The effect of geometrical presentation of multimodal cation-exchange ligands on selective recognition of hydrophobic regions on protein surfaces

James Woo, Siddharth Parimal, Matthew R. Brown, Ryan Heden, Steven M. Cramer†
Department of Chemical and Biological Engineering and Center for Biotechnology and
Interdisciplinary Studies, Rensselaer Polytechnic Institute

† corresponding author:

Steven M. Cramer

E-mail: crames@rpi.edu

Phone: (518) 276-6198

Fax: (518) 276-4030

Abstract

The effects of spatial organization of hydrophobic and charged moieties on multimodal (MM) cation-exchange ligands were examined by studying protein retention behavior on two commercial chromatographic media, Capto™ MMC and Nuvia™ cPrime™. Proteins with extended regions of surface-exposed aliphatic residues were found to have enhanced retention on the Capto MMC system as compared to the Nuvia cPrime resin. The results further indicated that while the Nuvia cPrime ligand had a strong preference for interactions with aromatic groups, the Capto MMC ligand appeared to interact with both aliphatic and aromatic clusters on the protein surfaces. These observations were formalized into a new set of protein surface property descriptors, which quantified the local distribution of electrostatic and hydrophobic potentials as well as distinguishing between aromatic and aliphatic properties. Using these descriptors, high-performing Quantitative Structure-Activity Relationship (QSAR) models ($R^2 > 0.88$) were generated for both the Capto MMC and Nuvia cPrime datasets at pH 5 and pH 6. Descriptors of electrostatic properties were generally common across the four models; however both Capto MMC models included descriptors that quantified regions of aliphatic-based hydrophobicity in addition to aromatic descriptors. Retention was generally reduced by lowering the ligand densities on both MM resins. Notably, elution order was largely unaffected by the change in surface density, but smaller and more aliphatic proteins tended to be more affected by this drop in ligand density. This suggests that modulating the exposure, shape and density of the hydrophobic moieties in multimodal chromatographic systems can alter the preference for surface exposed aliphatic or aromatic residues, thus providing an additional dimension for modulating the selectivity of MM protein separation systems.

Keywords: Multimodal chromatography; Hydrophobic interaction; Protein surface properties;
Quantitative Structure-Activity Relationship

1. Introduction

Recently, the use of multimodal chromatographic media has demonstrated that an effective combination of orthogonal modalities (i.e. ion-exchange (IEX), hydrophobicity, etc.) can provide unique process advantages over traditional chromatographic media with only a single modality. For example, the enhanced salt-tolerance over IEX operations has made these multimodal materials particularly useful for protein capture from harvested cell culture fluid [1-6]. Within the general category of multimodal materials, the combination of modalities into a single molecule (mixed-mode ligands) provides better control of the distribution of modalities than combining these moieties macroscopically such as in mixed-bed chromatography [7] or with mixed-ligand media [8]. An analysis of a comprehensive set of multimodal cation-exchange and anion-exchange ligands has demonstrated that ligands containing aromatic moieties possessed increased salt-tolerant adsorption as compared to MM ligands containing aliphatic chain groups [9,10]. The effect of hydrogen bonding has been less conclusive as different authors have reported that these groups promote or suppress protein adsorption at high salt concentrations [9-11]. Other mixed-mode ligands that are classified as hydrophobic charge induction chromatography (HCIC) media interact primarily via hydrophobic interactions and elute the protein by changing the pH to induce electrostatic repulsion between the protein and resin surfaces [12]. These ligands have proven to be effective separation media, particularly for the purification of antibodies [13-16].

This still leaves many variables in mixed-mode ligand design yet to be characterized, particularly the role of spatial organization (i.e. distance between moieties, steric accessibility of each group) which could enhance the specificity of the ligand by constraining the interaction potential of each modality. The effect of spatial organization on multimodal retention behavior is exemplified by two multimodal cation-exchange chromatographic media, Capto MMC and Nuvia cPrime, which consist of similar chemical moieties but differ primarily in the structural organization around the hydrophobic aromatic ring. Both have been shown to interact with proteins in a hydrophobic manner but exhibited different affinities for the model protein bovine serum albumin [17,18].

Several approaches have been developed to model protein retention in multimodal chromatography and have generally been based on methods previously developed for IEC and HIC systems. Nfor et al. utilized a generalized thermodynamic framework for IEC and HIC chromatography [19] to fit adsorption isotherms to a range of data for various proteins and multimodal ligands at different salt and pH conditions [20]. This approach was extended further with the use of linear free energy relationships for modeling the retention of tripeptides on Capto adhere [21]. Zhou et al. found that a stoichiometric displacement model was able to fit retention data on a multimodal resin when operated in both ion-exchange and HIC modes [22]. QSAR models were developed for a multimodal anion-exchange resin, Capto adhere, that predicted the retention of a library of proteins in the presence of several co-solutes (ethylene glycol, urea, arginine) [23]. Further, the steric mass-action model (originally developed for ion-exchange processes) was used to successfully predict retention behavior of two proteins on the Capto adhere resin using both NaCl and Arg-HCl linear gradients [24].

The current study probes the difference in adsorptive properties between the Capto MMC and Nuvia cPrime resins by measuring the retention of a range of proteins in salt-based linear gradient elution chromatography. From these data, a simple classification scheme is developed to predict the affinity of a protein to either resin based on the presence of clusters consisting of exposed aliphatic or aromatic residues. From this approach, novel protein surface property descriptors are developed to recognize clusters of electrostatic and hydrophobic potential on the protein surface that molecular dynamics simulations had indicated could form ideal binding sites with these ligands [25]. Quantitative Structure-Activity Relationship (QSAR) models are then employed to select relevant protein surface properties and predict the retention of a library of proteins on these chromatographic media. Finally, these retention experiments are repeated with resin material at several ligand densities to ascertain whether ligand chemistry is the main contributor to the selective differences in protein adsorption between the Nuvia cPrime and Capto MMC resins. This work demonstrates how the intermolecular and intramolecular organization of chemical moieties in multimodal systems might be used to modulate hydrophobic properties and to create novel selectivity for particular protein surface hydrophobicity features.

2. Material and Methods

2.1 Materials

Glacial acetic acid and guanidine hydrochloride were purchased from Thermo Fisher Scientific (Pittsburgh, PA). Recombinant human ubiquitin was purchased from Boston Biochem, Inc. (Cambridge, MA). All other chemicals and proteins were obtained from Sigma-Aldrich (St. Louis, MO). Capto MMC and Capto MMC ImpRes chromatography media were purchased from GE Healthcare (Uppsala, Sweden). Nuvia cPrime chromatography media at various ligand densities were gifts from Bio-Rad Laboratories (Hercules, CA). Glass columns (5 mm x 50 mm) and adapters were purchased from Pharmacia Biotech (Uppsala, Sweden).

2.2 Column Packing Procedure

Chromatographic resin was first equilibrated in deionized water and then was resuspended in a 50% (v/v) slurry in deionized water. Slurry (2.2 mL) was poured into a 5 mm (ID) x 50 mm column and flow-packed in deionized water at 0.5 mL/min for 30 minutes. The flow adapter was adjusted onto the surface of the resin bed and flow was adjusted to 1 mL/min and packed for another 30 minutes. The adapter was again adjusted onto the bed surface at the final bed volume of ~1 mL.

2.3 Protein Library Screening Experiments

Linear gradient experiments were performed on an Äkta Explorer 100 (Amersham Biosciences, Uppsala, Sweden). Running buffers for all experiments were prepared from a single 25x concentrate solution (500 mM acetate, pH 5 or 500 mM phosphate, pH 6) and diluted to the desired concentration without further pH adjustment. Packed columns were equilibrated at 1 column volume (CV)/min with 5 CV of 1% Buffer B (Buffer A + 1.5 M NaCl) in Buffer A prior to the start of each experiment. Proteins were dissolved in the equilibration buffer (1% Buffer B) to 3 mg/mL and pipetted into 96-well UV transparent well plates. Fifty µL of protein solution were aspirated by an A-905 autosampler (Amersham Biosciences, Uppsala, Sweden) and injected into the column. A linear salt gradient from 1 to 100% Buffer B was generated over 45 CV and held at 100% B for 8 mL (to account for the dead volume of the chromatography system). The column was then re-equilibrated with 7 CV prior to the next injection. Absorbance at the column effluent was measured at 280 nm and 215 nm using a 2 mm UV flow cell. Retention times were determined by calculating the center-of-mass for each peak. The conductivity in mS/cm was determined for that retention time and the conductivity was used to determine the elution salt concentration value.

2.4 Preparation of Protein 3D structures

All protein structures were obtained from the RCSB Protein Data Bank; the corresponding PDB codes can be found in Table 1. Water molecules and co-solutes were removed from the structure file and homology modeling (Molecular Operating Environment (MOE), Montreal, Québec, Canada) was performed to replace any segments of the polypeptide missing from the structural file. Structures were protonated at pH 7 using the Protonate3D function in MOE and subjected to three rounds of tethered energy minimization using the Amber99 forcefield.

2.5 Calculation of Residue Cluster Descriptors

Using the prepared protein structure file, the solvent-accessible surface area (ASA) of each atom was calculated using MOE. The surface area of all side chain atoms corresponding to a particular residue were summed together and % exposure of this residue was calculated as the ratio of ASA_{residue} to ASA_X for a Gly-X-Gly tripeptide. pKa values for all titratable groups were calculated using PROPKA 2.0 [26]. The location of each residue in the protein structure was recorded as the center-of-mass for the residue's side chain. Residue sidechains that were more than 15% solvent-exposed were considered to be on the protein surface. Basic residues where $pK_a > pH$ and acidic residues where $pK_a < pH$ were considered to be charged.

Clusters were calculated using same methodology as used by Hou et al. [27]. All of the inter-residue distances were calculated for a particular family of residues (i.e. charged acidic residues or exposed aliphatic residues), and any residue pairs falling within the 2-10 Å range were recorded into a text file. For binary property clusters (i.e. acidic-aromatic clusters), distances were calculated for residue pairs containing one residue of each property. Any residue pairs sharing a common residue were considered linked and grouped into a cluster. Three values are calculated for each cluster; # of clusters, largest cluster size (# of residues in the largest cluster), and total number of residues in clusters (# of residues within 2-10 Å of each other).

2.6 Calculation of Individual Property Map and Overlapping Clusters Descriptors

Using the prepared protein structure files, electrostatic potential maps were generated using the Adaptive Poisson-Boltzmann Solver (APBS) [28] at the desired pH using protonation states as determined by PROPKA [26]. Hydrophobic potential maps of the protein were also generated by assigning hydrophobic potentials to surface-exposed protein atoms based on exposed aliphatic or aromatic residues, the spatial-aggregation propensity (SAP) algorithm [29], or the Black and Mould hydropathy index [30]. Then, a uniform grid of points spaced 1 Å

apart was placed throughout the volume of the PDB file. At each grid point, a hydrophobic potential was assigned based on the hydrophobic potential (either SAP, hydropathy index, aromatic or aliphatic residue) of all nearby atoms weighted by a decay rate of $1/r$. Local maxima were identified based on a contour analysis of the potential map and all adjacent grid points to that maxima were assigned as a cluster. For each cluster, a center of mass (COM) and a potential-weighted average radius for the cluster were determined from the grid points within that cluster.

The sum of potential values of grid points that are assigned to a cluster and within a distance of r_{avg} of the COM was defined as the total strength of the cluster. Strong clusters were defined as the top 50% of clusters from a potential map of a protein. Four descriptors were generated for each potential map: the number of clusters, and the total strength, normalized strength and mean strength of the top 25% of clusters. The normalized strength was defined as the total strength divided by the volume of those clusters. The mean strength was defined as the total strength divided by the number of these clusters. Descriptors were also generated for the number of pairs of overlapping strong-strong, strong-weak and weak-weak clusters, defined as those clusters overlapping within r_{avg} from their COMs.

2.7 Preparation of QSAR models

First, a training set was assembled by compiling the various descriptors and response value (elution salt concentration) into a comma-delimited file (.csv). This training set was loaded into the “Yet Another Modeling Software” (YAMS) hosted by the Rensselaer Exploratory Center for Cheminformatics Research (RECCR) [31]. A recursive feature elimination approach was used to reduce the initial descriptor set to the minimum required for accurate prediction. Twelve iterations of selection were performed with intermediate PLS models, retaining the top 80% of descriptors after each iteration. The best of four final models (MLR, PLS, SVM and Random Forest) generated by the YAMS software was selected for each dataset. Because the size of each dataset was quite small (19 proteins), 10 rounds of 10-fold cross-validation were employed to evaluate the model in lieu of an external test set.

3. Results and Discussion

Two commercially-available, multimodal cation-exchange ligands, Capto MMC and Nuvia cPrime, form a natural pair for investigating the effect of ligand structure across a broad spectrum of protein chemistries. These two ligands (Table 3) contain similar chemical groups,

namely a phenyl ring and a carboxylate group that are separated by an amide linkage. Since these chromatographic media are synthesized from different manufacturers, they also differ in some other properties besides ligand structure. The attachment of the Nuvia cPrime ligand to the base matrix at the para-position on the phenyl ring via a secondary amine creates a steric barrier that limits the interaction potential to a single face of the ring. Since the Capto MMC ligand is attached via the α -carbon to an aliphatic thioether group and 3-PEG extender, the phenyl group is fully exposed to the solvent environment and thus has a wider range of motion around the site of attachment.

Although the Capto MMC and Nuvia cPrime ligands are attached to agarose and acrylamide base matrices, respectively, both matrices are considered to be relatively inert to most proteins [32-35]. The volumetric ligand density of the Capto MMC material is ~20% higher than for Nuvia cPrime, however the manufacturing ranges for ligand densities overlap between the two products. The porosity and mean pore radius were estimated from inverse size-exclusion chromatography (ISEC) data previously obtained for Sepharose 4 Fast Flow [36] and UNOsphere S [37] which were used as estimates of the properties of the Capto and Nuvia base matrices respectively. Based on the ISEC data, a rough estimate of specific surface area ($\text{m}^2/\text{mL}_{\text{resin}}$) and surface ligand density was obtained as described previously in the literature [36]. Both matrices were reported to have a similar average particle size and porosity; however, the average pore size was smaller for the agarose matrix. Despite having similar volumetric ligand densities and matrix porosities, the smaller pore radius of the Capto base matrix resulted in a higher specific surface area and a lower surface ligand density as compared to Nuvia cPrime. Since ligand density contributes to the adsorptive properties of resin materials [38,39], both ligands were tested with both high and low ligand density media to account for variations in surface ligand densities as will be discussed below.

Linear gradient experiments were carried out with a library of commercially-available proteins on Capto MMC and Nuvia cPrime at pH 5 and 6. Proteins tended to elute from both resins much later in the NaCl-based linear gradient as compared to the equivalent ion-exchange ligand (CM Sepharose Fast Flow) (Figure 1a and Figure 1b). This salt-tolerant behavior had been previously reported for Capto MMC [10,17]. In general, proteins with an acidic isoelectric point ($\text{pI} < 6$) tended to adsorb more weakly to these multimodal (MM) cation-exchange resins, particularly at higher pH. While the majority of proteins tested were similarly retained on both

resins, several proteins (α -lactalbumin, β -lactoglobulin A/B, human/bovine serum albumin, ubiquitin, α -chymotrypsin and avidin) showed markedly increased retention on Capto MMC as compared to Nuvia cPrime at both pH conditions. Interestingly, the proteins that exhibited selectively enhanced retention on Capto MMC spanned a range of pI values (5.1-9.7), which indicated that electrostatic forces alone could not explain this enhanced adsorption behavior and suggested a prominent role for hydrophobic interactions in defining the selectivity observed between these resin systems.

In another study, molecular dynamics (MD) simulations were performed which examined interactions between the protein ubiquitin and the Capto MMC and Nuvia cPrime ligands at the atomic level in order to examine differences in the regions of the protein surface probed by these molecules [25]. Binding maps were generated for the protein surface by plotting the ratio of local ligand density to ligand density in bulk solution (ρ/ρ_0). The Capto MMC/ubiquitin simulation revealed that ligand interactions were focused on a hydrophobic pocket adjacent to the positively-charged face of the protein where the phenyl ring strongly interacted with a dense cluster of aliphatic residues (Figure 2a). Interestingly, the Nuvia cPrime ligand did not have the same preference for that binding pocket and preferred to interact more weakly with the entire positively-charged face of the protein and with several exposed aromatic residues at the bottom of this region (Figure 2b).

This suggested that the Capto MMC and Nuvia cPrime ligands could be recognizing different hydrophobic surface features based on the type of side chain (aromatic or aliphatic) exposed on the protein surface. Based on this hypothesis, the proteins in the library employed in Figure 1 were classified according to the hydrophobic clusters present on their surfaces. Proteins were defined as 'aliphatic' or 'aromatic' if the mean size of the largest cluster of each hydrophobic sidechain type was in the top 75% of the library population. Additional categories were defined for proteins with both types of hydrophobic clusters or neither. Figure 3 reorganizes the retention results for the protein library based on this simple classification scheme. Interestingly, those proteins with enhanced retention on Capto MMC relative to Nuvia cPrime were all classified as 'aliphatic' or 'both', while those with strong retention on both resins (>1.0 M NaCl) were all classified as 'aromatic' proteins.

While proteins with low to moderate retention tended to display neither type of hydrophobic cluster, several proteins exhibiting this type of retention behavior were found in the other

classes (conalbumin, α -chymotrypsinogen A and trypsin proteins). This is not completely surprising since this simple classification approach only considers sites of hydrophobic interaction potential and ignores the fact that proteins adsorb to these multimodal surfaces via both hydrophobic and electrostatic interactions. Accordingly, various electrostatic properties of the protein surfaces were also calculated in an attempt to improve the prediction of retention behavior. Regions of strong attractive and repulsive electrostatic potentials were identified as well as the proximity of those regions to the previously identified hydrophobic clusters. The properties of these clusters were then quantified and converted into discrete numerical values or descriptors as described in Sections 2.5 and 2.6.

Using this dataset of cluster-based protein surface descriptors, quantitative structure-activity relationship (QSAR) models were generated as described in the Methods section to create predictive models of the linear gradient retention data on Capto MMC and Nuvia cPrime at pH 5 and pH 6. The initial set of 143 descriptors was reduced to a concise set of 5-6 relevant protein surface properties by using the sensitivity analysis feature selection tool, which employed a recursive feature elimination algorithm to remove the least relevant descriptors. The descriptors selected for each QSAR model and their performance statistics are presented in Table 4. In addition, the models for each pH and resin are presented in Figure 5. As can be seen, the performance of these QSAR models ranged from $R^2 = 0.88$ to $R^2 = 0.94$. Notably, clusters of positive EP, negative EP, and aromaticity on the protein surface were selected for all four final QSAR models. While none of these descriptors alone were directly correlated with the observed protein retention behavior in MM systems, the presence of these descriptors in all of the models indicates that these were important protein surface features recognized by both MM resins.

This set of model proteins were also evaluated for their retention behavior on CM Sepharose Fast Flow in order to further probe the contributions of electrostatic interactions. Interestingly, while the descriptor for positive EP (number of basic residue clusters) was correlated with the retention of basic proteins ($pI > 8.0$) at both pH 5 and pH 6 ($R^2 = 0.76$) on the ion exchange resin, (Figure 6), it was a poor predictor of retention for proteins with a low pI . One reason for this behavior is that the proteins with a pI above pH 8 had very few negative residues. This behavior extended to the MM QSAR models as well. At higher pH (pH 6) where these negative charges were more prominent, the number of basic residue clusters had a reduced weighting in

the models since the calculation ignored any adjacent negative residues and would have overrepresented the extent of electrostatic attraction. To compensate, negative EP descriptors (acidic clusters, acidic-aromatic clusters and negative EP clusters) were selected in all of the models to account for the repulsive effects of negative charges on the protein surface. The negative EP descriptors were generally weak contributors in both models and tended to have a negative weight, supporting the hypothesis that they accounted for electrostatic repulsion.

The hydrophobic descriptors selected in these models (Table 4) also appeared to corroborate the earlier classification approach (Figure 3). Models based on the Nuvia cPrime datasets were biased towards aromatic descriptors, while the Capto MMC models used descriptors for aliphatic residue clusters and for general hydrophobicity. This further supports our hypothesis that the Capto MMC resin has a strong affinity for hydrophobic proteins regardless of type, while the Nuvia cPrime resin is more selective towards exposed aromatic residues.

Another factor affecting the comparability between these two resins was that the Nuvia cPrime material had approximately a 50% higher ligand density (126 $\mu\text{mol/mL}_{\text{resin}}$) as compared the Capto MMC media ($\sim 80 \mu\text{mol/mL}_{\text{resin}}$). To ascertain the effect of ligand density on the observed selectivity differences between the Capto MMC and Nuvia cPrime material, protein retention was also measured on the low ligand density Capto MMC ImpRes media ($\sim 32 \mu\text{mol/mL}_{\text{resin}}$) and two low density Nuvia cPrime media (76 and 30 $\mu\text{mol/mL}_{\text{resin}}$) with equivalent ligand densities to the Capto MMC and Capto MMC ImpRes resins, respectively.

Figures 7a and 7b compare the elution salt concentration data of the proteins for various ligand densities on the Nuvia cPrime and Capto MMC systems, respectively. As can be seen in Figure 7a, essentially no difference in retention was observed between the high density Nuvia cPrime and the intermediate ligand density material that was comparable to the commercial Capto MMC media. On the other hand, all of the proteins eluted earlier on both the low-density Capto MMC ImpRes resin and the lowest density Nuvia cPrime media. Interestingly, while the retention decreased, the elution order for most of the proteins remained essentially the same at these lower ligand densities as evidenced by the high correlation coefficient (R^2). The reduced retention on both low density materials suggested that the ligand density may have dropped below a critical level which reduced the total number of ligands that were available to interact with the contacting surface of the adsorbed protein [40].

However, several proteins proved to be more sensitive to ligand density and were observed to elute earlier in the linear gradient at the lower ligand densities, with this effect being more pronounced at pH 5 than at pH 6. Alpha-lactalbumin and β -lactoglobulins A and B, which were classified as 'aliphatic' cluster proteins in Figure 3 and were hypothesized to adsorb more strongly via hydrophobic interactions, were only sensitive to a drop in ligand density in the Capto MMC system. A depiction of the MM ligands at different ligand densities are presented in Figures 8a and 8b. As can be seen in the figure, the Capto MMC ligands can aggregate on the resin surface due to their exposed aromatic rings. Reducing the ligand density would increase the average distance between immobilized ligands and reduce the frequency of these aggregated ligand complexes, thus lowering the hydrophobic potential of the resin. This does not occur with the Nuvia cPrime ligand as it is directly immobilized to the resin matrix via the hydrophobic moiety (Figure 7b), which greatly restricts aromatic-aromatic interactions between two ligands, regardless of surface density. Ubiquitin and aprotinin were found to be more sensitive to ligand density changes on both Capto MMC and Nuvia cPrime. These proteins are quite small (<9 kDa) and present a dense, positively charged surface that would likely be more sensitive to drops in ligand density because fewer ligands would now be able to interact with these protein surfaces.

4. Conclusions

In this work, the spatial organization of hydrophobic and charged moieties on multimodal cation-exchange ligands has been shown to have a substantial effect on the protein retention behavior in MM cation exchange systems. These results indicated that the sterically-constrained Nuvia cPrime ligand had a strong preference for retaining proteins with exposed aromatic groups, while the relatively exposed Capto MMC ligand appeared to retain proteins with both aliphatic and aromatic clusters on their surfaces. While retention was generally reduced by lowering the ligand densities on both MM resins, elution order was mostly unaffected by the change in surface density. Predominantly aliphatic proteins tended to be more affected by the drop in ligand density on the Capto MMC resin, while smaller proteins experienced stronger reductions in retention from a drop in ligand density on both Capto MMC and Nuvia cPrime resin materials. These observations were formalized into numerical descriptors that were used to identify proteins with prominent aliphatic or aromatic surface clusters and which could be used in a simple classification scheme to identify proteins that were sensitive to these changes in ligand chemistry, namely those proteins with aliphatic clusters. This demonstrates how the chemistry of hydrophobic surfaces on proteins can be

used as an additional dimension by which to separate proteins on these multimodal materials. From this approach, a new set of protein surface property descriptors was developed that quantified the local distribution of electrostatic and hydrophobic potentials as well as distinguishing between aromatic and aliphatic properties. Using these descriptors, high-performing QSAR models ($R^2 > 0.88$) were generated for both the Capto MMC and Nuvia cPrime datasets at pH 5 and pH 6. Descriptors of electrostatic properties were generally common across the four models; however both Capto MMC models included descriptors that quantified regions of aliphatic-based hydrophobicity in addition to aromatic descriptors. The selective differences observed between these similar multimodal ligands suggests the exposure, shape and density of the hydrophobic moieties found on multimodal ligands may be leveraged in order to develop a concise set of multimodal resins with orthogonal modes of selectivity to exploit for superior resolution in challenging protein separations. This hydrophobic selectivity could be further augmented by altering groups adjacent to the aromatic ring to increase its hydrophobic character or by adjusting the separation distance between the charged and hydrophobic moieties. Future work will expand upon these design concepts, creating a homologous set of MM ligands which systematically examines these geometrical and moiety effects to determine whether they may also create new dimensions for multimodal-based protein separations.

5. Acknowledgements

This work was supported by NSF Grant CBET 1150039 and Bio-Rad Laboratories Inc. The authors thank Prof. Curt Breneman and Dr. Michael Krein for their assistance in using the YAMS software and validating the final models. The authors also thank Mark Snyder and Hong Chen from Bio-Rad Laboratories for providing Nuvia cPrime materials at different ligand densities for this study.

6. References

- [1] J. Chen, J. Tetrault, Y. Zhang, A. Wasserman, G. Conley, M. DiLeo, E. Haines, A.E. Nixon, A. Ley, The distinctive separation attributes of mixed-mode resins and their application in monoclonal antibody downstream purification process, *Journal of Chromatography A* 1217 (2010) 216-224.
- [2] K.A. Kaleas, C.H. Schmelzer, S.A. Pizarro, Industrial case study: Evaluation of a mixed-mode resin for selective capture of a human growth factor recombinantly expressed in *E. coli*, *Journal of Chromatography A* 1217 (2010) 235-242.
- [3] T. Arakawa, Y. Kita, H. Sato, D. Ejima, MEP chromatography of antibody and Fc-fusion protein using aqueous arginine solution, *Protein Expression and Purification* 63 (2009) 158-163.
- [4] J. Paul, S. Jensen, A. Dukart, G. Cornelissen, Optimization of a preparative multimodal ion exchange step for purification of a potential malaria vaccine, *Journal of Chromatography A* 1366 (2014) 38-44.
- [5] K.A. Kaleas, M. Tripodi, S. Revelli, V. Sharma, S.A. Pizarro, Evaluation of a multimodal resin for selective capture of CHO-derived monoclonal antibodies directly from harvested cell culture fluid, *Journal of Chromatography B* 969 (2014) 256-263.
- [6] L.S. Wolfe, C.P. Barringer, S.S. Mostafa, A.A. Shukla, Multimodal chromatography: Characterization of protein binding and selectivity enhancement through mobile phase modulators, *Journal of Chromatography A* 1340 (2014) 151-156.
- [7] Z. el Rassi, C. Horvath, Tandem columns and mixed-bed columns in high-performance liquid chromatography of proteins, *J Chromatogr* 359 (1986) 255-264.
- [8] B. Buszewski, R.M. Gadzala-Kopiuch, M. Jaroniec, Chromatographic Properties of Mixed Chemically Bonded Phases with Alkylamide and Aminopropyl Ligands, *Journal of Liquid Chromatography & Related Technologies* 20 (1997) 2313-2325.
- [9] B.L. Johansson, M. Belew, S. Eriksson, G. Glad, O. Lind, J.L. Maloisel, N. Norrman, Preparation and characterization of prototypes for multi-modal separation media aimed for capture of negatively charged biomolecules at high salt conditions, *J Chromatogr A* 1016 (2003) 21-33.
- [10] B.L. Johansson, M. Belew, S. Eriksson, G. Glad, O. Lind, J.L. Maloisel, N. Norrman, Preparation and characterization of prototypes for multi-modal separation aimed for capture of positively charged biomolecules at high-salt conditions, *J Chromatogr A* 1016 (2003) 35-49.
- [11] T. Yang, C.M. Breneman, S.M. Cramer, Investigation of multi-modal high-salt binding ion-exchange chromatography using quantitative structure–property relationship modeling, *Journal of Chromatography A* 1175 (2007) 96-105.
- [12] S.C. Burton, D.R. Harding, Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aqueous buffers, *J Chromatogr A* 814 (1998) 71-81.
- [13] E. Boschetti, Antibody separation by hydrophobic charge induction chromatography, *Trends in Biotechnology* 20 (2002) 333-337.
- [14] S.J. Mountford, R. Daly, E.M. Campi, A. Bhagwat, A.J. Robinson, M.T.W. Hearn, N-Heterocyclic-based adsorbents for antibody purification—effect of ligand structure, *Journal of Molecular Recognition* 27 (2014) 285-295.
- [15] J. Pezzini, C. Cabanne, J.W. Dupuy, R. Gantier, X. Santarelli, A study on the nature of interactions of mixed-mode ligands HEA and PPA HyperCel using phenylglyoxal modified lysozyme, *Journal of Chromatography B* 960 (2014) 209-213.

- [16] G. Zhao, L. Zhang, S. Bai, Y. Sun, Analysis of hydrophobic charge induction displacement chromatography by visualization with confocal laser scanning microscopy, *Separation and Purification Technology* 82 (2011) 138-147.
- [17] M.A. Holstein, S. Parimal, S.A. McCallum, S.M. Cramer, Mobile phase modifier effects in multimodal cation exchange chromatography, *Biotechnology and Bioengineering* 109 (2012) 176-186.
- [18] J. Yan, Q.-L. Zhang, D.-Q. Lin, S.-J. Yao, Protein adsorption behavior and immunoglobulin separation with a mixed-mode resin based on p-aminohippuric acid, *Journal of Separation Science* 37 (2014) 2474-2480.
- [19] J.M. Mollerup, T.B. Hansen, S. Kidal, A. Staby, Quality by design—Thermodynamic modelling of chromatographic separation of proteins, *Journal of Chromatography A* 1177 (2008) 200-206.
- [20] B.K. Nfor, M. Noverraz, S. Chilamkurthi, P.D.E.M. Verhaert, L.A.M. van der Wielen, M. Ottens, High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents, *Journal of Chromatography A* 1217 (2010) 6829-6850.
- [21] S. Chilamkurthi, D.M. Sevilano, L.H.G. Albers, M.R. Sahoo, P.J.T. Verheijen, L.A.M. van der Wielen, J.L.d. Hollander, M. Ottens, Thermodynamic description of peptide adsorption on mixed-mode resins, *Journal of Chromatography A* 1341 (2014) 41-49.
- [22] K. Zhao, L. Yang, X. Wang, Q. Bai, F. Yang, F. Wang, Preparation of a novel dual-function strong cation exchange/hydrophobic interaction chromatography stationary phase for protein separation, *Talanta* 98 (2012) 86-94.
- [23] Y. Hou, S.M. Cramer, Evaluation of selectivity in multimodal anion exchange systems: A priori prediction of protein retention and examination of mobile phase modifier effects, *Journal of Chromatography A* 1218 (2011) 7813-7820.
- [24] H.S. Karkov, L. Sejergaard, S.M. Cramer, Methods development in multimodal chromatography with mobile phase modifiers using the steric mass action model, *Journal of Chromatography A* 1318 (2013) 149-155.
- [25] S. Parimal, Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York, 2014, p. 132.
- [26] H. Li, A.D. Robertson, J.H. Jensen, Very fast empirical prediction and rationalization of protein pKa values, *Proteins: Structure, Function, and Bioinformatics* 61 (2005) 704-721.
- [27] Y. Hou, C.J. Morrison, S.M. Cramer, Classification of protein binding in hydroxyapatite chromatography: synergistic interactions on the molecular scale, *Anal Chem* 83 (2011) 3709-3716.
- [28] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, *Proceedings of the National Academy of Sciences* 98 (2001) 10037-10041.
- [29] N. Chennamsetty, V. Voynov, V. Kayser, B. Helk, B.L. Trout, Prediction of Aggregation Prone Regions of Therapeutic Proteins, *The Journal of Physical Chemistry B* 114 (2010) 6614-6624.
- [30] S.D. Black, D.R. Mould, Amino acid scale: Hydrophobicity of physiological L-alpha amino acids, *Analytical biochemistry* 193 (1991) 72-82.
- [31] M. Krein, T.-W. Huang, L. Morkowchuk, D.K. Agrafiotis, C.M. Breneman, Developing Best Practices for Descriptor-Based Property Prediction: Appropriate Matching of Datasets, Descriptors, Methods, and Expectations, *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*, Wiley-VCH Verlag GmbH & Co. KGaA, 2012, p. 33-64.
- [32] B. Gelotte, Studies on gel filtration : Sorption properties of the bed material sephadex, *Journal of Chromatography A* 3 (1960) 330-342.

- [33] J.K. Inman, H.M. Dintzis, Derivatization of cross-linked polyacrylamide beads. Controlled introduction of functional groups for the preparation of special-purpose, biochemical adsorbents, *Biochemistry* 8 (1969) 4074-4082.
- [34] K. Sun, A.H. Sehon, THE USE OF POLYACRYLAMIDE GELS FOR CHROMATOGRAPHY OF PROTEINS, *Canadian Journal of Chemistry* 43 (1965) 969-976.
- [35] J.C. Janson, Adsorption phenomena on Sephadex, *J Chromatogr* 28 (1967) 12-20.
- [36] P. DePhillips, A.M. Lenhoff, Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography, *Journal of Chromatography A* 883 (2000) 39-54.
- [37] E.X.P. Almodóvar, Y. Tao, G. Carta, Protein adsorption and transport in cation exchangers with a rigid backbone matrix with and without polymeric surface extenders, *Biotechnology Progress* 27 (2011) 1264-1272.
- [38] H. Aono, I. Ilescu, D. Cecchini, S. Wood, J.T. McCue, Mitigation of chromatography adsorbent lot performance variability through control of buffer solution design space, *Journal of Chromatography A* 1318 (2013) 198-206.
- [39] T. Vicente, R. Fáber, 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, *Biotechnology and Bioengineering* 108 (2011) 1347-1359.
- [40] D. Wu, R.R. Walters, Effects of stationary phase ligand density on high-performance ion-exchange chromatography of proteins, *Journal of Chromatography A* 598 (1992) 7-13.

Figure 1 - Chromatographic retention data of the protein library on Capto MMC and Nuvia cPrime under linear salt gradient conditions (1.5 M sodium chloride over 45 column volumes). Data are arranged to group homologous proteins together and then order the library based on increasing elution salt concentration on the Nuvia cPrime resin. Non-eluting proteins are indicated by an asterisk. a) Retention data for pH 5. b) Retention data for pH 6.

Figure 2 – Ligand probability density maps for Capto MMC (C1) and Nuvia cPrime (N1 and N2) MD simulations with ubiquitin [25], and a diagram of hydrophobic residue clusters on ubiquitin.

Figure 3 - Chromatographic retention data for Capto MMC and Nuvia cPrime, reorganized based on the dominant type of hydrophobic residue clusters on each protein. (a) pH 5. (b) pH 6.

Figure 4 - Various surface potential maps of ubiquitin used for descriptor calculation.

Figure 5 - Best and cross-validated QSAR models for protein retention on Capto MMC resin at (a) pH 5 and (b) pH 6, and on Nuvia cPrime at (c) pH 5 and (d) pH 6.

521

522 Figure 6 - Correlation of basic residue property descriptors with protein retention on CM
523 Sepharose Fast Flow.

524

525 Figure 7 - Effect of surface ligand density on (a) Capto MMC resins and (b) Nuvia cPrime
526 resins.

527

528 Figure 8 – Diagram of potential adjacent ligand interactions on (a) Capto MMC resins and (b)
529 Nuvia cPrime resins.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

A.

Capto MMC

$\sim 80 \mu\text{mol/mL}_{\text{resin}}$
 0.7 ligands/nm^2

$\sim 32 \mu\text{mol/mL}_{\text{resin}}$
 0.3 ligands/nm^2

B.

Nuvia cPrime

$126 \mu\text{mol/mL}_{\text{resin}}$
 1.4 ligands/nm^2

$76 \mu\text{mol/mL}_{\text{resin}}$
 0.9 ligands/nm^2

$30 \mu\text{mol/mL}_{\text{resin}}$
 0.3 ligands/nm^2

Table 1: Summary of proteins used in the linear gradient retention studies.

Protein	PDB Code	pI	Size (kDa)	Protein	PDB Code	pI	Size (kDa)
Aprotinin	1PIT	10.5	6.5	Ribonuclease A	1RBX	9.45	13.7
Avidin	1VYO	9.69	28.7	Ribonuclease B	1RBB	8.9	13.7
Bovine Serum Albumin	3V03	5.82	66.3	Bovine Trypsin	1S0Q	10.3	23.3
Conalbumin	1AIV	6.69	75.8	Porcine Trypsin	1S81	10.5	23.5
Horse Cytochrome C	1HRC	10.25	11.7	Ubiquitin	1UBQ	6.79	8.6
Human Serum Albumin	1AO6	5.67	66.4	α -Chymotrypsin	5CHA	9.17	25.2
Lysozyme	1AKI	11.35	14.3	α -chymotrypsinogen A	2CGA	8.52	25.7
Ovalbumin	1OVA	4.9	42.7	α -lactalbumin	1F6S	5	14.1
Papain	9PAP	8.88	23.4	β -lactoglobulin A	1B0O	5.1	18.2
				β -lactoglobulin B	1BSQ	5.1	18.3

Table 2: Summary of protein surface property descriptors as described in Sections 2.5 and 2.6.

Types of Residue Clusters		Residue Cluster Descriptors
Residue Properties	Property Groups	
Acidic	Acidic-Aromatic	Number of Clusters
Basic	Acidic-Aliphatic	Size of Largest Cluster
Aromatic	Basic-Aromatic	Total Number of Residues in Clusters
Surface Potential	Types of Surface Potential Descriptors	
	Single Property	Overlapping Properties
Negative EP	Number of Clusters	Number of Strong-Strong Pairs
Positive EP	Number of Strong Clusters (Top 25% of clusters ranked by intensity)	Number of Strong-Weak Pairs
Hydrophobicity	Total Strength (Strongest 25% of clusters)	Number of Weak-Strong Pairs
Hydrophilicity	Normalized Strength of Clusters	Number of Weak-Weak Pairs
Aromatic	Mean Strength of Clusters	
Aliphatic		

Table 3: Summary of CM Sepharose Fast Flow, Capto MMC and Nuvia cPrime resin properties.

Resin	CM Sepharose Fast Flow	Capto MMC	Capto MMC ImpRes	Nuvia cPrime	Nuvia cPrime (Low Density Prototype)	Nuvia cPrime (Ultra-low Density Prototype)
Ligand
Manufacturer	GE Healthcare			Bio-Rad Laboratories		
Base Matrix	Agarose			Acrylamido Gel		
Resin Linker	Diglycerol			None		
Average Particle Size (μm)	90	75	40	70		
Ligand Density ($\mu\text{mol}/\text{mL}_{\text{resin}}$)	90-130	70-90	25-39	126	76	30
Average Pore Radius (nm)	26 [37]			68 [38]		
Total Porosity	0.84 [37]			0.85 [38]		
Average Surface Ligand Density (ligands/ nm^2)	0.8-1.1	0.6-0.8	0.2-0.3	2.8	1.8	0.6

Table 4: Summary of QSAR models for protein retention on Capto MMC and Nuvia cPrime at pH 5 and 6.

		pH 5		pH 6	
		Capto MMC	Nuvia cPrime	Capto MMC	Nuvia cPrime
Model Type		PLS	MLR	PLS	MLR
R^2		0.94	0.91	0.91	0.88
Cross-validated R^2		0.87	0.79	0.8	0.75
Selected Descriptors					
Single Property	Mean Strength		Negative EP	Hydrophobic	Aromatic
	Normalized Strength		Negative EP		
Overlapping Properties	Weak/Weak	Positive EP / Negative EP	Aromatic / Aliphatic	Aromatic / Aliphatic	Aromatic / Aliphatic
Residue Clusters	# of Clusters	Basic	Basic	Basic	Basic
	Largest Cluster Size	Aromatic	Aromatic	Aromatic	Aromatic
		Aliphatic-Acidic		Acidic	Acidic
		Aromatic-Acidic		Aromatic-Acidic	