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Parameter-by-parameter method for steric mass action model of ion exchange chromatography: Theoretical considerations and experimental verification

Yu-Cheng Chen, Shan-Jing Yao, Dong-Qiang Lin*

Zhejiang Key Laboratory of Smart Biomaterials, Key Laboratory of Biomass Chemical Engineering of Ministry of Education, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 7 June 2022
Revised 8 August 2022
Accepted 10 August 2022
Available online 13 August 2022

Keywords:
Ion-exchange chromatography
Steric mass action model
Parameter estimation
Mechanistic model

ABSTRACT

Ion exchange chromatography (IEC) is one of the most widely-used techniques for protein separation and has been characterized by mechanistic models. However, the time-consuming and cumbersome model calibration hinders the application of mechanistic models for process development. A new methodology called "parameter-by-parameter method (PbP)" was proposed with mechanistic derivations of the steric mass action (SMA) model of IEC. The protocol includes four steps: (1) first linear regression (LR1) for characteristic charge; (2) second linear regression (LR2) for equilibrium coefficient; (3) linear approximation (LA) for shielding factor; (4) inverse method (IM) for kinetic coefficient. Four SMA parameters could be one-by-one determined in sequence, reducing the number of unknown parameters per species from four to one, and predicting almost consistent retention. Numerical single-component experiments were investigated firstly, and the PbP method showed excellent agreement between experiments and simulations. The effects of loadings on the PbP and Yamamoto methods were compared. It was found that the PbP method had higher accuracy and robustness than the Yamamoto method. Moreover, a five-experiment strategy was suggested to implement the PbP method, which is straightforward to reduce the cost of calibration experiments. Finally, a real-world multi-component separation was challenged and further confirmed the feasibility of the PbP method. In general, the proposed method can not only reliably estimate the SMA parameters with comprehensive physical understanding but also accurately predict retention over a wide range of loading conditions.

1. Introduction

The separation and purification of proteins with liquid chromatography have progressed over recent decades. Process development and optimization are critical for ion-exchange chromatography (IEC). Many researchers prefer to develop mechanistic models rather than repeat lots of conventional experiments in order to improve process efficiency [1,2]. Mechanistic models are useful in process development [3,4], resin characterization [5], process optimization [6,7] and model predictive control [8]. However, the determination of model parameters is essential for the application and generalization of modeling tools. The steric mass action (SMA) isotherm introduced by Brooks and Cramer [9] with four parameters (characteristic charge ν, equilibrium coefficient k_{eq}, shielding factor σ, and kinetic coefficient k_{kin}) is a popular extension to the stoichiometric displacement model (SDM) for IEC. These models can describe in a good way how the operating conditions affect IEC [10]. Both the SMA and SDM isotherms share similar profiles under diluted conditions, but they are radically different in the nonlinear region. There are two main available ways to calibrate the SMA model, the inverse method (IM) and the Yamamoto method [11–17].

IM is a standard practice to fit the observed experimental data to a specific model directly. Given an objective function, the calibration of an isotherm model is converted into an optimization problem [18]. A faulty optimization algorithm would make results unscientific and unreasonable or trap them in a local optimum. In practice, the calibration of SMA parameters may be not only an ill-conditioned problem but also an ill-posed problem, which means there are lots of solutions that satisfy the objective function. Moreover, the iterative process is time-consuming and dependent on the available computational power. The choice of appropriate initial guesses is troublesome without full physical understanding. The fewer parameters to be calibrated, the more reliable results of IM.

* Corresponding author.
E-mail address: lindq@zju.edu.cn (D.-Q. Lin).

https://doi.org/10.1016/j.chroma.2022.463418
0021-9673/

In contrast to IM, the Yamamoto method is a straightforward method due to physical cognition, which has been successfully adopted for process development and optimization of IEC in recent years [14,15,19–22]. Through a few linear gradient elution experiments (LGEs), this method can efficiently figure out two SMA parameters (ν and k_{eq}) by a linear relationship between the salt concentration at retention time and the normalized gradient slope. Moreover, the Yamamoto method is derived from a retention model, so it can also be applied to predict retention [23,24]. However, Creasy et al. [25] reported that the Yamamoto method can only predict retention when the loading is less than 5–10% of the column binding capacity. Saleh et al. [15] also employed the Yamamoto method to calculate initial guesses of SMA parameters at low loadings. Operating under diluted conditions has strict requirements for analytical equipment. However, the loadings for industrial-level processes are often high to ensure the process productivity. Experiments with various loadings are often carried out with different equipment, and the cross-device risk may lead to model failure and recalibration of model parameters [26]. Therefore, the Yamamoto method has some inherent limitations for industrial application. This might be that it is based on the SDM, which can not be transplanted directly to the nonlinear part of the SMA isotherm.

To better estimate the SMA parameters and accurately predict the retention of IEC, a new parameter-by-parameter (PbP) method would be developed in the present work that comes from some theoretic derivations and simplifying assumptions. Numerical experiments with different gradient lengths and loadings were conducted for validation and evaluation of the proposed method as well as real-world experiments. The Yamamoto method was used as a benchmark for comparison. The performance of these two methods was evaluated by parameter estimation and retention prediction. The quantitative bias analysis was employed to assess the sensitivity of specific variables and reveal the application scope of the developed method. Then a dedicated experimental strategy for better implementation of the PbP method was suggested.

2. Theory

2.1. Mechanistic model

The equilibrium dispersive model (EDM) is used as the column model of chromatography, including the description of convection, dispersion, and adsorption [27]. This model assumes that there is an instantaneous equilibrium between the solid and liquid phases, which in case of slow rates of mass transfer would possibly need additional adjustments. Hereby, the protein concentration in the bulk phase is equal to the average concentration in the intraparticular mobile phase [28]. The model lumps all effects caused by the axial dispersion and other mass transfer effects into one parameter, the apparent axial dispersion coefficient D_{app}, to describe peak broadening [29,30]. The mass balance for the mobile phase is given by:

$$\frac{\partial c_i}{\partial t}(z,t) = -\frac{u}{\varepsilon_t}\frac{\partial c_i}{\partial z}(z,t) + D_{app}\frac{\partial^2 c_i}{\partial z^2}(z,t) - \frac{1-\varepsilon_t}{\varepsilon_t}\frac{\partial q_i}{\partial t}(z,t) \quad (1)$$

where c_i and q_i represent the concentration of component i in the mobile phase and the stationary phase (mol/L), respectively. z and t are the axial position (m) and time (s), respectively. u denotes the superficial velocity of liquid phase (m/s). ε_t is the total porosity.

The initial conditions (ICs) and boundary conditions (BCs) are defined as follows. A rectangular pulse injection is considered as:

$$c_{inj,i}\left(t_{inj}\right) = \begin{cases} c_{inj}x_i & 0 < t \le t_{inj} \\ 0 & t > t_{inj} \end{cases} \quad (2)$$

where $c_{inj,i}$ and x_i are the injection concentration (mol/L) and the mole fraction of protein i, respectively. c_{inj} and t_{inj} represents the injection concentration (mol/L) and injection time (s), respectively. The Danckwerts BCs are applied as:

$$D_{app}\frac{\partial c_i}{\partial z}(0,t) = \frac{u}{\varepsilon_t}\left[c_i(0,t) - c_{inj,i}(t)\right] \quad (3)$$

$$D_{app}\frac{\partial c_i}{\partial z}(L,t) = 0. \quad (4)$$

The kinetic SMA isotherm is described as [31]:

$$k_{kin,i}\cdot\frac{\partial q_i}{\partial t}(z,t) = k_{eq,i}\left[\Lambda - \sum_{j=1}^{n}\left(\nu_j+\sigma_j\right)\cdot q_j(z,t)\right]^{\nu_i}\cdot c_i(z,t) - q_i(z,t)\cdot c_s^{\nu_i}(z,t) \quad (5)$$

$$q_s(z,t) = \Lambda - \sum_{j=1}^{n}\nu_j\cdot q_j(z,t) \quad (6)$$

where $k_{kin,i}$, $k_{eq,i}$, ν_i and σ_i are the kinetic coefficient, equilibrium coefficient, characteristic charge, and shielding factor of protein i, respectively. $k_{eq,i}$ is the ratio of adsorption and desorption constants affecting the elution peak position, while $k_{kin,i}$ is the reciprocal of desorption constant characterizing the influences of desorption rate and affecting the peak width [32]. Λ represents the total ion-exchange capacity for binding (mol/L). c_s and q_s are the salt concentration (mol/L) in the mobile phase and the stationary phase, respectively. $\sum_{j=1}^{n}(\nu_j+\sigma_j)q_j$ means the nonlinear term caused by the steric shielding of protein.

At the adsorption equilibrium, the left term of Eq. (5) is equal to zero, and it can be simplified as:

$$\frac{q_i(z,t)}{c_i(z,t)} = k_{eq,i}\cdot c_s^{-\nu_i}(z,t)\left[\Lambda - \sum_{j=1}^{n}\left(\nu_j+\sigma_j\right)\cdot q_j(z,t)\right]^{\nu_i}. \quad (7)$$

2.2. Parameter-by-parameter method

The core of the PbP method is a retention model obtained from the linearization of $\sum_{j=1}^{n}(\nu_j+\sigma_j)q_j$. The retention model can be utilized to determine the SMA parameters in sequence and predict the retention if the model parameters are known.

2.2.1. Retention model

Under diluted conditions ($\Lambda \gg \sum_{j=1}^{n}(\nu_j+\sigma_j)q_j$), the SMA model of component j at retention time of protein i ($t_{R,i}$) is approximate to the linear form with the Henry coefficient $H_{ij} = k_{eq,j}\cdot c_{s,R,i}^{-\nu_j}\cdot\Lambda^{\nu_j}$ as [12]:

$$q_{ij}(t_{R,i}) \approx H_{ij}(t_{R,i})\cdot c_{ij}(t_{R,i}) = \frac{H_{ij}(t_{R,i})\cdot c_{ij}(t_{R,i})}{\bar{c}_{inj}\left(t_{inj}\right)}\cdot\bar{c}_{inj}\left(t_{inj}\right) = k_{ij}(t_{R,i})\cdot\bar{c}_{inj}\left(t_{inj}\right). \quad (8)$$

where c_{ij} and q_{ij} represent the mean concentration in the mobile and stationary phase (mol/L), respectively, $\bar{c}_{inj} = c_{inj}\cdot u\cdot t_{inj}/L$ denotes the relative injection concentration (mol/L).

Using Eq. (8), the nonlinear term at $t_{R,i}$ can be related to the nonlinear coefficient a_i as:

$$\sum_{j=1}^{n}\left(\nu_j+\sigma_j\right)\cdot q_{ij}(t_{R,i}) \approx \sum_{j=1}^{n}k_{ij}(t_{R,i})\cdot\left(\nu_j+\sigma_j\right)\cdot\bar{c}_{inj}\left(t_{inj}\right) = a_i(t_{R,i})\cdot\bar{c}_{inj}\left(t_{inj}\right) \quad (9)$$

which is the most important feature of the PbP method.

The retention factor k_i' of protein i is given by [33]:

$$k_i'(t) = \frac{1-\varepsilon_t}{\varepsilon_t}\frac{q_i(L,t)}{c_i(L,t)}. \tag{10}$$

Using Eqs. (7) and (9), Eq. (10) can be rewritten as:

$$k_i'(t) = \frac{1-\varepsilon_t}{\varepsilon_t}\cdot k_{eq,i}\cdot c_s^{-\nu_i}(L,t)\cdot\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i}. \tag{11}$$

As for gradient elution of IEC, k_i' and c_s can be associated with the retention volume $V_{R,i}$ and the column volume V_{col} as [34]:

$$\int_0^{V_{R,i}} \frac{dV(t)}{k_i'(t)\cdot V_{col}\cdot\varepsilon_t} = 1 \tag{12}$$

$$V(t) = \frac{(1-\varepsilon_t)[c_s(L,t)-c_{initial}]}{GH}\cdot V_{col.} \tag{13}$$

The normalized gradient slope can be defined by:

$$GH = \frac{(1-\varepsilon_t)(c_{final}-c_{initial})}{CV_G} \tag{14}$$

where CV_G is the elution volume in column volume (CV) units. c_{final} and $c_{initial}$ denote the final and initial salt concentration of gradient elution at the column inlet, respectively. Using Eqs. (11), (13), and (14) related to $c_{s,R,i}$ (the salt concentration at $t_{R,i}$), Eq. (12) can be rewritten as:

$$\int_{c_{initial}}^{c_{s,R,i}} c_s^{\nu_i} dc_s = k_{eq,i}\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i} GH. \tag{15}$$

Then, integrating Eq. (15), we have

$$c_{s,R,i}^{\nu_i+1} = c_{initial}^{\nu_i+1} + k_{eq,i}(\nu_i+1)\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i} GH. \tag{16}$$

If the condition $(c_{s,R,i}/c_{initial})^{\nu_i+1} \gg 1$ is satisfied in most cases, the first term of the right of Eq. (16) is neglectable. As a result, Eq. (16) can be expressed as the logarithmic form as:

$$\lg GH = (\nu_i+1)\lg c_{s,R,i} - \lg\left[k_{eq,i}(\nu_i+1)\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i}\right]. \tag{17}$$

For the specific component in this retention model, there are three variables (GH, $c_{s,R}$, and $\bar{c}_{inj}$), and three undetermined coefficients (ν, k_{eq}, and a). In fact, the property of protein $c_{s,R}$ is dependent on the operating conditions (GH and $\bar{c}_{inj}$), so $c_{s,R}$ is the dependent variable while GH and $\bar{c}_{inj}$ are the independent variables. Using this retention model, we can calculate three undetermined coefficients (ν, k_{eq}, and a) by changing one of the operating conditions (GH or $\bar{c}_{inj}$) parameter-by-parameter as follows.

2.2.2. Estimation of characteristic charge

According to mechanisms of the SMA isotherm, ν is a fundamental feature to characterize the adsorption strength. It is associated with the order of retention but has no obvious significance in loadings [14]. Therefore, several LGEs with different gradient lengths and consistent loadings can be used to calibrate ν. For this, Eq. (17) can be rewritten as the linear equation with slope $m_{1,i} = \nu_i + 1$ and intercept $n_{1,i} = -\lg[k_{eq,i}(\nu_i+1)(\Lambda - a_i\bar{c}_{inj})^{\nu_i}]$ as:

$$\lg GH = m_{1,i}\lg c_{s,R,i} + n_{1,i}. \tag{18}$$

Eq. (18) is named as the first linear regression (LR1) of the PbP method. ν_i can be obtained as:

$$\nu_i = m_{1,i} - 1. \tag{19}$$

2.2.3. Estimation of equilibrium coefficient

One of two independent variables (GH) has been used to determine ν. At the same time, both $\bar{c}_{inj}$ and k_{eq} appear in the formula of $n_{1,i}$. Thus, the rest independent variable $\bar{c}_{inj}$ is employed to estimate k_{eq} in this section. $n_{1,i}$ can be further expressed as the linear function with slope $m_{2,i} = -k_{eq,i}^{-\nu_i^{-1}}\cdot a_i$ and intercept $n_{2,i} = k_{eq,i}^{-\nu_i^{-1}}\cdot\Lambda$ as:

$$10^{-n_{1,i}\cdot\nu_i^{-1}}(\nu_i+1)^{-\nu_i^{-1}} = m_{2,i}\bar{c}_{inj} + n_{2,i}. \tag{20}$$

Eq. (20) is named as the second linear regression (LR2) of the PbP method. Combining $m_{2,i}$ and $n_{2,i}$, a_i and $k_{eq,i}$ can be represented as:

$$a_i = -\frac{m_{2,i}}{n_{2,i}}\Lambda \tag{21}$$

$$k_{eq,i} = \left(\frac{n_{2,i}}{\Lambda}\right)^{\nu_i}. \tag{22}$$

To complete the LR2, the left-hand side of Eq. (20) should be calculated at every loading. If the same number of LGEs were conducted at every loading, it requires extra effort of experiments. A reasonable approach is derived as follows to reduce the experiment cost.

It is assumed that ν is identical at varying loadings, thereby Eq. (17) is a group of parallels. If the intercepts n_1 of these parallels are known, the left-hand side of Eq. (20) can be obtained owing to the calibrated ν by LR1. The point-slope formula can figure out the intercepts based on multicollinearity as:

$$n_{1,i} = -(\nu_i+1)\lg c_{s,R,i} + \lg GH. \tag{23}$$

2.2.4. Estimation of shielding factor

Eq. (9) can be rewritten as:

$$\sum_{j=1}^{n} k_{ij}(t_{R,i})\cdot\left(\nu_j+\sigma_j\right) = a_i(t_{R,i}) \tag{24}$$

where ν and a of individual protein i can be calculated by LR1 and LR2 as mentioned above, respectively. Therefore, the linear system of equations, $\mathbf{K}(\boldsymbol{\nu}+\boldsymbol{\sigma}) = \boldsymbol{a}$, can be shown as:

$$\begin{bmatrix} k_{11} & \cdots & k_{1j} \\ \vdots & \ddots & \vdots \\ k_{i1} & \cdots & k_{ij} \end{bmatrix}\begin{bmatrix} \nu_1+\sigma_1 \\ \vdots \\ \nu_j+\sigma_j \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_i \end{bmatrix} \tag{25}$$

where the element $k_{ij} = H_{ij}\cdot c_{ij}/\bar{c}_{inj}$ in $\mathbf{K}$ is valid using Eq. (8).

If the coefficient matrix $\mathbf{K}$ is invertible and positive definite, the linear system can be calculable as:

$$\boldsymbol{\sigma} = \mathbf{K}^{-1}\boldsymbol{a} - \boldsymbol{\nu}. \tag{26}$$

For the single-component system, Eq. (26) can be simplified as:

$$\sigma = \frac{\bar{c}_{inj}}{k_{eq}\cdot c_{s,R}^{-\nu}\cdot\Lambda^{\nu}\cdot c_{ij}}a - \nu. \tag{27}$$

The above method to calculate σ is named as the linear approximation (LA) of the PbP method.

2.2.5. Estimation of kinetic coefficient

As mentioned above, three parameters (ν, k_{eq}, and σ) can be directly calibrated by LR1, LR2, and LA, respectively. However, the straightforward estimation of the fourth parameter k_{kin} is extremely difficult, because it does not show up explicitly in the static SMA isotherm as Eq. (7). k_{kin} should be estimated by IM with elution curves of LGEs. The objective function is expressed as:

$$\min_{k_{kin}} J(c_h; k_{kin}) := \min_{k_{kin}} \sum_{j=1}^{m}\sum_{i=1}^{n} \geq \frac{\left\|c_i - c_{h,i}(z,t;k_{kin,i})\right\|_{L^2}^2}{\|c_i\|_{L^2}^2} \tag{28}$$

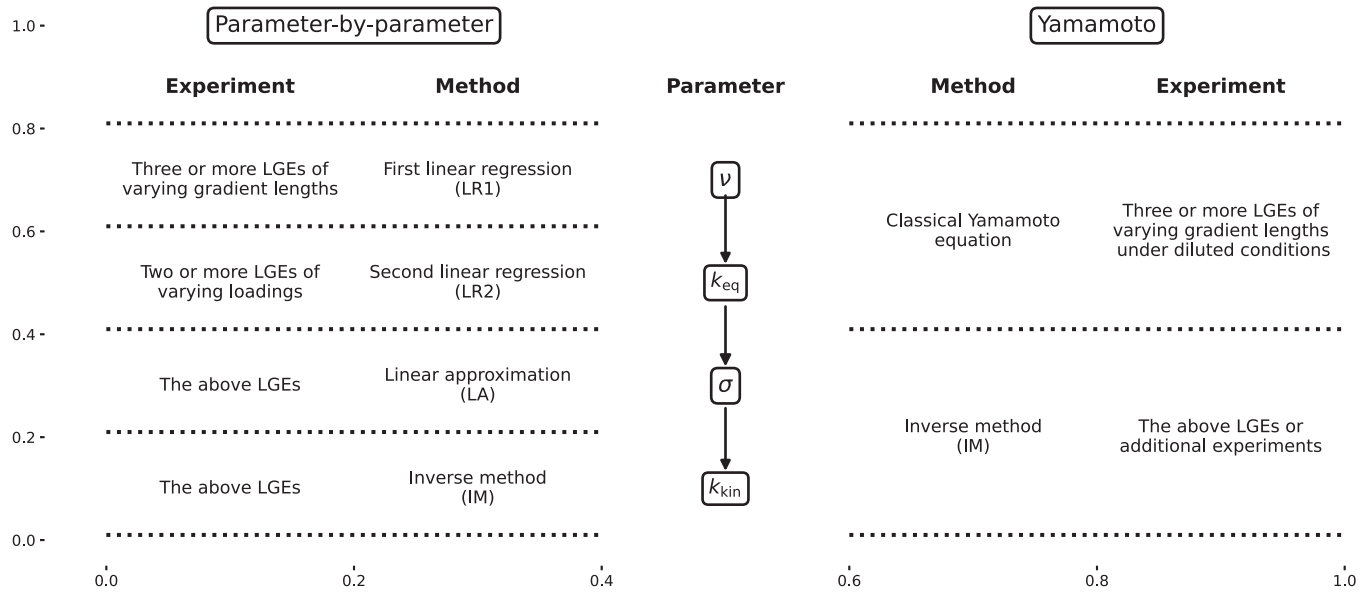

Fig. 1. Scheme of the PbP and Yamamoto method for parameter estimation.

where m and n are the total number of LGEs and components, respectively. For specific component i, c_i and $c_{h,i}$ represent the measurement and numerical approximation, respectively. $\|\cdot\|_{L^2}$ denotes the L^2-norm. The above optimization problem is solved with a combination of deterministic and heuristic methods, after which k_{kin} can be estimated.

2.2.6. Parameter estimation

Based on the above derivation, a parameter-by-parameter method is proposed to estimate the SMA parameters by several LGEs in different gradient lengths and loading conditions. The method includes four steps as follows:

(1) ν is estimated by LR1 with $c_{s,R}$ of LGEs in varying gradient lengths.

(2) k_{eq} is determined by LR2 with $c_{s,R}$ of LGEs under different loading conditions, and the point-slope formula is used to measure the intercepts of LR1.

(3) σ is calculated by solving a set of linear equations based on LA.

(4) k_{kin} is calibrated by IM with elution curves of LGEs.

The flowchart of the PbP method is shown in Fig. 1.

2.2.7. Retention prediction

The known parameters can also be applied to predict the retention by the above retention model. Using Eq. (17), $c_{s,R,i}$ can be derived as:

$$c_{s,R,i} = \left[k_{eq,i}(\nu_i+1)\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i}\mathrm{GH}\right]^{1/(\nu_i+1)}. \tag{29}$$

Combining Eqs. (14) and (17), the retention of protein i in CV units $\mathrm{CV}_{R,i}$ can be described by:

$$\mathrm{CV}_{R,i} = \frac{1-\varepsilon_t}{\mathrm{GH}}\left\{\left[k_{eq,i}(\nu_i+1)\left(\Lambda - a_i\bar{c}_{inj}\right)^{\nu_i}\mathrm{GH}\right]^{\frac{1}{\nu_i+1}} - c_{initial}\right\} \tag{30}$$

which can be calculated by a combination of operating conditions (GH and $\bar{c}_{inj}$), the nonlinear coefficient a using LR2, and the SMA parameters (ν and k_{eq}).

2.3. Yamamoto method

If $\bar{c}_{inj}$ tends to zero, Eq. (17) can be simplified as:

$$\lg\mathrm{GH} = (\nu_i+1)\lg c_{s,R,i} - \lg\left[k_{eq,i}(\nu_i+1)\Lambda^{\nu_i}\right] \tag{31}$$

which is identified to the classical Yamamoto equation [13]. It is worth noting that the simplification condition is an infinitesimal loading, so the Yamamoto method can be used at low loadings only. Eq. (31) contains two parameters (ν_i and $k_{eq,i}$), and can be rewritten with the slope $m_i = \nu_i + 1$ and intercept $n_i = -\lg[k_{eq,i}(\nu_i+1)\Lambda^{\nu_i}]$ as:

$$\lg\mathrm{GH} = m_i\lg c_{s,R,i} + n_i. \tag{32}$$

Then, ν_i and $k_{eq,i}$ can be obtained by several LGEs in varying gradient lengths as:

$$\nu_i = m_i - 1 \tag{33}$$

$$k_{eq,i} = \frac{10^{-n_i}}{m_i\Lambda^{\nu_i}}. \tag{34}$$

The other two parameters (σ and k_{kin}), are typically solved by combining IM with elution curves of LGEs [13]. The comparison of the Yamamoto and the PbP method for parameter estimation and experiment requirements is presented in Fig. 1. The Yamamoto method can also be used to predict the retention. If $\bar{c}_{inj}$ in Eq. (30) approaches zero, it can be simplified as:

$$\mathrm{CV}_{R,i} = \frac{1-\varepsilon_t}{\mathrm{GH}}\left\{\left[k_{eq,i}(\nu_i+1)\Lambda^{\nu_i}\mathrm{GH}\right]^{\frac{1}{\nu_i+1}} - c_{initial}\right\}. \tag{35}$$

3. Materials and methods

3.1. Numerical solution and software

To solve the mechanistic model of IEC, the discontinuous Galerkin finite element method (DG-FEM) was adopted [35,36], which is an efficient numerical approach for differential algebraic equations (DAEs). These codes were implemented by Python 3.10 with many functional packages as follows: the fundamental package NumPy for matrix calculations [37], a stiff solver in the scientific framework SciPy with the Jacobian matrix for solving the initial problem [38], the powerful module scikit-opt for heuristic algorithms, and the 2D graphics toolkit Matplotlib for generating publication-quality images [39].

Table 1
Parameters of column geometry, operating condition, mass transport, and binding of proteins in numerical and real-world experiments.

Catalog	Parameter	Symbol	Numerical experiment [18]		Real-world experiment [25,40]	
			Value	Unit	Value	Unit
Geometry	Column length	L	25.0	cm	5.0	cm
	Total porosity	ε_t	0.58	–	0.86	–
	Particle porosity†	ε_p	0.30	–	0.77	–
	External porosity	ε_e	0.40	–	0.38	–
	Ionic capacity	Λ	0.57‡	mol/L	0.20	mol/L
Operating	Initial salt concentration	$c_{initial}$	0.05	mol/L	0.02	mol/L
	Final salt concentration	c_{final}	0.55	mol/L	0.32	mol/L
	Flow rate		2.88E-8	m^3/s	0.50	mL/min
Transport	Apparent axial dispersion (external volume)	$D_{app,e}$	0.30[P]	mm^2/s	0.79	mm^2/s
	Apparent axial dispersion§	D_{app}	0.21	mm^2/s	0.35	mm^2/s
Binding‖	Characteristic charge	ν	7.00	–	–	–
	Equilibrium coefficient	k_{eq}	0.200	–	–	–
	Shielding factor	σ	50.0	–	–	–
	Kinetic coefficient	k_{kin}	0.100[P]	sM^{ν}	–	sM^{ν}

†: ε_t was calculated by $\varepsilon_e + (1-\varepsilon_e)\varepsilon_p$.
‡: Λ was modified to satisfy the relationship of $\Lambda/(\nu+\sigma)$ =0.01 in numerical experiments.
§: D_{app} of EDM was estimated by $D_{app}\varepsilon_t = D_{app,e}\varepsilon_e$ [30].
[P]: $D_{app,e}$ and k_{kin} were selected for validation of the PbP method in numerical experiments, although the column model used in [18] is different from this work.
‖: The ground truths of four SMA parameters are unknown in real-world experiments.

Table 2
A summary of all numerical experiments.

No.	Purpose	CV_G (CV)	LF_{col} (% column)
1	LR1	8	5
2	LR1	16	5
3	LR1	32	5
4	LR2	8	1
5	LR2	16	2
6	LR2	32	10

3.2. Numerical experiment

To validate the accuracy and feasibility of the PbP method, numerical experiments were performed based on a presumed protein with the known SMA parameters according to Heymann et al. [18]. The column-specific and SMA parameters are available from Heymann et al. [18] as listed in Table 1. The DAEs that include the EDM and the SMA isotherm were implemented to conduct virtual bind-and-elute experiments under different conditions and obtain elution curves.

To evaluate the effects of loading conditions, the loading factor LF was used as a comparative benchmark, which is defined by the ratio of the injected amount $\bar{c}_{inj}$ to the saturated adsorption capacity q_{max} as [27]:

$$\mathrm{LF} = \frac{\bar{c}_{inj}}{q_{max}}. \tag{36}$$

For the single-component system with the SMA isotherm ($q_{max} = \Lambda/(\nu+\sigma)$), the loading factor can be expressed as:

$$\mathrm{LF} = \frac{\nu+\sigma}{\Lambda} \cdot \bar{c}_{inj}. \tag{37}$$

The loading factor of column $\mathrm{LF}_{col} = \mathrm{LF}/\varepsilon_t$ is applied in numerical experiments. Six LGEs with varying gradient lengths and loadings were generated, as listed in Table 2. The experiments were classified into two groups: the first three in different gradient lengths for LR1 and the remaining three at different loadings for LR2. The molecular weight of the presumed protein was 15 kDa. The sample contained 0.05 mol/L NaCl and the injection volume was 1.0 CV. After sample injection, the column was washed with 0.05 mol/L NaCl and then eluted with the linear salt gradient from 0.05 to 0.55 mol/L NaCl.

3.3. Real-world experiment

A series of LGEs on the separation of monoclonal antibody (mAb) monomer-dimer mixtures with IEC were taken from Reck et al. [40] and Creasy et al. [25]. A stock solution contained 72% mAb and 28% dimer. The parameters of column geometry, operating condition, mass transport, and binding of proteins are summarized in Table 1. All experiments are divided into two groups, as displayed in Table 3. The loading buffer contained 0.02 mol/L Na^+. After the sample injection, the column was washed with 0.02 mol/L Na^+ and then eluted with the linear salt gradient from 0.02 to 0.32 mol/L Na^+.

3.4. Model qualification

Numerical experiments were used to calibrate the SMA parameters. The agreement between the calibration $\hat{p}$ and ground truth p could be measured by a L^2-error as:

$$\sqrt{\frac{\left\| p - \hat{p} \right\|_{L^2}^2}{\|p\|_{L^2}^2}} \tag{38}$$

which can also be applied to evaluate the agreement of retention.

Afterwards, the estimated parameters were employed to simulate elution profiles. In the same way, the L^2-error of elution profiles was used as:

$$\sqrt{\frac{J(c_h; k_{kin})}{mn}}. \tag{39}$$

4. Results and discussion

4.1. Evaluating parameter-by-parameter method by numerical experiment

The numerical experiment is a useful tool to evaluate if the model prediction and the virtual experiments give the same results in chromatography [18,41,42]. Zhang et al. [41] performed numerical experiments to evaluate the regularization to solve the inverse problem of competitive Langmuir isotherm. It is more suitable than real-world experimental data to test and verify the

Table 3
A summary of all real-world experiments.

No.	Purpose	CV_G (CV)	Loadings (g/L column)	Loading volume (mL)	Ref.
1	LR1	5	2	0.1	Reck et al. [40]
2	LR1	10	2	0.1	Reck et al. [40]
3	LR1	15	2	0.1	Reck et al. [40]
4	LR1	25	2	0.1	Reck et al. [40]
5	LR1	40	2	0.1	Reck et al. [40]
6	LR2	25	4	0.97	Creasy et al. [25]
7	LR2	25	10	2.4	Creasy et al. [25]
8	LR2	25	20	4.9	Creasy et al. [25]
9	LR2	25	40	9.7	Creasy et al. [25]

method of parameter estimation as a preliminary attempt because the agreement of parameters can be measured. EDM is a well-known column model, where the lumped parameter D_{app} includes peak broadening effects caused by axial dispersion as well as by all other mass transfer [30]. Kaczmarski [43] reported that the solution of EDM is equivalent to that of the general rate model (GRM) that is the most detailed column model combining the mass transfer in various kinds when D_{app} is calculated reasonably. Due to the fact that the PbP method was developed by the retention model, EDM was sophisticated enough to conduct numerical experiments according to the survey that the ignorance of other mass transfer (film diffusion, pore diffusion) nearly did not contribute to offset of retention [29].

As presented in Table 1, four SMA parameters of numerical experiments were provided from Heymann et al. [18]. Six LGEs were set with different gradient lengths and loadings as shown in Table 2. Gradient lengths increased in an exponential way from 8 to 32 CV. Creasy et al. [25] suggested that the Yamamoto method could only represent the retention when the loading is less than 5–10% of the column binding capacity. Therefore, the first three LGEs for LR1 were set at 5% LF_{col}, while the second three LGEs for LR2 would be conducted at different loadings from 1 to 10% LF_{col}.

With the parameters listed in Tables 1 and 2, elution curves are generated, as shown in Fig. 2. For three LGEs at 5% LF_{col} in different gradient lengths (Fig. 2a–c), it can be seen that $c_{s,R}$ declines steadily from 0.38 to 0.35, then to 0.32 as the gradient length increases from 8 to 16, then to 32 CV. When comparing $c_{s,R}$ in the same gradient length but different loadings (Fig. 2d to a, e to b, c to f), it can be found that $c_{s,R}$ decreases slightly (from 0.40 to 0.38, 0.36 to 0.35, 0.32 to 0.30) as the loading increases. These two trends have been reported by prior research qualitatively [25,44,45], which can be explained by Eq. (30) from the viewpoint of mathematics in this study. $c_{s,R}$ is a decreasing function of the gradient length but an increasing function of the loading. This formula brings it into correspondence with the qualitative analysis of the effects of loadings and gradient lengths in the previous study. No obvious differences are observed between the real-world and numerical experiments.

4.1.1. Parameter estimation of numerical experiment

GH of all experiments is calculated by Eq. (14), and $c_{s,R}$ is determined by the moment analysis [26]. Three LGEs injected with 5% LF_{col} in different gradient lengths (Fig. 2a–c) are used to establish the LR1 with GH and $c_{s,R}$ as shown in Fig. 3a. The point-slope formula is applied to measure the intercepts of three LGEs (Fig. 2d–f), which are used to build the LR2 as shown in Fig. 3b. The use of the point-slope formula leads to a group of parallel lines, which effectively saves the experimental efforts.

It can be found that there is a significant positive correlation between GH and $c_{s,R}$. The correlation coefficient is close to 1.00, which means that the derivation and assumption of LR1 are valid under the experimental conditions. In fact, there is an equivalence formula between Eqs. (18) and (32), which allows the line in Fig. 3 to represent the Yamamoto method. That is, the Yamamoto method is equivalent to the first step of the PbP method (LR1). The outcomes of LR2 are presented in Fig. 3b. A positive correlation is discovered in LR2 with the correlation coefficient of 1.00, which indicates that the linearization of the nonlinear term is tenable for the scope of loadings in this picture.

Table 4
SMA parameters of numerical experiments estimated by different methods.

Parameter	Unit	Ground Truth	Parameter-by-parameter		Yamamoto	
			Value	L^2-error	Value	L^2-error
ν	–	7.00	6.92	0.011	6.92	0.011
k_{eq}	–	0.200	0.188	0.062	0.109	0.457
σ	–	50.0	44.81	0.104	–	–
k_{kin}	sM^{ν}	0.100	0.103	0.032	–	–

For the PbP method, three parameters (ν, k_{eq}, and σ) can be estimated by the slopes and intercepts of LR1 and LR2 in Fig. 3. ν is calibrated by Eq. (19) based on the slope of LR1 (m_1). k_{eq} is determined by Eq. (22) related to the result of LR1 and the intercept of LR2 (n_2). σ is calculated by Eq. (27) with LR1 and LR2. k_{kin} is figured out by IM with elution curves. For the Yamamoto method, two parameters (ν and k_{eq}) are determined by Eqs. (33) and (34) according to the slope and intercept of LR1, respectively. All of the SMA parameters estimated by the PbP and Yamamoto methods are summarized in Table 4.

For the first parameter ν, the calibration of the proposed method stacks up with the ground truths with the L^2-error of 0.011. A similar value is obtained by the Yamamoto method. As mentioned above, the first step of the PbP method (LR1) is equivalent to the Yamamoto method owing to the consistent function between Eqs. (18) and (32).

For the second parameter k_{eq}, the proposed method can almost precisely retrieve the values of k_{eq} with the L^2-error of 0.062. By contrast, the Yamamoto method leads to an unfavorable L^2-error of 0.457, which is seven times higher than that of the PbP method. The results confirm that the main limitation of the Yamamoto method is the paucity of support at high loadings [25]. The PbP method can work well at high loadings may be due to the introduction of the loading term.

For the third parameter σ, the proposed method shows good agreement with the L^2-error of 0.104. Traditionally, it has been argued that σ is the most intangible of four parameters. This parameter accounts for the nonlinear part of the isotherm, thereby the determination of σ requires LGEs at high loadings [17], or some additional breakthrough experiments [46].

For the last parameter k_{kin}, it is the only parameter to be calculated by IM with elution profiles. An acceptable L^2-error is 0.032, due to the reasonable estimation of the former three parameters. This parameter should be considered in EDM, but it can be elim-

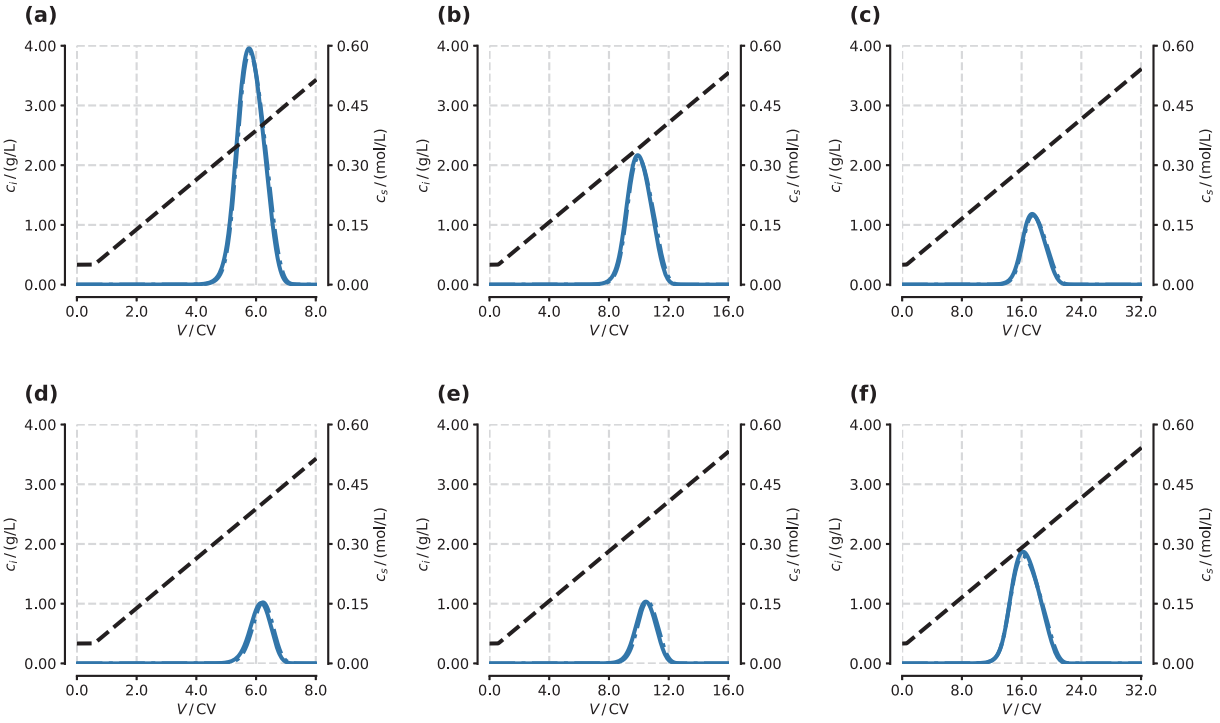

Fig. 2. Linear gradient elution curves of numerical experiments (dotted-dashed lines) and model simulation with the PbP method (solid lines). Dashed lines: salt gradients at the column outlet. (a) 8 CV gradient and 5% LF_{col}, (b) 16 CV gradient and 5% LF_{col}, (c) 32 CV gradient and 5% LF_{col}, (d) 8 CV gradient and 1% LF_{col}, (e) 16 CV gradient and 2% LF_{col}, (f) 32 CV gradient and 10% LF_{col}.

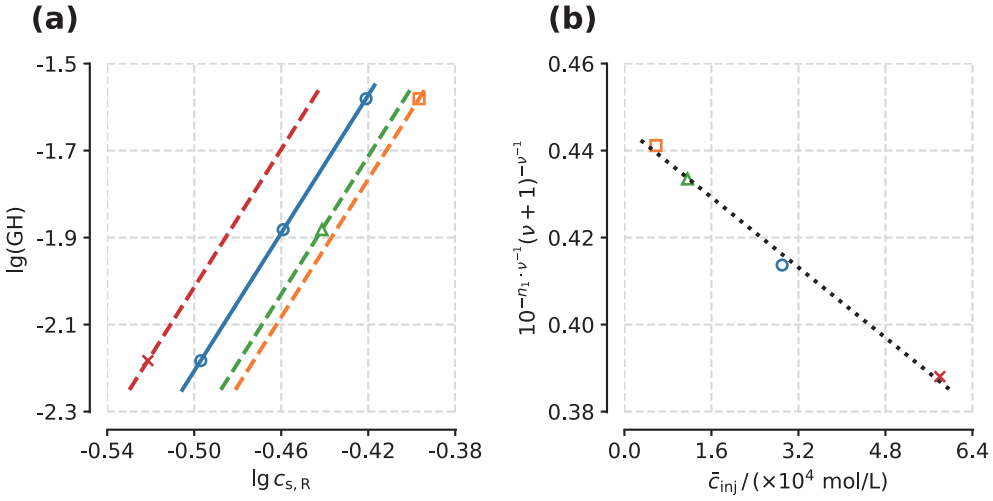

Fig. 3. First linear regression (a) and second linear regression (b) of numerical experiments. Solid line: LGEs with LF_{col} of 5% to estimate ν by LR1; dashed lines: LGEs with different loadings to calculate the intercepts by the point-slope formula; dotted lines: LR2 to determine k_{eq} and α. LF_{col}: □: 1%, Δ: 2%, ○: 5%, ×: 10%.

inated if the column model is the lumped kinetic model (LKM), where the mass transfer resistance in the solid phase is limited.

Above all, the feasibility of one-by-one estimation of the SMA parameters is verified. The proposed method is conducive to reducing the number of parameters with IM as much as possible and enhancing our physical understanding of the SMA parameters. Elution profiles with the parameters estimated by PbP method are compared in Fig. 2. The excellent agreement can be found between experiments and model simulations, both in terms of peak height and peak width. The retention predicted by the proposed method is slightly earlier than that of numerical experiments, which may be accounted for a reduced equilibrium coefficient k_{eq}. To a great extent, the elution curve prediction depends on the calibrated parameters.

4.1.2. Retention prediction of numerical experiment

The retention can be calculated directly by the retention model of the PbP method rather than complete elution profiles by solving the mechanistic models of IEC. That is, Eq. (30) can be employed to predict the retention easily by a combination of the operating conditions (GH and $\bar{c}_{inj}$) and the parameters (ν, k_{eq} and a). The same parameters except for a are used to measure the retention by the Yamamoto method with Eq. (35). The results are shown in Fig. 4.

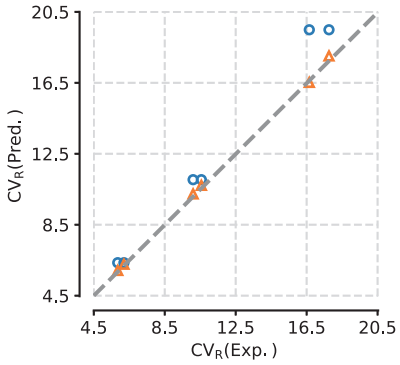

Fig. 4. A comparison of the PbP (Δ) and Yamamoto (o) method for retention prediction of numerical experiments.

The retentions predicted by the PbP method are in excellent agreement (L^2-error is 0.013) with ground truths in all tested gradient lengths and loading conditions, although there are some subtle differences. At high loadings, the proposed method can also predict consistent retention, owing to the introduction of the loading term. The slight systematic biases might be related to the approach to determining the retention in the present work. The exponentially modified-Gaussian (EMG) functions [47] may be a better choice than the moment analysis [26].

However, the Yamamoto method shows a high L^2-error of 0.124, which predicts obviously later elution. Moreover, the predicted retentions are unaffected by loadings. It is because that the Yamamoto method has no regarding loadings. The disparity becomes more significant with the increase of loadings. Therefore, the Yamamoto method cannot provide an adequate representation of retention, especially for high loadings. The most likely reason is that the simplification condition of Eq. (31) is an infinitesimal loading, resulting in the retention prediction with the Yamamoto method is only suitable under diluted conditions.

The comparison of two approaches reveals that the retention prediction with the PbP method is superior to the Yamamoto method. These disparities can be explained by the formula of Eqs. (30) and (35). The only difference between two equations is that the ionic capacity of the developed method is always less than that of the Yamamoto method due to the introduction of the loading term. That is, the ionic capacity of the developed method is operating capacity, while that of the Yamamoto method is total capacity. Under diluted conditions, the operating capacity is equal to the total capacity, which means that the Yamamoto method is the special case of the PbP method when the loading approaches zero. The introduction of the loading term can account for the reduction of ionic capacity caused by the increase of loadings. The detailed impacts of loadings on these two methods will be discussed in the following section.

4.2. Effect of loading on the performance of parameter-by-parameter method

There are many assumptions and approximations during the development of the PbP method, especially for loadings, which have a substantial impact on the method's performance. In the above cases, the developed method has been verified at specified loadings. To confirm the feasibility and robustness of the developed method at different loadings, the effects of loadings would also be investigated by numerical experiments. In this section, all column-specific and SMA parameters are consistent with those in numerical experiments, but LF_{col} and the number of LGEs are different. The LF_{col} increases from 0.2 to 51.2% taking a power of 2 exponentially. The column binding capacity of 51.2% is sufficiently representative of the upper bound because the classical SMA isotherm cannot comprehensively represent nonlinear adsorption behavior as explained above. At least two different loadings are required in the PbP method, while one in the Yamamoto method. Therefore, the L^2-error of the former method is visualized by contour maps, while that of the latter method is represented by line charts. On this contour map, the x-axis is the first loading, and the y-axis is the second loading. The shape of this map is upper-triangular since the two loadings are invertible.

4.2.1. Effect of loading on parameter estimation

The effects of loadings on the parameter estimation with the Yamamoto method are shown in Fig. 5. For the parameter ν (Fig. 5a), the estimation is approximately equal to the ground truth at different loadings. The results indicate that the Yamamoto method is suitable to estimate ν precisely over a wide range of loadings. For the parameter k_{eq} (Fig. 5b), the L^2-error increases rapidly from 0.078 to 0.906 with the increase of LF_{col} from 0.2 to 51.2%. When LF_{col} is less than 0.4%, the L^2-error can be kept below 0.100 and the estimation of k_{eq} is reasonable, which corroborates the findings in the literatures [15,20]. If LF_{col} increases to 6.4%, a high L^2-error beyond 0.5 would be found. A possible explanation may be the lack of an adequate term for the description

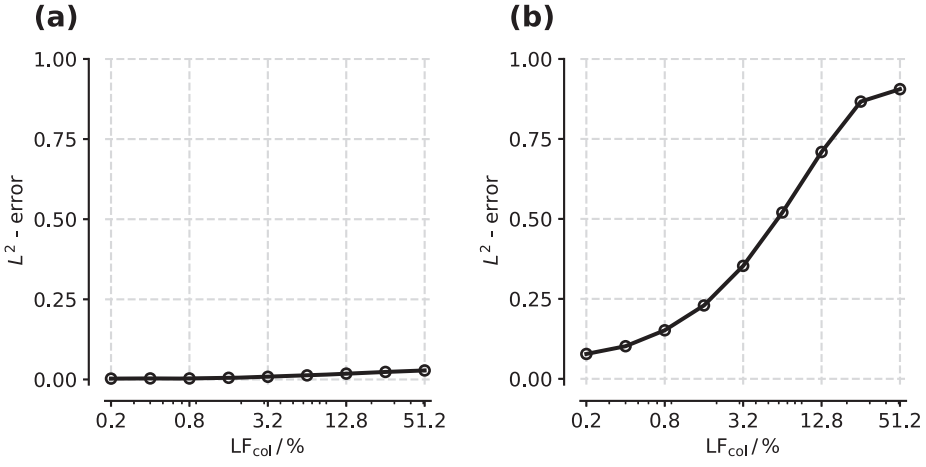

Fig. 5. Effects of loadings on the estimation of ν (a) and k_{eq} (b) with the Yamamoto method.

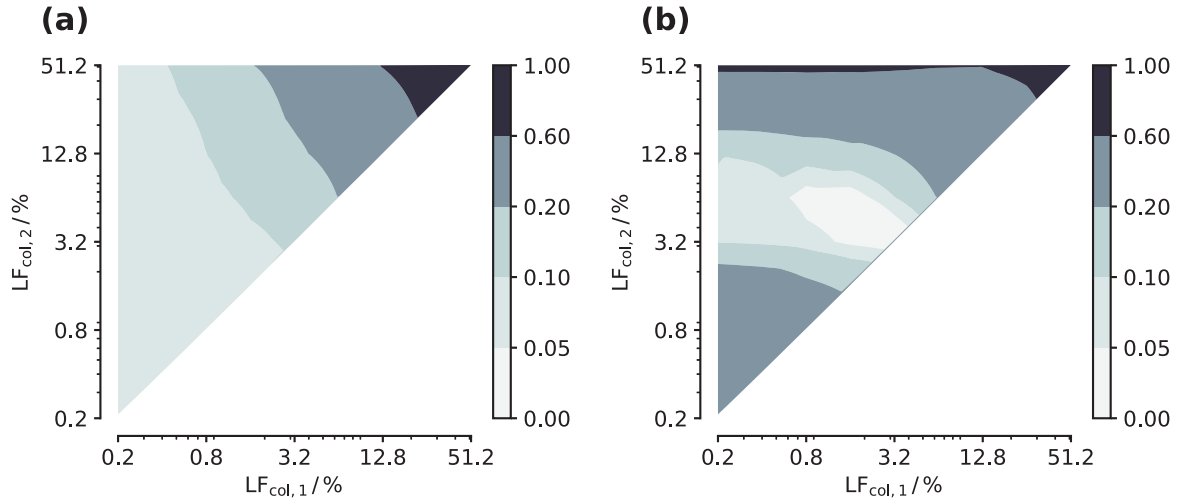

Fig. 6. Effects of loadings on the estimation of k_{eq} (a) and σ (b) with the PbP method. Colormap: visualization of L^2-error; x-axis: first loading; y-axis: second loading.

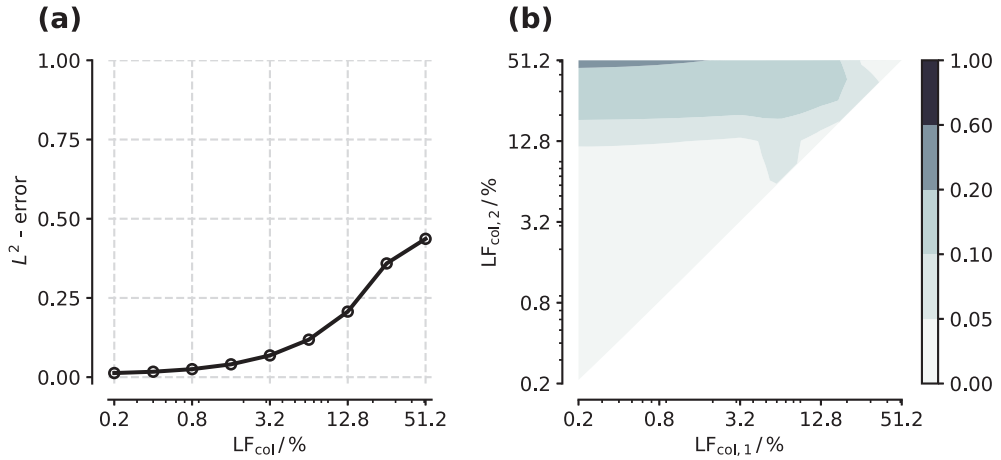

Fig. 7. Effects of loadings on the retention prediction with the Yamamoto (a) and PbP (b) method. Colormap: visualization of L^2-error; x-axis: first loading; y-axis: second loading.

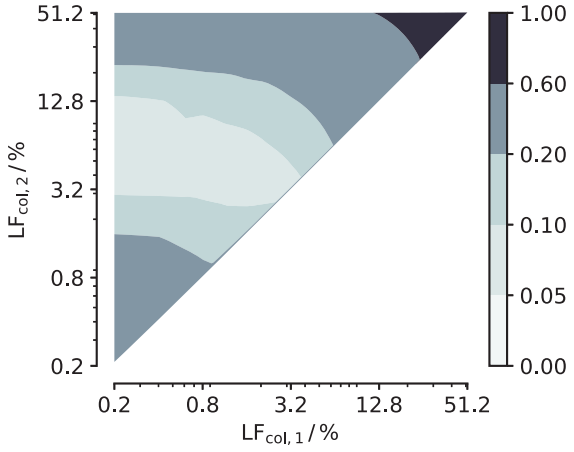

Fig. 8. Comprehensive effects of loadings on the estimation of both k_{eq} and σ with the PbP method. Colormap: visualization of L^2-error; x-axis: first loading; y-axis: second loading.

of varying loadings in the Yamamoto method. Additionally, the Yamamoto method comes from the SDM, which cannot be directly transplanted to the nonlinear region of the SMA isotherm.

The effects of loadings on the parameter estimation with the PbP method are presented in Fig. 6.

For the parameter ν, the derivation and assumption of the PbP method are similar to the Yamamoto method as previously explained, so the effects of loadings on the estimation of ν are same as shown in Fig. 5a.

For the parameter k_{eq} (Fig. 6a), all loadings are literally fulfilled with the L^2-error less than 0.6, while only LF_{col} less than 7.5% roughly are satisfied for the Yamamoto method. The results indicate that the PbP method broadens the application range. It is found that LR2 is tendency to be valid at low loadings corresponding to the assumption of the retention model in Eq. (8).

For the parameter σ (Fig. 6b), a high L^2-error is found at bottom left of the contour map ($LF_{col,1}$ and $LF_{col,2}$ are less than 1.6%), which means that it is difficult to estimate reasonably the parameter σ at low loadings compared to the above linear parameters (ν and k_{eq}). It is interesting to find that there is a narrow band ($LF_{col,1}$<6.4% and 2.0%<$LF_{col,2}$<20%) on the center of the contour map with very low L^2-error where σ can be estimated precisely. This desired loading condition may be the transition region or nonlinear region of the isotherm. Actually, the parameter σ accounts for the nonlinear adsorption. Therefore, the determination of σ may require high loadings, which are consistent with literature reports [15]. On the contrary, a high L^2-error about 0.6 distributes on

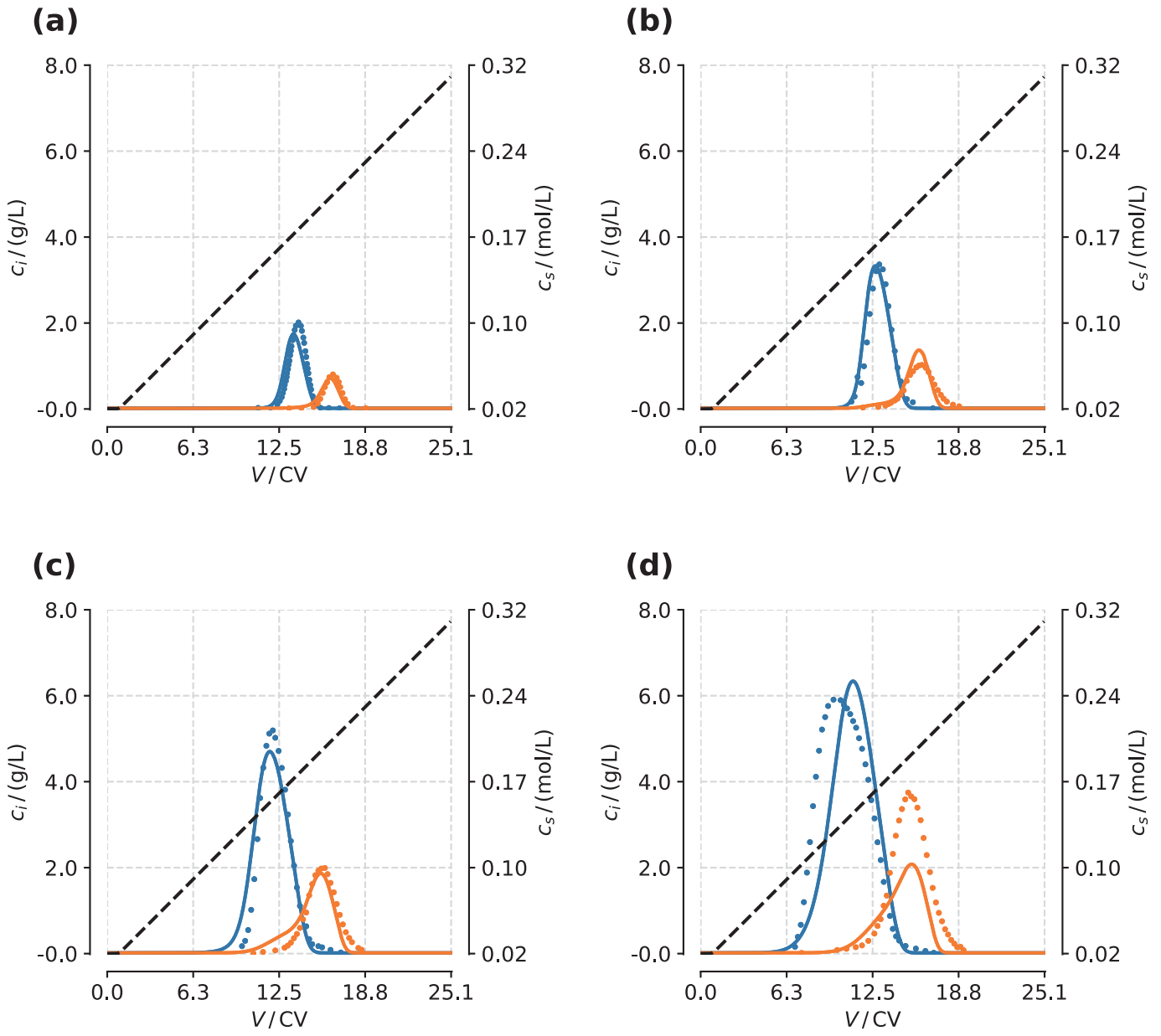

Fig. 9. Linear gradient elution curves of real-world experiments (dotted lines) and model simulation with the PbP method (solid lines). Dashed lines: salt gradients at the column outlet. Loadings: (a) 4, (b) 10, (c) 20, and (d) 40 g/L of column.

the top right, where the loadings are quite high. This discrepancy might be attributed to that the developed retention model cannot be valid at high loadings. Therefore, it is important to find an appropriate loading to estimate σ as accurately as possible.

4.2.2. Effect of loading on retention prediction

The results of the effects of loadings on the retention prediction are shown in Fig. 7.

For the Yamamoto method (Fig. 7a), a remarkable increase of L^2-error on the retention prediction is found as the loading increases. This disagreement between prediction and experiments can be explained by a straight transplant from SDM to SMA. The Yamamoto method is better for the SDM calibration rather than SMA, and it has to be adjusted before being used for SMA.

For the PbP method (Fig. 7b), the retention prediction is quite impressive. In general, the L^2-error under all conditions are less than 0.2, demonstrating that the proposed method has pinpoint accuracy and high robustness for the retention prediction. An area of very low L^2-error (<0.05) can be found at loading factors less than 12.8%. The proposed method has superior performance in the linear region of isotherm, which is due to multiple linear assumptions during the derivation of LR1, LR2 and LA.

In general, the proposed method in the present work has more application prospects because it works at both low and high loadings, while the Yamamoto method cannot be extrapolated to all conditions.

4.3. Experiment design for parameter-by-parameter method

Based on the above analysis, the loading is an important factor for the PbP method. The estimation of k_{eq} and σ have different requirements for the loading condition. It is necessary to consider

Table 5
SMA parameters of real-world experiments estimated by the PbP method.

Parameter	Unit	Monomer	Dimer
$\nu^{\dagger}$		10.2	14.8
k_{eq}	–	3.48	13.91
σ	–	30.4	62.3
k_{kin}	sM^{ν}	2.51E-07	9.67E-10
L^2-error	–	0.314	

†: Parameters are taken from Reck et al. [40].

an optimal loading to apply the PbP method. Therefore, a total L^2-error containing the effects of loadings on the estimation of k_{eq} and σ is shown in Fig. 8, which also means a stacking chart built on Fig. 6a and b. This contour map is similar to Fig. 6b, indicating that the loading for determining σ should be considered firstly.

From this chart, a straightforward experimental strategy is proposed for better application and generalization of this method with five LGEs as follows: three LGEs in different gradient lengths to estimate ν (fitting to LR1); two LGEs at higher loadings to determine k_{eq} and σ (fitting to LR2 and solving LA); elution profiles of five LGEs to correlate k_{kin} by IM.

For the first three LGEs, a loading factor less than 3.2% is recommended considering the sample consumption in calibration experiments and symmetric peaks to the benefit of determining retention time easily, although LR1 is not limited by the loading. It is advisable to conduct at least two LGEs at loadings higher than the first three LGEs for LR1 to accurately estimate k_{eq} and σ. The first loading factor is also less than 3.2% but it should be different from the above three LGEs. The second loading is kept above 3.2% but less than 12.8%. If the error of the calculated results is tiny enough, the calculation will be utilized as a determined value instead of an initial guess of the deterministic algorithm. In this way, the parameters determined by IM will be reduced from four to one. Compared to the previous study [15], the present method requires only one additional experiment at high loadings, but all of four SMA parameters can be estimated one at a time in sequence with reasonable accuracy.

4.4. Evaluating parameter-by-parameter method by real-world experiment

In the above sections the PbP method has been successfully confirmed by numerical experiments, a real-world experiment on the separation of mAb and dimer was considered further in this section. Four elution profiles taken from Creasy et al. [25] are summarized in Fig. 9. Noted that the injection area in Fig. 9d is little higher than the integral area of monomer and dimer in gradient elution, which means that there were some protein losses in the injection phase.

4.4.1. Parameter estimation of real-world experiment

All computational procedures are similar to those of numerical experiments. Elution curves approach symmetric and the fraction data is almost continuous in this case, so $c_{s,R}$ can be determined by the peak maximum for simplicity. Note that the values of monomer (ν_M =10.2) and dimer (ν_D =14.8) from Reck et al. [40] were directly adopted in this work. Perfect positive correlations of LR2 are found as presented in Fig. 10. All of four estimated parameters are summarized in Table 5. In general, the one-by-one estimation of the SMA parameters in real-world multi-component experiments is feasible and effective.

The simulated elution profiles based on the estimated parameters are compared in Fig. 10. Overall, a good agreement with the L^2-error of 0.314 is found between experiments and model simulations. These results are contrary to that of Creasy et al. [25] who

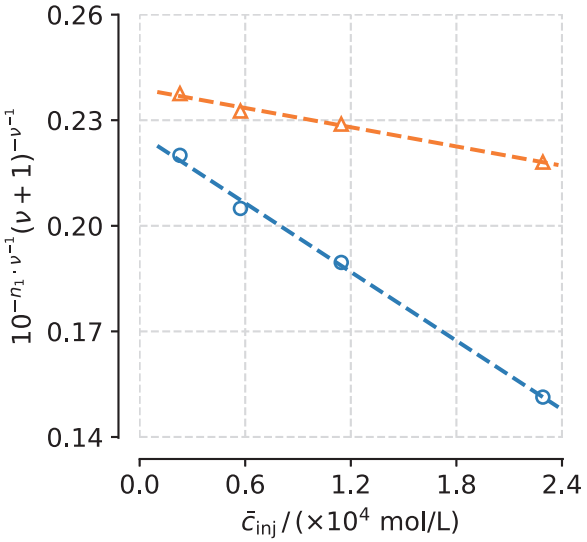

Fig. 10. Second linear regression of monomer (o) and dimer (Δ) of mAb.

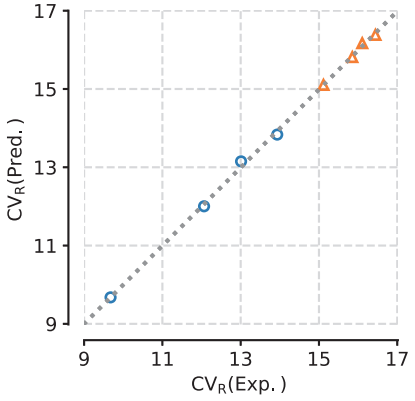

Fig. 11. A comparison of retention prediction of monomer (o) and dimer (Δ) of mAb with the PbP method.

argued the SMA model could not accurately predict the elution behavior of monomer and dimer on the same dataset. The cause might be the unreasonable SMA parameters estimated by the Yamamoto method. A smaller k_{eq} in their work would make proteins elute earlier. Most importantly, the results of σ produced by IM are roughly seven times as those in the present work, making the predictions more tailing and highly asymmetrical than experiments. The results demonstrate that the PbP method in this work may be more tenable.

As shown in Fig. 10, it could be found that the differences between model simulations and experiments become more significant at high loadings, especially for dimer. Although the predicted retention accords with the experiments, there is a disparity in peak shape at loading of 40 g/L column. This discrepancy could be attributed to an oversimplified model. As for the column model, EDM is not sophisticated enough to provide a comprehensive description of IEC when the mass transfer residence cannot be ignorable [29]. Therefore, the transport dispersive model or general rate model may be a better choice for high loadings. As for the isotherm model, the classical SMA isotherm may not provide an accurate enough description of nonlinear adsorption behavior reported by a large number of published studies [40,48,49]. The

modified SMA like self-association SMA [50] or multi-state SMA model [51] may be more appropriate for overloading conditions.

4.4.2. Retention prediction of real-world experiment

The method as stated in the section of numerical experiments is applied in the real-world case as well. An excellent agreement (Fig. 11) between model predictions and experiments can be found with the L^2-error of 0.005 at all loadings. The results indicate that the developed retention model can provide an adequate representation of retention in this case. The imperceptible errors might be due to the determined approach of the retention in this work. The suitable methods to calculate retention, including peak maximum, moment analysis, EMG functions and so on, will be explored in our further investigation to improve the performance of the PbP method.

5. Conclusions

A PbP method was proposed to estimate the SMA parameters and predict the retention with LGEs of IEC, which is based on the linear assumptions and approximate derivations with in-depth mechanistic understanding of the chromatographic model. Numerical experiments of six LGEs with varying gradient lengths and loadings were used firstly to evaluate the feasibility of parameter estimation with the proposed method compared to the Yamamoto method. Three parameters (ν, k_{eq}, and σ) can be estimated step-by-step with the proposed method, reducing the number of undetermined parameters from four to one. In contrast, only two parameters (ν and k_{eq}) can be calibrated by the Yamamoto method. Two methods obtain the same ν, but a significantly different k_{eq}. The proposed method can estimate k_{eq} with high accuracy, whose error is seven times lower than that of the Yamamoto method. The nonlinear parameter σ can also be calculated with a reasonable precision directly by elution curves without any additional protein breakthrough experiment. As for the retention prediction, the developed method is suitable in all gradient lengths and loading conditions, while the Yamamoto method can only be applied under diluted conditions. Moreover, the effects of loadings on the robustness of two methods for parameter estimation and retention prediction were assessed. The results indicated that the PbP method was more reliable than the Yamamoto method under undiluted conditions. An experiment design strategy based on the above analysis was proposed for more efficient implementation of the PbP method.

In addition, real-world experiments on the separation of mAb monomer-dimer mixtures in different gradient lengths and loading conditions were used to further confirm the feasibility of the PbP method in multi-component system. The results demonstrated that the PbP method developed in the present work can not only reliably estimate the SMA parameters with high precision but also accurately predict retention in both numerical single-component and real-world multi-component experiments. The following studies would be focused on the applications of the PbP method for the modeling and optimization of IEC.

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.

CRediT authorship contribution statement

Yu-Cheng Chen: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Shan-Jing Yao:** Writing – review & editing. **Dong-Qiang Lin:** Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing.

Acknowledgments

This work was supported by the National Key R&D Program of China (2021YFE0113300) and National Natural Science Foundation of China (22078286).

References

[1] H. Narayanan, M. Sponchioni, M. Morbidelli, Integration and digitalization in the manufacturing of therapeutic proteins, Chem. Eng. Sci. 248 (2022), doi:10.1016/j.ces.2021.117159.
[2] B.K. Nfor, P. Verhaert, L.A.M. van der Wielen, J. Hubbuch, M. Ottens, Rational and systematic protein purification process development: the next generation, Trends Biotechnol. 27 (2009) 673–679, doi:10.1016/j.tibtech.2009.09.002.
[3] C. Shi, Z.Y. Gao, Q.L. Zhang, S.J. Yao, N.K.H. Slater, D.Q. Lin, Model-based process development of continuous chromatography for antibody capture: a case study with twin-column system, J. Chromatogr. A 1619 (2020), doi:10.1016/j.chroma.2020.460936.
[4] Y.N. Sun, C. Shi, Q.L. Zhang, S.J. Yao, N.K.H. Slater, D.Q. Lin, Model-based process development and evaluation of twin-column continuous capture processes with protein A affinity resin, J. Chromatogr. A 1625 (2020), doi:10.1016/j.chroma.2020.461300.
[5] Y.N. Sun, C. Shi, Q.L. Zhang, N.K.H. Slater, A. Jungbauer, S.J. Yao, D.Q. Lin, Comparison of protein A affinity resins for twin-column continuous capture processes: process performance and resin characteristics, J. Chromatogr. A 1654 (2021), doi:10.1016/j.chroma.2021.462454.
[6] C. Shi, B. Jiao, X.J. Chen, R. Chen, W. Gong, S.J. Yao, D.Q. Lin, Process development and optimization of continuous capture with three-column periodic counter-current chromatography, Biotechnol. Bioeng. 118 (2021) 3313–3322, doi:10.1002/bit.27689.
[7] C. Shi, Q.L. Zhang, S. Vogg, D.Q. Lin, M. Sponchioni, M. Morbidelli, Analysis and optimal design of batch and two-column continuous chromatographic frontal processes for monoclonal antibody purification, Biotechnol. Bioeng. 118 (2021) 3420–3434, doi:10.1002/bit.27763.
[8] D.J. Karst, F. Steinebach, M. Soos, M. Morbidelli, Process performance and product quality in an integrated continuous antibody production process, Biotechnol. Bioeng. 114 (2017) 298–307, doi:10.1002/bit.26069.
[9] C.A. Brooks, S.M. Cramer, Steric mass-action ion-exchange - displacement profiles and induced salt gradients, AIChE J. 38 (1992) 1969–1978, doi:10.1002/aic.690381212.
[10] 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, doi:10.1002/bit.260250605.
[11] S. Yamamoto, M. Nomura, Y. Sano, Adsorption chromatography of proteins - determination of optimum conditions, AIChE J. 33 (1987) 1426–1434, doi:10.1002/aic.690330903.
[12] L. Pedersen, J. Mollerup, E. Hansen, A. Jungbauer, Whey proteins as a model system for chromatographic separation of proteins, J. Chromatogr. B 790 (2003) 161–173, doi:10.1016/s1570-0232(03)00127-2.
[13] M. Rudt, F. Gillet, S. Heege, J. Hitzler, B. Kalbfuss, B. Guelat, Combined Yamamoto approach for simultaneous estimation of adsorption isotherm and kinetic parameters in ion-exchange chromatography, J. Chromatogr. A 1413 (2015) 68–76, doi:10.1016/j.chroma.2015.08.025.
[14] D. Saleh, R. Hess, M. Ahlers-Hesse, N. Beckert, M. Schonberger, F. Rischawy, G. Wang, J. Bauer, M. Blech, S. Kluters, J. Studts, J. Hubbuch, Modeling the impact of amino acid substitution in a monoclonal antibody on cation exchange chromatography, Biotechnol. Bioeng. (2020) 11, doi:10.1002/bit.27798.
[15] D. Saleh, G. Wang, B. Muller, F. Rischawy, S. Kluters, J. Studts, J. Hubbuch, Straightforward method for calibration of mechanistic cation exchange chromatography models for industrial applications, Biotechnol. Prog. 36 (2020) 12, doi:10.1002/btpr.2984.
[16] T. Vicente, R. Fabar, P.M. Alves, M.J.T. Carrondo, J.P.B. Mota, Impact of ligand density on the optimization of ion-exchange membrane chromatography for viral vector purification, Biotechnol. Bioeng. 108 (2011) 1347–1359, doi:10.1002/bit.23058.
[17] V. Natarajan, S. Ghose, S.M. Cramer, Comparison of linear gradient and displacement separations in ion-exchange systems, Biotechnol. Bioeng. 78 (2002) 365–375, doi:10.1002/bit.10231.
[18] W. Heymann, J. Glaser, F. Schlegel, W. Johnson, P. Rolandi, E. von Lieres, Advanced score system and automated search strategies for parameter estimation in mechanistic chromatography modeling, J. Chromatogr. A 1661 (2022), doi:10.1016/j.chroma.2021.462693.
[19] S. Yamamoto, Plate height determination for gradient elution chromatography of proteins, Biotechnol. Bioeng. 48 (1995) 444–451, doi:10.1002/bit.260480506.
[20] W.R. Keller, S.T. Evans, G. Ferreira, D. Robbins, S.M. Cramer, Use of Mini-Columns for linear isotherm parameter estimation and prediction of benchtop column performance, J. Chromatogr. A 1418 (2015) 94–102, doi:10.1016/j.chroma.2015.09.038.

[21] D. Saleh, G. Wang, B. Mueller, F. Rischawy, S. Kluters, J. Studts, J. Hubbuch, Cross-scale quality assessment of a mechanistic cation exchange chromatography model, Biotechnol. Prog. 37 (2021) 13, doi:10.1002/btpr.3081.
[22] D. Saleh, G. Wang, F. Rischawy, S. Kluters, J. Studts, J. Hubbuch, In silico process characterization for biopharmaceutical development following the quality by design concept, Biotechnol. Prog. 37 (2021) e3196, doi:10.1002/btpr.3196.
[23] T. Ishihara, T. Kadoya, H. Yoshida, T. Tamada, S. Yamamoto, Rational methods for predicting human monoclonal antibodies retention in protein A affinity chromatography and cation exchange chromatography - structure-based chromatography design for monoclonal antibodies, J. Chromatogr. A 1093 (2005) 126–138, doi:10.1016/j.chroma.2005.07.077.
[24] T. Ishihara, S. Yamamoto, Optimization of monoclonal antibody purification by ion-exchange chromatography - application of simple methods with linear gradient elution experimental data, J. Chromatogr. A 1069 (2005) 99–106, doi:10.1016/j.chroma.2004.10.040.
[25] A. Creasy, J. Reck, T. Pabst, A. Hunter, G. Barker, G. Carta, Systematic interpolation method predicts antibody monomer-dimer separation by gradient elution chromatography at high protein loads, Biotechnol. J. 14 (2019), doi:10.1002/biot.201800132.
[26] W.R. Keller, S.T. Evans, G. Ferreira, D. Robbins, S.M. Cramer, Understanding the effects of system differences for parameter estimation and scale-up of high throughput chromatographic data, J. Chromatogr. A 1661 (2022) 462696, doi:10.1016/j.chroma.2021.462696.
[27] M.Z. Elfallah, G. Guiochon, Prediction of a protein band profile in preparative reversed-phase gradient elution chromatography, Biotechnol. Bioeng. 39 (1992) 877–885, doi:10.1002/bit.260390810.
[28] L.J. Benedini, D. Figueiredo, J. Cabrera-Crespo, V.M. Goncalves, G.G. Silva, G. Campani, T.C. Zangirolami, F.F. Furlan, Modeling and simulation of anion exchange chromatography for purification of proteins in complex mixtures, J. Chromatogr. A 1613 (2020), doi:10.1016/j.chroma.2019.460685.
[29] D.E. Cherrak, S. Khattabi, G. Guiochon, Adsorption behavior and prediction of the band profiles of the enantiomers of 3-chloro-1-phenyl-1-propanol - influence of the mass transfer kinetics, J. Chromatogr. A 877 (2000) 109–122, doi:10.1016/s0021-9673(00)00189-8.
[30] A. Seidel-Morgenstern, H. Schmidt-Traub, M. Schulte, A. Seidel-Morgenstern, Modeling of chromatographic processes, in: Preparative Chromatography, Wiley, 2020, pp. 311–354.
[31] T.C. Huuk, T. Hahn, A. Osberghaus, J. Hubbuch, Model-based integrated optimization and evaluation of a multi-step ion exchange chromatography, Sep. Sci. Technol. 136 (2014) 207–222, doi:10.1016/j.seppur.2014.09.012.
[32] 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, doi:10.1002/elsc.201400247.
[33] S.R. Gallant, A. Kundu, S.M. Cramer, Modeling nonlinear elution of proteins in ion-exchange chromatography, J. Chromatogr. A 702 (1995) 125–142, doi:10.1016/0021-9673(94)00992-i.
[34] E.S. Parente, D.B. Wetlaufer, Relationship between isocratic and gradient retention times in the high-performance ion-exchange chromatography of proteins - theory and experiment, J. Chromatogr. 355 (1986) 29–40, doi:10.1016/s0021-9673(01)97301-7.
[35] K. Meyer, J.K. Huusom, J. Abildskov, High-order approximation of chromatographic models using a nodal discontinuous Galerkin approach, Comput. Chem. Eng. 109 (2018) 68–76, doi:10.1016/j.compchemeng.2017.10.023.
[36] K. Meyer, S. Leweke, E. von Lieres, J.K. Huusom, J. Abildskov, ChromaTech: a discontinuous Galerkin spectral element simulator for preparative liquid chromatography, Comput. Chem. Eng. 141 (2020), doi:10.1016/j.compchemeng.2020.107012.
[37] S.V.D. Walt, S.C. Colbert, G. Varoquaux, The NumPy array: a structure for efficient numerical computation, Comput. Sci. Eng. 13 (2011) 22–30, doi:10.1109/MCSE.2011.37.
[38] 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, 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 Miranda 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, C. SciPy, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020) 261–272, doi:10.1038/s41592-019-0686-2.
[39] J.D. Hunter, Matplotlib: a 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95, doi:10.1109/MCSE.2007.55.
[40] J.M. Reck, T.M. Pabst, A.K. Hunter, X.Y. Wang, G. Carta, Adsorption equilibrium and kinetics of monomer-dimer monoclonal antibody mixtures on a cation exchange resin, J. Chromatogr. A 1402 (2015) 46–59, doi:10.1016/j.chroma.2015.05.007.
[41] Y. Zhang, G.L. Lin, P. Forssen, M. Gulliksson, T. Fornstedt, X.L. Cheng, A regularization method for the reconstruction of adsorption isotherms in liquid chromatography, Inverse Probl. 32 (2016), doi:10.1088/0266-5611/32/10/105005.
[42] S. Osterroth, P. Menstell, A. Schwammle, J. Ohser, K. Steiner, Adjoint optimization for the general rate model of liquid chromatography, Comput. Chem. Eng. 133 (2020) 15, doi:10.1016/j.compchemeng.2019.106657.
[43] K. Kaczmarski, On the optimization of the solid core radius of superficially porous particles for finite adsorption rate, J. Chromatogr. A 1218 (2011) 951–958, doi:10.1016/j.chroma.2010.12.093.
[44] S. Yamamoto, K. Nakanishi, R. Matsuno, T. Kamikubo, Ion exchange chromatography of proteins-prediction of elution curves and operating conditions. II. Experimental verification, Biotechnol. Bioeng. 25 (1983) 1373–1391, doi:10.1002/bit.260250516.
[45] S.R. Gallant, S. Vunnum, S.M. Cramer, Optimization of preparative ion-exchange chromatography of proteins: linear gradient separations, J. Chromatogr. A 725 (1996) 295–314, doi:10.1016/0021-9673(95)00909-4.
[46] J. Morgenstern, G. Wang, P. Baumann, J. Hubbuch, Model-based investigation on the mass transfer and adsorption mechanisms of mono-pegylated lysozyme in ion-exchange chromatography, Biotechnol. J. 12 (2017), doi:10.1002/biot.201700255.
[47] E. Grushka, Characterization of exponentially modified Gaussian peaks in chromatography, Anal. Chem. 44 (1972) 1733 &, doi:10.1021/ac60319a011.
[48] T. Briskot, T. Hahn, T. Huuk, J. Hubbuch, Protein adsorption on ion exchange adsorbers: a comparison of a stoichiometric and non-stoichiometric modeling approach, J. Chromatogr. A 1653 (2021) 462397, doi:10.1016/j.chroma.2021.462397.
[49] T.C. Huuk, T. Hahn, K. Doninger, J. Griesbach, S. Hepbildikler, J. Hubbuch, Modeling of complex antibody elution behavior under high protein load densities in ion exchange chromatography using an asymmetric activity coefficient, Biotechnol. J. 12 (2017), doi:10.1002/biot.201600336.
[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, doi:10.1002/ceat.200800082.
[51] J. Diedrich, W. Heymann, S. Leweke, S. Hunt, R. Todd, C. Kunert, W. Johnson, E. von Lieres, Multi-state steric mass action model and case study on complex high loading behavior of mAb on ion exchange tentacle resin, J. Chromatogr. A 1525 (2017) 60–70, doi:10.1016/j.chroma.2017.09.039.