Journal of Chromatography A 1754 (2025) 466027

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Molecular insights into chromatography: Automated workflows for the virtual design of methacrylate-based chromatography resins

Tim Ballweg [a,1], Modan Liu [a,1], Ahmed Mama [a], Wolfgang Wenzel [b], Matthias Franzreb [a,*]

[a] Institute of Functional Interfaces, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
[b] Institute of Nanotechnology, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

ARTICLE INFO

Keywords:
Chromatography
Virtual design
Workflow
Molecular modeling

ABSTRACT

Computational chemistry provides invaluable insights into the behaviors and properties of various materials at the molecular level. This capability is of particular interest in chromatography where adsorbents engage with target molecules through intricate interactions. However, the broad integration of molecular simulations into the field of chromatography has been notably limited, despite significant achievements in previous studies. One potential reason is the requirement for considerable expertise to effectively configure these simulations, presenting a significant barrier to entry. In this context, workflow management systems (WMSs) provide a viable solution by allowing experts to automate complex simulation tasks, making them accessible to the wider research community without necessitating in-depth knowledge of the simulation process. This manuscript outlines the creation and application of two automated workflows designed to generate comprehensive all-atom models of methacrylate-based chromatography resin surfaces and to rapidly calculate binding free energies with peptides as target molecules. These innovations represent a significant advancement in the field by streamlining the simulation process, enhancing predictive accuracy, and making complex molecular modeling more accessible to researchers across disciplines. By publishing these workflows, we aim to catalyze molecular modeling in the field of chromatography by encouraging scientists to utilize and build upon our work.

1. Introduction

The field of chromatography is fundamental to the purification and analysis of biomolecules [1,2]. In this context, molecular simulations have emerged as a powerful tool, offering detailed insights into the interactions between biomolecules and resin surfaces at an atomic level [3–9]. However, despite their potential and continued progress, the widespread integration of these simulations into chromatographic research has been limited. A primary obstacle is the technical complexity of setting up and running these simulations, which necessitates specialized expertise, and significant computational resources [10,11]. These barriers have historically restricted the use of molecular simulations to a niche within the chromatographic community, limiting their potential impact. This situation highlights the urgent need for more accessible and user-friendly methodologies.

Addressing these challenges, we suggest the utilization of workflow management systems (WMS) as a viable solution. WMSs are sophisticated software environments designed to automate simulations and data management, especially on supercomputing platforms [12–14]. These programs enable the scheduling and parallel processing of individual tasks, as well as the flexible design of workflows [15,16]. Additionally, WMSs simplify the execution of complex simulations, offering user-friendly interfaces that allow researchers, irrespective of their computational background, to engage with advanced molecular modeling [17,18]. Thus, WMSs open up sophisticated simulation methodologies to a broader range of researchers thereby expanding their application in scientific investigations.

In this study, we showcase the use of two WMSs, SimStack [12] and KNIME [19,20], to automate complex multistage modeling workflows, aimed at streamlining molecular modeling in chromatography. These workflows build upon our previous publication in which we successfully introduced and validated a virtual design approach for generating all-atom models of methacrylate-based chromatography resin surfaces, inclusive of the backbone structure [8]. These models proved pivotal for

* Corresponding author.
E-mail address: matthias.franzreb@kit.edu (M. Franzreb).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.chroma.2025.466027
Received 8 February 2025; Received in revised form 14 April 2025; Accepted 6 May 2025
Available online 8 May 2025

accurately predicting the binding affinity of linear peptides. The first presented workflow of this study utilizes SimStack to automate the creation of bespoke all-atom resin surface models (AAM). By integrating commonly used ion-exchange ligands, short spacer molecules, and methacrylate-based monomers into our workflow, we enable users to simulate a broad spectrum of chromatography resins and modes. The second workflow, automated in KNIME, can be utilized to rapidly calculate binding poses and binding free energies between linear peptides and the previously generated AAM. Together, these automated workflows enhance the capability to simulate a wide range of biomolecular interactions, offering a comprehensive and adaptable tool.

By making these workflows publicly available, we want to encourage the chromatography community to engage with, and further develop our contributions. This initiative aims to showcase how the implementation of WMSs can accelerate the implementation of molecular simulation in the field of chromatography. By utilizing the workflows, users can design novel resin materials, more accurately predict complex interactions with various biomolecules and refine separation processes through a streamlined, "one-click" operation. This paper delves into the methodology behind the development of these automated workflows, their potential applications, and the anticipated impact on the field of chromatography. As we look to the future, we discuss the potential for these workflows to be adapted and expanded, broadening their utility across various aspects of chromatographic research and thereby solidifying molecular simulations as a cornerstone of modern chromatographic analysis. For readers seeking a more comprehensive introduction to modeling and its practical applications in chromatography, we refer to our previous work [8], which provides an overview of simulation fundamentals and their utility in chromatographic studies.

2. Material and methods

2.1. Automated workflow for the virtual design of methacrylate-based chromatography resins

In our approach to the virtual design of chromatography resins, we employ a multiscale modeling strategy that combines coarse-grained and atomistic molecular dynamics simulations to develop detailed all-atom models of methacrylate-based resin surfaces. This methodology was detailed in our preceding publication, showcasing the simulation of a commercial multimodal chromatography resin's all-atom structure, inclusive of its polymer backbone [8]. In this section, we provide an in-depth description of the simulation techniques employed, their implementation within the workflow management system SimStack, and the latest updates and enhancements, serving as a guide for scientists intending to build upon our work. Furthermore, we demonstrate how automating consecutive simulation tasks in WMSs can benefit users with limited computational expertise. By leveraging the workflow's flexible design, simulations can be independently set up through SimStack's graphical user interface (GUI) to generate virtual representations of methacrylate-based chromatography resin surfaces. In the SimStack engine, a comprehensive workflow is divided into multiple sub-components known as Workflow Active Nodes (WaNos) [12]. Each WaNo typically corresponds to a specific computational simulation software and its associated simulation protocol — for example, VASP [21] for density functional theory (DFT) calculations or LAMMPS [22] for all-atom (AA) or coarse-grained (CG) molecular dynamics simulations. Data is transferred between WaNos such that the output from a predecessor WaNo serves as the input to a successor WaNo, facilitating a streamlined simulation pipeline.

To generate molecular models of methacrylate-based chromatography resin surfaces, we first use a CG model, inspired by Sedghamiz at el [23], to generate the polymer network structure. This coarse-grained model is then converted to an all-atom representation, which is subsequently relaxed under ambient conditions to yield an atomistic polymer network of the chromatography resin. Additionally, a stand-alone input WaNo gathers critical parameters that define the building blocks of the chromatography resin, while a stand-alone output WaNo enables easy downloading of the final all-atom representation in LAMMPS datafile format and conversion to MOL2 format for interfacing with docking simulations in the Schrödinger software package. Fig. 1 illustrates the conceptual layout, a schematic representation of intermediate data, as well as a snapshot of the workflow in Simstack.

From the developer's perspective the workflow consists of five blocks, each corresponding to a WaNo:

1. The "Input_Polymer-Building-Blocks" WaNo establishes the chemical composition of the chromatography resin, which is constructed from three building blocks: methacrylate-based monomers and crosslinkers, as well as ligands. Users can select the desired molecule types and specify their quantities within the simulation box using dropdown menus and text fields. In the current version of the workflow, the monomer options are 2-hydroxyethyl methacrylate (HEMA) and dihydroxypropyl methacrylate (DHPMA), the crosslinker is ethylene glycol dimethacrylate (EGDMA), and the ligands available are tryptophan (Trp), sulfonic acid (S), and tetramethylammonium (Q). Fig. 2 displays the structures of all molecules available for selection. In our latest update, we have introduced the option to incorporate linear alkyl spacers between the monomer and the ligand, allowing users to further customize the resin's structural properties. The length of the spacer can be specified by setting the number of carbon atoms in the alkyl chain. The default number of porogen molecules in the initial binding box is defined by the combined total of the monomer and crosslinker molecules. Besides the types of building blocks, the ratios of monomer to crosslinker to ligand, as well as the absolute density of individual building blocks, are determinant factors during adsorption. For instance, altering either the total number of ligands or the simulation box dimensions allows control over the ligand density. In the current workflow, in the initial state two-thirds of the simulation box is filled with neutral, dummy porogens to generate a three-dimensional interfacial structure of the chromatography resin along the z-direction.
2. Originally developed for 3D printing materials, the "CG-Polymerization" WaNo simulates the dynamic polymerization process of a Kremer-Grest [24] type coarse-grained polymer network. A set of radical-based polymerization reactions incorporate individual monomers and predefined crosslinker into the polymer network. During the polymerization process, activation of monomers and propagation of polymer chains are homogeneously distributed across the simulation box, simulating an "always-on" exposure. Initiators and quenchers are not explicitly included in the coarse-grained model; instead, activation and quenching rates correspond to stochastic swapping between inactivated monomers and activated monomers. All reaction rates are derived from Sedghamiz at el. representing a generic methacrylate-based resin [23], and are identical for HEMA- or DHPMA-based polymer networks. After the polymerization reactions, ligands are attached to monomers through a stochastic process: bond formation reactions occur only when ligands are located close to monomers. After a fixed time, neutral porogens, unused monomers and crosslinkers, and unbound ligands are removed from the coarse-grained representation to mimic rinsing. An important set of input parameters in the "CG-Polymerization" WaNo are the simulation box dimensions. The packing densities of individual building blocks are determined by both the simulation box size and the absolute numbers of molecules specified in the "Input_Polymer-Building-Blocks" WaNo. The initial configuration of the all-atom representation is also scaled based on the coarse-grained simulation box size. For high-throughput simulations, it is practical to fix the simulation box size and modify the absolute numbers in the "Input_Polymer-Building-Blocks" WaNo, as the all-atom relaxation protocol is optimized for moderate simulation box sizes around $20 \times 20 \times 30$ nm^3. The output of the

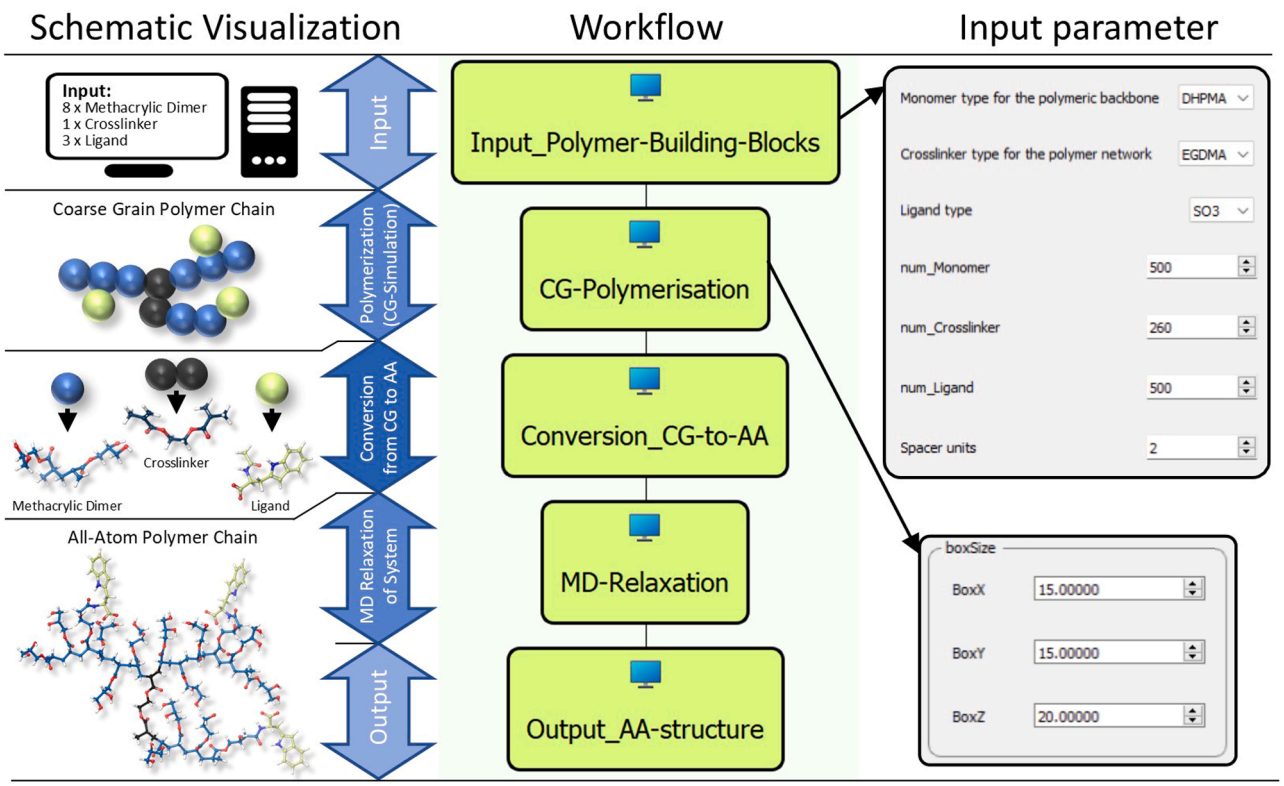

Fig. 1. Workflow for the virtual design of chromatography resins – The figure illustrates a schematic representation of the intermediate data between the WaNos and includes a snapshot of the SimStack GUI with the required used input for each Wano.

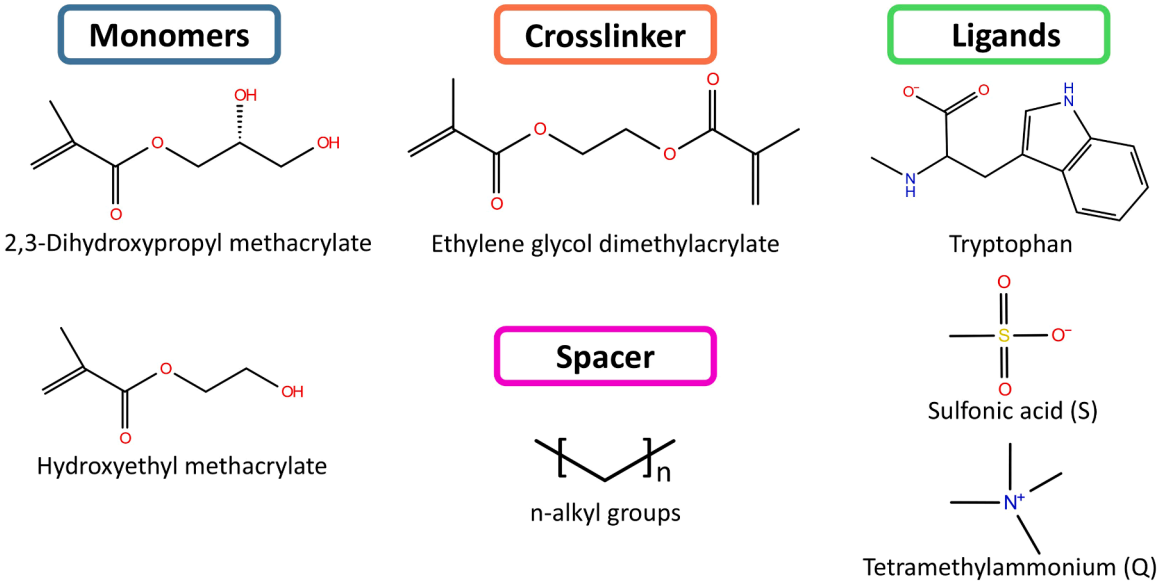

Fig. 2. Available polymer building blocks in the workflow - The figure displays the molecular structures of all selectable molecules within the current version of the workflow, organized into four categories: monomers, crosslinker, ligands, and spacer.

"CG-Polymerization" WaNo is a coarse-grained polymer network saved as a LAMMPS datafile.

3. The "Conversion_CG-to-AA" WaNo converts the coarse-grained representation to an all-atom representation. All-atom templates of the specified building blocks (e.g., HEMA and DHPMA as dimer units) are inherited from the "Input_Polymer-Building-Blocks" WaNo. The conversion process includes placing the all-atom molecules, connecting them according to the topology derived from the coarse-grained polymer network, and making corrections to force-field parameters:
   a) All building block variants have limited dimensions, typically $<1 \times 1 \times 1$ nm^3. An empty all-atom simulation box is first scaled up by 110 % based on settings from the "CG-Polymerization" WaNo to allow extra padding and avoid atomic clashes. Individual all-atom building blocks are then placed corresponding to coordinates from the coarse-grained polymer network. If spacer

units are specified between the monomer and the ligand, the linear spacer segment is placed near the ligand molecule with a (–0.5 nm, –0.5 nm, –0.5 nm) offset to prevent clashes. A short dynamic deformation of the all-atom simulation box is performed to recover the original dimensions.

b) Bonding information from the coarse-grained polymer network is used to connect all-atom building blocks. A hydrogen atom is removed from the bond formation site, and partial charges are re-equilibrated. For monomers with asymmetric bond formation sites, relative positions are used to assign "left" or "right" ends — for example, carbons labeled 1′ or 4′ are marked for connection to other building blocks. In the all-atom representation, a coarse-grained entity A with an x-coordinate smaller than entity B will have an A–B bond on the 4′-site of A and the 1′-site of B.

c) As new bonds are introduced in the all-atom model, missing bonding parameters and torsional parameters are added according to the OPLS-AA [25] force field.

After updating atomic coordinates, partial charges, molecular topology, and force-field parameters, the output of the "Conversion_CG-to-AA" WaNo is a comprehensive LAMMPS datafile serving as the initial snapshot for the "MD-Relaxation" WaNo.

4. The "MD-Relaxation" WaNo performs a molecular dynamics simulation. The advanced script mode is used to run the simulation with a predefined protocol, along with the auto-generated LAMMPS datafile from the "Conversion_CG-to-AA" WaNo. The all-atom relaxation incorporates multiple stages, with the simulation timestep incrementally adapted to prevent numerical divergence due to atomic clashes.
5. The "Output_AA-strucutre" WaNo saves the relaxed polymer network structure of the all-atom representation from the "MD-Relaxation" WaNo as a LAMMPS datafile. By utilizing bonding information from the "Conversion_CG-to-AA" WaNo, bond types such as tetrahedral C–C, aromatic C–C, C = C, and C = O are assigned during the conversion of the LAMMPS data to MOL2 file format. The MOL2 file is then used as input for the Schrödinger software suite for automated binding pose sampling and binding energy calculations.

This workflow has been utilized in a previous publication [8], and we are continuously adding new features, such as additional building blocks, incorporation of spacer units, and optimization of relaxation protocols. From the user's perspective, generating a single polymer network structure involves selecting the molecule types for each building block via drop-down menus, specifying the number of molecules within the simulation box (default dimensions of $15 \times 15 \times 30$ nm^3), and executing the workflow. For high-throughput simulations, SimStack enables users to employ for-loops that read parameter lists from YAML-based configurations, assigning them to the "Input_Polymer-Building-Blocks" WaNo. As the workflow processes each parameter set, it generates a corresponding MOL2 datafile, which can subsequently be used in binding energy screenings. If customized molecules are required for the building blocks, the "Input_Polymer-Building-Blocks" WaNo needs to be replaced, and the workflow starts from the "CG-Polymerization" WaNo. Individual setups for monomer, crosslinker, and ligand structures can be specified in the "CG-Polymerization" WaNo and the "Conversion_CG-to-AA" WaNo. Additionally, a concatenated manifest including all force-field parameters for the three components is needed in the "Conversion_CG-to-AA" WaNo. The workflow can be downloaded from the following Github repository:

https://github.com/Ahmedkhalil-Mama/Automated-Workflows-for-All-Atom-Resin-and-Peptide-Adsorption-Modeling with a complete example data set located in:

https://bwsyncandshare.kit.edu/s/Zt4Lei3AjPxxrDi

All required software, along with the specific software versions and hardware employed to run the workflow, are listed in the Software section of the Supplementary Information.

2.2. Use case: automation of binding pose screenings and binding free energy calculations

One practical application of our virtually designed chromatography resins involves calculating binding free energies (ΔG) with target molecules. In our previous work, we established the utility of these ΔG calculations for the accurate prediction of Langmuir constants (K_L) for linear peptides. This section outlines the automation of these intricate simulation sequences and data management steps within the WMS KNIME, highlighting the latest enhancements that introduce the capability to process multiple adsorbent structures and target molecules, thereby enhancing the workflow's versatility. For an in-depth rationale behind selecting these specific simulation steps for this use case, we direct readers to our preceding publication [8].

Fig. 3a presents a flowchart of the automated process, delineating all essential steps required for the rapid computation of binding energies with peptides as target molecules. The diagram uses red to denote input and output nodes, orange for nodes involving tasks for molecular simulation software, and yellow for nodes dedicated to data management. The percentile difference (PD) between the binding free energy from MM-GBSA calculations ($dG_{MM\text{-}GBSA}$) [26] and the binding free energy from GFN2-xTB single-point energy calculations ($dG_{GFN2\text{-}xTB}$) [27] was used to determine possible false positive binding poses. KNIME was chosen for this automation since it provides, with the Schrödinger extensions, pre-built nodes for several key simulation steps, including Glide LigPrep, Glide Docking, MacroModel Minimization, and MM-GBSA calculations [28]. Additionally, the "Volume Overlap Matrix" node was employed to eliminate poses that overlap with the most recently saved pose. We employed the "Run Maestro Command" node to execute a script within Maestro, enabling the combination of the target molecule and adsorbent structure into one file for the MacroModel minimization. Subsequently, the same node was used to separate them again, truncate the model by removing atoms >15 Å away from the target molecule and hydrogen-saturate the cut sides. The absence of a preexisting node for the semiempirical single-point energy calculations led us to use the "Chemistry External Tool" node for setting up GFN2-xTB simulations. The simulation outcomes, including the calculated potential energies of the target molecule, adsorbent, and their complex, are stored in a designated subfolder. A custom Python script, executed through the "Python Script" node, imports these energies and processes them to compute the binding free energy.

To further streamline ΔG screenings across various adsorbent structures and target molecules, our latest updates have incorporated several loops, with Figure S1 providing an expanded view of the simulation workflow, including key loops. The enhanced functionality of our workflow now permits users to input multiple docking grids of virtually designed chromatography resins, alongside SMILES strings of a multiple of target molecules. Within the KNIME environment, the workflow processes each resin sequentially through a looping mechanism, and for each resin, it iteratively loops through all the target molecules. This iterative approach facilitates the calculation of ΔGs for every possible combination of target molecule and chromatography resin. Upon completion, the workflow compiles all calculated ΔG values into an Excel spreadsheet, systematically organized by the corresponding docking grid and target molecule names from which each ΔG value was derived. The calculated binding poses of the target molecules on the resin surfaces are imported into a Schrödinger project file. Fig. 3b showcases a screenshot of the developed KNIME workflow. The grey nodes, termed metanodes, aggregate other nodes to streamline the workflow. These metanodes replicate the structure of the flowchart outlined in Fig. 2a, facilitating a straightforward understanding of the computational steps within the workflow. Users with advanced skills have the option to explore each metanode, examining the underlying sub-nodes for additional insights or modifications to the workflow's configuration.

To initiate their own ΔG calculations, users need to configure only

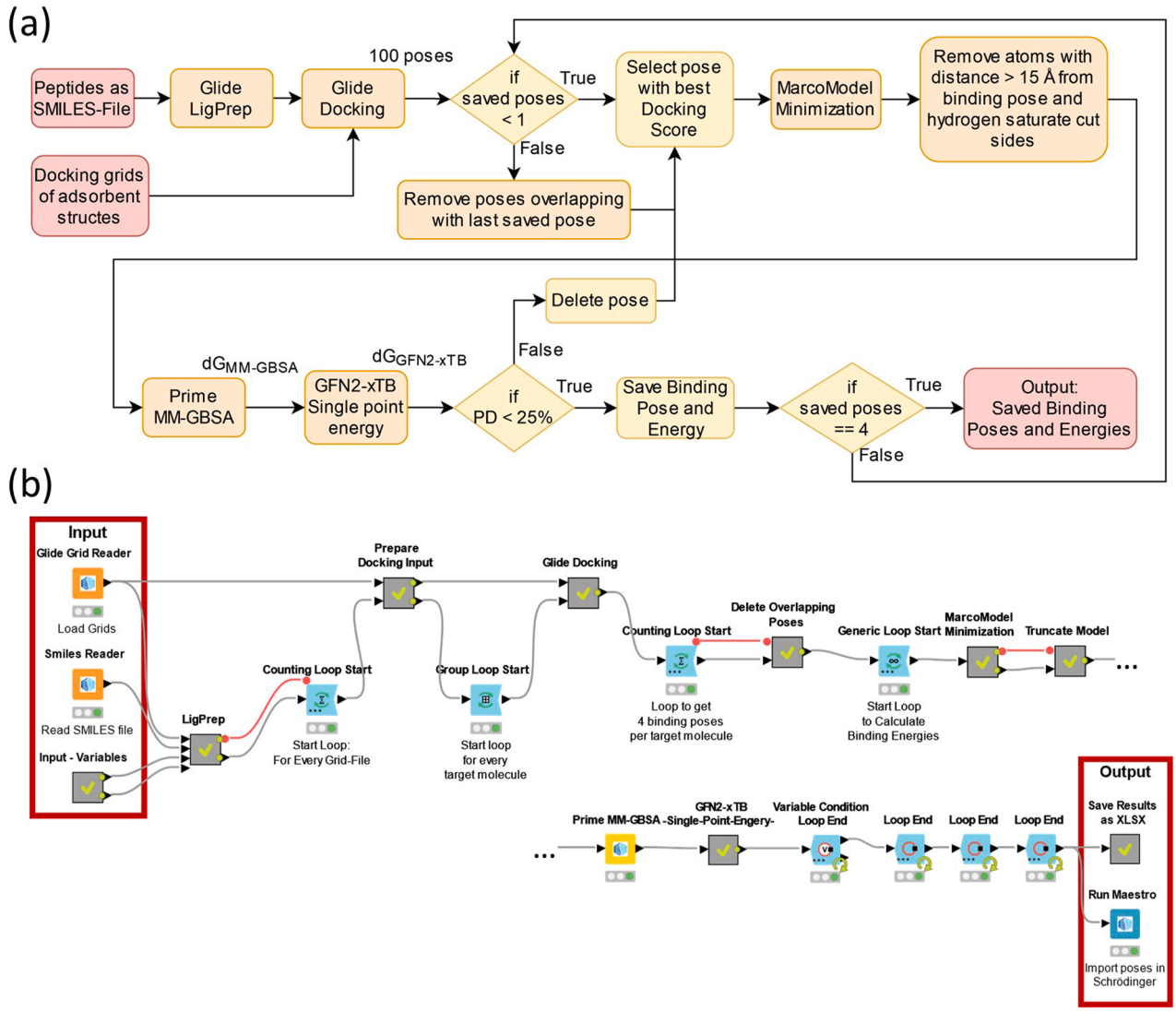

Fig. 3. Workflow automation with KNIME - Panel (a) illustrates the flowchart of the simulation and data management steps involved in calculating binding energies with peptides. Input and output nodes are marked in red, tasks related to simulation software in orange, and data management steps in yellow. Panel (b) shows the realization of this workflow in KNIME, with metanodes (grey) corresponding to flowchart steps for straightforward understanding or customization of the workflow.

three input nodes. The "Glide Grid Reader" node is for uploading the docking grids of their virtually designed resin surfaces and the target molecules' SMILES strings are uploaded through the "Smiles Reader" node. Additionally, the "Input Variables" Metanode allows users to set the pH for the target molecule preparation and to specify required system paths, including the main directory for saving results and the location of a file in SBC format for each docking grid, which details the atoms to be immobilized during the MacroModel minimization. Once these inputs are configured, the workflow is ready to be executed, demonstrating its adaptability and user-friendly design for conducting ΔG calculations across numerous resin and target molecule combinations. The workflow is available on GitHub: https://github.com/Ahmedkhalil-Mama/Automated-Workflows-for-All-Atom-Resin-and-Peptide-Adsorption-Modeling

It can also be directly imported into a locally installed KNIME environment via the following link: https://hub.knime.com/s/bPcNwd0wgw5zvWH3

All required software, along with the specific software versions and hardware employed to run the workflow, are listed in the Software section of the Supplementary Information.

3. Results and discussion

3.1. Automated workflow for virtual design of methacrylate-based chromatography resins

In this section, we demonstrate the application of our automated workflow to virtually design a methacrylate-based chromatography resin surface. This case study highlights the workflow's ability to construct a detailed resin model from initial input parameters, featuring newly implemented options for ion-exchange ligands and alkyl spacers of adjustable length. For this demonstration, the following input parameters were selected: DHPMA was chosen as the monomer, EGDMA as the crosslinker, butane (C4) as the spacer, and tetramethylammonium as the Q-type ligand. The simulation box dimensions were set to 15 × 15 × 20 nm, containing 500 monomer molecules, 260 crosslinker molecules, and 300 ligand molecules. These values were selected to represent a realistic and functional composition for a chromatography resin [8,29].

Fig. 4a illustrates the polymer network in CG resolution, which is the output of the "CG-Polymerization" WaNo in the SimStack workflow. This intermediate structure results from the dynamic polymerization of the methacrylic backbone followed by a ligand attachment step. The CG

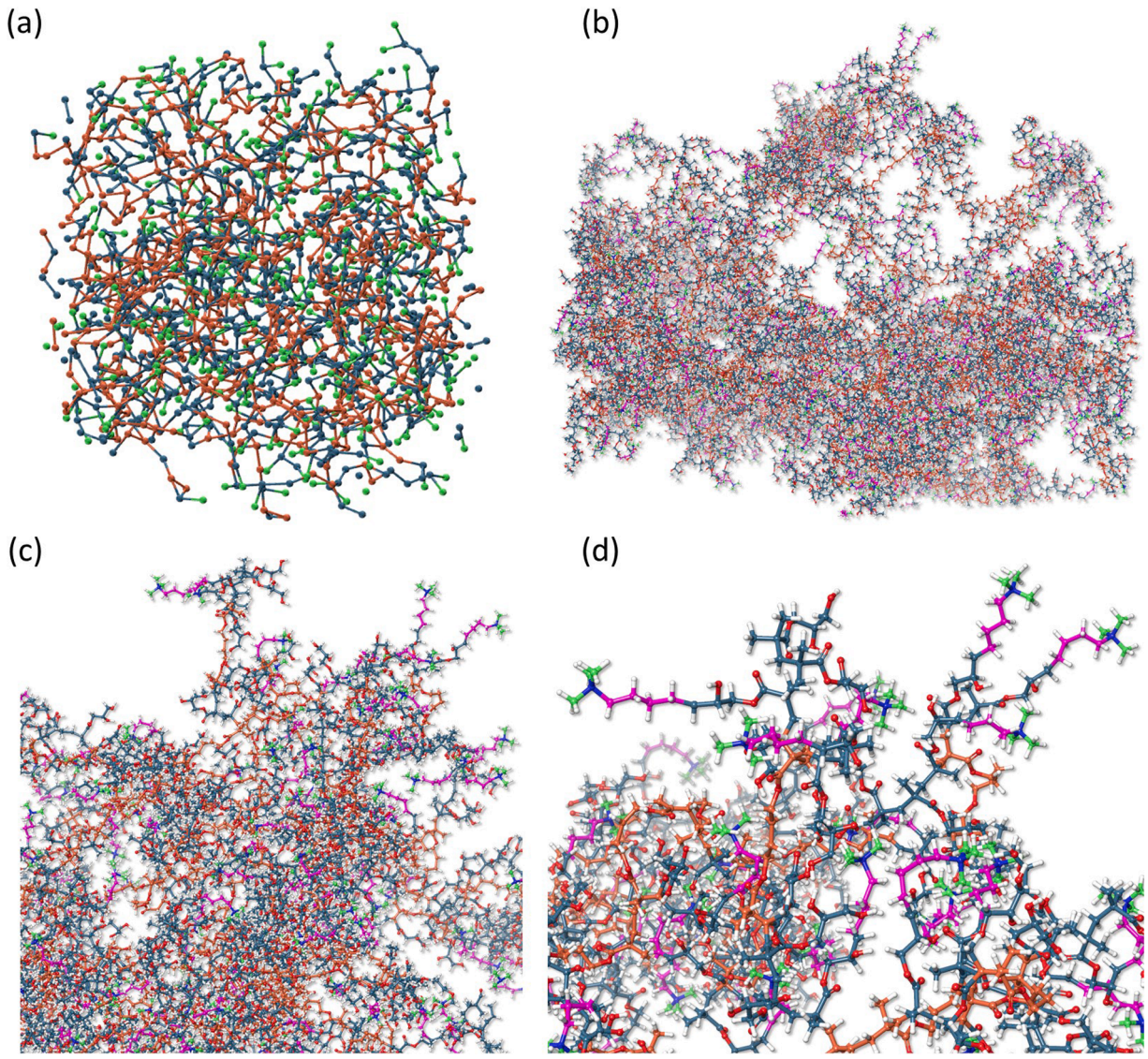

Fig. 4. Virtually designed methacrylate-based ion-exchange resin – Panel (a) illustrates the polymer network in CG resolution. Panel (b) shows the final structure of the surface model after conversion to AA resolution. Panels (c) and (d) provide close-up views of the resin surface, highlighting the dense crosslinked polymer backbone and Q-type ligands, protruded by C4 spacers. The CG spheres and carbon atoms are color-coded according to the building blocks of the resin: DHPMA monomers are light blue, EGDMA crosslinker are orange, C4 spacer are pink, and Q-type ligands are depicted in green.

spheres are color-coded based on their origin: light blue for the DHPMA monomer, orange for EGDMA, and green for the Q-type ligand. The upper region of the model represents the inside of a pore, resembling the interface between the mobile phase and the stationary phase. In the subsequent steps, the CG structure is converted to an all-atom representation, smaller fragments not incorporated into the main polymer network are removed, and the system is relaxed to ensure a realistic polymeric structure. Fig. 4b shows the refined all-atom model, with carbon atoms color-coded according to the same scheme used in the CG structure. Additionally, pink carbon atoms highlight the newly introduced C4 spacers. The workflow fully automatically produced a highly refined molecular model from minimal user input. Through a seamlessly managed transition from coarse-grained to all-atom resolution, the structural fidelity of the polymer network was preserved, while detailed atomic information were added. The Panels 3c and 3d provide a closer view of the resin surface, revealing two key insights into its structural features. First, the interspersed orange crosslinker segments indicate a densely crosslinked polymer backbone, a critical factor for mechanical stability in real chromatographic materials [29,30]. Second, the C4 spacer positions the Q-type ligands so that they protrude distinctly from the surface, potentially enhancing their accessibility for target molecules. Despite this protrusion, the surface remains highly heterogeneous, featuring polymer chains of varying lengths that extend into the interior of the pore and create a complex interfacial environment.

By automating each major simulation task, the workflow minimizes manual intervention and reduces the risk of human error, providing a reliable, reproducible, and rapid means of generating detailed resin models. Beyond expediting model generation, this approach also promotes consistency and accuracy across different simulation runs. The resulting all-atom models serve as an excellent foundation for further investigations, facilitating the optimization of chromatographic process conditions and the development of innovative resin materials.

3.2. *Use cases for virtual design of chromatography resins*

The advancements described in our virtual design methodology open up new possibilities for enhancing chromatographic separation processes. This section explores the versatile applications of the developed resin surface models, underscoring their potential to enrich both the fundamental understanding of chromatography and the practical

optimization of separation conditions. One of the most notable capabilities of our resin surface models is the ability to compute binding free energies for target molecules, as detailed in Section 2.2. In previous studies with more simplified resin surface models, such ΔG values were used to predict isocratic and gradient elution behaviors [6,31,32], as well as Langmuir constants [8,33]. Such predictive power is highly valuable for optimizing chromatographic parameters, enabling researchers to select the most suitable resin and fine-tune critical factors like pH [34] and ionic strength [4,6] to maximize yield and purity. By efficiently forecasting molecule-resin interactions, virtual design expands the range of possible resin candidates and reduces the need for extensive high-throughput experimentation. By publishing our validated workflow for ΔG calculations with linear peptides, we provide a fully automated pipeline for high-throughput screening of peptide–resin interactions in a virtual environment. This approach drastically reduces manual input, minimizes user error, and supports batch processing of multiple resin and peptide candidates. Consequently, researchers can more efficiently identify optimal resin chemistries and operating conditions with fewer experimental trials.

Additionally, our virtual design approach facilitates rational design in resin development. By leveraging detailed molecular insights, researchers can design novel resin materials with attributes tailored to specific separation challenges. For instance, the latest workflow update introduces the option to incorporate linear alkyl spacers of variable lengths between the ligand and the polymer backbone, thereby allowing detailed exploration of how spacer length influences ligand accessibility and adsorption behavior. This concept can be further extended to examine different ligand chemistries and backbone compositions, either to minimize non-specific interactions or to deliberately exploit them for particular separations. Integrating this virtual resin design strategy with optimization techniques, such as genetic algorithms or machine learning, could accelerate the discovery of new resin formulations, each optimized for unique performance criteria. Furthermore, the introduced resin models can serve as powerful educational tools. By offering students and newcomers to the field a visual and interactive platform for understanding chromatographic separations at the molecular level, they bridge theoretical knowledge with practical insights. Observing how specific modifications to resin architecture or operating conditions affect molecular adsorption dynamics provides a hands-on learning experience that can substantially deepen comprehension of core chromatographic principles.

Overall, the virtual design approach not only enhances traditional workflows for resin screening but also sets the stage for the next generation of chromatography research and education. Through its predictive capabilities and adaptability, it promises to streamline development pipelines, guide rational design efforts, and broaden the scope of chromatographic separations.

4. Conclusion and further directions

In this study, we introduced and detailed two automated multistage simulation workflows within the WMSs SimStack and KNIME. The first workflow, automated in SimStack, can be utilized to generate detailed all-atom models of methacrylate-based chromatography resin surfaces from a set of user-defined input parameters. Our approach is presently optimized for methacrylate-based resins - a choice justified by their widespread use in chromatography, their well-defined radical polymerization mechanisms, and the critical influence of their hydrophobic backbones on separation performance. Although the principle of workflow-driven automation is inherently general, extending our methodology to other synthetic polymers, like polystyrene-based resins, would require only modest modifications. In contrast, adapting the workflow to materials such as agarose or silica would necessitate substantial methodological revisions or the development of alternative modeling strategies due to their fundamentally different formation mechanisms.

The second workflow represents a practical use case of resin surface models that employs KNIME for high-throughput binding energy calculations with linear peptides. The workflow accepts peptides in SMILES format, allowing the incorporation of nonstandard residues or post-translational modifications provided that the overall peptide size remains within these constraints of Schrödinger's Glide Docking. This module imposes a limit of 500 atoms or 100 rotatable bonds.

These workflows represent a significant advance for the field of chromatography, providing a unified platform for constructing all-atom models of methacrylate-based chromatography resin surfaces and calculating binding free energies with various target molecules. Beyond demonstrating their technical feasibility, we also explored several potential applications of our virtually designed resins. These include the virtual screening of binding and elution conditions, as well as the rational design of novel resins tailored to specific separation challenges. We posit that making these workflows publicly available provides a compelling example of how WMSs can accelerate the broader adoption of molecular simulations in chromatography. By continuing to enhance and expand these methodologies, we aim to facilitate more accurate and accessible molecular modeling across the discipline.

CRediT authorship contribution statement

Tim Ballweg: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Modan Liu:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Conceptualization. **Ahmed Mama:** Software, Methodology. **Wolfgang Wenzel:** Writing – review & editing, Supervision. **Matthias Franzreb:** Writing – review & editing, Supervision, 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.

Acknowledgments

The authors extend their gratitude to the "Virtual Materials Design" (VirtMat) initiative at the Karlsruhe Institute of Technology (KIT), funded by the Helmholtz Association, within the Joint Lab VMD, for their financial support of this work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2025.466027.

Data availability

The manuscript contains links to download the developed workflows.

References

[1] G. Guiochon, L.A. Beaver, Separation science is the key to successful biopharmaceuticals, J. Chromatogr. A 1218 (49) (2011) 8836–8858, https://doi.org/10.1016/j.chroma.2011.09.008. Bd.Nr.SDez.
[2] H. Schmidt-Traub, M. Schulte, A. Seidel-Morgenstern, Hrsg., *Preparative Chromatography*, 1. Aufl, Wiley, 2012, https://doi.org/10.1002/9783527649280.
[3] F. Dismer, J. Hubbuch, A novel approach to characterize the binding orientation of lysozyme on ion-exchange resins, J. Chromatogr. A 1149 (2) (2007) 312–320, https://doi.org/10.1016/j.chroma.2007.03.074. Bd.Nr.SMai.
[4] F. Dismer, M. Petzold, J. Hubbuch, Effects of ionic strength and mobile phase pH on the binding orientation of lysozyme on different ion-exchange adsorbents, J. Chromatogr. A 1194 (1) (2008) 11–21, https://doi.org/10.1016/j.chroma.2007.12.085. Bd.Nr.SJuni.

[5] S S. Parimal, Garde, S.M. Cramer, Interactions of multimodal ligands with proteins: insights into selectivity using molecular dynamics simulations, Langmuir 31 (27) (2015) 7512–7523, https://doi.org/10.1021/acs.langmuir.5b00236. Bd.Nr.SJuli.
[6] S S. Banerjee, 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. Bd.S.
[7] R. Sardella, E. Camaioni, A. Macchiarulo, A. Gioiello, M. Marinozzi, A. Carotti, Computational studies in enantioselective liquid chromatography: forty years of evolution in docking- and molecular dynamics-based simulations, TrAC Trends Anal. Chem. 122 (2020) 115703, https://doi.org/10.1016/j.trac.2019.115703. Bd.S.
[8] T. Ballweg, M. Liu, J. Grimm, E. Sedghamiz, W. Wenzel, M. Franzreb, All-atom modeling of methacrylate-based multi-modal chromatography resins for Langmuir constant prediction of peptides, J. Chromatogr. A 1730 (2024) 465089, https://doi.org/10.1016/j.chroma.2024.465089. Bd.S.
[9] L. Zhang, Y. Sun, Molecular simulation of adsorption and its implications to protein chromatography: a review, Biochem. Eng. J. 48 (3) (2010) 408–415, https://doi.org/10.1016/j.bej.2009.12.003. Bd.Nr.S.
[10] I. Kumari, P. Sandhu, M. Ahmed, Y. Akhter, Molecular dynamics simulations, challenges and opportunities: a biologist's prospective, Curr. Protein Pept. Sci. 18 (11) (2017) 1163–1179, https://doi.org/10.2174/1389203718666170622074741.
[11] R.S. Katiyar, P.K. Jha, Molecular simulations in drug delivery: opportunities and challenges, WIREs Comput. Mol. Sci. 8 (4) (2018) e1358, https://doi.org/10.1002/wcms.1358. Bd.Nr.S.
[12] C.R.C. Rêgo, SimStack: an intuitive workflow framework, Front. Mater. 9 (2022), https://doi.org/10.3389/fmats.2022.877597. Bd.Mai.
[13] A. Jain, FireWorks: a dynamic workflow system designed for high-throughput applications, Concurr. Comput. Pract. Exp. 17 (27) (2015) 5037–5059, https://doi.org/10.1002/cpe.3505. Bd.Nr.S.
[14] S M. Uhrin, P. Huber, J. Yu, N. Marzari, G. Pizzi, Workflows in AiiDA: engineering a high-throughput, event-based engine for robust and modular computational workflows, Comput. Mater. Sci. 187 (2021) 110086, https://doi.org/10.1016/j.commatsci.2020.110086. Bd.S.
[15] M. Roozmeh, I. Kondov, Workflow generation with wfGenes, in: 2020 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS), 2020, pp. 9–16, https://doi.org/10.1109/WORKS51914.2020.00007. Nov.S.
[16] W.A. Warr, Scientific workflow systems: pipeline Pilot and KNIME, J. Comput. Aided Mol. Des. 26 (7) (2012) 801–804, https://doi.org/10.1007/s10822-012-9577-7. Bd.Nr.SJuli.
[17] J. Schaarschmidt, Workflow engineering in Materials design within the BATTERY 2030+ project, Adv. Energy Mater. 17 (12) (2022) 2102638, https://doi.org/10.1002/aenm.202102638. Bd.Nr.S.
[18] R. Joshi, KNIME workflows for applications in medicinal and computational chemistry, Artif. Intell. Chem. 2 (1) (2024) 100063, https://doi.org/10.1016/j.aichem.2024.100063. Bd.Nr.SJuni.
[19] M.R. Berthold, KNIME - the Konstanz information miner: version 2.0 and beyond, SIGKDD Explor Newsl 11 (1) (2009) 26–31, https://doi.org/10.1145/1656274.1656280. Bd.Nr.S.
[20] A. Fillbrunn, C. Dietz, J. Pfeuffer, R. Rahn, G.A. Landrum, M.R. Berthold, KNIME for reproducible cross-domain analysis of life science data, J. Biotechnol. 261 (2017) 149–156, https://doi.org/10.1016/j.jbiotec.2017.07.028. Bd.S.
[21] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B 47 (1) (1993) 558–561, https://doi.org/10.1103/PhysRevB.47.558. Bd.Nr.S.
[22] A.P. Thompson, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. 271 (2022) 108171, https://doi.org/10.1016/j.cpc.2021.108171. Bd.S.
[23] E. Sedghamiz, M. Liu, W. Wenzel, Challenges and limits of mechanical stability in 3D direct laser writing, Nat. Commun. 13 (1) (2022) 2115, https://doi.org/10.1038/s41467-022-29749-9. Bd.Nr.S.
[24] K. Kremer, G.S. Grest, Dynamics of entangled linear polymer melts: A molecular-dynamics simulation", J. Chem. Phys. 92 (8) (1990) 5057–5086, https://doi.org/10.1063/1.458541. Bd.Nr.S.
[25] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-Atom Force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (45) (1996) 11225–11236, https://doi.org/10.1021/ja9621760. Bd.Nr.S.
[26] S.K. Tripathi, R. Muttineni, S.K. Singh, Extra precision docking, free energy calculation and molecular dynamics simulation studies of CDK2 inhibitors, J. Theor. Biol. 334 (2013) 87–100, https://doi.org/10.1016/j.jtbi.2013.05.014. Bd.SOkt.
[27] S C. Bannwarth, Ehlert, S. Grimme, GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions", J. Chem. Theory Comput. 15 (3) (2019) 1652–1671, https://doi.org/10.1021/acs.jctc.8b01176. Bd.Nr.SMärz.
[28] J. Bhachoo, T. Beuming, Investigating protein–Peptide interactions using the Schrödinger Computational Suite", in: O. Schueler-Furman, N. London (Eds.), Modeling Peptide-Protein Interactions: Methods and Protocols, Springer, New York, NY, 2017, pp. 235–254, https://doi.org/10.1007/978-1-4939-6798-8_14. Hrsg.S.
[29] M.J. Beneš, D. Horák, F. Svec, Methacrylate-based chromatographic media, J. Sep. Sci. 28 (15) (2005) 1855–1875, https://doi.org/10.1002/jssc.200500186. Bd.Nr.S.
[30] M. Andersson, M. Ramberg, B.-L. Johansson, The influence of the degree of cross-linking, type of ligand and support on the chemical stability of chromatography media intended for protein purification, Process Biochem 33 (1) (1998) 47–55, https://doi.org/10.1016/S0032-9592(97)00068-X. Bd.Nr.S.
[31] K.M.H. Lang, J. Kittelmann, F. Pilgram, A. Osberghaus, J. Hubbuch, Custom-tailored adsorbers: a molecular dynamics study on optimal design of ion exchange chromatography material, J. Chromatogr. A 1413 (2015) 60–67, https://doi.org/10.1016/j.chroma.2015.08.021. Bd.S.
[32] K.M.H. Lang, J. Kittelmann, C. Dürr, A. Osberghaus, J. Hubbuch, A comprehensive molecular dynamics approach to protein retention modeling in ion exchange chromatography, J. Chromatogr. A 1381 (2015) 184–193, https://doi.org/10.1016/j.chroma.2015.01.018. Bd.S.
[33] P.S. Ghosal, A.K. Gupta, Determination of thermodynamic parameters from Langmuir isotherm constant-revisited, J. Mol. Liq. 225 (2017) 137–146, https://doi.org/10.1016/j.molliq.2016.11.058. Bd.S.
[34] F. Dismer, J. Hubbuch, 3D structure-based protein retention prediction for ion-exchange chromatography, J. Chromatogr. A 1217 (8) (2010) 1343–1353, https://doi.org/10.1016/j.chroma.2009.12.061. Bd.Nr.S.