

# Are Unfolding Pathways of Multi-Domain Proteins on Chromatographic Surfaces Related to Solution Phase Unfolding Pathways?

A Dissertation Presented to the Faculty  
of the School of Engineering and Applied Science  
University of Virginia

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy (Chemical Engineering)

By  
Adrian Gospodarek  
April 2013

## Abstract

Hydrophobic interaction chromatography (HIC) purifies proteins based on their apparent hydrophobicities, making it a valuable tool in downstream purification. However, interactions between buried hydrophobic residues and HIC surfaces can cause conformational changes that lead to compromised chromatographic behavior and yield losses. Recent observations with proteins on ion-exchange chromatography (IEC) surfaces demonstrate that electrostatic interactions may also drive conformational changes. Unfolding of proteins on HIC and IEC surfaces has particular implications for multi-mode chromatography (MMC), which involves ligands with both hydrophobic and charged functionalities. This dissertation extends previous investigations of protein unfolding in chromatography, identifies how protein structure on HIC surfaces change under increasing denaturing conditions, relates these structures to solution unfolded states, characterizes the self-association behavior of a therapeutic protein identified in IEC, and investigates how a model protein unfolds on an IEC and MMC surface.

The thermal unfolding of three model multi-domain proteins on HIC surfaces of differing hydrophobicities was investigated with hydrogen-exchange mass spectrometry (HXMS). The melting temperature of the proteins is lowered by different amounts when the proteins are adsorbed on the different surfaces. Further, the structures of the proteins on the surface share similar characteristics to proteins partially unfolded thermally and with chemical denaturants in solution. This suggests protein unfolding on HIC surfaces can be related to how a protein unfolds in solution.

The utility of HXMS was also demonstrated in identifying the regions involved in the self-association of a therapeutic protein (Protein X). Protein X displays split peak behavior during IEC whereas an analogous protein similar to Protein X does not. HXMS was used to show that a modified region not present on the analogous protein is involved in the intermolecular contact between associated molecules. However, a conformational change was also detected in a binding pocket of the unmodified region suggesting self-association may involve insertion of the modified region from one molecule into this binding pocket of another.

Finally, protein unfolding of a model protein on IEC and MMC surfaces was investigated with HXMS. Unfolding was found to depend on pH and the surface type with unfolding only occurring on the IEC and MMC surfaces with cationic moieties. Not all regions of the protein are affected the same way as certain regions unfold more than others. No evidence of unfolding was observed on the MMC surfaces under conditions where electrostatics were minimized, suggesting the cationic moieties of the ligands and not the hydrophobic ones are responsible for unfolding.

## Acknowledgments

I would like to first and foremost acknowledge my family. This includes my parents: Marta and Jurek Gospodarek, my grandparents: Jadwiga and Stefan Ciolek, and two of the best people I have ever met in Charlottesville (or anywhere else): Sara Davis and Joe Kozlowski. All of you provided invaluable company, entertainment, and support that if nonexistent would have made my thesis difficult to complete.

This thesis of course would not have been possible without the guidance and support of my advisor, Erik Fernandez, and co-advisor, Mr. O'Connell. Their valuable perspective and insight into research was only matched by their genuine interest and effort in improving my abilities as a researcher and thinker, but most importantly: helping me become a better person.

I would also like to thank Professor Shirts and Professor Acton for being on my committee. Professor Shirts: your "strong" line of questioning during my proposal helped me transform my research efforts from just a presentation of data to a compelling story. Professor Acton: Thank you for providing your unique insight for my dissertation. A special thanks to Dr. André Dumetz for not only agreeing to be on my committee and providing the collaboration resulting in Chapter 3 but for also mentoring me during my summer internship at GlaxoSmithKline. You certainly went above and beyond what was required in your role as supervisor.

This thesis would also not have been possible without the support and experiences with past lab members: Shin Ahn, Joeseeph Costanzo, Jing Guo, Jacob Jordan, Wei Qi, Aming Zhang, and Shaozie Zhang. A special acknowledgment to Rob Deitcher: one of the most intelligent and insightful persons I have ever met.

Thank you also to all the undergraduate assistants who assisted me in lab and beared with me while I was developing my own mentoring skills: Diana Hiser, Marissa Smatlak, and Weitong Sun.

Finally, I would like to acknowledge and thank all others who supported me and helped make my time enjoyable in Charlottesville: Brittany Boser, Todd Bowman, Oliver Daniel, Derek Falcone, Rosemary and Lucas Cox-Galhotra, Angie Hasseman, David Hibbitts, Laura King, Arnaldo Marrero, Ernie Perez, Carolina Ramoa, Jason Reck, Xiapou Wang, and many more...

## Table of Contents

Abstract.....	i
Acknowledgments .....	iii
Table of Contents .....	v
List of Figures and Tables .....	vii
List of Symbols .....	xvii
1. Background and Introduction .....	1
2. Unfolding Structures of Multi-domain Proteins on a Series of HIC Surfaces .....	5
2.1. Background and Introduction.....	5
2.2. Materials and Methods .....	8
2.2.1. Materials .....	8
2.2.2. Temperature studies.....	8
2.2.3. Guanidine studies .....	10
2.2.4. HPLC-MS .....	10
2.2.5. Size exclusion chromatography .....	13
2.3. Theory .....	15
2.3.1 Isotopic (Deuterium) Labeling.....	15
2.3.2. Measuring and Calculating Fractions of Labeled Peptide.....	16
2.3.3. Unfolding Free Energies and Rates.....	17
2.4. Results and Discussion .....	22
2.4.1. Unfolding effects of chromatographic surface chemistry and temperature on transferrin .....	22
2.4.2. Unfolding effects of chromatographic surface chemistry and temperature on antitrypsin.....	28
2.4.3. Relating partially unfolded states in solution and on chromatographic surfaces .....	32
2.4.4 Unfolding effects of surface and GdnHCl .....	37
2.4.5. Solvent protection increase on the chromatographic surface .....	50
2.4.6. Superposition of changes in solvent protection on the surface .....	65
2.5. Conclusions .....	79
3. Protein Self-Association: Identification of Association Sites and Related Conformational Changes by Hydrogen Exchange .....	80
3.1. Background and Introduction.....	80
3.2. Materials and Methods .....	82
3.2.1 Materials .....	82
3.2.2. Hydrogen Exchange Mass Spectrometry.....	83
3.2.3. Isothermal Titration Calorimetry.....	85
3.2.4. Size-Exclusion Chromatography.....	86
3.3. Results and Discussion .....	87
3.3.1 HX-MS Measurements of Conformational Changes.....	87
3.3.2. Isothermal Titration Calorimetry of Protein X Samples.....	100
3.3.3. Size-Exclusion Chromatography.....	102
3.4. Conclusions .....	104

4. Unfolding of a Model Protein on Ion Exchange and Multi-Mode Chromatography Surfaces .....	105
4.1. Background and Introduction.....	105
4.2. Materials and Methods .....	109
4.2.1. Materials .....	109
4.2.2. pH Studies.....	109
4.2.3. Salt effect studies.....	112
4.2.4. HPLC-MS .....	112
4.3. Theory .....	115
4.3.1. Measuring and Calculating Fractions of Labeled Peptide.....	115
4.3.2. Fast, medium, and slow exchangers.....	115
4.4. Results and Discussion .....	118
4.4.1. Whole protein studies.....	118
4.4.2. Digestion studies on SP Sepharose FF and Capto MMC at pH 4.5.....	127
4.4.3. Digestion studies on SP Sepharose FF at pH 4.0.....	137
4.4.4. Relating BSA acid-induced denaturation in solution and on the surface .....	144
4.4.5. Digestion studies on Capto MMC at pH 5.5 with varying ammonium sulfate concentrations.....	146
4.5. Conclusions .....	150
5. Conclusions and Future Work.....	152
5.1. Conclusions .....	152
5.2. Future Work.....	155
6. References.....	159
Appendix A. Supplementary Data.....	165

## List of Figures and Tables

- Figure 2.1. Representation of Millipore’s Ultrafree-MC centrifugal filter unit taken from Vesely et al. (2012). ..... 9
- Figure 2.2. Schematic of valve and column setup used to deliver sample to MS. (1) Sample is injected manually with glass syringe into a 200  $\mu$ L sample loop. (2) Sample is digested with Porcine pepsin, resulting peptides are trapped on a C8 column, and salt is removed. (3) Flow path in valve changes so that HPLC (high-performance liquid chromatography) elutes peptides off C8 column to deliver to C18 column for further resolution before delivery to the MS. Blue lines and arrows depict manual sample injection. Green lines and arrows depict flowpath of sample/peptides through each step. Red lines and arrows depict inactive flow at each step. Dashed box depicts temperature controlled region of 1.0°C..... 12
- Figure 2.3. Determination of  $k_u$ ,  $m$ , and  $\Delta G_{u,o}$  for a reporter peptide. (a)  $M_N$  and  $M_U$  are measured along with their peak intensities from the mass spectrum to calculate  $M_{NU}$  and  $D/N$ . (b)  $D/N$  (diamonds) vs. GdnHCl concentration is plotted and fitted to a four-parameter ( $N_{D/N}$ ,  $U_{D/N}$ ,  $C_{1/2}$ , and  $m$ ) sigmoidal equation (Equation 2.9). (c)  $[N]$  (circles) is calculated at various labeling times and GdnHCl concentrations from the sigmoidal equation. (d) For each GdnHCl concentration,  $[N]$  is regressed vs. labeling time (Equation 2.11). Slope of regressed line is  $-k_u$  and y-intercept is  $\Delta G_u$  (squares). (e)  $\Delta G_u$  is regressed vs. GdnHCl concentration (Equation 2.13). Slope of regressed line is  $m$  and y-intercept is  $\Delta G_{u,o}$ . 95% confidence intervals of all regressed parameters are propagated to  $m$  and  $\Delta G_{u,o}$ ..... 18
- Figure 2.4. Left: Native structure of transferrin (PDB ID 2HAV) highlighting the N-terminal domain (blue) and C-terminal domain II (green). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts Cysteines of transferrin. Grey depicts missing sequence coverage..... 23
- Figure 2.5. Left: Thermal unfolding of different reporter peptides of transferrin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM  $PO_4$ , 1.5M  $(NH_4)_2SO_4$ , pH 7.0 buffer at 22 to 52°C. Samples were labeled with deuterated buffer for 10 minutes. The  $D/N$  for Residues 67-81 is shown in panel (a). The  $D/N$  for Residues 125-131 is shown in panel (b). Right: Location of residues 67-81 and the other residues that follow the general unfolding pattern are shown in red on the native structure (PDB ID 2HAV). The exception to the general unfolding pattern, Residues 125-131, is shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C. .... 25
- Figure 2.6. Gradient elution of transferrin on Phenyl 650M (blue), Butyl 650M (red), and Hexyl 650C (green). 500  $\mu$ g of sample were loaded onto 0.5 mL columns at a flow rate of 1 mL/min with 1.5M  $(NH_4)_2SO_4$ , 25mM  $PO_4$  (high salt) at pH 7.0 and room



temperature. Columns were washed with 2 CVs of high salt and protein was eluted with a 25 CV gradient from high salt to 25mM PO<sub>4</sub> at pH 7.0 and room temperature.. 27

Table 2.1. DBC and retention volume comparison on Tosoh HIC media ..... 27

Figure 2.7. Native structure of transferrin (PDB ID 2QUG) and location of reporter peptides obtained in this study (highlighted in green). Grey depicts missing sequence coverage. .... 29

Figure 2.8. Left: Labeling patterns of different reporter peptides of antitrypsin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 92°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 185-189 is shown in panel (a). The *D/N* for Residues 266-273 is shown in panel (b). Right: Location of residues 185-189 and the other residues that unfold more on Butyl 650M and Hexyl 650C than on Phenyl 650M are shown in blue on the native structure (PDB ID 2QUG). Residues 266-273 and the other residues that unfold to statistically equal extents on all three surfaces are shown in red. Residues that were not observed with enhance labeling on the surfaces are shown in green. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C..... 30

Figure 2.9. Difference in *D/N*,  $\Delta D/N$ , between human serum transferrin adsorbed on media (Phenyl 650M – blue, Hexyl 650C – green, Butyl 650M – red) and in solution in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer.  $\Delta D/N$  values represent the average reporter peptide unfolding for each domain weighted by the number of residues in each peptide. Data is averaged and weighted for all reporter peptides in the N-terminal domain (left) and C-terminal domain (right). Error bars represent sample propagated 95% confidence intervals for  $\Delta D/N$  from triplicate data points collected on Phenyl 650M at 22°C..... 33

Figure 2.10. Folded regions of antitrypsin in the guanidine denatured intermediate (left) and when adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22°C (right). Green represents those regions folded in the solution intermediate (left) and on all media (right). Yellow represents those regions that unfold on Hexyl 650C only. Blue arrows designate regions that show similar stability between the two cases..... 35

Figure 2.11. Unfolded regions of antitrypsin in the guanidine denatured intermediate (left) and when adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22°C (right). Red represents those regions unfolded in the solution intermediate (left) and on all media (right). Orange represents those regions that unfold on Butyl 650M and Hexyl 650C to a greater extent than on Phenyl 650M. Blue arrows designate regions that show similar stability between the two cases..... 36

Figure 2.12. Unfolding of residues 223 to 259 in the N-terminal domain of transferrin in solution (blue, diamonds), and on Hexyl 650C (red, squares) in 25mM PO<sub>4</sub>, 1.5M

$(\text{NH}_4)_2\text{SO}_4$ , 0 to 5M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 2HAV) are shown in red. The N-terminal and C-terminal domain of transferrin are shown in blue and green, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines..... 39

Table 2.2. Characteristics of labeling patterns for transferrin Residues 223-259 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9). ..... 40

Figure 2.13. Unfolding of residues 554 to 589 in the C-terminal domain of transferrin in solution (blue, diamonds), and on Hexyl 650C (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 5M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 2HAV) are shown in red. The N-terminal and C-terminal domain of transferrin are shown in blue and green, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines..... 41

Table 2.3. Characteristics of labeling patterns for transferrin Residues 554-589 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9). ..... 41

Figure 2.14. Unfolding of residues 127 to 137 in Domain I of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines..... 43

Table 2.4. Characteristics of labeling patterns for BSA Residues 127-137 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9). ..... 44

Figure 2.15. Unfolding of residues 333 to 340 in Domain II of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines..... 45

Table 2.5. Characteristics of labeling patterns for BSA Residues 333-340 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9). ..... 45

Figure 2.16. Unfolding of residues 529 to 543 in Domain III of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues

on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines..... 47

Table 2.6. Characteristics of labeling patterns for BSA Residues 529-543 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9)..... 47

Table 2.7. Unfolding Rates and Free Energies of BSA on Phenyl 650M..... 49

Figure 2.17. Left: Native structure of BSA (PDB ID 3V03) highlighting Domain I (red), Domain II (green), and Domain III (blue). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts Cystines of BSA. Grey depicts missing sequence coverage..... 51

Figure 2.18. Thermal unfolding and solvent protection increase of different regions of BSA in solution (blue, diamonds), on Butyl 650M (red, squares), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 82°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 307-313 is shown in panel (a). The *D/N* for Residues 3-14 is shown in panel (b). Right: Location of Residues 307-313 and the other residues with similar unfolding pattern are shown in red on the native structure (PDB ID 3V03). Residues 3-14 and the other residues where no increased labeling in solution is observed are shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Butyl 650M at 22°C..... 53

Figure 2.19. SEC chromatograms of BSA in solution (left) and eluted from Butyl 650M (right) at temperatures of 52°C (a), 62°C (b), 72°C (c), and 82°C (d). Samples were incubated for 1 hour, eluted with 200 μL of desorption buffer, diluted with 500 μL of working buffer, and loaded on a TSK gel G3000SWXL column for analysis. .... 56

Figure 2.20. SEC chromatograms of BSA eluted from Phenyl 650M at temperatures of 52°C (top left), 62°C (top right), 72°C (bottom left), and 82°C (bottom right). Samples were incubated for 1 hour, eluted with 200 μL of desorption buffer, diluted with 500 μL of working buffer, and loaded on a TSK gel G3000SWXL column for analysis. .... 57

Figure 2.21. Overlap of solvent-protected BSA residues and predicted aggregation hot spots. Residue groups of BSA that unfold on Butyl 650M (*D/N* > 0.5) at lower temperatures and become more solvent-protected (*D/N* < 0.1) at 80°C are highlighted in red. Residues predicted to be aggregation hot spots by AGGRESCAN and PASTA are shown with solid and dashed lines, respectively..... 59

Figure 2.22. Effect of temperature on different reporter peptides of transferrin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 52°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 251-259 is shown in panel (a). The *D/N* for Residues 180-188 is shown in

panel (b). Right: Location of Residues 251-259 and the other residues with increased solvent protection on Phenyl 650M are shown in red on the native structure (PDB ID 2HAV). Residues 180-188 and the other residues with increased solvent protection on all three surfaces are shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C..... 62

Figure 2.23. SEC chromatograms of transferrin eluted from Butyl 650M (left) and Hexyl 650C (right) at temperatures of 22°C (a), 32°C (b), and 52°C (c). Sample were incubated for 1 hour, eluted with 200  $\mu$ L of desorption buffer, diluted with 500  $\mu$ L of working buffer and loaded on a TSK gel G3000SWXL column for analysis. .... 64

Table 2.8.  $D/N$  differences for different comparisons to solution labeling at various temperature offsets. .... 67

Table 2.9.  $D/N$  differences for different surface-surface labeling comparisons at various temperature offsets. .... 68

Figure 2.24. Difference in  $D/N$ ,  $\Delta D/N$ , between transferrin adsorbed on Butyl 650M at 32°C and Hexyl 650C at 32°C (a), adsorbed on Butyl 650M at 42°C and Hexyl 650C at 42°C (b), adsorbed on Hexyl 650C at 32°C and on Butyl 650M at 22°C (c), and adsorbed on Hexyl 650C at 42°C and on Butyl 650M at 32°C (d)..... 69

Figure 2.25. Difference in  $D/N$ ,  $\Delta D/N$ , between human serum transferrin adsorbed on Butyl 650M at 32°C and Phenyl 650M at 52°C (a), adsorbed on Butyl 650M at 42°C and in solution at 72°C (b), adsorbed on Phenyl 650M at 52°C and in solution at 72°C (c), and adsorbed on Hexyl 650C at 52°C and in solution at 72°C (d)..... 70

Table 2.10.  $D/N$  differences for different comparisons to solution labeling at various temperature offsets. .... 72

Table 2.11.  $D/N$  differences for different surface-surface labeling comparisons at various temperature offsets. .... 73

Figure 2.26. Difference in  $D/N$ ,  $\Delta D/N$ , between antitrypsin adsorbed on Phenyl 650M at 52°C and on Hexyl 650C at 32°C (left), and adsorbed on Phenyl 650M at 72°C and on Hexyl 650C at 52°C (right)..... 74

Figure 2.27. Difference in  $D/N$ ,  $\Delta D/N$ , between antitrypsin adsorbed on Butyl 650M at 72°C and on Hexyl 650C at 52°C (left), and adsorbed on Butyl 650M at 92°C and on Hexyl 650C at 72°C (right)..... 75

Table 2.12.  $D/N$  differences for different solution-surface and surface-surface labeling comparisons at various temperature offsets..... 76

Figure 2.28. Difference in  $D/N$ ,  $\Delta D/N$ , between BSA adsorbed on Hexyl 650C at 42°C and on Butyl 650M at 22°C (a), adsorbed on Hexyl 650C at 52°C and Butyl 650M at 32°C (b),

- adsorbed on Hexyl 650C at 62°C and on Butyl 650M at 42°C (c), and adsorbed on Butyl 650M at 72°C and on Hexyl 650C at 52°C (d)..... 77
- Figure 3.1. Protein X representation. The unmodified region is represented by a blue square and the modified region is represented by a brown triangle..... 82
- Table 3.1. Sample Conditions for HX-MS Studies ..... 83
- Figure 3.2.  $\Delta D/N$  values between concentrations of 0.8 and 16.6 mg/mL for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 2 minutes (top) and 10 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences..... 89
- Figure 3.3.  $\Delta D/N$  values between concentrations of 0.8 and 16.6 mg/mL for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 30 minutes (top) and 120 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences..... 90
- Figure 3.4.  $\Delta D/N$  values between concentrations of 0.5 and 18.9 mg/mL for Protein X reporter peptides in pH 5.2, 30 mM sodium succinate at labeling times of 2 minutes (top) and 10 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences..... 92
- Figure 3.5.  $\Delta D/N$  values between concentrations of 0.5 and 18.9 mg/mL for Protein X reporter peptides in pH 5.2, 30 mM sodium succinate at labeling times of 30 minutes (top) and 120 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences..... 93
- Figure 3.6.  $D/N$  differences between concentrations of 0.8 and 16.6 mg/mL at pH 6.4 (top) and between concentrations of 0.5 and 18.8 mg/mL at pH 5.2 (bottom) for reporter peptides in the modified region and beginning of unmodified region. Red, yellow, green, or blue represent peptides that at high concentrations show a decrease in  $D/N$  of 0.15 or greater, from 0.10 to 0.15, from 0.05 to 0.10, and from 0 to 0.05, respectively..... 95
- Figure 3.7. Schematic representation of different association mechanisms based on changes in solvent protection. Shown are possible mechanisms of dimerization without and with insertion of the modified region into a binding pocket. Also shown is a possible structure for oligomers..... 97
- Figure 3.8.  $\Delta D/N$  values between Protein X at 16.6 mg/mL with and without modified region for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 30 seconds (top) and 2 minutes (bottom). Error bars represent 95%

- confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences. .... 98
- Figure 3.9.  $\Delta D/N$  values between Protein X at 16.6 mg/mL with and without modified region for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 10 minutes (top) and 30 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences. .... 99
- Table 3.2. Thermodynamic Parameters of Protein X Dissociation ..... 100
- Figure 3.10. Calorimetric data for the dissociation of Protein X. Raw data (a) and integrated injection heats (c) for injections of 21.5 mg/mL Protein X, 30 mM sodium succinate, pH 6.4 solution. Raw data (b) and integrated injection heats (d) for injections of 20.2 mg/mL Protein X, 30 mM sodium succinate, pH 5.2 solution. .... 101
- Figure 3.11. SEC chromatograms of Protein X at various concentrations below (a) and above (b) 0.50 mg/mL in 10 mM sodium succinate, pH 5.8 solution. .... 103
- Figure 4.1. Solvent exposure changes of BSA in solution (diamonds), adsorbed on SP Sepharose FF (left, squares), and adsorbed on Q Sepharose FF (right, triangles) in 50 mM citrate at pHs 3.0 to 4.5 and in 50 mM phosphate at pHs 6.5 to 8.0. In order to obtain measurable levels of deuterium uptake at each pH different labeling times were used for pH 3.0 (100 minutes), 3.5 (50 minutes), 4.0 (10 minutes), 4.5 (5 minutes), and 6.5 to 8.0 (5 minutes) as described in Methods (Section 4.2.2). Error bars represent sample 95% confidence intervals estimated based on triplicate data points collected on SP Sepharose FF at pH 4.5. .... 119
- Figure 4.2. Solvent exposure changes of BSA under various conditions in 50 mM citrate at pHs 3.0 to 4.5 and in 50 mM phosphate at pHs 5.0 to 8.5. Top: BSA in solution (diamonds), adsorbed on Capto MMC (left, circles), and adsorbed on Capto Adhere (right, horizontal lines). Bottom: Solvent exposure changes of BSA adsorbed on Capto MMC (left, circles), adsorbed on SP Sepharose FF (left, squares), adsorbed on Capto Adhere (right, horizontal lines), and adsorbed on Q Sepharose FF (right, triangles). In order to obtain measurable levels of deuterium uptake at each pH different labeling times were used for pH 3.0 (100 minutes), 3.5 (50 minutes), 4.0 (10 minutes), 4.5 (5 minutes), and 6.5 to 8.0 (5 minutes) as described in Methods (Section 4.2.2). Error bars represent sample 95% confidence intervals estimated based on triplicate data points collected on SP Sepharose FF at pH 4.5. .... 122
- Figure 4.3. Solvent exposure changes of BSA in solution (diamonds), adsorbed on Capto MMC (left, squares), and adsorbed on Capto Adhere (right, triangles) in 50 mM phosphate and varying concentrations of ammonium sulfate at pH 5.5. Samples were labeled with deuterated buffer for 20 minutes. Error bars represent sample 95% confidence intervals of triplicate data points collected on SP Sepharose FF at pH 4.5. .... 124

Figure 4.4. Solvent exposure changes of BSA in solution (diamonds), adsorbed on SP Sepharose FF (left, squares), and adsorbed on Capto MMC (right, triangles) in 50 mM citrate at pH 4.5. Samples were labeled with deuterated buffer for 5 to 120 minutes. Error bars represent estimated 95% confidence intervals based on triplicate data points collected on SP Sepharose FF at pH 4.5 and 5 minutes labeling. ....126

Figure 4.5. Left: Native structure of BSA (PDB ID 3V03) highlighting Domain I (red), Domain II (green), and Domain III (blue). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts cysteines of BSA. Grey depicts missing sequence coverage.....128

Table 4.1. Four classes of unfolding behavior of BSA on SP Sepharose FF and Capto MMC at pH 4.5 .....129

Figure 4.6. Class 1 peptides in which BSA unfolds on SP Sepharose FF and Capto MMC to equal extents. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The *D/N* for Residues 127-137 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The *D/N* for Residues 200-209 is shown in panel (b) and their location on the native structure is shown in orange on the right.....130

Figure 4.7. Class 2 peptides in which unfolding occurs on SP Sepharose FF and Capto MMC but more on SP Sepharose FF. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The *D/N* for Residues 39-45 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The *D/N* for Residues 307-313 is shown in panel (b) and their location on the native structure is shown in orange on the right.....131

Figure 4.8. Class 3 peptides in which unfolding occurs only on SP Sepharose FF. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The *D/N* for Residues 18-32 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The *D/N* for Residues 49-70 is shown in panel (b) and their location on the native structure is shown in orange on the right. Also shown in blue on the native structure are the only other residues (529-541) that follow this behavior.....132

Figure 4.9. Class 4 peptide in which less labeling is observed on the surfaces. Unfolding of Residues 575-583 of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The *D/N* for

Residues 575-583 is shown on the left and their location on the native structure (PDB ID 3V03) is shown in red on the right. ....	134
Table 4.2. Change in fast, medium, and slow exchanging residues and unfolding rate from solution to surface for the 4 classes of peptides.....	135
Table 4.3. Three classes of unfolding behavior of BSA on SP Sepharose FF at pH 4.0.....	138
Figure 4.10. Labeling patterns of the three different classes of BSA peptides in solution (blue, diamonds) and on SP Sepharose FF (red, squares) in 50mM Citrate, pH 4.0 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The $D/N$ for Residues 18-32 is shown in panel (a) and their location on the native structure (PDB ID 3V03) along with other Class A peptides are shown in red on the right. The $D/N$ for Residues 39-45 is shown in panel (b) and their location on the native structure along with other Class B peptides are shown in blue on the right. The $D/N$ for Residues 575-583 is shown in panel (c) and their location on the native structure along with other Class C peptides are shown in green on the right. ....	140
Figure 4.11. Difference in $D/N$ from pH 4.0 to 4.5 for different regions of BSA in solution (blue) and on SP Sepharose FF (red) in 50mM Citrate buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. Positive values indicate more labeling at pH 4.0 while negative values indicate less labeling. The difference in $D/N$ for Residues 2-7 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The difference in $D/N$ for Residues 219-226 is shown in panel (b) and their location on the native structure is shown in orange on the right. The difference in $D/N$ for Residues 324-329 is shown in panel (c) and their location on the native structure is shown in green on the right. All other residues that follow 1 of the 3 patterns are also shown on the native structure in their representative color..	142
Table 4.4. Three classes of unfolding behavior of BSA on Capto MMC at pH 5.5.....	146
Figure 4.12. Three labeling patterns observed for different regions of BSA in solution (blue, diamonds), and on Capto MMC (green, triangles) in 50mM phosphate, pH 5.5 buffer at varying ammonium sulfate concentrations. Samples were labeled with deuterated buffer for 20 minutes. Panel (a) shows the $D/N$ for an example, Residues 18-32, of the general pattern where no unfolding on surface relative to solution and no effect of ammonium sulfate is observed. Panel (b) shows the $D/N$ for an example, Residues 2-7, of the first exceptional pattern in which unfolding is observed on the surface but without any effect from ammonium sulfate. Panel (c) shows the $D/N$ for the one group of residues, Residues 219-226, of the second exceptional pattern in which both unfolding on the surface and ammonium sulfate effects are observed. The location of the groups of residues in the general pattern, and the first and second exceptional labeling pattern are shown on the native structure (PDB ID 3V03) on the right in green, red, and blue, respectively.....	148
Table A1. Transferrin Reporter Peptides .....	165



Table A2. Antitrypsin Reporter Peptides.....	166
Table A3. BSA Reporter Peptides.....	167
Table A4. Unfolding Rates and Free Energies of Transferrin on Hexyl 650C.....	168

## List of Symbols

$A_f$	Fraction of fast exchangers
$A_m$	Fraction of medium exchangers
$A_s$	Fraction of slow exchangers
$C_{1/2}$	Concentration of GdnHCl at midpoint of denaturation curve
$C_{dimer}$	Concentration of dimers
$C_{monomer}$	Concentration of monomers
$C_{species}$	Concentration of the different species in self-association
D	Deuterium molecule
$D$	Number of deuterated amides in a peptide
$D/N$	$D/N$ value of a peptide
$D/N(t_{longest})$	$D/N$ value at longest labeling time
$D/N(t_{shortest})$	$D/N$ value at shortest labeling time
$\Delta G_{diss}$	Free energy of dissociation
$\Delta G_u$	Unfolding free energy at a given GdnHCl concentration
$\Delta\Delta G_u$	Effect of surface on unfolding free energy
$\Delta G_{u,ads}$	Unfolding free energy when adsorbed
$\Delta G_{u,sol}$	Unfolding free energy in solution
$\Delta G_{u,o}$	Unfolding free energy in absence of GdnHCl
H	Hydrogen molecule
$\Delta H_{diss}$	Enthalpy of dissociation

$k_f$	Rate constant for protein folding
$k_m$	Rate constant for unfolding of medium exchangers
$k_s$	Rate constant for unfolding of slow exchangers
$k_{int}$	Rate constant for intrinsic exchange of hydrogen for deuterium
$k_{obs}$	Observed first order rate constant of exchange
$k_u$	Rate constant for protein unfolding
$\Delta k_u$	Effect of surface on rate constant for protein unfolding
$k_{u,ads}$	Rate constant for protein unfolding when adsorbed
$k_{u,sol}$	Rate constant for protein unfolding in solution
$K_{diss}$	Dissociation equilibrium constant
$m$	Effect of denaturant concentration on stability
$m_0$	Non-deuterated mass of a peptide
$m_{100}$	Fully-deuterated mass of a peptide
$m_t$	Mass of a peptide at a given labeling time
$M_N$	Mass of native peak under the EX1 exchange regime
$M_U$	Mass of unfolded peak under the EX1 exchange regime
$M_{NU}$	Weighted average mass of native and unfolded peaks
$MW_{species}$	Molecular weight of the different species in self-association
$[N]$	Concentration or fraction of native peptide at a given labeling time
$[N_0]$	Concentration or fraction of native peptide at time zero
$N$	Residue or protein in native state
$N$	Number of exchange-competent residues in a peptide
$N_{D/N}$	$D/N$ value of native peptide in EX1 exchange regime

$p$	Slope of unfolding transition in denaturation curve
$R$	Universal gas constant
$\Delta S_{diss}$	Entropy of dissociation
$T$	Temperature
$U$	Residue or protein in unfolded state

# 1. Background and Introduction

The past two decades in the pharmaceutical industry have seen a new growth towards therapeutic proteins and antibodies. Although the industry is still dominated by small molecules, companies are increasingly investing resources into these biopharmaceuticals (biopharm). Pharmaceutical pipelines are increasingly becoming more populated with these therapeutic proteins and antibodies.

With this shift towards biopharm, new challenges emerge for companies in developing cost-effective techniques that can mass produce therapeutic proteins while ensuring product quality and safety for patients. The production of therapeutic proteins requires expression and growth in bacterial or mammalian cells. While these natural “bioreactors” are the most efficient means of producing the drug targets, a large number of byproducts are formed during the production. These byproducts, host cell proteins (HCPs) specifically, which are necessary for cell growth and target protein production, need to be reduced to the parts per million level to ensure no immunogenic response when the target protein is delivered to a patient. In addition, costly and advanced purification methods with well-defined operational windows are required because HCPs are difficult to separate from the target protein. Finally, viral clearance and reduction of high-molecular weight aggregates need to be addressed to further ensure product quality.

Chromatography is the method of choice for achieving the purification and is often used in multiple steps. Much of research has been dedicated to optimizing the different operational windows to achieve target purities and yields at a minimum cost.

Unfortunately, one of the more proven techniques, hydrophobic interaction chromatography (HIC), can cause the target protein to irreversibly unfold and lose its therapeutic function. Much of the research has also been dedicated to better understand the phenomena and to develop models for mobile and stationary properties that affect unfolding.

It would be ideal if protein unfolding in HIC could be predicted only from knowledge of the protein's solution stability. However, the "hydrophobicity" of the HIC surface has also been shown to affect unfolding of the protein. In addition there is increasing evidence that unfolding occurs on surfaces other than those used in HIC. For example, unexpected elution profiles have been observed in ion exchange chromatography (IEC).

It is important to isolate solution effects from surface effects to properly determine the extent to which the surface is responsible for any unexpected elution behavior. The overall goal of this thesis is to advance the understanding of how protein behavior in solution and properties of the surface can be related to predicting unfolding in chromatographic separations.

In Chapter 2, I hypothesize that the stability of multi-domain proteins in solution is similar to that on HIC surfaces. Specifically, proteins that unfold via multi-state mechanisms in solution will have partially unfolded states on HIC surfaces where the regions less stable in solution are unfolded while those more stable retain native structure. This hypothesis is tested by monitoring changes in solvent exposure of a set of model proteins as they unfold. Hydrogen-exchange mass spectrometry (HXMS) is used to detect deuterium uptake levels in solution and on a series of increasingly hydrophobic surfaces as temperature or chemical denaturant concentration is increased.

In Chapter 3, I hypothesize that HXMS can help characterize unexpected self-association behavior of a therapeutic protein that has important implications in IEC separations. In this study, pH, protein concentration, and ionic strength are found to have effects on self-association prior to any IEC. Size-exclusion chromatography (SEC) and isothermal titration calorimetry (ITC) can help determine how each of these variables affect self-association, but do not provide information about the regions involved and the inherent mechanism for association. The hypothesis that HXMS can provide useful information in this case is investigated by identifying the residues involved in the association sites along with any conformational changes related to self-association. A mechanism for the self-association of the protein is then proposed.

Finally in Chapter 4, I investigate the occurrence of protein unfolding on IEC and multi-modal chromatography (MMC) surfaces. In this chapter, I hypothesize that electrostatic interactions between charged residues of proteins and the ionic moieties of these surfaces can lead to protein unfolding on the surface. This hypothesis is tested by observing the unfolding behavior of a model multi-domain protein on different IEC and MMC surfaces with HXMS. The effect of electrostatic interactions on unfolding is examined by varying pH and ionic strength.

Through the studies in this thesis, the overall objective of increasing the knowledge of the relationship between the stability of a protein in solution and its unfolding behavior on chromatographic surfaces will be advanced. These studies will help advance the goal of developing the modeling needed for predicting if and how specific proteins will unfold on chromatographic surfaces with minimal experiments. Ultimately achieving this lofty goal

will require many additional steps. In my conclusions and future works section, I will outline recommendations I believe will be best suited in achieving this.



## **2. Unfolding Structures of Multi-domain Proteins on a Series of HIC Surfaces**

### ***2.1. Background and Introduction***

Hydrophobic interaction chromatography (HIC) purifies proteins based on their apparent hydrophobicities, making it a valuable tool in downstream purification (Queiroz 2001; Hjerten 1973; Mccue 2009). HIC stationary phases come in a variety of backbone and hydrophobic ligand chemistries leading to differing degrees of hydrophobicity and protein binding capabilities. Industrial applications focus toward the more hydrophobic HIC resins due to their ability to effectively bind proteins with lower salt concentration requirements. However, these same resins can lead to greater protein unfolding and compromised recoveries (Ingraham et al. 1985; Kato et al. 1984; Fausnaugh et al. 1984).

HIC resins target the hydrophobic amino acids of proteins during the adsorption process, many of which tend to populate the interior of the protein, or its hydrophobic core. This interaction can cause a change in protein conformation, and if significant and irreversible, the protein can lose function, and therapeutic efficacy. Furthermore, yield losses can also occur even with reversible unfolding, as unfolded molecules will elute differently from the folded molecules (Wu et al. 1986; Benedek 1988). Designing HIC processes that optimize selectivity while preventing the effects of unfolding remains a challenge for many proteins, especially considering the many variables that influence adsorption, including resin, salt type and concentration, pH, and temperature (Gagnon et al. 1995). It would be valuable to have predictive approaches capable of identifying proteins that are too unstable for HIC. This, however, requires knowledge of the relationship

between how proteins unfold in solution and on HIC surfaces. Knowledge of this relationship and how it relates to measured solution stability data would help screen media selection for protein purification.

On the other hand, protein unfolding has been used to manipulate retention and selectivity (Lindahl & Vogel 1984). Previous studies have demonstrated that HIC can separate a single protein with different conformations (Wada et al. 1983; Valliere-Douglass et al. 2008; Deitcher et al. 2009). Recent work at Amgen has demonstrated that HIC can resolve antibody variants that differ by just 4 kDa (Valliere-Douglass et al. 2008). Clippings of the N-or C-terminus provide sufficient conformational differences for separation by HIC whereas SEC or IEC cannot sufficiently or completely resolve these variants.

Further, HIC can separate protein variants with similar native conformations but different stabilities. Carefully chosen processing conditions can selectively unfold less stable protein variants to alter their conformation and retention behavior. Studies with a disulfide-reduced variant of bovine  $\alpha$ -lactalbumin (BLA) showed that reduction of a specific disulfide bond alters the retention time and adsorption strength ( $\sim 6-8$  kJ/mol) of the variant (Deitcher et al. 2009). These studies suggest that HIC can effectively separate variants that do not differ in size or charge but do differ in conformational stability. Thus, an understanding of protein stability on a HIC surface not only might prevent yield losses, but also could allow for potential retention and selectivity manipulations amongst variants.

To optimize resin selection it is important to understand why proteins unfold on certain resins and not others. There is evidence that resins of higher hydrophobicity destabilize proteins more than lesser ones (Jungbauer et al. 2005; Deitcher et al. 2010).

However, there is currently no empirical or theoretical framework to identify the specific resins that destabilize a protein.

In this work, we have tested the hypotheses that adsorbed protein unfolding increases as the hydrophobicity of the resin surface increases, and that regions that unfold are dependent on their relative solution stability. We have also hypothesized that the thermodynamic effects of different HIC surfaces on the free energy of unfolding can be determined. To test these hypotheses, we have examined three protein systems that undergo multidomain unfolding on a series of increasingly hydrophobic resins over a range of temperatures and of concentrations of guanidine hydrochloride. We have demonstrated that similar conformational changes occur, but to different extents, on different resins. We have also demonstrated that the quantitative effects of a HIC surface on the free energy of unfolding can be determined from guanidine denaturation monitored with HXMS. Finally, we have also been able to relate the partially unfolded state at the surface of chromatographic media and in solution.

## **2.2. Materials and Methods**

### **2.2.1. Materials**

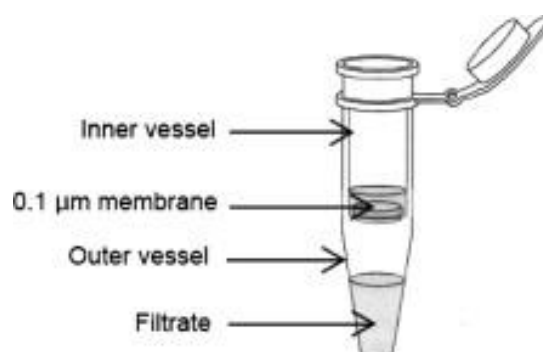
Human serum transferrin (transferrin), human  $\alpha$ -antitrypsin (antitrypsin), and bovine serum albumin (BSA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Potassium phosphate, ammonium sulfate, calcium chloride, ethylenediaminetetraacetic acid (EDTA), citric acid, formic acid, trifluoroacetic acid (TFA), and guanidine hydrochloride (GdnHCl) were purchased from Fisher Scientific (Houston, TX, USA) and were of HPLC-grade quality or better. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Thermo Scientific (Rockford, IL, USA).

The Tosoh HIC resins used in this study, Phenyl 650M, Butyl 650M, and Hexyl 650C were purchased from Fisher Scientific (Houston, TX, USA). Ultrafree®-MC centrifugal filter units were purchased from Fisher Scientific (Houston, TX, USA) for the separation of supernatant liquid from resin particles.

### **2.2.2. Temperature studies**

For no-surface control experiments, 5  $\mu$  L of 20 mg/mL protein solution were mixed with 45  $\mu$  L of deuterated buffer at room temperature and at 10 degree increments in temperature up to 82°C and equilibrated for 2 hours. The protein solutions and labeling buffers were prepared at pH 7.0, 25 mM potassium phosphate, and 1.5M ammonium sulfate (working buffer). Labeling times of 10, 20, 30, 40, and 60 minutes were examined. After labeling, 5  $\mu$  L of quench buffer (150 mM potassium phosphate, pH 1.5) kept in an ice bath was added, bringing the final solution pH to 2.6, near the pH minimum of the hydrogen-deuterium exchange reaction. Samples were kept at room temperature for 40 seconds

before 147  $\mu$  L of desorption buffer was added to the solution. The desorption buffer contained 100 mM citric acid, 8M GdnHCl, 100 mM TCEP, and 27 mM EDTA in H<sub>2</sub>O at pH 2.6. After addition of desorption buffer, the sample was placed on ice for 2 minutes before being put at room temperature for 40 seconds. Then, 600  $\mu$ L of 95% H<sub>2</sub>O, 5% acetonitrile, 0.1% formic acid, and 0.01% trifluoroacetic acid was added to dilute the protein and GdnHCl concentration. Solution phase samples were placed at room temperature for 40 seconds to replicate the time between sample quenching and introduction into the MS for the adsorbed phase experiments that involved two additional centrifugation steps (40 seconds each).



**Figure 2.1. Representation of Millipore's Ultrafree-MC centrifugal filter unit taken from Vesely et al. (2012).**

For adsorbed phase experiments, 35  $\mu$ L of 5 mg/mL protein solution in working buffer was added to 65  $\mu$ L of resin slurry (50:50 dry resin:working buffer) in the inner vessel of an Ultrafree®-MC centrifugal filter unit (see Figure 2.1). The samples were allowed to equilibrate overnight to ensure adsorption equilibrium. Prior to labeling, the sample was centrifuged at 7400 rcf for 30 seconds to form a filtrate. To initiate labeling, 90  $\mu$ L of deuterated buffer and 10  $\mu$ L of non-deuterated buffer were added to the inner vessel at room temperature and at 10 degree increments in temperature to 82°C. Labeling times

of 10, 20, 30, 40, and 60 minutes were examined. After labeling, 10  $\mu\text{L}$  of quench buffer was added to the inner vessel and the filter unit was immediately centrifuged at 7400 rcf for 30 seconds. The inner vessel was transferred to a new empty filter unit on ice and 200  $\mu\text{L}$  of desorption buffer was added. The sample was placed on ice for 2 minutes and then centrifuged at 7400 rcf for 30 seconds. Finally, 600  $\mu\text{L}$  of 95% deionized, distilled (dd) $\text{H}_2\text{O}$ , 5% acetonitrile, 0.05% formic acid, and 0.01% trifluoroacetic acid solution (sample pump solution) were added to the filtrate upon completion.

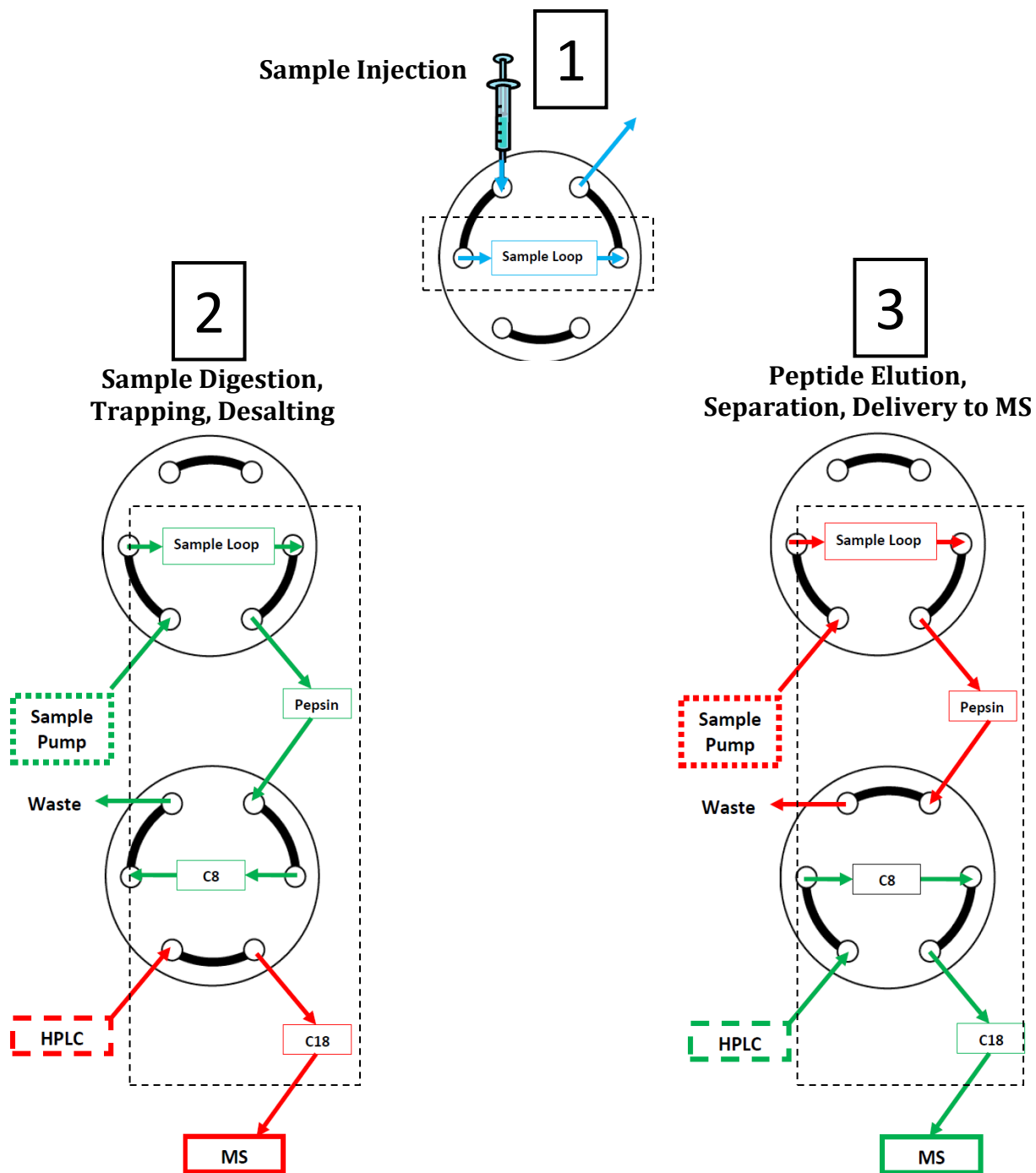
### **2.2.3. Guanidine studies**

Solution and adsorbed phase experiments for the guanidine studies were done with similar protocols to the temperature studies. In these studies the temperature was held constant at 22°C while the concentration of GdnHCl in the protein, protein-resin samples, and deuterated buffer varied from 0 to 4.5M in increments of 0.5M. Samples were equilibrated overnight prior to labeling. Data were collected in increments of 0.1M GdnHCl for the concentration range in which the transition from folded to unfolded structure occurred.

### **2.2.4. HPLC-MS**

A schematic of the valve and column setup used to deliver samples to the MS is shown in Figure 2.2. Samples were injected into a 200  $\mu\text{L}$  stainless steel sample loop using a 500  $\mu\text{L}$  glass syringe. A sample pump (LabAlliance™ Series I) pumped sample pump solution and the injected sample at 100  $\mu\text{L}/\text{min}$  through the sample loop and into an immobilized pepsin column (2.1 mm inner diameter by 60 mm length) where proteolytic

digestion took place. Pepsin preferentially cleaves at the C-terminal side of Phenylalanine, Leucine, Tryptophan, Tyrosine, Alanine, Glutamic Acid, and Glutamine allowing consistent peptide fragments to be generated for different runs. Peptides exiting the column were trapped, desalted, and concentrated on a C<sub>8</sub>-desalting column (TR1/25109/02, 1 mm inner diameter by 8 mm length, Michrom Bioresources, Inc, shown as "C8" in Figure 2.2). After this desalting step (6 minutes), flow was switched from the sample pump to the Surveyor MS HPLC pump (shown as "HPLC" in Figure 2.2) to elute the peptides off the C<sub>8</sub> column. An XBridge C<sub>18</sub> column (186003563, 2.1 inner diameter by 50 mm length, 3.5  $\mu$  m pore size, Waters, Inc, shown as "C18" in Figure 2.2) downstream of the C<sub>8</sub> column was used for improved resolution for the large number of peptides.



**Figure 2.2.** Schematic of valve and column setup used to deliver sample to MS. (1) Sample is injected manually with glass syringe into a 200  $\mu\text{L}$  sample loop. (2) Sample is digested with Porcine pepsin, resulting peptides are trapped on a C8 column, and salt is removed. (3) Flow path in valve changes so that HPLC (high-performance liquid chromatography) elutes peptides off C8 column to deliver to C18 column for further resolution before delivery to the MS. Blue lines and arrows depict manual sample injection. Green lines and arrows depict flowpath of sample/peptides through each step. Red lines and arrows depict inactive flow at each step. Dashed box depicts temperature controlled region of 1.0°C.



For the solution and adsorbed phase studies, a short gradient run was employed to minimize back-exchange but still effectively resolve peptides. The treatment for peptide desorption was a 17 minute gradient of 70% solvent A (ddH<sub>2</sub>O, 0.1% formic acid, 0.01% TFA) and 30% solvent B (acetonitrile, 0.8% formic acid) to 40% solvent A, followed by a 2 minute gradient from 40% solvent A to 10% solvent A, followed by 4 minutes at 10% solvent A.

Peptides were eluted directly to a LTQ linear ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA). Data were collected in a positive ion, profile mode with an ESI voltage of 4.3 kV, a capillary temperature of 250°C, and sheath gas flow rate of 15 units. The peptides identified in MS/MS experiments are shown in the supporting information (Tables A1, A2, and A3).

### **2.2.5. Size exclusion chromatography**

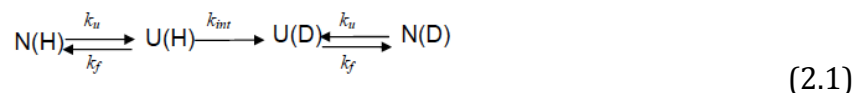
Size exclusion chromatography (SEC) was performed to evaluate aggregate formation on the surface. Protein samples at 5 mg/mL in working buffer were equilibrated for 1 hour on the three different HIC media in a Ultrafree®-MC centrifugal filter unit within a 1.5 mL microcentrifuge tube and at temperatures from 22 to 82°C in increments of 10°C. The sample was then centrifuged at 7400 rcf for 1 minute to remove supernatant. The filter unit was transferred to a new microcentrifuge tube on ice and 200 µL of desorption buffer was added. The sample was placed on ice for 10 minutes and then centrifuged at 7400 rcf for 1 minute. The flowthrough was immediately collected, diluted with 500 µL of working buffer and injected into a 500 µL sample loop of an AKTA Explorer.

A TSK gel G3000SWXL column (TOSOH Bioscience, 7.8mm ID \* 30cm, 5  $\mu$ m) was equilibrated with 5 column volumes (CVs) of working buffer at 0.5 mL/min prior to sample loading on the column. Sample loading was done at 0.5 mL/min with working buffer and continued until all monomer and oligomer peaks (where applicable) were detected and the 215 nm signal returned to baseline.

## 2.3. Theory

### 2.3.1 Isotopic (Deuterium) Labeling

The exchange of a buried amide hydrogen, H, for a deuterium, D, can be modeled by the following (Zhang & Smith 1993)



where  $k_u$ ,  $k_f$ , and  $k_{int}$  are the unfolding, folding, and intrinsic exchange rates and N and U represent the amide in its native and unfolded state, respectively. For very stable regions of tertiary structure

$$k_f \gg k_{int} \quad (2.2)$$

and the overall first order rate constant for exchange,  $k_{obs}$ , is

$$k_{obs} = k_u \times k_{int} \quad (2.3)$$

This limiting regime of exchange kinetics is commonly referred to as the EX2 limit.

For very unstable regions of tertiary structure

$$k_f \ll k_{int} \quad (2.4)$$

and the overall rate constant for exchange is

$$k_{obs} = k_u \quad (2.5)$$

This limiting regime of exchange kinetics is commonly referred to as the EX1 limit.

### 2.3.2. Measuring and Calculating Fractions of Labeled Peptide

The extent to which a peptide has been labeled with deuterium is determined by (Zhang & Smith 1993)

$$\frac{D}{N} = \frac{m_t - m_0}{m_{100} - m_0} \quad (2.6)$$

where  $D$  is the number of deuterated amides and  $N$  is the total number of exchange-competent residues in a peptide. Also,  $m_t$  is the mass of a peptide after a given labeling time,  $m_0$  is the non-deuterated mass of that peptide, and  $m_{100}$  is the fully-deuterated mass of that peptide.  $m_t$  is determined from the first moment analysis of a peptide mass spectrum. In this study, HDExaminer (Sierra Analytics, Modesto, CA) was used to determine  $m_t$  for all peptides.

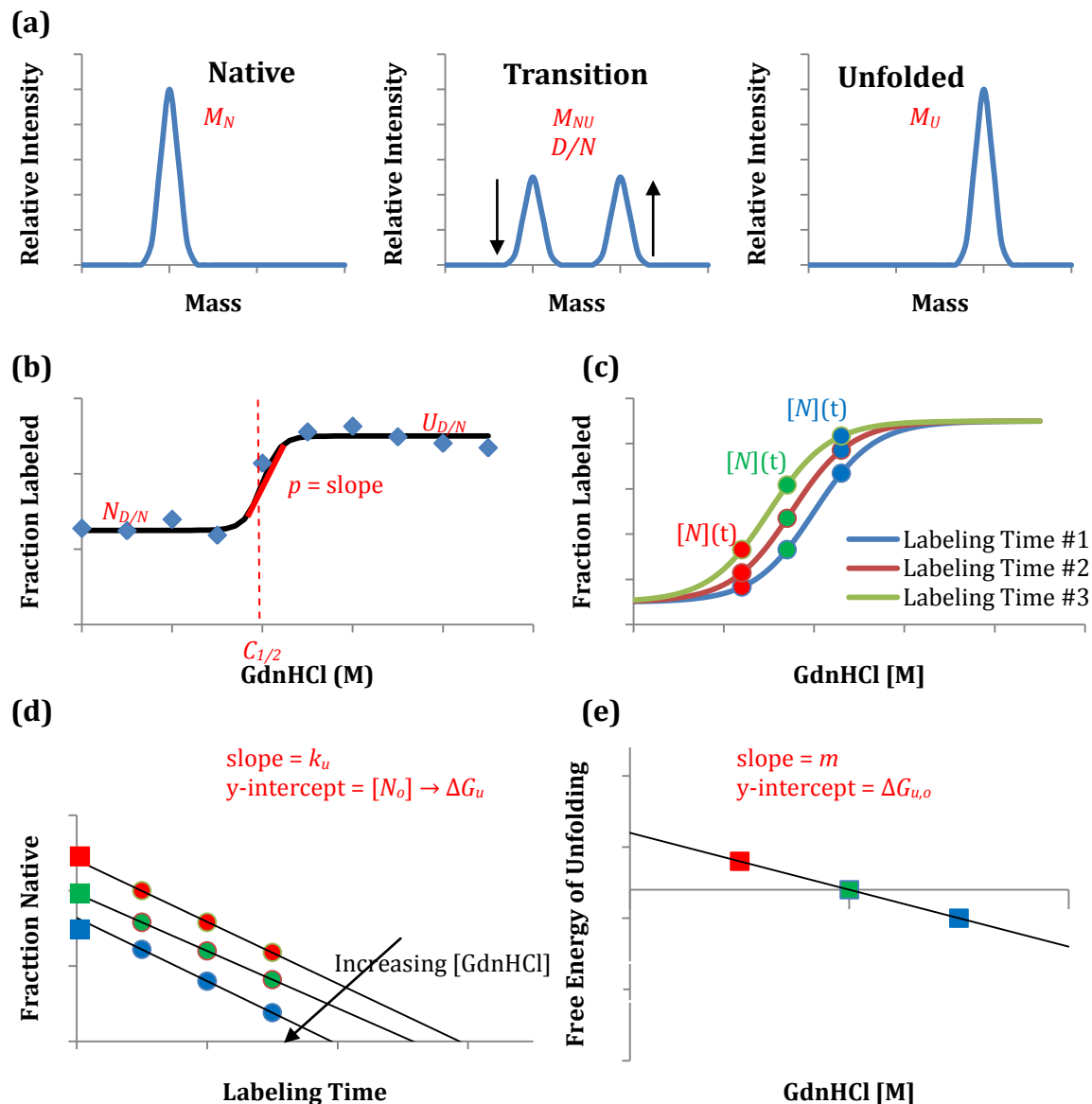
In the present cases, exchange-competent refers to all peptide residues except Proline, which does not have amide hydrogen and the N-terminal residue of the peptide, which does not have a backbone amide. Back-exchange of deuterium for hydrogen occurs when the deuterated protein molecules are introduced back into H<sub>2</sub>O-solvents used in high performance liquid chromatography (HPLC) to resolve peptides. Thus, the residue immediately after the N-terminal residue of a peptide is also not counted; the back-exchange of this residue is unusually high (Bai et al. 1993). Equation (2.6) also accounts for back-exchange experienced during the time,  $\approx 12$ -16 minutes, in-between sample quenching and introduction into the MS (Zhang & Smith 1993).

### 2.3.3. Unfolding Free Energies and Rates

Solution unfolding free energies and rates can be determined from experimental GdnHCl denaturation curves monitored by HXMS (Powell & Fitzgerald 2003; Tang et al. 2007; West et al. 2008). It is proposed the same can be done for proteins adsorbed on surfaces. Typically, destabilization of protein folding equilibria leads to an increase in the population of fully (or partially) unfolded molecules relative to native ones. An overview of the analysis required to relate this shift in population to unfolding free energies and rates is shown in Figure 2.3. First, mass spectra are interpreted to generate the denaturation curves (Figure 2.3a). For apparent EX1 behavior, the refolding rate is slow compared to the labeling rate, and molecules that unfold have time to completely label before they become protected again. For this situation, the observed decrease in  $D/N$  protein is limited by the unfolding rate,  $k_u$

$$\frac{dN}{dt} = -k_u[N] \quad (2.7)$$

where  $[N]$  is the concentration or fraction of native peptide at a given labeling time. In the mass spectrum, the molecular mass of this peak,  $M_N$ , is representative of the unlabeled mass plus 1 Da for each residue that is exchange competent under native conditions (e.g. located on surface of protein).



**Figure 2.3.** Determination of  $k_u$ ,  $m$ , and  $\Delta G_{u,0}$  for a reporter peptide. (a)  $M_N$  and  $M_U$  are measured along with their peak intensities from the mass spectrum to calculate  $M_{NU}$  and  $D/N$ . (b)  $D/N$  (diamonds) vs. GdnHCl concentration is plotted and fitted to a four-parameter ( $N_{D/N}$ ,  $U_{D/N}$ ,  $C_{1/2}$ , and  $m$ ) sigmoidal equation (Equation 2.9). (c)  $[M]$  (circles) is calculated at various labeling times and GdnHCl concentrations from the sigmoidal equation. (d) For each GdnHCl concentration,  $[N]$  is regressed vs. labeling time (Equation 2.11). Slope of regressed line is  $-k_u$  and y-intercept is  $\Delta G_u$  (squares). (e)  $\Delta G_u$  is regressed vs. GdnHCl concentration (Equation 2.13). Slope of regressed line is  $m$  and y-intercept is  $\Delta G_{u,0}$ . 95% confidence intervals of all regressed parameters are propagated to  $m$  and  $\Delta G_{u,0}$ .

Under the EX1 limit and at concentrations of GdnHCl where the protein molecules transition from all native to a mixture of native and unfolded, a second peak appears at high molecular mass, representative of unfolded molecules. The molecular mass of this unfolded peak,  $M_U$ , also stays constant but grows in area with labeling time at the expense of the native peak until no native peak remains. The weighted mass (centroid) of the two peaks in the transition region,  $M_{NU}$ , can then be used to determine  $[N]$  from

$$[N] = 1 - \frac{M_{NU} - M_N}{M_U - M_N} \quad (2.8)$$

In the limit where  $M_{NU}=M_U$ ,  $[N]$  is 0 and in the limit where  $M_{NU}=M_N$ ,  $[N]$  is 1. In this work, experimental  $D/N$  values have been measured for different concentrations of GdnHCl using Equation 2.6 and  $m_t=M_{NU}$  to generate the denaturation curves.

Next, the denaturation curves are fitted to a sigmoidal equation to allow generation of additional data points by interpolation in the transition region (Figure 2.3b).  $D/N$  values at different GdnHCl concentrations and a fixed labeling time are fitted with LSQCURVEFIT, a non-linear least squares solver, from MATLAB® to a four-parameter sigmoidal equation

$$\frac{D}{N} = N_{D/N} + \frac{U_{D/N} - N_{D/N}}{1 + e^{-p([GdnHCl] - C_{1/2})}} \quad (2.9)$$

where  $N_{D/N}$  is the  $D/N$  value of the native peak,  $U_{D/N}$  is the  $D/N$  value of the unfolded peak,  $p$  is a parameter describing the slope of the unfolding transition, and  $C_{1/2}$  is the midpoint concentration of the denaturation curve.

Next, fitted  $D/N$  values generated in MATLAB® from Equation 2.9 can be used to calculate  $M_{NU}$  at given GdnHCl concentrations with rearrangement of Equation 2.6 and  $m_t=M_{NU}$

$$M_{NU} = \frac{D}{N} \times (m_{100} - m_0) + m_0 \quad (2.10)$$

which can then be used to calculate  $[N]$  using Equation 2.8 (Figure 2.3c). This analysis is repeated for each labeling time studied.

Next, generated values of  $[N]$  at different labeling times and GdnHCl concentrations can be used to calculate unfolding free energies and rates. Equation 2.7 is integrated and solved to obtain

$$\ln[N] = -k_u t + \ln[N_o] \quad (2.11)$$

where  $[N_o]$  is the fraction of native reporter peptide at zero labeling time. Regression of Equation 2.11 with  $[N]$  measured at different labeling times yields  $k_u$  and  $[N_o]$  (Figure 2.3d). Assuming ideal solution, the standard state unfolding free energy,  $\Delta G_u$ , for a given reporter peptide of a protein can then be determined from

$$\Delta G_u = -RT \ln \left( \frac{[N_o]}{1 - [N_o]} \right) \quad (2.12)$$

where  $R$  is the gas constant, and  $T$  is temperature.

Finally, assuming linear variation of unfolding free energy with guanidine concentration (Aune & Tanford 1968), the values of  $\Delta G_u$  calculated at different GdnHCl concentrations can be regressed to determine the unfolding free energy in the absence of guanidine,  $\Delta G_{u,o}$ , (Figure 2e) from

$$D G_u = m[\text{GdnHCl}] + D G_{u,o} \quad (2.13)$$

where  $m$  is the coefficient for the effect of denaturant on stability (Greene & Pace 1974). Regressed values of  $\Delta G_{u,o}$  in solution,  $\Delta G_{u,sol}$ , and on the surface,  $\Delta G_{u,ads}$ , can be compared to determine the thermodynamic effect,  $\Delta \Delta G_u$ , the surface has on the protein from



$$\Delta\Delta G_u = \Delta G_{u,ads} - \Delta G_{u,sol} \quad (2.14)$$

Regressed values of  $k_u$  in solution,  $k_{u,sol}$ , and on the surface,  $k_{u,ads}$ , can also be compared the kinetic effect,  $\Delta k_u$ , the surface has on the protein from

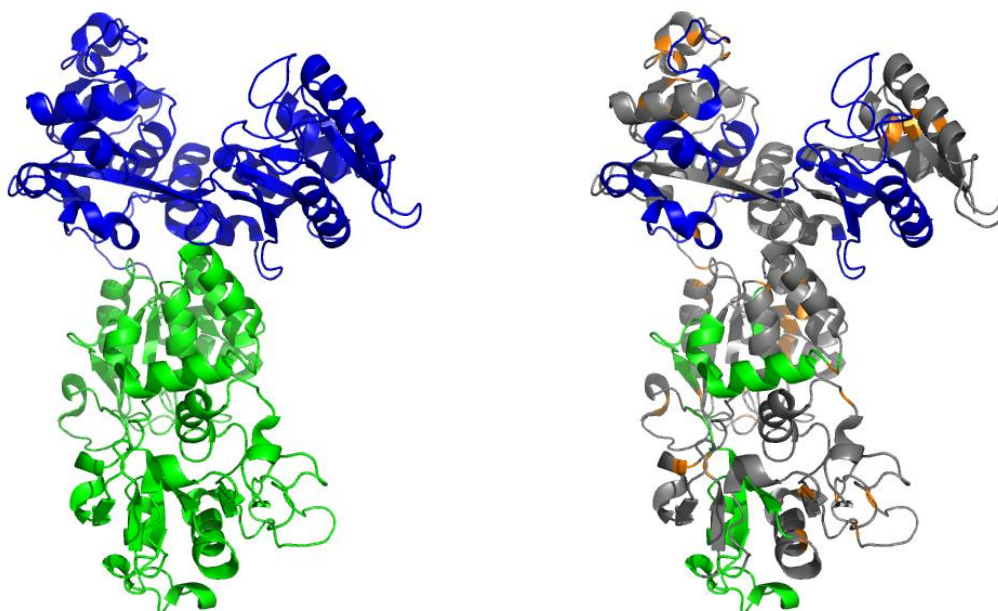
$$\Delta k_u = k_{u,ads} - k_{u,sol} \quad (2.15)$$

The different values calculated from Equations (2.10) to (2.15) originate from the fitting of experimental data with Equation (2.9).

## ***2.4. Results and Discussion***

### **2.4.1. Unfolding effects of chromatographic surface chemistry and temperature on transferrin**

The effects of HIC surface type and temperature on the unfolding of transferrin were studied with HXMS. MS spectra were collected and compared for transferrin labeled in solution and while adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C at different temperatures. A large amount of precipitation at temperatures above 52°C prevented analysis of higher temperature conditions. The mass spectra for 32 reporter peptides were converted to  $D/N$  values using Equation (2.6). A total of 33% sequence coverage was obtained for the 679 residues of transferrin. Figure 2.4 shows the native structure of transferrin and its two domains (left) and the location of the reporter peptides and Cysteines of transferrin (right). Many of the areas with missing sequence coverage are populated with Cysteines involved in disulfide bridges, which can impact proteolytic digestion. Although reporter peptides are available for both domains, the most coverage is obtained in the N-terminal domain, 40%, while it is lower in the C-terminal domain, 24%.



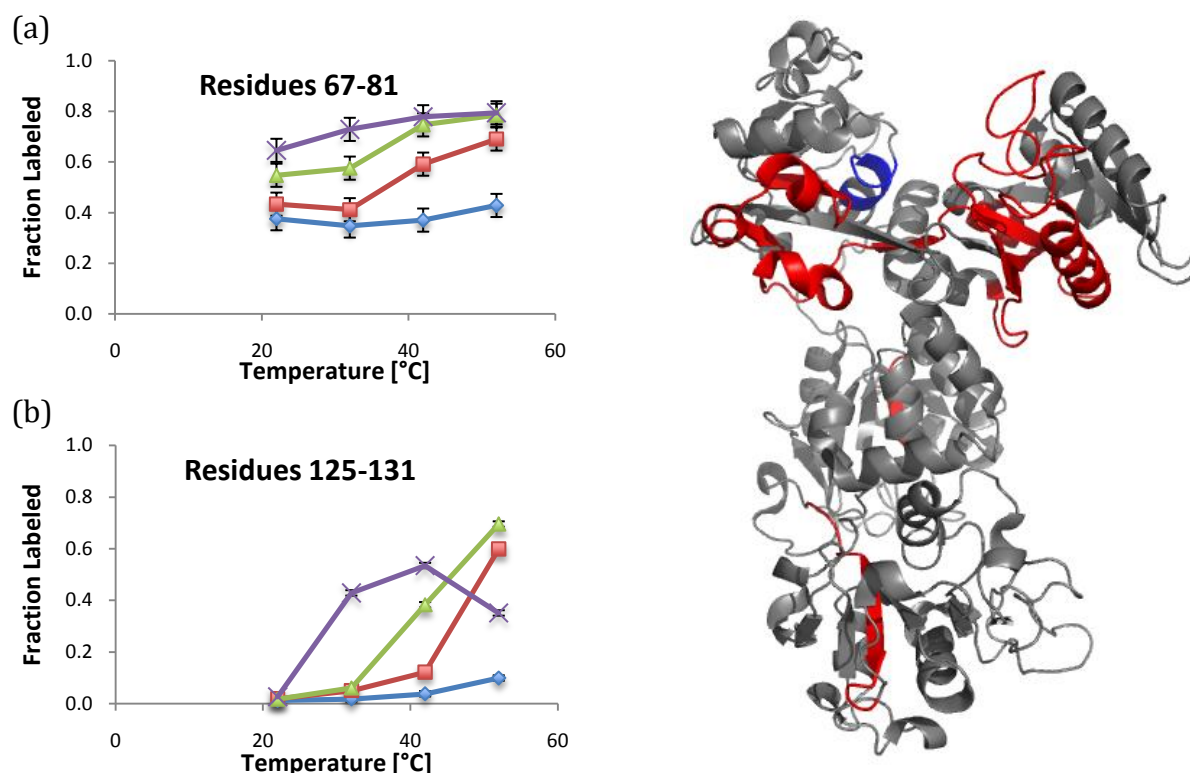
**Figure 2.4. Left: Native structure of transferrin (PDB ID 2HAV) highlighting the N-terminal domain (blue) and C-terminal domain II (green). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts Cysteines of transferrin. Grey depicts missing sequence coverage.**

In general, the reporter peptides of transferrin that unfold on the surfaces follow a similar pattern. Unfolding in solution is minimal and occurs at the highest temperatures while on the chromatographic surfaces unfolding occurs at lower temperatures. Of the three surfaces, unfolding on Butyl 650M occurs at the lowest temperature range and reaches the highest  $D/N$  values. The unfolding on Phenyl 650M and Hexyl 650C occur at higher temperatures and of these two, higher  $D/N$  values were observed for Hexyl 650C. All reporter peptides follow this general pattern with variations only in the degree of unfolding. There was only one exception to this pattern that will be discussed later.

The  $D/N$  values for a reporter peptide representative of the general pattern in solution and on the three surfaces are shown in Figure 2.5a. Residues 67-81, shown in panel (a), are representative of the unfolding pattern of all but 1 reporter peptide for transferrin. In solution, no large increases in solvent exposure are observed for these residues over the temperature range studied. The largest increase is between 42 and 52°C from 0.37 to 0.43  $D/N$ . This is consistent with studies that report transferrin starts to unfold in solution only at temperatures above 60°C (Shen et al. 1992). In contrast, on Phenyl 650M, a large increase in solvent exposure from 0.41 to 0.59 was observed between 32 and 42°C, indicating significant unfolding. On Hexyl 650C, the largest solvent exposure increase also occurs between 32 and 42°C, from 0.58 to 0.78. As observed for this peptide, other reporter peptides also labeled more completely on Hexyl 650C than on Phenyl 650M. Finally, on Butyl 650M, a significant portion of these residues are already labeled at 22°C. The labeling of this reporter peptide along with 19, 1, and 1 reporter peptides spanning residues 202-310, 386-392, and 424-435, respectively, demonstrate that all three chromatographic surfaces induced significant unfolding at lower temperatures than in solution. Further, the extent of unfolding varies among the three surfaces, but the ordering of their impact is consistent. The locations of these residues are shown in red on the native structure of transferrin in Figure 2.5 on the right.

Residues 125-131, shown in Figure 2.5(b), are from the one reporter peptide that behaved somewhat differently. Like all of the other reporter peptides, the chromatographic surfaces induce more labeling than solution alone, and the ordering of the effects was the same as the other reporter peptides. However, for Residues 125-131, there were two distinguishing aspects to the labeling behavior. First, no labeling is observed at

22°C in solution or on any of the three surfaces. The complete burial of these residues in the native state suggests that for their exposure to occur stronger denaturing conditions are needed compared to the other reporter peptides of transferrin. Second, the labeling of transferrin on Butyl 650M displayed a maximum vs. temperature and decreased between 42 and 52°C to sub-32°C levels as shown in Figure 2.5(b). Although the labeling of several reporter peptides from the general trend also decreased between 42 and 52°C on Butyl 650M, none of those decreased to sub-32°C labeling as with Residues 125-131. Intermolecular interactions or aggregation may be responsible for this, as discussed later.

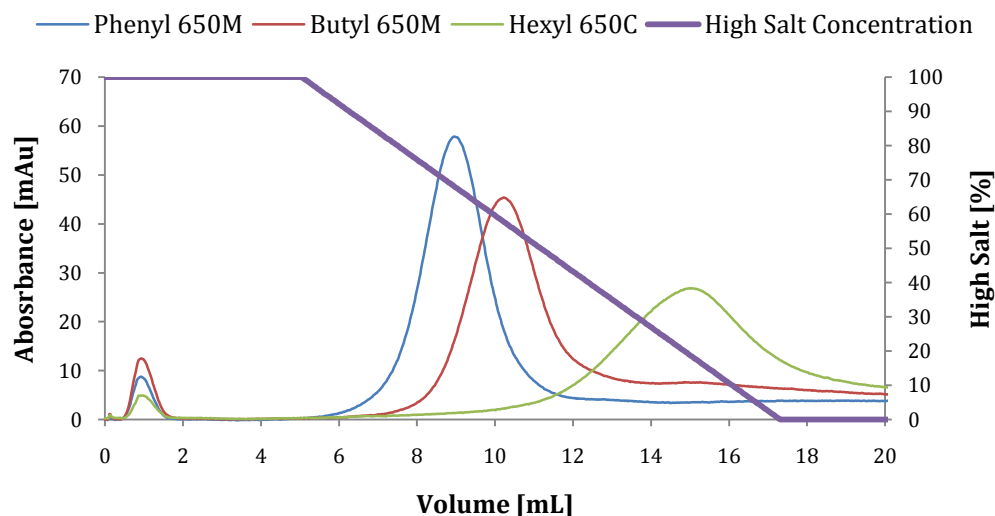


**Figure 2.5. Left: Thermal unfolding of different reporter peptides of transferrin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 52°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 67-81 is shown in panel (a). The *D/N* for Residues 125-131 is shown in panel (b). Right: Location of residues 67-81 and the other residues that follow the general unfolding pattern are shown in red on the native structure (PDB ID 2HAV). The exception to the general unfolding pattern, Residues 125-131, is shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C.**

The apparent melting curves shown in Figures 2.5b demonstrate that adsorption to all three surfaces lowers the melting temperature relative to solution for this reporter peptide of transferrin. However, the amount it is lowered depends on the type of surface. It should be noted that these “melting curves” obtained with HXMS are not identical to typical thermal denaturation curves. In thermal denaturation measured by other spectroscopic techniques such as circular dichroism or fluorescence, the unfolding curves of a protein will depend on its stability and heat capacity. With HXMS there is an added time element where the curves will be affected by the labeling time and intrinsic labeling rate. HXMS denaturation curves are shifted to the left due to this effect. A more detailed explanation of this effect on denaturation curves for the case of EX2 exchange is available elsewhere (Ghaemmaghami et al. 2000).

The degree of unfolding of human serum transferrin on the surfaces does not completely follow their reported relative hydrophobicities. Tosoh uses dynamic binding capacities (DBC) for lysozyme as a hydrophobicity measure (See Table 2.1). One appealing aspect to this measure is the high stability of lysozyme as a probe protein; thus, we would not expect unfolding to influence the measurement. Gradient elution experiments at low protein loadings were performed in this study with human serum transferrin as a simple alternative hydrophobicity measure. Figure 2.6 shows how the gradient elution time varied on the three different media. Table 2.1 compares the two hydrophobicity measures with unfolding rank of each surface measured by the increase in  $D/N$  for the reporter peptides of transferrin. The hydrophobicity rank of the resins based both on DBC and retention volume is as follows: Hexyl 650C > Butyl 650M > Phenyl 650M. The extent of

transferrin unfolding, however, follows the order: Butyl 650M > Hexyl 650C > Phenyl 650M.



**Figure 2.6.** Gradient elution of transferrin on Phenyl 650M (blue), Butyl 650M (red), and Hexyl 650C (green). 500  $\mu\text{g}$  of sample were loaded onto 0.5 mL columns at a flow rate of 1 mL/min with 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 25mM  $\text{PO}_4$  (high salt) at pH 7.0 and room temperature. Columns were washed with 2 CVs of high salt and protein was eluted with a 25 CV gradient from high salt to 25mM  $\text{PO}_4$  at pH 7.0 and room temperature.

**Table 2.1.** DBC and retention volume comparison on Tosoh HIC media

	Phenyl 650M	Butyl 650M	Hexyl 650C
<b>DBC [mg/mL]*</b>	27.5	32.2	33.2
<b>Retention Volume [mL]</b>	8.8	9.9	14.6
<b>Hydrophobicity Rank</b>	3	2	1
<b>Unfolding Rank<sup>1</sup></b>	3	1	2

\*DBC as reported by Tosoh

<sup>1</sup>Based on increase in  $D/N$  for the reporter peptides of transferrin

A similar discrepancy between adsorption and unfolding order has been observed previously in a study with bovine  $\alpha$ -lactalbumin unfolding on Phenyl Sepharose 6FF and Butyl Sepharose 4FF (Fogle et al. 2006). In that study, more unfolding was observed on Butyl Sepharose 4FF despite higher retention factors for bovine  $\alpha$ -lactalbumin on Phenyl Sepharose 6FF. The authors hypothesized that differences in binding orientation or ligand density between the two resins may also influence unfolding. The different ligand densities could influence whether a protein molecule interacts with one or multiple ligands on the surface. Different binding orientations might also occur, as has been demonstrated for lysozyme on ion exchange surfaces of differing ligand densities (Dismer & Hubbuch 2007). Since ligand densities are not provided by Tosoh, it is unknown if the variations follow ligand densities or not.

#### **2.4.2. Unfolding effects of chromatographic surface chemistry and temperature on antitrypsin**

As a second protein system, the effects of HIC surface type and temperature on the unfolding of antitrypsin were also studied with HXMS. MS spectra were collected and compared for antitrypsin labeled in solution and while adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C at different temperatures. Unlike transferrin, samples were capable of incubation up to 92°C without precipitation. The mass spectra for 28 reporter peptides were converted to  $D/N$  values using Equation (2.6). A total of 53% sequence coverage was obtained for the 392 residues of antitrypsin. Figure 2.7 shows the native structure of antitrypsin and the location of the reporter peptides. Unlike transferrin, antitrypsin does



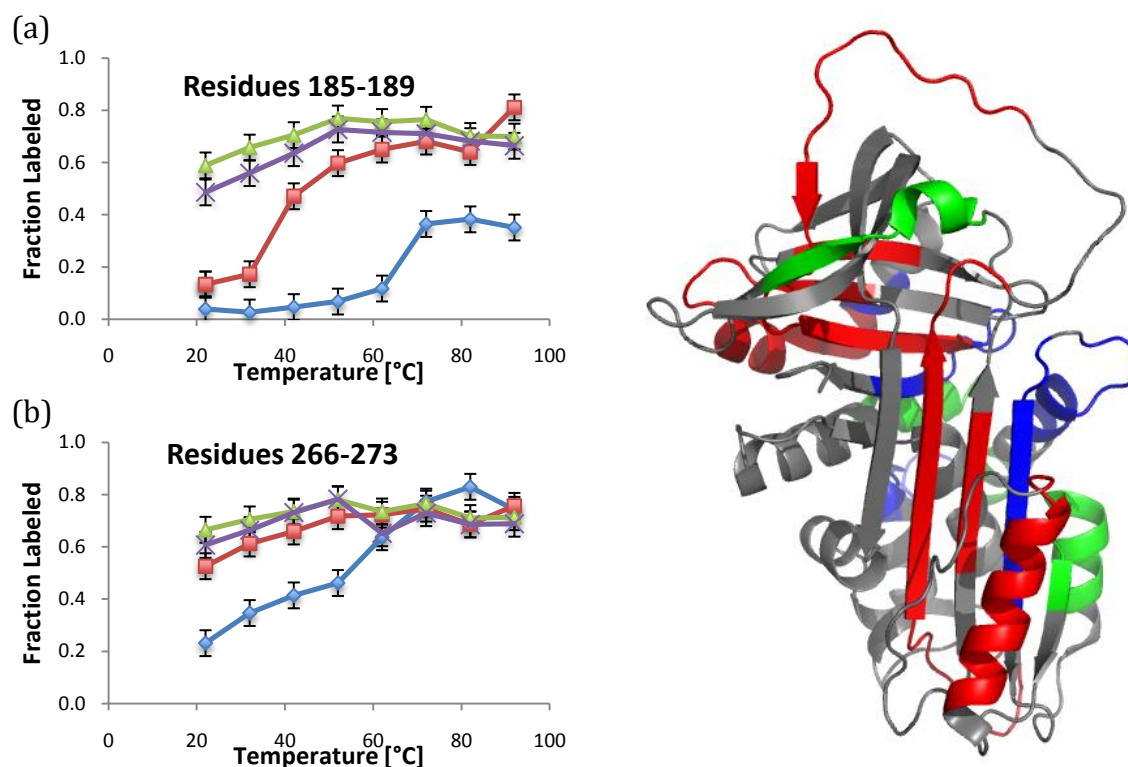
not contain any disulfide bridges which may explain the greater sequence coverage compared to transferrin.



**Figure 2.7. Native structure of transferrin (PDB ID 2QUG) and location of reporter peptides obtained in this study (highlighted in green). Grey depicts missing sequence coverage.**

The labeling trends of antitrypsin are similar to those of transferrin but differences exist. As with transferrin, contact with the hydrophobic chromatographic surface enhances labeling for most but not all reporter peptides. Also, Phenyl 650M results in equal or less labeling enhancement than the other two surfaces for most reporter peptides of antitrypsin. However, for antitrypsin, the different denaturing effects of Butyl 650M and Hexyl 650C are not as apparent as they are for transferrin. Hexyl 650C resulted in more labeling than Butyl 650M for many reporter peptides although it was statistically significant for only 6 groups of residues: 78-87, 98-108, 110-119, 278-288, and 304-317. The other difference in labeling patterns between antitrypsin and transferrin is that most antitrypsin reporter peptides are observed with statistically equal enhanced labeling on all

3 surfaces: 151-166, 185-189, 238-241, 252-270, 318-343, and 353-375. Two peptides showing this general set of trends are shown in Figure 2.8. Four groups of residues: 24-32, 88-92, 131-142, and 197-207, were not observed with enhanced labeling on the surfaces.



**Figure 2.8. Left: Labeling patterns of different reporter peptides of antitrypsin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 92°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 185-189 is shown in panel (a). The *D/N* for Residues 266-273 is shown in panel (b). Right: Location of residues 185-189 and the other residues that unfold more on Butyl 650M and Hexyl 650C than on Phenyl 650M are shown in blue on the native structure (PDB ID 2QUG). Residues 266-273 and the other residues that unfold to statistically equal extents on all three surfaces are shown in red. Residues that were not observed with enhanced labeling on the surfaces are shown in green. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C.**

For antitrypsin, the extent of unfolding for residues on the surface correlates more directly with increasing resin hydrophobicity (see Table 2.1). In the cases (25% of antitrypsin coverage) where there are statistically significant differences in reporter peptide labeling between all 3 surfaces (as determined by Student t-test with Bonferroni corrected  $\alpha=0.006$ ), more labeling is observed on Hexyl 650C followed by Butyl 650M and then Phenyl 650M. It is not clear why these reporter peptides of antitrypsin follows this trend and transferrin does not.

Many reporter peptides (44% of antitrypsin coverage) that unfold when adsorbed do so to statistically equal extents on all three surfaces. Ignoring statistics however, the  $D/N$  for these residues is consistently the highest on Hexyl 650C, followed by Butyl 650M, and then Phenyl 650M. Regardless, the different denaturing ability of the three surfaces is more apparent for the reporter peptides where there are statistically significant differences in labeling between the three surfaces.

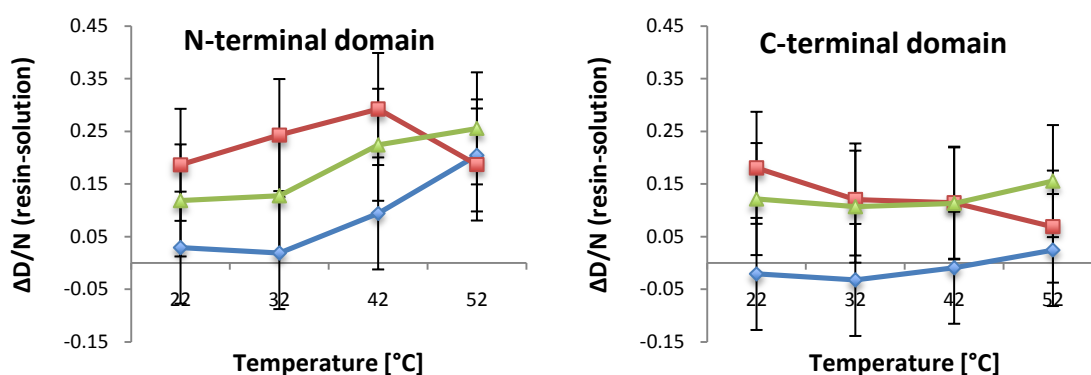
The varying degree of unfolding observed for antitrypsin and the presence of an exception to the general unfolding pattern of transferrin suggests that local stability is important in determining how different groups of residues or domains of a protein will unfold on a surface. The knowledge and use of a transferrin and antitrypsin local stability in solution to explain unfolding trends on the surfaces is presented in the next section.

### 2.4.3. Relating partially unfolded states in solution and on chromatographic surfaces

Partial unfolding of proteins can occur from changes of different conditions. Partial unfolding in solution has been observed from increased temperature from adding denaturants such as urea or GdnHCl, and at extremes in pH (Santra et al. 2005; Ahmad & Khan 2005). In addition to the present work, partially unfolded states have been observed on HIC surfaces (Mcnay & Fernandez 1999; Engel et al. 2004; Gospodarek et al. 2011). Chemical denaturation and the mechanism of surface unfolding are not as well understood as thermal or pH-induced denaturation. Recent molecular simulations of urea-induced unfolding provide some insight into the complexity of this mechanism (Horinek & Netz 2011). However, it is not known how similar are the partially unfolded structures from different denaturation mechanisms.

Transferrin was chosen for this study because it is known to unfold via a multi-state mechanism in solution. Although the structural transition has not been fully characterized, it is known that the N-terminal domain is less stable than the C-terminal domain and unfolds first as denaturant concentration is increased (Tang et al. 2007; Shen et al. 1992). Figure 2.9 shows the effect each chromatographic surface has on the unfolding of these two domains. The difference in  $D/N$  between solution and surface,  $\Delta(D/N)$ , represents the average reporter peptide unfolding for each domain weighted by the number of residues in each peptide. In general, almost all  $\Delta(D/N)$  values are positive, indicating adsorption increases solvent exposure. Within the uncertainties, at every temperature and for every chromatographic surface, a more positive  $\Delta(D/N)$  is observed for the N-terminal domain peptides compared to the C-terminal domain peptides. The greater adsorption-induced

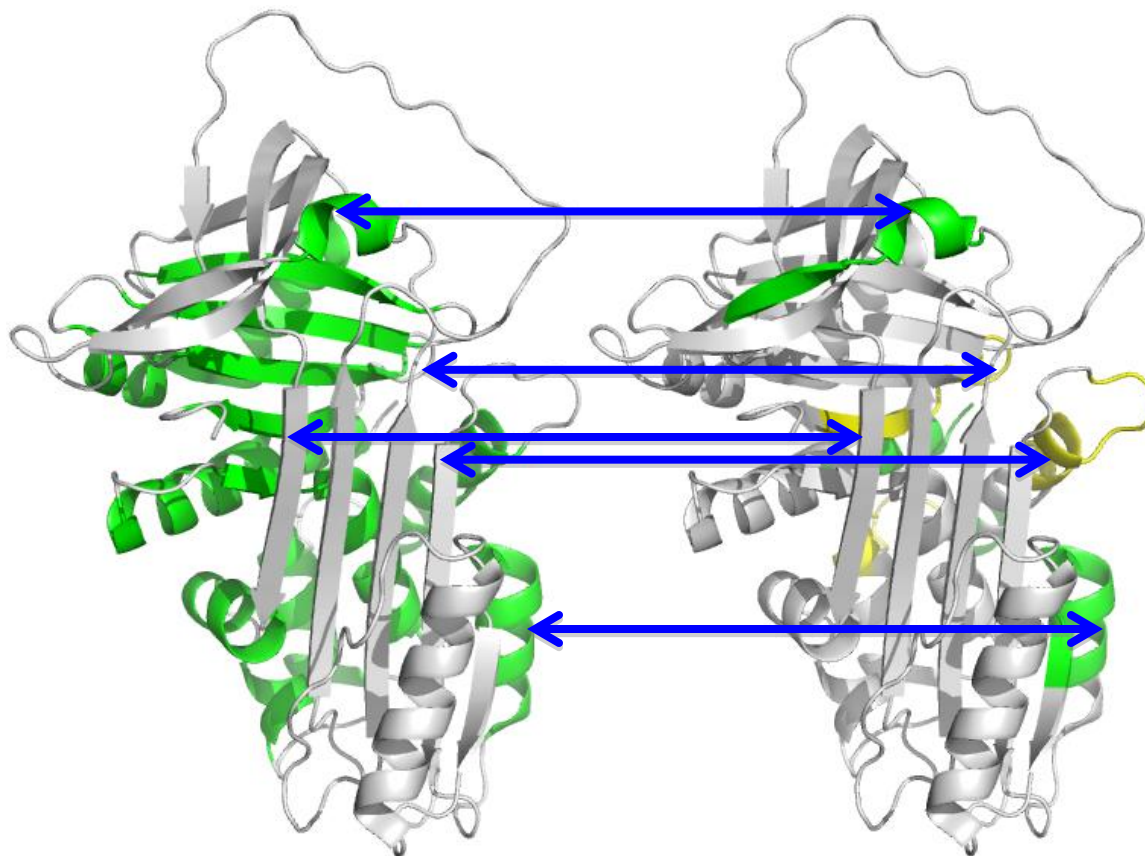
unfolding in the N-terminal domain is consistent with solution stability studies where the N-terminal domain unfolds at milder chemical denaturing conditions (Tang et al. 2007; Shen et al. 1992).



**Figure 2.9.** Difference in  $D/N$ ,  $\Delta D/N$ , between human serum transferrin adsorbed on media (Phenyl 650M - blue, Hexyl 650C - green, Butyl 650M - red) and in solution in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , pH 7.0 buffer.  $\Delta D/N$  values represent the average reporter peptide unfolding for each domain weighted by the number of residues in each peptide. Data is averaged and weighted for all reporter peptides in the N-terminal domain (left) and C-terminal domain (right). Error bars represent sample propagated 95% confidence intervals for  $\Delta D/N$  from triplicate data points collected on Phenyl 650M at 22°C.

Antitrypsin was also chosen for this study because it is also known to unfold via a multi-state mechanism in solution. In the presence of low concentrations of GdnHCl, antitrypsin has a well-characterized unfolding intermediate structure. At low concentrations of GdnHCl (~1M), several  $\beta$ -sheet strands and an  $\alpha$ -helix unfold. The remaining regions stay folded until higher concentrations of GdnHCl, where the protein molecule is fully unfolded (Krishnan & Gierasch 2011). This partially unfolded intermediate is critical for antitrypsin biological function and exists as a small, but significant, subpopulation under physiological conditions (Yamasaki et al. 2010).

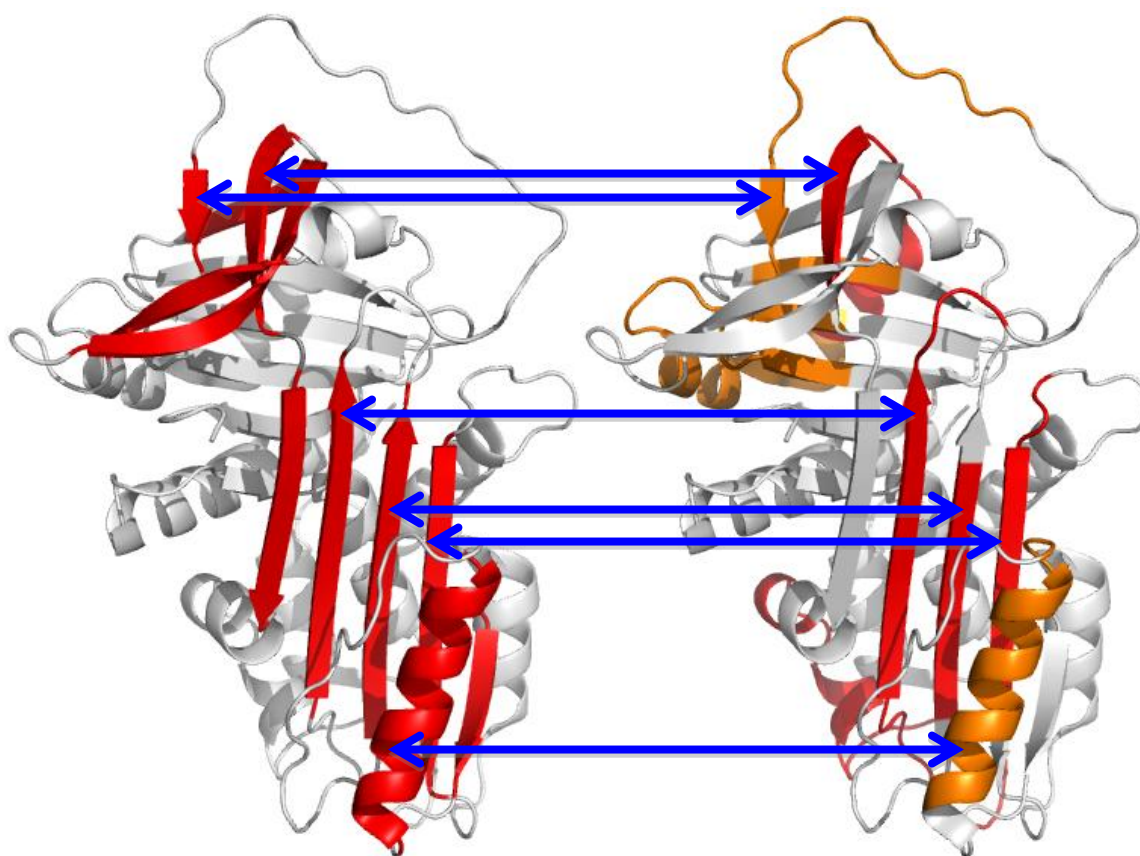
The folded regions of antitrypsin of the solution intermediate are similar to those observed on the HIC surfaces of this work as shown in Figure 2.10. The residues that remain in their native structure in the solution intermediate and do not unfold on the HIC surfaces are shown in green. Three reporter peptides identified as unfolding more on Hexyl 650C than on Butyl 650M and Phenyl 650M (highlighted blue in Figure 2.8) only have statistically significant (as determined by Student t-test with Bonferroni corrected  $\alpha=0.006$ ) enhanced labeling on Hexyl 650C compared to solution. These reporter peptides with lower stability that unfold only on Hexyl 650C are shown in yellow. The regions for which no sequence coverage was obtained are shown in white. The blue arrows identify 6 regions that remain folded in the solution intermediate as well as when adsorbed on Phenyl 650M or Butyl 650M. Of these 6, 3 also do not unfold on Hexyl 650C. This suggests that there is a correspondence between stability in solution and during adsorption; domains of greater local stability in folding intermediates appear to be among the most resistant to unfolding on a denaturing hydrophobic chromatographic surface.



**Figure 2.10. Folded regions of antitrypsin in the guanidine denatured intermediate (left) and when adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22°C (right). Green represents those regions folded in the solution intermediate (left) and on all media (right). Yellow represents those regions that unfold on Hexyl 650C only. Blue arrows designate regions that show similar stability between the two cases.**

The unfolded regions of antitrypsin in the solution intermediate are also similar to the reporter peptides whose  $D/N$  are not affected by adsorption on the HIC surfaces, as shown in Figure 2.11. The regions that unfold in the solution intermediate and are more solvent-exposed on all HIC media are shown in red. The reporter peptides that are more solvent-exposed on Butyl 650M and Hexyl 650C and to a lesser extent on Phenyl 650M are shown in orange. The regions for which no sequence coverage was obtained are shown in

white. The blue arrows identify 5 regions that unfold in the solution intermediate and also unfold when adsorbed on Butyl 650M and Hexyl 650C. Of these 5, 3 also do not unfold on Phenyl 650M.



**Figure 2.11. Unfolded regions of antitrypsin in the guanidine denatured intermediate (left) and when adsorbed on Phenyl 650M, Butyl 650M, and Hexyl 650C in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , pH 7.0 buffer at 22°C (right). Red represents those regions unfolded in the solution intermediate (left) and on all media (right). Orange represents those regions that unfold on Butyl 650M and Hexyl 650C to a greater extent than on Phenyl 650M. Blue arrows designate regions that show similar stability between the two cases.**

The combined results of HDX mapping for both of these two domain protein systems support the idea that the domains of a protein which are more stable in solution will also be more stable on hydrophobic surfaces. A relation between solution unfolding



and adsorption unfolding should be considered important for developing methods to predict protein unfolding on HIC surfaces. The similarity between the folded and unfolded residues of antitrypsin between the solution intermediate and on the HIC surfaces supports the hypothesis that a protein will have similar conformational states under different denaturing conditions. Further, the consistency of the transferrin N-terminal domain having lower stability in solution and on the surface supports the hypothesis that individual domain stability in solution is also significant in anticipating what groups of residues or domains will be most affected by a chromatographic surface.

#### **2.4.4 Unfolding effects of surface and GdnHCl**

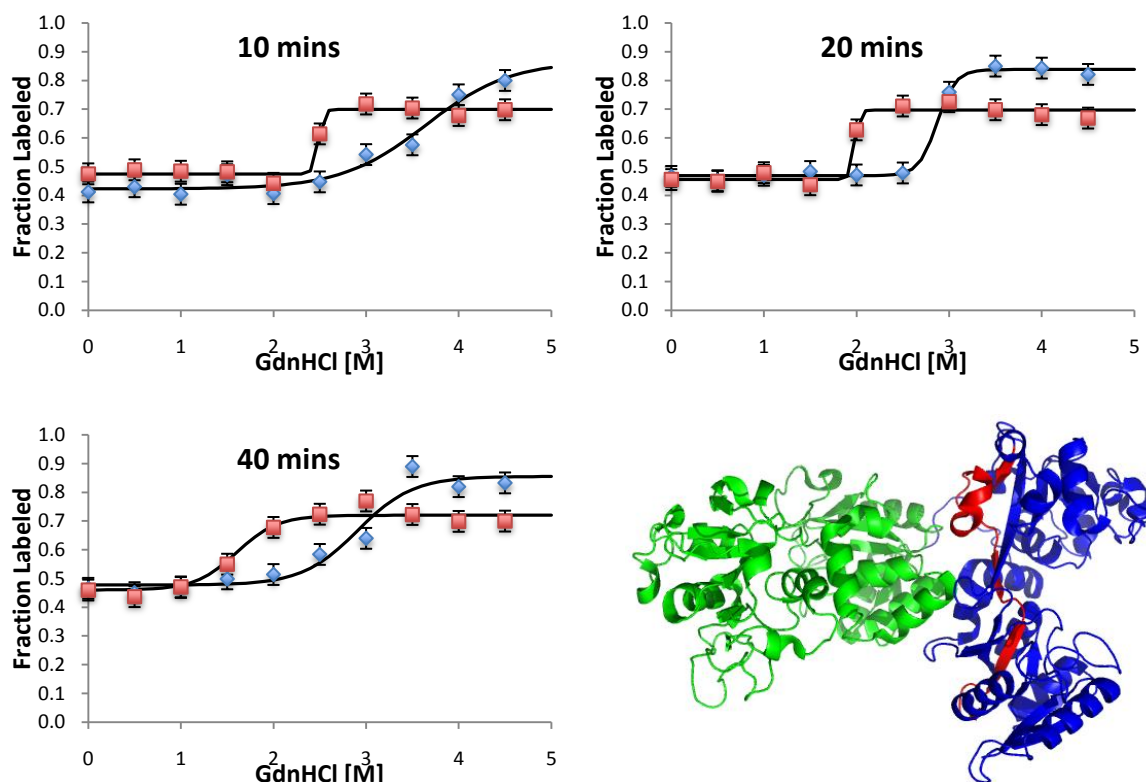
Although the above results show that (1) the most stable domain of a two domain protein tends to be the more stable domain when adsorbed on a hydrophobic surface, and (2) more hydrophobic surfaces tend to be more denaturing. However, up to now, there is no simple theory which can predict whether a protein will unfold.

Knowledge of the thermodynamic effect of a chromatographic surface on the free energy of unfolding,  $\Delta\Delta G_u$ , and its kinetic effect,  $\Delta k_u$ , together with knowledge of  $\Delta G_u$  and  $k_u$  of a solution protein could theoretically be used to determine if that protein will unfold on the surface. Here, we attempt to determine these parameters with HX-MS for transferrin labeled in solution and while adsorbed on Hexyl 650C in increasing concentrations of GdnHCl. This method has been used previously to determine the  $\Delta G_u$  of the two domains of transferrin in solution (Tang et al. 2007).

The sequence coverage, 33%, is the same as the temperature studies with transferrin. Many of the reporter peptides, however, do not have large increases in  $D/N$

over the concentration range of GdnHCl studied. This lack of change results from reporter peptides that are already mostly solvent-exposed in solution and/or on the surface in the absence of GdnHCl ( $D/N$  values greater than 0.8). The transition conditions for these peptides are more difficult to discern, as solvent exposure cannot increase very much more. Nine peptides with well-defined transition regions were chosen for analysis: four from the N-terminal domain and five from the C-terminal domain. Of these nine, the peptides with lowest  $\Delta G_u$  from the two domains are presented in detail below. The lowest (smallest positive value)  $\Delta G_u$  were chosen as they best represent the slowest exchanging residues which have been previously identified as the best measure of the global unfolding reaction (Bai et al. 1995)

The first reporter peptide presented here illustrates how the surface affects labeling of the N-terminal domain of transferrin in the presence of varying concentrations of GdnHCl. The  $D/N$  values for Residues 223 to 259 are shown in Figure 2.12. The  $D/N$  values at 0M GdnHCl are already greater than 0, as a number of these residues are solvent accessible on the protein surface. In general, the  $D/N$  value is constant as the concentration of GdnHCl is increased from 0M until the concentration where unfolding begins. Within the transition region, the  $D/N$  value increases with increasing GdnHCl concentration. This trend continues until these residues are fully unfolded and the  $D/N$  value reaches its highest point and remains constant with increasing GdnHCl concentration.



**Figure 2.12.** Unfolding of residues 223 to 259 in the N-terminal domain of transferrin in solution (blue, diamonds), and on Hexyl 650C (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 5M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 2HAV) are shown in red. The N-terminal and C-terminal domain of transferrin are shown in blue and green, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines.

At the low concentrations of 0 to 1.5M GdnHCl, the  $D/N$  of Residues 223 to 259 are statistically the same when adsorbed on Hexyl 650C as in solution. This suggests that Hexyl 650C does not affect the stability of this region (represented by Residues 223 to 259) of the protein. However, as the GdnHCl concentration increases, the transition to unfolding occurs at lower concentrations for adsorbed transferrin. Table 2.2 lists the midpoints,  $C_{1/2}$ , and slopes,  $p$ , obtained from fitting Equation (2.9) for the unfolding transitions in solution and when adsorbed on Hexyl 650C for the different labeling times. Equation (2.9) was also

used a convenient interpolation tool to provide more data points in the transition region.

The large uncertainties of  $p$  are due to the lack of data points in the transition region.

**Table 2.2. Characteristics of labeling patterns for transferrin Residues 223-259 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9).**

	10 minutes	20 minutes	40 minutes
<b>Solution <math>C_{1/2}</math></b>	3.6±0.6	2.9±0.1	2.9±0.4
<b>Hexyl 650C <math>C_{1/2}</math></b>	2.4±0.2	2.0±0.5	1.6±0.4
<b>Solution <math>p</math></b>	2.2±2.0	10±8.2	3.4±3.0
<b>Hexyl 650C <math>p</math></b>	34.9±31.2	39±34.5	4.5±3.1

The second reporter peptide presented here illustrates how the surface affects the labeling of the C-terminal domain of transferrin in the presence of varying concentrations of GdnHCl. The  $D/N$  values for Residues 554 to 589 are shown in Figure 2.13. The behavior of these residues is similar to those of Residues 223 to 259. The  $D/N$  for this reporter peptide in solution and when adsorbed is constant at lower concentrations of GdnHCl, but then increases until it reaches a new constant value that does not vary with increasing concentration of GdnHCl. Again, as the GdnHCl concentration increases, the transition to unfolding occurs at lower concentrations for adsorbed transferrin. Table 2.3 lists the characteristics of the denaturation curve in solution and while adsorbed.

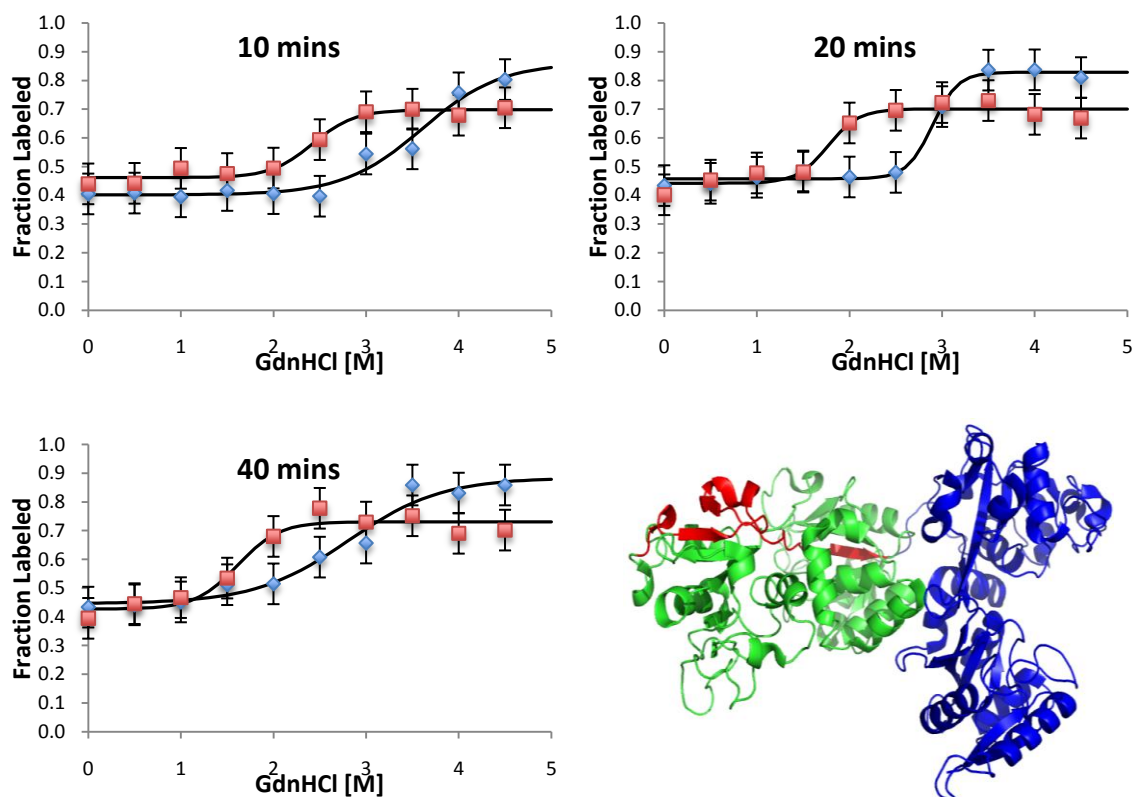


Figure 2.13. Unfolding of residues 554 to 589 in the C-terminal domain of transferrin in solution (blue, diamonds), and on Hexyl 650C (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 5M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 2HAV) are shown in red. The N-terminal and C-terminal domain of transferrin are shown in blue and green, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines.

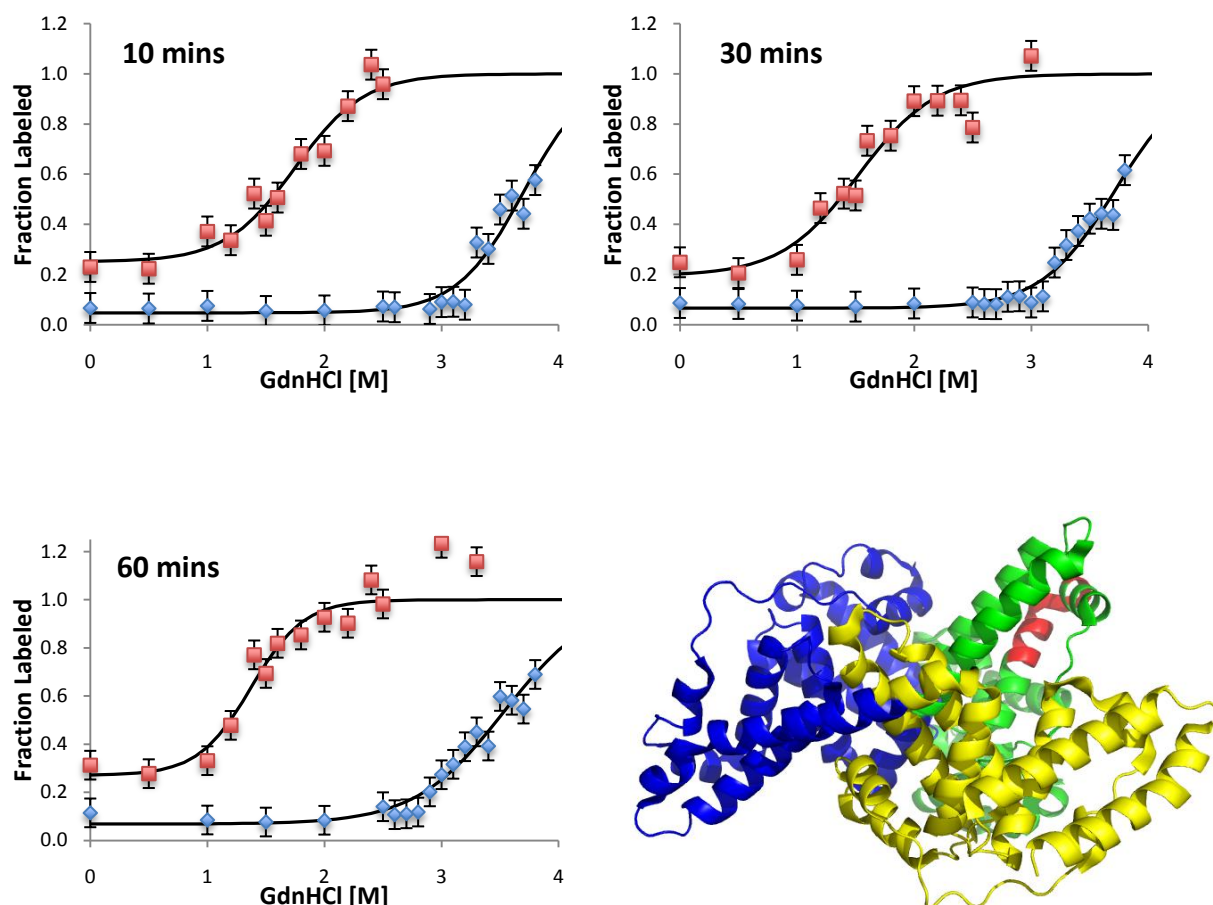
Table 2.3. Characteristics of labeling patterns for transferrin Residues 554-589 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9).

	10 minutes	20 minutes	40 minutes
<b>Solution <math>C_{1/2}</math></b>	$3.6 \pm 0.6$	$2.9 \pm 0.3$	$2.8 \pm 0.4$
<b>Hexyl 650C <math>C_{1/2}</math></b>	$2.4 \pm 0.2$	$1.8 \pm 0.3$	$1.6 \pm 0.3$
<b>Solution <math>p</math></b>	$2.3 \pm 2.0$	$7.3 \pm 7.0$	$2.0 \pm 1.8$
<b>Hexyl 650C <math>p</math></b>	$4.4 \pm 3.8$	$6.0 \pm 5.0$	$4.4 \pm 4.0$

Although  $\Delta G_u$  and  $k_u$  were calculated for these 2 reporter peptides and the other 7 with well-defined transition regions (see Table A4), the uncertainties in these values are too large to make statistically meaningful comparisons. To address this issue of large uncertainties, BSA-Phenyl 650M was investigated as a second system with an emphasis on data points in the transition region as discussed below.

MS spectra were collected and compared for BSA in solution and while adsorbed on Phenyl 650M in increasing concentrations of GdnHCl. Phenyl 650M was chosen for this study as Butyl 650M and Hexyl 650C cause most of the protein to unfold even at low concentrations of GdnHCl, as will be presented and discussed in Section 2.4.5. Again, not all reporter peptides are suitable for analysis since those mostly solvent-exposed ( $D/N > 0.8$ ) with no GdnHCl will not have well-defined transition curves. The reporter peptides with the lowest  $\Delta G_u$  from each of the three domains are presented here. The lowest (smallest positive value)  $\Delta G_u$  were chosen as they best represent the slowest exchanging residues which have been previously identified as the best measure of global unfolding.

The  $D/N$  values for Residues 127 to 137 from Domain I of BSA in the presence of varying concentrations of GdnHCl are shown in Figure 2.14. In general, the labeling trends for these residues are similar to those of transferrin. In solution, at low concentrations of GdnHCl, the  $D/N$  values remain constant with increasing GdnHCl concentration until unfolding begins. Then, the  $D/N$  values increase with increasing GdnHCl concentration. In solution, the  $D/N$  is limited to 0.69 at 3.8M GdnHCl and 60 minutes labeling. For this region, higher concentrations of GdnHCl are probably needed for full unfolding. The low solubility of BSA at GdnHCl concentrations above 4M prevented data collection at higher concentrations.



**Figure 2.14.** Unfolding of residues 127 to 137 in Domain I of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines.

The destabilizing effect of Phenyl 650M is observed by comparing its labeling patterns to that in solution. On Phenyl 650M, the unfolding transition region occurs at lower GdnHCl concentrations similar to transferrin on Hexyl 650C. Further, the  $D/N$  on the surfaces reaches full solvent exposure value of 1.0. It is noteworthy that the midpoint of the transition region is affected significantly by adsorption to Phenyl 650M but the slope is not. Also noteworthy is that unlike transferrin, labeling at 0M GdnHCl is already higher for

adsorbed BSA than in solution. The lower folding stability of BSA (compared to transferrin) likely causes BSA to partially unfold on Phenyl 650M even in the absence of chemical denaturants. Calculated  $\Delta G_u$  and  $k_u$  in this case is applicable only to the unfolding from state at 0M GdnHCl to the fully unfolded state at higher GdnHCl. Table 2.4 lists the  $C_{1/2}$  and  $p$ -values obtained from fitting Equation (2.9) to the labeling data in solution and when adsorbed on Phenyl 650M for the different labeling times.

**Table 2.4. Characteristics of labeling patterns for BSA Residues 127-137 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9).**

	10 minutes	30 minutes	60 minutes
<b>Solution <math>C_{1/2}</math></b>	3.7±0.1	3.7±0.1	3.5±0.1
<b>Phenyl 650M <math>C_{1/2}</math></b>	1.7±0.3	1.5±0.4	1.4±0.2
<b>Solution <math>p</math>-value</b>	3.6±1.3	3.2±0.8	2.6±0.7
<b>Phenyl 650M <math>p</math>-value</b>	3.4±2.9	3.0±3.7	4.4±4.1

The  $D/N$  values for Residues 333 to 340 from Domain II of BSA in the presence of varying concentrations of GdnHCl are shown in Figure 2.15. The trends are similar to those of Residues 127 to 137. On the chromatographic surface, the unfolding transition occurs at lower concentrations of GdnHCl. In solution, the unfolding transition does not occur until higher concentrations of GdnHCl with the  $D/N$  value limited to 0.73. Again, labeling at 0M GdnHCl is already higher for adsorbed BSA than in solution limiting the application of calculated  $\Delta G_u$  and  $k_u$  to unfolding of the 0M adsorbed state to the fully-unfolded state. Table 2.5 lists the characteristics of the labeling patterns in solution and while adsorbed.



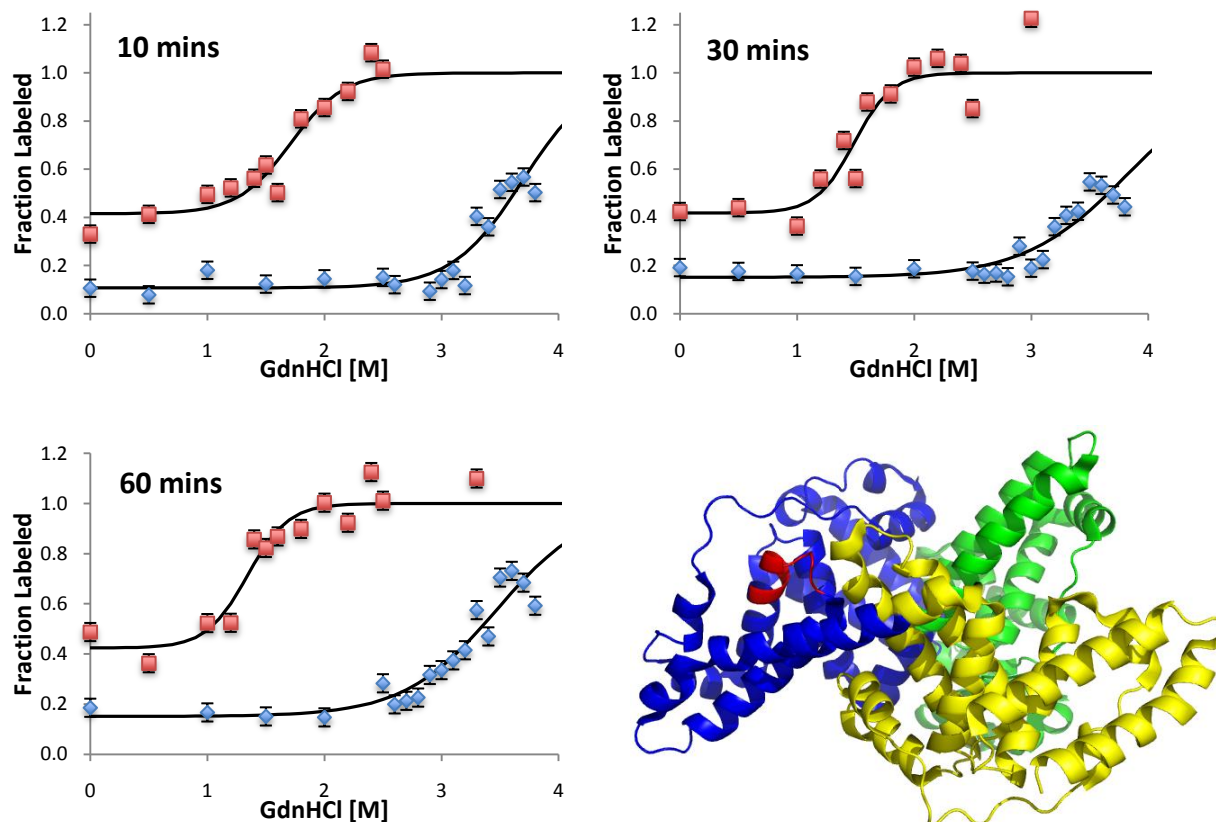


Figure 2.15. Unfolding of residues 333 to 340 in Domain II of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines.

Table 2.5. Characteristics of labeling patterns for BSA Residues 333-340 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9).

	10 minutes	30 minutes	60 minutes
<b>Solution <math>C_{1/2}</math></b>	$3.7 \pm 0.1$	$3.8 \pm 0.2$	$3.5 \pm 0.1$
<b>Phenyl 650M <math>C_{1/2}</math></b>	$1.7 \pm 0.3$	$1.5 \pm 0.3$	$1.3 \pm 0.2$
<b>Solution <math>p</math>-value</b>	$3.3 \pm 1.6$	$2.3 \pm 1.1$	$2.5 \pm 0.9$
<b>Phenyl 650M <math>p</math>-value</b>	$4.4 \pm 4.5$	$6.2 \pm 12.4$	$5.8 \pm 7.0$

The  $D/N$  values for Residues 529 to 543 from Domain III of BSA in the presence of varying GdnHCl concentrations are shown in Figure 2.16. Compared to the two previous reporter peptides, this reporter peptide reaches  $D/N$  values of 1.0 in solution. It is noteworthy that at 60 minutes labeling on the surface (bottom left of Figure 2.16) the  $D/N$  (0.9) is within experimental error between 1.0 and 2.5M GdnHCl but increases to 1.2 and 1.1 at 3.0 and 3.3M GdnHCl suggesting three-state unfolding as a possibility. Although three-state unfolding monitored with HXMS has been observed previously for the case of transferrin in solution (Tang et al. 2007), the unusual  $D/N$  values of  $>1.0$  suggests these data points may be outliers. For the purposes of this study, these two data points were ignored with  $\Delta G_u$  and  $k_u$  calculated from interpolation of the curve between only 0 and 2.5M GdnHCl. The other trends are similar to the two previous reporter peptides. Table 2.6 lists the characteristics of the labeling patterns in solution and while adsorbed.

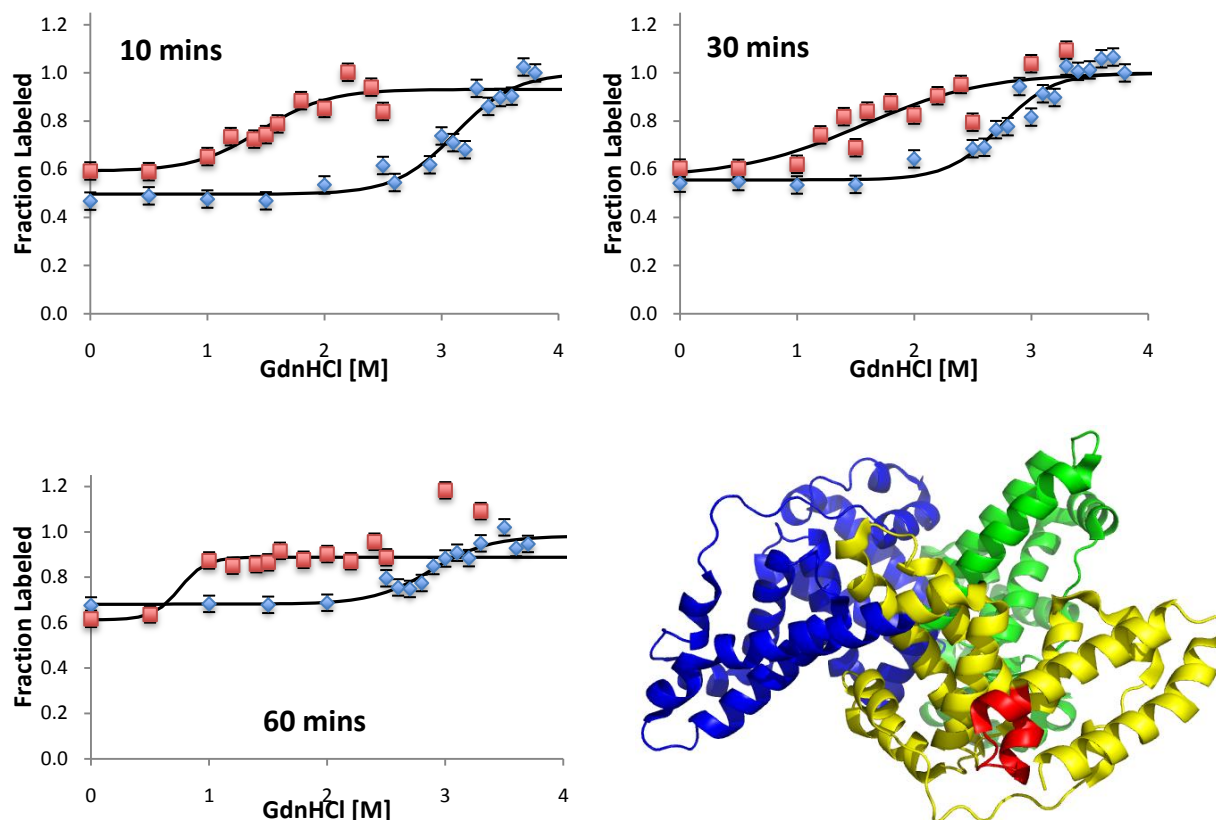


Figure 2.16. Unfolding of residues 529 to 543 in Domain III of BSA in solution (blue, diamonds), and on Phenyl 650M (red, squares) in 25mM  $\text{PO}_4$ , 1.5M  $(\text{NH}_4)_2\text{SO}_4$ , 0 to 4M GdnHCl, pH 7.0 buffer at 22°C and various labeling times. The location of the residues on the native structure (PDB ID 3V03) are shown in red. Domain I, II, and III are shown in green, blue, and yellow, respectively. Fits of the experimental data with Equation (2.9) are shown as solid lines.

Table 2.6. Characteristics of labeling patterns for BSA Residues 529-543 in the presence of varying GdnHCl concentrations. Uncertainties presented were determined by the MATLAB program used to fit the data to Equation (2.9).

	10 minutes	30 minutes	60 minutes
<b>Solution <math>C_{1/2}</math></b>	3.1±0.2	2.7±0.1	2.9±0.2
<b>Phenyl 650M <math>C_{1/2}</math></b>	1.5±0.3	1.6±0.7	0.75±0.32
<b>Solution <math>p</math>-value</b>	3.9±3.2	4.5±2.9	4.0±2.7
<b>Phenyl 650M <math>p</math>-value</b>	3.5±4.4	1.9±3.1	8.9±10.6

Interpolated values for  $[N]$  were calculated using Equations (2.9), (2.10), and (2.8) as described in the methods. These  $[N]$  values were fitted vs. labeling time to determine  $N_o$ , using Equation (2.11).  $N_o$  values were then used to calculate  $\Delta G_u$  using Equation (2.12), and the denaturant dependence of  $\Delta G_u$  in Equation (2.13) was used to determine  $m$  and  $\Delta G_u$  in the absence of denaturant,  $\Delta G_{u,o}$ . Solution and adsorbed  $\Delta G_u$ ,  $m$ -values, and  $k_u$  for Residues 127-137, 333-340, and 529-543 are shown in Table 2.7. In solution, domain stability based on  $\Delta G_u$  magnitude follows the order: I > II > III, although all three are the same within experimental uncertainty. There are conflicting studies in the literature as to which domain is most stable in solution. However, there is general consensus that Domain III is the least stable and unfolds first in chemical denaturation (Khan et al. 1987; Tayyab et al. 2000; Togashi et al. 2010; Wu et al. 2011). This order of domain stability is also seen on Phenyl 650M. Adsorption to the surface lowers the stability,  $\Delta\Delta G_u$ , of Domains I, II, and III by 4.5, 4.0, and 4.0 kcal/mol, respectively. This is consistent with the hypothesis that a surface has the same effect on all domains of a protein regardless of its solution stability. If true, then differences in how individual domains unfold or do not unfold on the surface then arise solely from differences in solution stability.

**Table 2.7. Unfolding Rates and Free Energies of BSA on Phenyl 650M**

	Domain I <sup>a</sup>	Domain II <sup>b</sup>	Domain III <sup>c</sup>
$\Delta G_{u,sol}$ [kcal/mol] <sup>1</sup>	10.9±5.5	7.8±3.4	7.2±3.0
$\Delta G_{u,ads}$ [kcal/mol] <sup>1</sup>	6.4±3.1	4.8±2.6	3.2±1.2
$\Delta\Delta G_u$ [kcal/mol] <sup>2</sup>	4.5±6.3	4.0±4.3	4.0±3.2
$m_{sol}$ [kcal/mol•M] <sup>1</sup>	2.3±1.1	2.1±1.1	2.4±1.0
$m_{ads}$ [kcal/mol•M] <sup>1</sup>	2.0±1.1	2.7±1.3	1.8±0.7
$m_{ref}$ [kcal/mol•M] <sup>3</sup>	3.7±1.1	3.5±1.3	3.5±1.3
$k_{u,sol}$ [min <sup>-1</sup> ] <sup>4</sup>	0.012±0.008	0.012±0.008	0.030±0.012
$k_{u,ads}$ [min <sup>-1</sup> ] <sup>4</sup>	0.015±0.010	0.017±0.009	~
$\Delta k_u$ [min <sup>-1</sup> ] <sup>5</sup>	0.003±0.012	0.005±0.012	~

<sup>a</sup> Data from GdnHCl denaturation curve of Residues 127-137 used

<sup>b</sup> Data from GdnHCl denaturation curve of Residues 333-340 used

<sup>c</sup> Data from GdnHCl denaturation curve of Residues 529-543 used

<sup>1</sup> Calculated with Equations 2.11, 2.12, and 2.13

<sup>2</sup> Calculated with Equation 2.14

<sup>3</sup> Calculated with Equation 12 from (Myers et al. 1995):

$$m=(0.28\pm 0.03)\times[\Delta ASA-(792\pm 780)(\# \text{ crosslinks})]$$

<sup>4</sup> Calculated with Equation 2.11

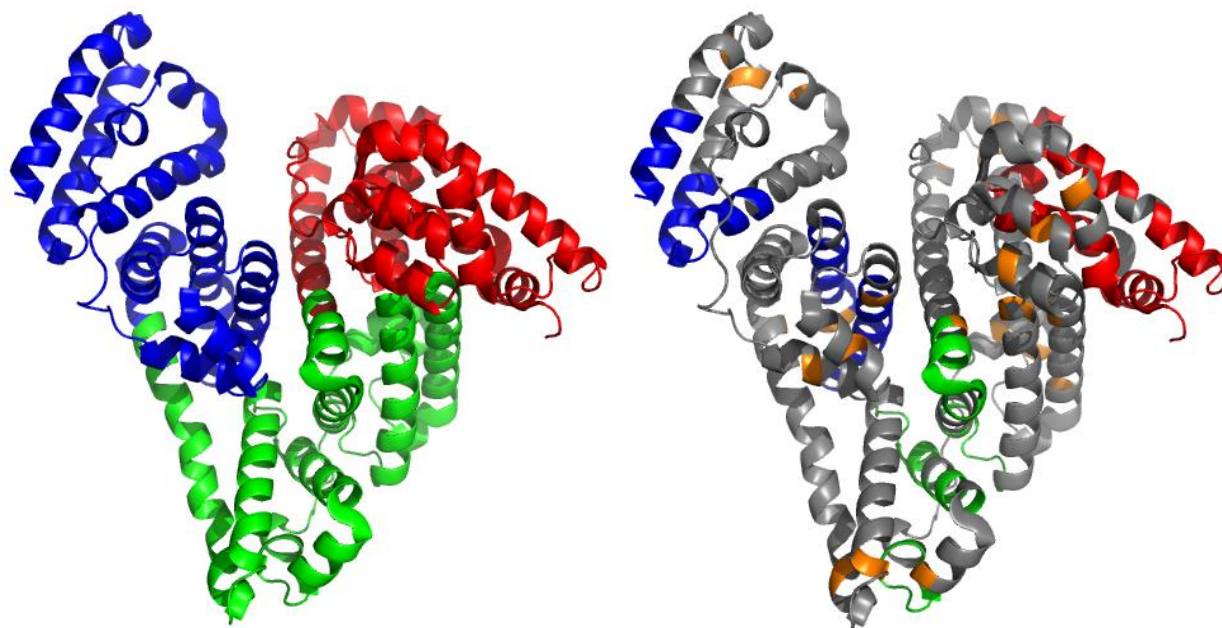
<sup>5</sup> Calculated with Equation 2.15

The different destabilizing effects of Hexyl 650C and Phenyl 650M are observed from comparison of the transferrin and BSA GdnHCl denaturation curves. The largest effect Hexyl 650C has on a domain's free energy of unfolding is 6.4 kcal/mol (Table A4), larger than Phenyl 650M's largest effect, 4.5 kcal/mol (Table 2.8). The effect on the unfolding rate is also greater for Hexyl 650C, 0.094 min<sup>-1</sup> (Table A4), than for Phenyl 650M, 0.005 min<sup>-1</sup> (Table 2.7). The larger  $\Delta\Delta G_u$  and  $\Delta k_u$  for Hexyl 650C seems consistent with our general observation that Hexyl 650C is more destabilizing to a protein than Phenyl 650M. In the temperature studies, the reporter peptides of antitrypsin and transferrin are generally more labeled when adsorbed on Hexyl 650C than on Phenyl 650M (see Figures

2.5 and 2.8) which is consistent with the larger  $\Delta\Delta G_u$  and  $\Delta k_u$  for Hexyl 650C. Finally, the  $m$ -values for BSA are similar in solution and on the surface as they are for transferrin. The  $m$ -values determined here compare favorably with theoretical values determined by the domain-specific change in accessible surface area,  $\Delta ASA$ , upon unfolding and by the number of disulfide crosslinks present (Myers et al. 1995). Interestingly, the  $m$ -value does not change significantly on the surface as might be expected from a different  $\Delta ASA$  compared to in solution. The  $\Delta ASA$  is expected to be less if some of the protein surface area interacts with the chromatographic surface.

#### **2.4.5. Solvent protection increase on the chromatographic surface**

As a third protein system, MS spectra were also collected and compared for BSA labeled in solution and while adsorbed on Butyl 650M and Hexyl 650C at different temperatures. The mass spectra for the reporter peptides were converted to  $D/N$  values using Equation (2.6). A total of 34% sequence coverage was obtained for the 585 residues of BSA. Figure 2.17 shows the native structure of BSA and its three domains (left) and the location of the reporter peptides and cysteines of BSA (right). Many of the areas with missing sequence coverage are populated with cysteines involved in disulfide bridges which can impact proteolytic digestion. Although reporter peptides are available for all three domains, the most coverage is obtained in Domain I, 40%, followed by Domain III, 37%. Domain II has the lowest coverage at 25%.



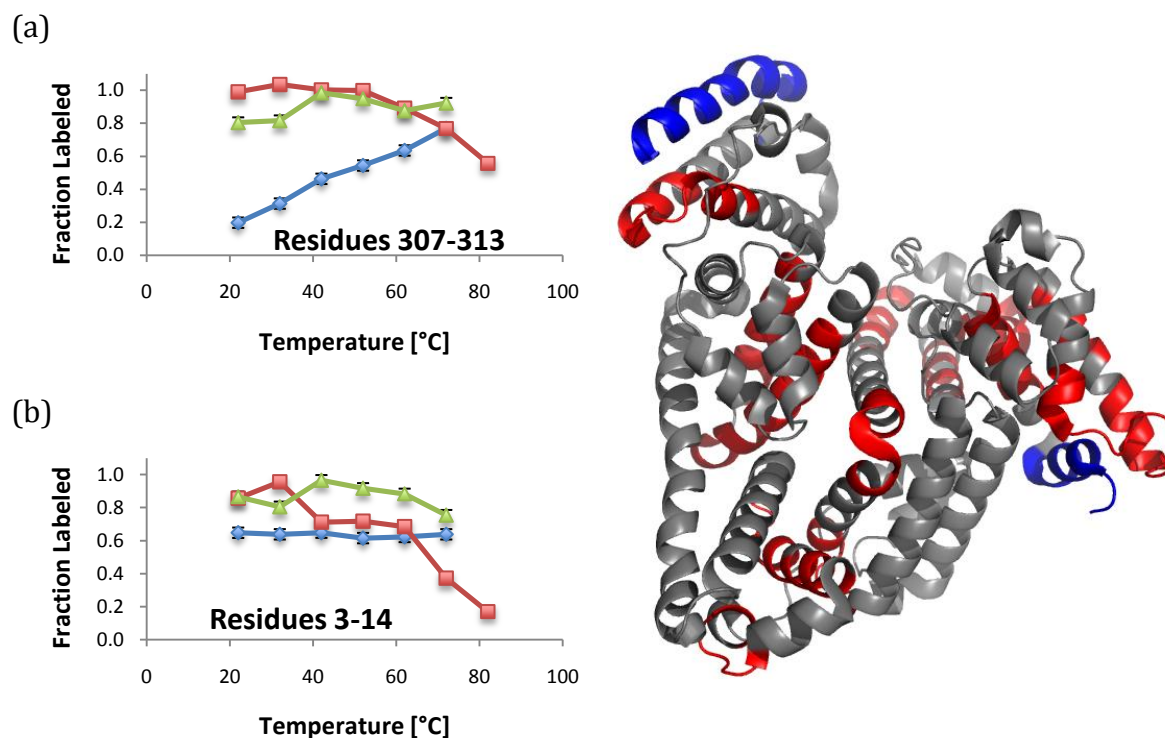
**Figure 2.17. Left: Native structure of BSA (PDB ID 3V03) highlighting Domain I (red), Domain II (green), and Domain III (blue). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts Cystines of BSA. Grey depicts missing sequence coverage.**

Unlike the two previous proteins, a new labeling pattern emerges on the surfaces for the reporter peptides of BSA where solvent protection increases with temperature. In solution, the  $D/N$  either increases with temperature or stays constant. Although minor variations exist, the other reporter peptides of BSA generally follow this pattern.

The  $D/N$  values for two reporter peptides where solvent protection on the surface increases with temperature are shown in Figure 2.18. Residues 307-313, shown in panel (a), are representative of reporter peptides in that (1) there was substantially more labeling on all surfaces at low temperature, (2) labeling increased in solution with an increase in temperature but become more solvent-protected on the surface at high temperatures. In solution, the solvent exposure of Residues 307-313 increases with temperature starting at 22°C. This result is consistent with the low stability of BSA that has

a reported melting temperature of 42°C (Lin & Koenig 1976). At 22°C on Butyl 650M, this region is fully unfolded on the surface while on Hexyl 650C it is mostly unfolded with a  $D/N$  value of 0.81. On the Hexyl 650C surface, the solvent exposure increases between 32 and 42°C then decreases between 42 and 62°C. On the Butyl 650M surface, the solvent exposure decreases progressively between 52 and 82°C; the reverse of what occurs in solution. In addition to Residues 307-313, this trend is also observed for Residues 20-31, 39-45, 49-70, 127-137, 154-164, 169-182, 200-209, 219-226, 307-313, 324-329, 331-340, 422-436, 440-460, and 530-547. These residues are shown in red on the native structure on the right of Figure 2.18.





**Figure 2.18. Thermal unfolding and solvent protection increase of different regions of BSA in solution (blue, diamonds), on Butyl 650M (red, squares), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 82°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 307-313 is shown in panel (a). The *D/N* for Residues 3-14 is shown in panel (b). Right: Location of Residues 307-313 and the other residues with similar unfolding pattern are shown in red on the native structure (PDB ID 3V03). Residues 3-14 and the other residues where no increased labeling in solution is observed are shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Butyl 650M at 22°C.**

The two remaining reporter peptides also showed decreased labeling with increasing temperature on Butyl 650M, but did not show an increase in solution labeling with temperature. The *D/N* values for Residues 3 to 14 are shown in Figure 2.18b and their location on the native structure are shown in blue on the right. In solution, this region has a *D/N* value of 0.65 at 22°C that does not vary with temperature. On Butyl 650M and Hexyl 650C, this region is more solvent-exposed at 22°C but not fully unfolded. Above 62°C, the *D/N* values decrease on both surfaces. The decrease is the largest on Butyl 650M with the

$D/N$  value reduced to 0.17 at 82°C. In addition to Residues 3-14, this trend is also observed for Residues 559-584. Both groups of residues are located at the termini of BSA, which may be related to why no increase in  $D/N$  with increasing temperature is observed.

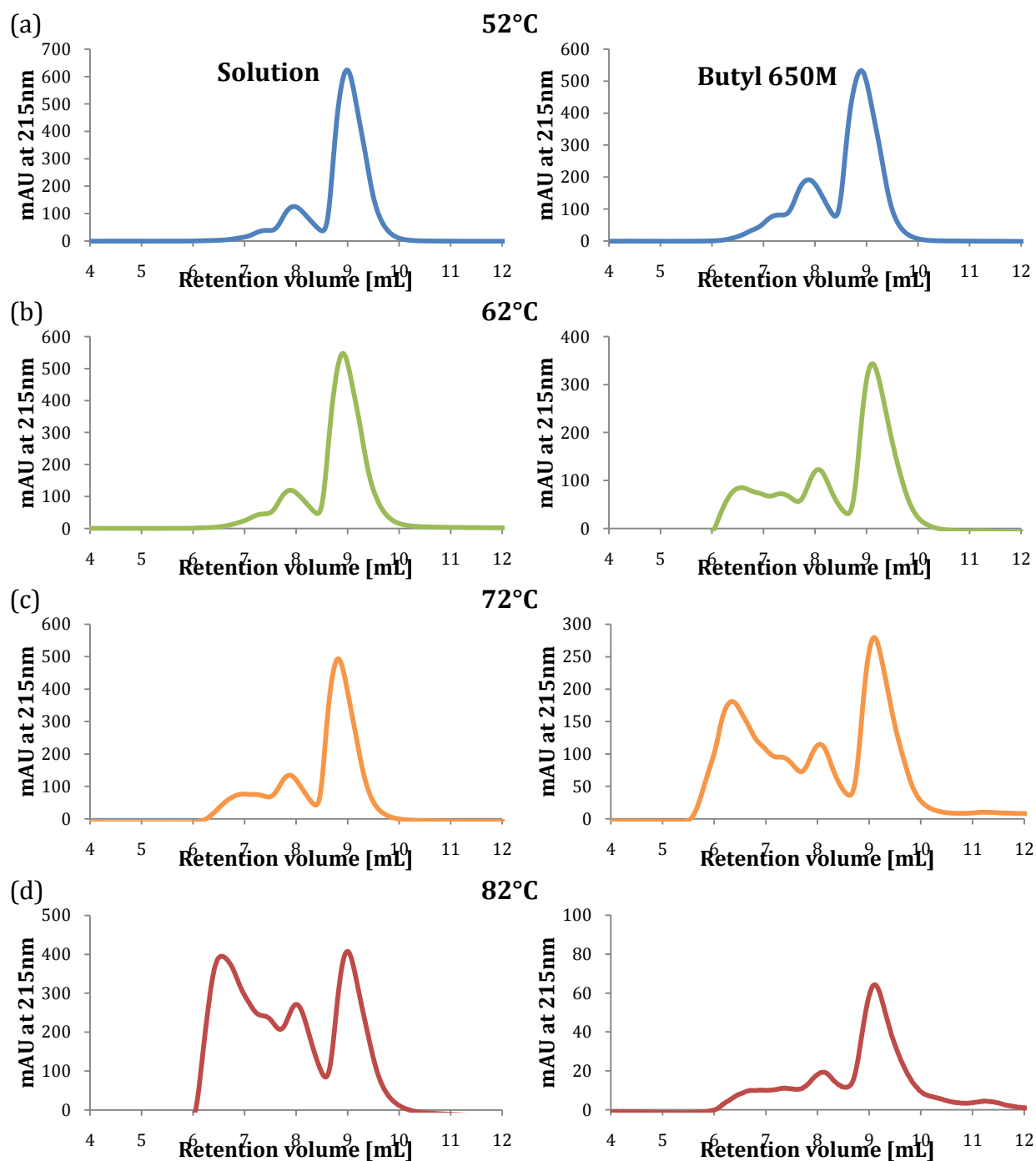
As mentioned in Section 2.4.4, Domain III of BSA is frequently cited as the least stable domain of BSA. The lower value of  $\Delta G_u$  in this study for Domain III (relative to the other domains) in solution and on Phenyl 650M is consistent with these observations. However, no distinct patterns are observed for the different domains of BSA during its thermal denaturation on Butyl 650M and Hexyl 650C. Instead a decrease in labeling is observed for certain reporter peptides in each of the three domains. Two scenarios may explain this decrease in labeling on the surface relative to solution. First, a decrease in labeling on the surface may occur from enhanced protection for the regions of the protein that bind to the ligands as has been observed previously. In one study, hydrogen-deuterium exchange rates were measured by NMR for bovine  $\alpha$ -lactalbumin in solution and adsorbed on hydrophobic, polystyrene nanospheres (Engel et al. 2004). In addition to identifying residues with less protection, several residues of bovine  $\alpha$ -lactalbumin were identified with enhanced protection when adsorbed on the polystyrene nanospheres. The authors hypothesized that these residues with enhanced protection, located in one region on lactalbumin, directly interact with the polystyrene surface leading to the increase in protection. Such an interaction with the surface would explain why certain regions of transferrin are less labeled on the surface than when in solution, as will be shown and discussed later.

A second possibility is that protein-protein interactions on the surface may reduce solvent accessibility for residues of the protein involved in such contacts. Indeed, hydrogen

deuterium exchange rates have been shown to be reduced by intermolecular interactions formed in bovine insulin dimers (Tokihiro et al. 1996), diphtheria toxin oligomers (Man et al. 2010), and aggregation of amyloid beta peptides (Qi et al. 2008).

SEC was performed to evaluate if irreversible aggregation of BSA on the surface might be leading to these decreases in solvent exposure. Figure 2.19 shows SEC chromatograms of BSA in solution (left) and adsorbed on Butyl 650M (right) incubated at temperatures of 52, 62, 72, and 82°C. At temperatures of 52°C (Figure 2.19a) and lower (data not shown) 2 distinct peaks are observed in solution (left) and eluted from Butyl 650M (right). These larger and smaller peak have been characterized previously as monomers and dimers, respectively (Suda et al. 2009). Although the same peaks are observed both in solution and from protein eluted from Butyl 650M, the dimer peak is larger for the protein sample exposed to the HIC surface. As temperature is increased, two additional higher molecular weight peaks are on Butyl 650M at 62C (Figure 2.19b, right) while not such peaks do not appear until 82C in solution (Figure 2.19d, left). On the surface the two additional peaks become larger at 72C (Figure 2.19c, right). They become smaller at 82C (Figure 2.19d, right) but this is likely due to poor recovery from irreversibly bound protein as judged by the comparatively low 215nm signal.

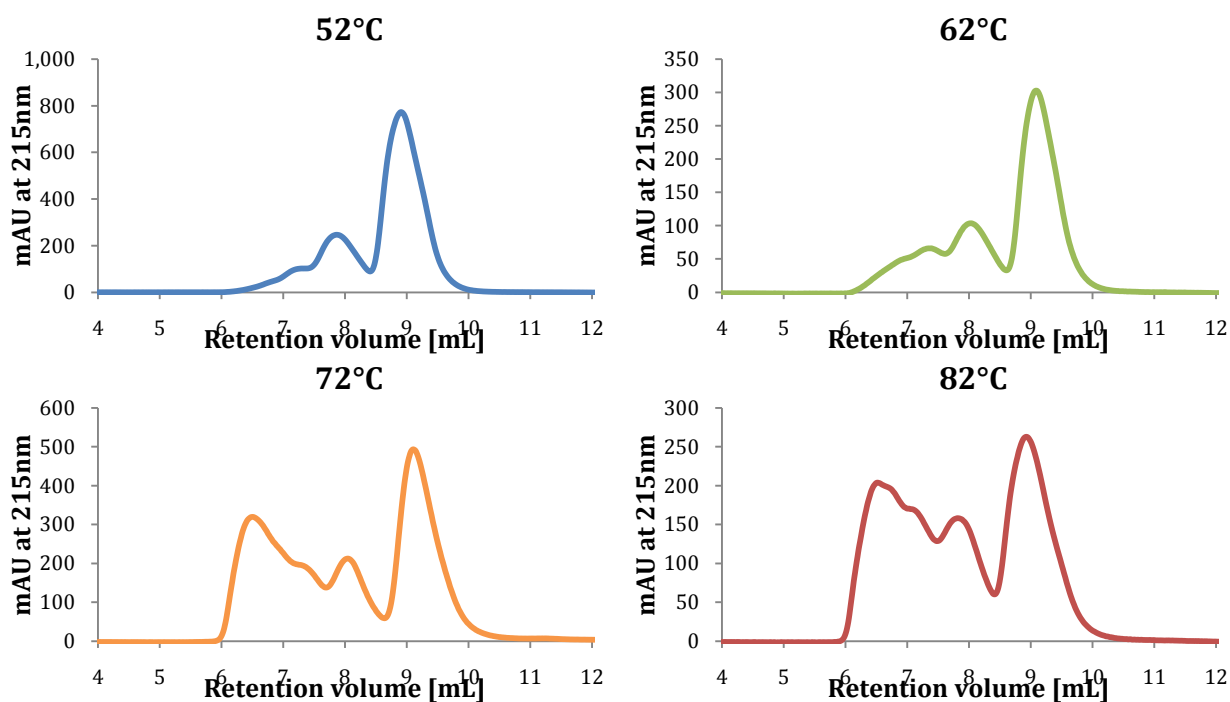
It is likely that the two additional peaks are oligomers and/or aggregates, as their SEC retention times are less than those of the monomer and dimer peaks. Thus, the presence of the HIC surface shifts the distribution of the species present towards the oligomers/aggregates at elevated temperatures. Aggregation on the surface is likely nonnative as increases in solvent exposure relative to the solution are observed before solvent protection begins to decrease at higher temperatures (see Figure 2.18).



**Figure 2.19.** SEC chromatograms of BSA in solution (left) and eluted from Butyl 650M (right) at temperatures of 52°C (a), 62°C (b), 72°C (c), and 82°C (d). Samples were incubated for 1 hour, eluted with 200  $\mu$ L of desorption buffer, diluted with 500  $\mu$ L of working buffer, and loaded on a TSK gel G3000SWXL column for analysis.

SEC chromatograms were also obtained for BSA adsorbed on Phenyl 650M at various temperatures. Figure 2.20 shows SEC chromatograms adsorbed on Phenyl 650M

and incubated at 52°C (top left), 62°C (top right), 72°C (bottom left), and 82°C (bottom right). Similar to BSA on Butyl 650M, two peaks are observed at temperatures of 52°C and lower. At 62°C on Phenyl 650M, additional peak(s) with similar retention to the oligomer/aggregate peaks on Butyl 650M at 62°C are observed. As with BSA on Butyl 650M, these peaks then increase in magnitude as temperature increases to 72 and then 82°C.



**Figure 2.20.** SEC chromatograms of BSA eluted from Phenyl 650M at temperatures of 52°C (top left), 62°C (top right), 72°C (bottom left), and 82°C (bottom right). Samples were incubated for 1 hour, eluted with 200  $\mu$ L of desorption buffer, diluted with 500  $\mu$ L of working buffer, and loaded on a TSK gel G3000SWXL column for analysis.

The reporter peptides with the largest increases in solvent protection are consistent with predicted areas of aggregation. The residue groups with  $D/N$  values less than 0.1 at

80°C are highlighted in red in Figure 2.21 while those regions predicted to be aggregation “hot spots” by the aggregation calculators AGGRESCAN (Conchillo-Solé et al. 2007) and PASTA (Trovato et al. 2007) are shown with solid and dashed lines, respectively. Residues in grey are regions with missing sequence coverage, while those in black are covered but do not show large increases in solvent protection (to below  $D/N$  of 0.2). Seven residue groups of BSA that unfold on Butyl 650M ( $D/N > 0.5$ ) at lower temperatures become mostly solvent-protected ( $D/N < 0.1$ ) at 80°C. Of these 7, 5 are correctly predicted by AGGRESCAN to be aggregation hot spots. However, three groups predicted to be hot spots by AGGRESCAN did not show large increases in solvent protection (to below  $D/N$  of 0.2). Of the 2 regions predicted to be aggregation hot spots by PASTA, 1 region becomes mostly solvent-protected at 80°C. The high degree of overlap between the hot spots and solvent-protected regions supports the idea that the regions with  $D/N$  less than 0.1 are directly involved in the nonnative aggregation of BSA.

DTHKSEIAHR	FKDLGEEHFK	GLVLIAFSQY	LQQCPFDEHV	1-40
KLVNELTEFA	KTCVADESHA	GCEKSLHTLF	GDELCKVASL	41-80
RETYGDMADC	CEKQEPERNE	CFLSHKDDSP	DLPKLPDPN	81-120
TLCDEFKADE	KKFWGKLYE	IARRHPYFYA	PELLYYANKY	121-160
NGVFQEECCQA	EDKGACLLPK	IETMREKVL	SSARQRLRCA	161-200
SIQKFGERAL	KAWSVARLSQ	KFPKAEFVEV	TKLVTDLTKV	201-240
HKECCHGDLL	ECADDRADLA	KYICDNQDTI	SSKLKECCDK	241-280
PLLEKSHCIA	EVEKDAIPEN	LPPLTADFAE	DKDVCKNYQE	281-320
AKDAFLGSFL	YEYSRRHPEY	AVSVLLRLAK	EYEATLEECC	321-360
AKDDPHACYS	TVFDKCLKHLV	DEPQNLIKQN	CDQFEKLG	361-400
GFQNALIVRY	TRKVPQVSTP	TLVEVSRSLG	KVGTRCCTKP	401-440
ESERPCTED	YLSLILNRLC	VLHEKTPVSE	KVTKCCTESL	441-480
VNRRPCFSAL	TPDETYVPKA	FDEKLFTFHA	DICTLPDTEK	481-520
QIKKQPTALVE	LLKHKPKATE	EQLKTVMENF	VAFVDKCCAA	521-560
DDKEACFAVE	GPKLVVSTQT	ALA		561-583

Figure 2.21. Overlap of solvent-protected BSA residues and predicted aggregation hot spots. Residue groups of BSA that unfold on Butyl 650M ( $D/N > 0.5$ ) at lower temperatures and become more solvent-protected ( $D/N < 0.1$ ) at 80°C are highlighted in red. Residues predicted to be aggregation hot spots by AGGRESCAN and PASTA are shown with solid and dashed lines, respectively.

Surface-induced aggregation has been observed previously although a clear mechanism has not been determined (Sluzky et al. 1991; Vermonden et al. 2001; Norde & Giacomelli 2000; Sethuraman & Belfort 2005).

One of two possibilities (or a combination of both) may explain surface-induced aggregation. First, the destabilization of the protein on the surface may remove the precursory requirement of unfolding for nonnative aggregation. In solution, an equilibrium exists between protein molecules in the native conformation and the unfolded state. Under physiological (non-denaturing) conditions the population consists mostly of native molecules ( $\Delta G_u > 0$ ). An increase in temperature will shift the equilibrium towards the unfolded molecules necessary for nonnative aggregation. Adsorption to hydrophobic surfaces can also shift the equilibrium towards unfolded molecules as observed with transferrin on Hexyl 650C and BSA on Phenyl 650M (see Tables A4 and 2.7). In these cases a larger population of unfolded molecules already exists at low temperatures. If the unfolded state of the adsorbed protein is conducive to aggregation, then adsorption may facilitate aggregation by increasing the population of molecules available for aggregation.

From a kinetic standpoint, a hydrophobic surface can also increase the unfolding rate of a protein. For nonnative aggregation, unfolding may be the rate-limiting step in the reaction mechanism. Then, if the unfolding rate is increased on the surface as with transferrin and BSA (see Tables A4 and 2.7), the overall rate will increase.

The other possibility for surface-promoted aggregation is from the surface affecting the rate at which protein molecules interact to form dimers and higher order oligomers. Collisions between protein molecules are expected to be more frequent on the 2-dimensional adsorption surface than in a 3-dimensional solution.

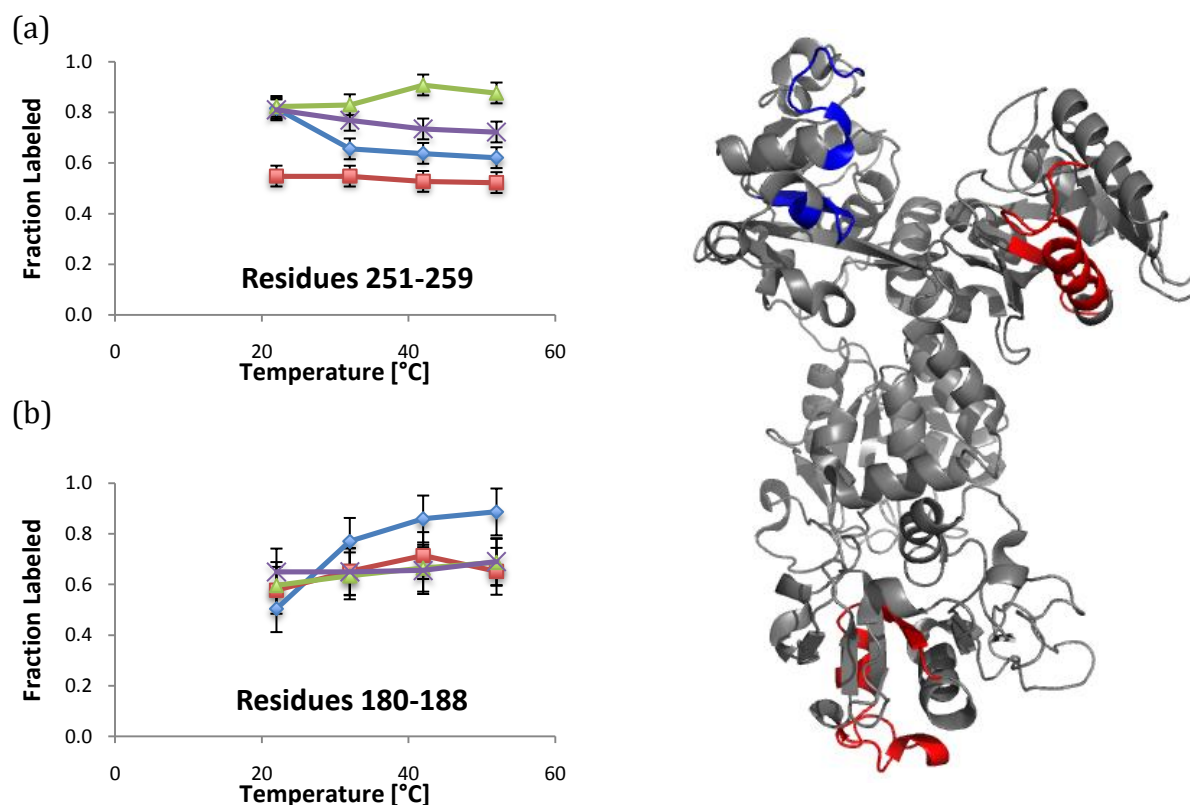
An increase in solvent protection on the surface is also observed for transferrin and antitrypsin, but not to the extent with BSA. Although most reporter peptides of transferrin and antitrypsin are more solvent-exposed on the surfaces, a few show an increase in



solvent protection when adsorbed. The increase in solvent protection is not necessarily observed on all three surfaces. Cases exist where an increase in solvent protection is observed on only Phenyl 650M. Figure 2.22 shows two such regions of transferrin where the solvent exposure is decreased on one or more surfaces. Residues 251-259, shown in panel (a), are representative of reporter peptides that unfold on Butyl 650M and Hexyl 650C, but show increased solvent protection on Phenyl 650M when compared to solution. Except at 22°C in solution, this region is more solvent-exposed when adsorbed on Butyl 650M and Hexyl 650C. Interestingly, less labeling occurs on Phenyl 650M than in solution at all temperatures. As with BSA, two possible explanations for this behavior are protection may be enhanced for the residues that bind to the HIC ligand or protein-protein interactions on the surface may limit solvent accessibility for the contacting residues. In addition to Residues 251-259, increased protection on Phenyl 650M is also observed for Residues 263-281, and 529-555. The location of these residues is shown in red on the native structure of transferrin on the right in Figure 2.22.

Two reporter peptides of transferrin showed an increase in solvent protection on all three surfaces. Residues 180-188, shown in Figure 2.22b, compose one of these two reporter peptides where solvent protection increase is not limited to just Phenyl 650M. As temperature increases, these residues unfold in solution and become more solvent-exposed with the  $D/N$  increasing from 0.50 at 22°C to 0.89 at 52°C. Yet, as temperature increases, no statistically significant increase occurs in the surface  $D/N$  value. While the labeling in solution and on the surfaces is the same at 22 and 32°C, statistically less labeling occurs on Butyl 650M and Hexyl 650C at 42°C, and less labeling occurs on all three surfaces at 52°C. Of all the regions of transferrin for which coverage was obtained, this one shows

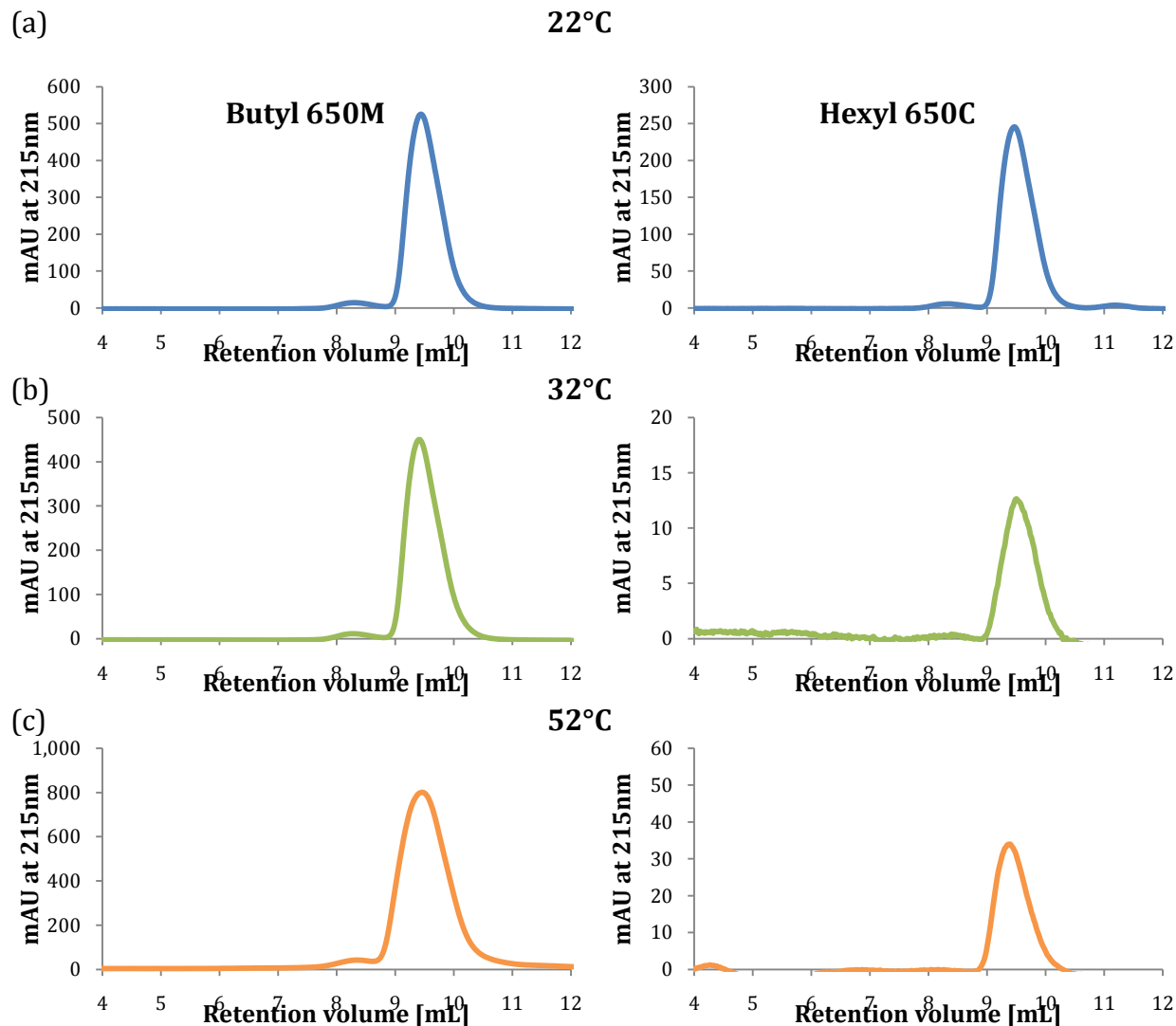
the greatest decrease in solvent exposure on the surface. In addition to Residues 180-188, an increase in solvent protection on all three surfaces is also observed for Residues 202-211. The location of these residues is shown in blue on the native structure of transferrin in Figure 2.22 on the right.



**Figure 2.22.** Effect of temperature on different reporter peptides of transferrin in solution (blue, diamonds), on Phenyl 650M (red, squares), on Butyl 650M (purple, Xs), and on Hexyl 650C (green, triangles) in 25mM PO<sub>4</sub>, 1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0 buffer at 22 to 52°C. Samples were labeled with deuterated buffer for 10 minutes. The *D/N* for Residues 251-259 is shown in panel (a). The *D/N* for Residues 180-188 is shown in panel (b). Right: Location of Residues 251-259 and the other residues with increased solvent protection on Phenyl 650M are shown in red on the native structure (PDB ID 2HAV). Residues 180-188 and the other residues with increased solvent protection on all three surfaces are shown in blue. Error bars represent sample 95% confidence intervals of triplicate data points collected on Phenyl 650M at 22°C.

SEC was also performed for transferrin samples to evaluate if irreversible aggregation on the surface might be leading to these decreases in solvent exposure. Figure 2.23 shows SEC chromatograms of transferrin adsorbed and incubated at temperatures of 22, 32, and 52°C on Butyl 650M (left) and on Hexyl 650C (right). At all temperatures examined, only 1 peak was observed. The retention time of this peak is consistent with that of SEC chromatograms with transferrin without adsorption to the HIC media (data not shown). The existence of only 1 peak demonstrates that no detectable irreversible or slow reversible association or aggregation is occurring on the surfaces. This lack of irreversible aggregation seems consistent with the small decreases in labeling observed for this protein. However, the smaller peak areas (lower recovery) of transferrin from Hexyl 650C suggest some irreversible adsorption is occurring. The SEC chromatograms are still representative of what is analyzed by the MS as both techniques use the same solution conditions to elute protein for analysis.

Although there is no evidence of irreversible aggregation, it should be noted though that weaker, reversible self-association may exist on the surface that increases solvent protection. These interactions may disappear upon elution, preventing detection by SEC. All experiments, however, were performed under low loading conditions at 5 mg/mL where protein-protein interactions are not expected to be significant. However under industrial-relevant operating conditions where loadings can exceed 20 mg/mL, these protein-protein interactions may be more significant. Also, considering aggregates that were observed for BSA on Butyl 650M and Hexyl 650C at loading of 5 mg/mL, aggregation may become a more important issue under the high loading conditions commonly used in industrial applications.



**Figure 2.23.** SEC chromatograms of transferrin eluted from Butyl 650M (left) and Hexyl 650C (right) at temperatures of 22°C (a), 32°C (b), and 52°C (c). Sample were incubated for 1 hour, eluted with 200  $\mu$ L of desorption buffer, diluted with 500  $\mu$ L of working buffer and loaded on a TSK gel G3000SWXL column for analysis.

Several groups of residues (24-32, 86-92, and 199-207) of antitrypsin also show an increase in protection when adsorbed on the HIC surfaces, though only at high temperatures (data not shown). SEC was performed for antitrypsin to determine if irreversible aggregation might occur on the surfaces. Similar to the transferrin SEC

chromatograms, only 1 peak is observed (data not shown) providing no direct evidence of aggregation.

It was shown here that higher-order-molecular weight species can form during HIC processes at conditions in which the protein is native in solution. In addition, the significantly reduced hydrogen-deuterium exchange for BSA on the surface suggests these species are aggregates that are not readily reversible in nature. The residue groups with the most solvent protection were predicted to be aggregation prone by AGGRESCAN and PASTA. These results have important implications in the design of HIC processes for proteins that are relatively unstable or aggregation prone in solution. Proteins predicted to be aggregation prone by these calculators could be flagged early in development. Proteins of higher stability, however, such as transferrin and antitrypsin (based on the results of these studies) are more amenable to higher denaturing process conditions. Although unfolding is observed for both proteins, no detectable aggregates are observed. In these cases, as long as unfolding is reversible more potentially denaturing process conditions could be considered for processing.

#### **2.4.6. Superposition of changes in solvent protection on the surface**

Temperature and type of HIC surface both have significant effects on the solvent protection on the proteins studied. As seen in the  $D/N$  values of Figures 2.5 and 2.8, the thermal unfolding of transferrin and antitrypsin are affected differently on Phenyl 650M, Butyl 650M, and Hexyl 650C. In addition, the effects are not uniform across all reporter peptides as discussed in Sections 2.4.1 and 2.4.2. Further, the effects of temperature and

surface are different for BSA. As seen in the labeling patterns of Figure 2.18 and SEC chromatograms of Figures 2.19 and 2.20, BSA unfolds at low temperatures on the surface and then aggregates at higher temperatures. One of the hypotheses of this work is that proteins unfold in similar ways under different denaturing environments. A relation between the unfolded state of proteins on the surface and their solution intermediates was already discussed in Section 2.4.3. Here we investigate whether the effects of temperature can produce the same labeling patterns as different chromatographic surfaces.

Solution-surface comparisons for the reporter peptides of transferrin show that the differences in labeling and variability are lower at temperature offsets,  $\Delta T$ s, higher than  $0^{\circ}\text{C}$ . Table 2.8 shows the average  $\Delta D/N$  magnitudes and the standard deviation,  $\sigma$ , in  $D/N$  over 32 reporter peptides for solution-surface comparisons at different  $\Delta T$ . Larger values of  $\Delta D/N$  represent poorer comparisons between labeling patterns, while larger values of  $\sigma$  reflect more variability amongst  $\Delta D/N$  for different reporter peptide comparisons. For solution-phenyl comparisons no specific  $\Delta T$  gives the best  $\Delta D/N$ , although  $\Delta T$  of 10 and  $20^{\circ}\text{C}$  work better than  $\Delta T$  of  $30^{\circ}\text{C}$ . Further,  $\sigma$  is lowest at  $\Delta T$  of  $20^{\circ}\text{C}$  indicating less variability for  $\Delta D/N$  across the reporter peptides at this temperature offset. A more distinct trend is observed for solution-butyl and solution-hexyl comparisons as both average  $\Delta D/N$  and  $\sigma$  decrease with an increase in temperature offset. Both have the most favorable comparisons at  $\Delta T$  of  $30^{\circ}\text{C}$  where  $\Delta D/N$  and  $\sigma$  are at a minimum. Other representations of  $\Delta D/N$  that account for superficial  $\Delta D/N$  values of  $\sim 0$  from fully labeled reporter peptides do not alter the outcome that surface-solution comparisons work better at  $\Delta T$ s  $> 0$ .

**Table 2.8.  $D/N$  differences for different comparisons to solution labeling at various temperature offsets.**

<b>Comparison</b>	<b><math>\Delta T</math> [°C]</b>	<b><math>\Delta D/N</math></b>	<b><math>\sigma</math></b>
<b>Phenyl</b>	0	0.10	0.084
<b>Phenyl</b>	10	0.10	0.084
<b>Phenyl</b>	20	0.10	0.070
<b>Phenyl</b>	30	0.15	0.085
<b>Butyl</b>	0	0.17	0.129
<b>Butyl</b>	10	0.15	0.115
<b>Butyl</b>	20	0.13	0.099
<b>Butyl</b>	30	0.10	0.082
<b>Hexyl</b>	0	0.14	0.121
<b>Hexyl</b>	10	0.13	0.110
<b>Hexyl</b>	20	0.10	0.082
<b>Hexyl</b>	30	0.09	0.078

The differences in labeling and variability for surface-surface comparisons are also smaller at  $\Delta T$ s greater than 0°C. Table 2.9 shows the average  $\Delta D/N$  magnitudes and  $\sigma$  in  $D/N$  for surface-surface comparisons at different  $\Delta T$ . Phenyl-butyl comparisons are best at a 20°C offset with an average  $\Delta D/N$  of 0.07 and  $\sigma$  of 0.065. Phenyl-hexyl comparisons do not work as well with an best  $\Delta D/N$  of 0.09 and  $\sigma$  of 0.075 at a 10°C offset. Of all comparisons, hexyl-butyl work best with the comparison at a 20°C offset having the lowest  $\Delta D/N$  and  $\sigma$  of any comparison: 0.03 and 0.033, respectively. A more detailed analysis for individual peptides for select solution-surface and surface-surface comparisons is presented below.

**Table 2.9.  $D/N$  differences for different surface-surface labeling comparisons at various temperature offsets.**

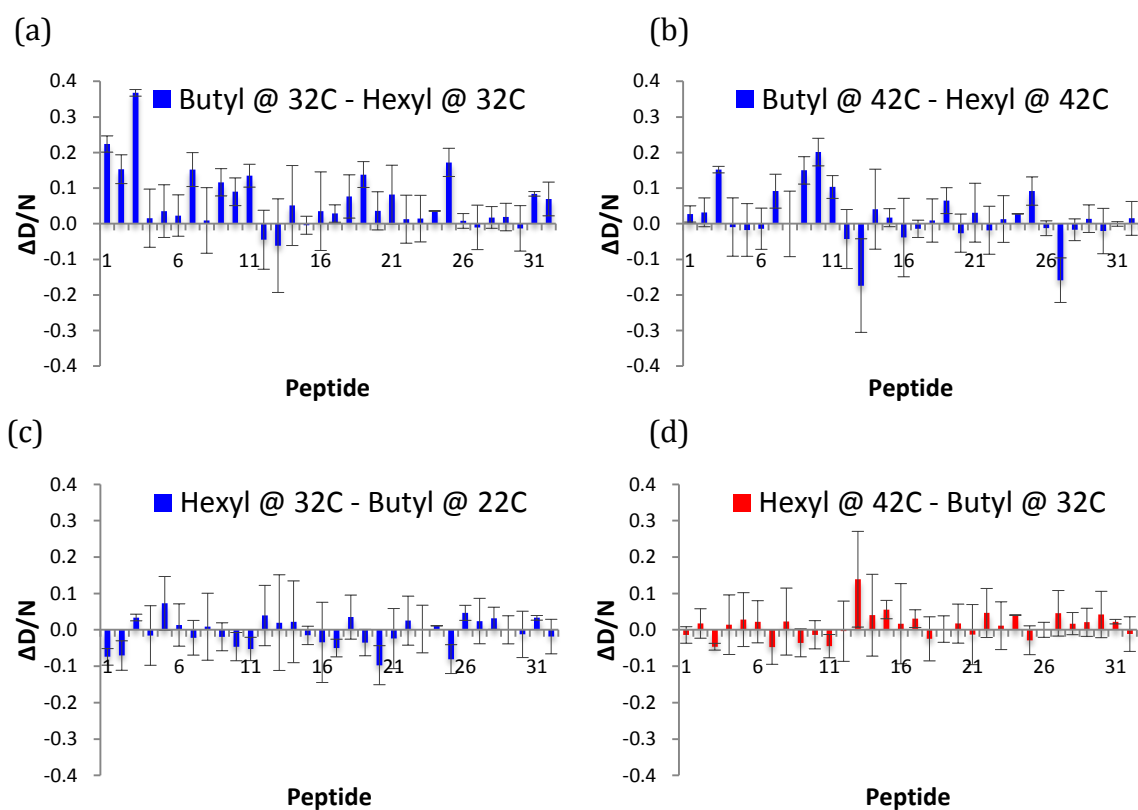
<b>Comparison</b>	<b><math>\Delta T</math> [°C]</b>	<b><math>\Delta D/N</math></b>	<b><math>\sigma</math></b>
<b>Phenyl-Butyl</b>	0	0.14	0.098
<b>Phenyl-Butyl</b>	10	0.11	0.079
<b>Phenyl-Butyl</b>	20	0.07	0.065
<b>Phenyl-Butyl</b>	30	0.10	0.117
<b>Phenyl-Hexyl</b>	0	0.12	0.079
<b>Phenyl-Hexyl</b>	10	0.09	0.075
<b>Phenyl-Hexyl</b>	20	0.10	0.097
<b>Phenyl-Hexyl</b>	30	0.14	0.123
<b>Hexyl-Butyl</b>	0	0.06	0.061
<b>Hexyl-Butyl</b>	10	0.03	0.033
<b>Hexyl-Butyl</b>	20	0.07	0.058
<b>Hexyl-Butyl</b>	30	0.12	0.118

Figure 2.24 shows how the differences in  $D/N$  values,  $\Delta D/N$ , for most transferrin reporter peptides absorbed on Butyl 650M and Hexyl 650C go to zero when compared at a 10°C offset. Error bars represent propagated uncertainties in  $\Delta D/N$  for individual reporter peptides. First, when compared at the same temperature most regions have positive  $\Delta D/N$  values corresponding to more unfolding on Butyl 650M. At 32°C (Figure 2.24a), 15 of the 32 reporter peptides have statistically significant  $\Delta D/N$  values greater than 0, with 8 values greater than 0.1.

However, when Hexyl 650C is compared to Butyl 650M at a temperature 10°C higher, as shown on the bottom of Figure 2.24, the differences between deuterium labeling patterns for most regions are zero within uncertainty. For the comparison of Hexyl 650C at 32°C to Butyl 650M at 22°C (Figure 2.24c), 22 of the 32 reporter peptides have differences of zero. Of the 10 that do not, only 1 peptide (reporter Peptide 20) has a  $\Delta D/N$  magnitude equal or greater than 0.1. For the comparison of Hexyl 650C at 42°C to Butyl 650M at 32°C (Figure 2.24d), 25 of the 32 reporter peptides have differences of zero. As with previous



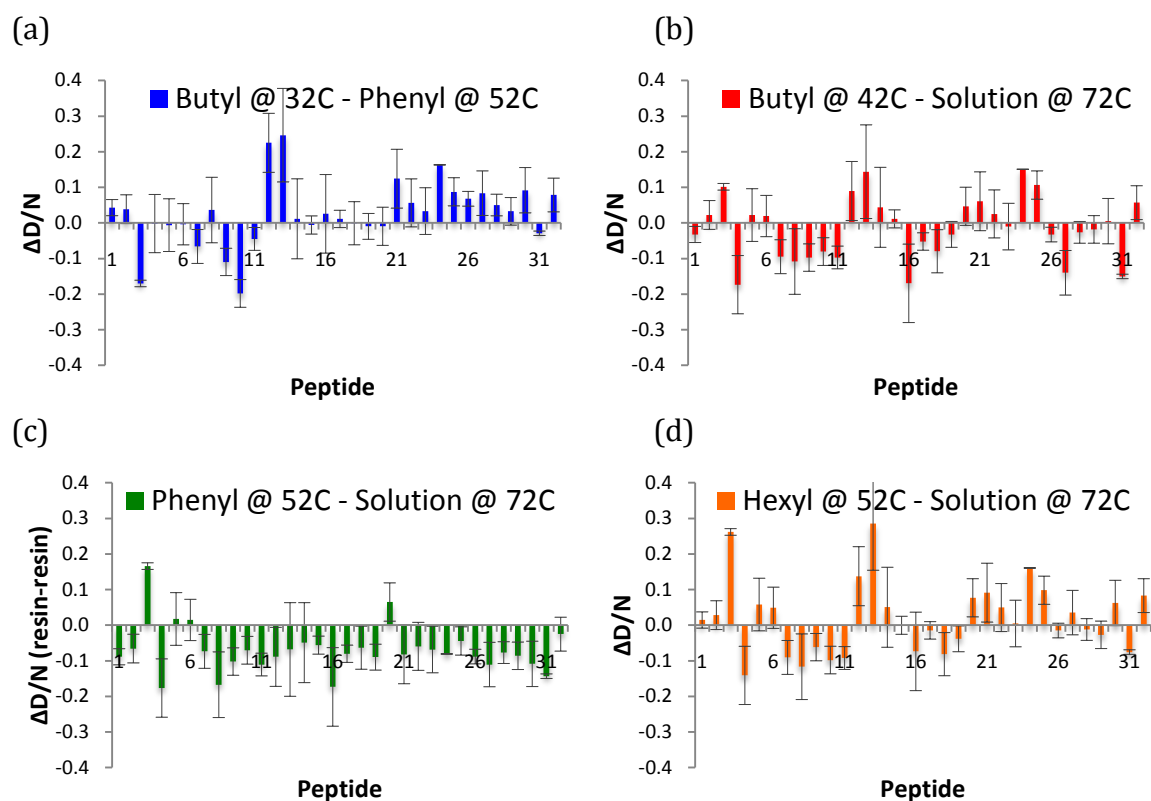
comparisons, only 1 peptide (reporter Peptide 13) has a  $\Delta D/N$  magnitude equal or greater than 0.1. This suggests that effect of changing from one alkyl surface to another on protein stability can be mimicked by a single shift in temperature. Thus, transferrin would seem to be denaturing on the alkyl surfaces in the same general way that it does with increasing temperature.



**Figure 2.24.** Difference in  $D/N$ ,  $\Delta D/N$ , between transferrin adsorbed on Butyl 650M at 32°C and Hexyl 650C at 32°C (a), adsorbed on Butyl 650M at 42°C and Hexyl 650C at 42°C (b), adsorbed on Hexyl 650C at 32°C and on Butyl 650M at 22°C (c), and adsorbed on Hexyl 650C at 42°C and on Butyl 650M at 32°C (d).

However, other comparisons are not as successful in superpositioning. For example, the labeling of transferrin on Butyl 650M at 32°C and on Phenyl 650M at 52°C is compared in Figure 2.25a. For this comparison only 15 of the 32 reporter peptides have differences

of zero. Of the 17 that do not, 7 have magnitudes greater than 0.1. Superposition patterns were also made for differences in surface and solution labeling at different temperatures, as shown in Figure 2.25b-d. Only 13, 10, and 16 peptides have differences of zero when comparing Butyl 650M, Phenyl 650M, and Hexyl 650C, respectively, to solution. Eleven, 9, and 9 peptides in the respective comparisons have  $\Delta D/N$  values with magnitudes greater than 0.1.



**Figure 2.25. Difference in  $D/N$ ,  $\Delta D/N$ , between human serum transferrin adsorbed on Butyl 650M at 32°C and Phenyl 650M at 52°C (a), adsorbed on Butyl 650M at 42°C and in solution at 72°C (b), adsorbed on Phenyl 650M at 52°C and in solution at 72°C (c), and adsorbed on Hexyl 650C at 52°C and in solution at 72°C (d).**

The unfolding patterns for antitrypsin on the HIC surfaces are also similar when compared different temperatures. Again, comparisons between surface-surface are better than solution-surface as more reporter peptides are observed with  $\Delta D/N$ s of zero.

Solution-surface comparisons for the reporter peptides of antitrypsin show that the differences in labeling and variability are lower at  $\Delta T$ s higher than 0°C. Table 2.10 shows the average  $\Delta D/N$  magnitudes and standard deviation,  $\sigma$ , in  $D/N$  over 31 reporter peptides for solution-surface comparisons at different  $\Delta T$ . Larger values of  $\Delta D/N$  represent poorer comparisons while larger values of  $\sigma$  reflect more variability amongst the reporter peptides for a given comparison. The best comparisons of solution to phenyl and hexyl occur at  $\Delta T$  of 20 and 40°C, respectively. These offsets represent minima for both  $\Delta D/N$  and  $\sigma$ . A specific minima is not observed for solution-butyl but instead  $\Delta T$ s of 40, 50, and 60°C provide the best comparison.

**Table 2.10.  $D/N$  differences for different comparisons to solution labeling at various temperature offsets.**

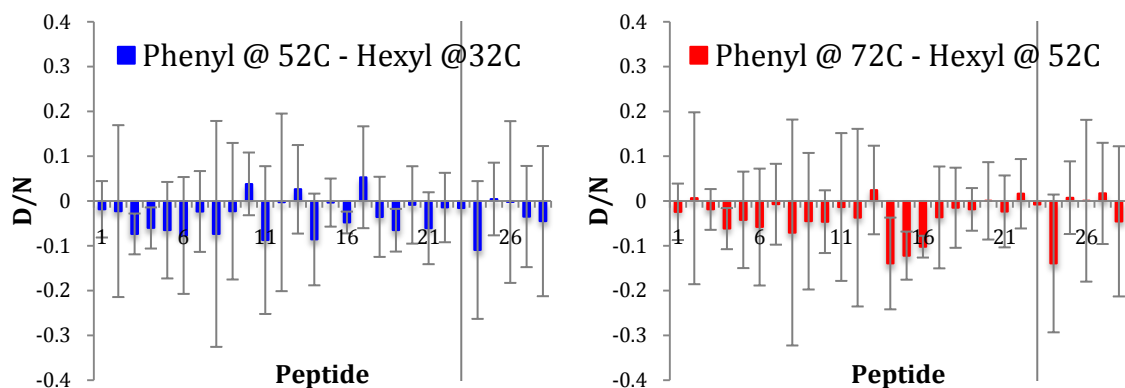
<b>Comparison</b>	<b><math>\Delta T</math> [°C]</b>	<b><math>\Delta D/N</math></b>	<b><math>\sigma</math></b>
<b>Phenyl</b>	0	0.15	0.13
<b>Phenyl</b>	10	0.13	0.11
<b>Phenyl</b>	20	0.12	0.11
<b>Phenyl</b>	30	0.14	0.12
<b>Phenyl</b>	40	0.16	0.14
<b>Phenyl</b>	50	0.18	0.14
<b>Phenyl</b>	60	0.18	0.15
<b>Butyl</b>	0	0.19	0.15
<b>Butyl</b>	10	0.16	0.14
<b>Butyl</b>	20	0.15	0.13
<b>Butyl</b>	30	0.14	0.13
<b>Butyl</b>	40	0.12	0.10
<b>Butyl</b>	50	0.12	0.10
<b>Butyl</b>	60	0.12	0.10
<b>Hexyl</b>	0	0.23	0.16
<b>Hexyl</b>	10	0.20	0.15
<b>Hexyl</b>	20	0.17	0.15
<b>Hexyl</b>	30	0.15	0.14
<b>Hexyl</b>	40	0.14	0.14
<b>Hexyl</b>	50	0.15	0.14
<b>Hexyl</b>	60	0.17	0.16

The differences in labeling and variability for surface-surface comparisons are also smaller at  $\Delta T$ s greater than 0°C. Table 2.11 shows the average  $\Delta D/N$  and  $\sigma$  in  $D/N$  for surface-surface comparisons at different  $\Delta T$ . Phenyl-butyl comparisons are best at a 20°C offset with an average  $\Delta D/N$  of 0.077 and  $\sigma$  of 0.083. Butyl-hexyl comparisons in terms of  $\Delta D/N$  are best at a 20°C offset but the lowest variability occurs at a 40°C offset. Phenyl-hexyl comparisons work best with an optimal  $\Delta D/N$  of 0.066 and  $\sigma$  of 0.059 also at a 20°C offset. A more detailed analysis for individual peptides for select solution-surface and surface-surface comparisons is presented below.

**Table 2.11.  $D/N$  differences for different surface-surface labeling comparisons at various temperature offsets.**

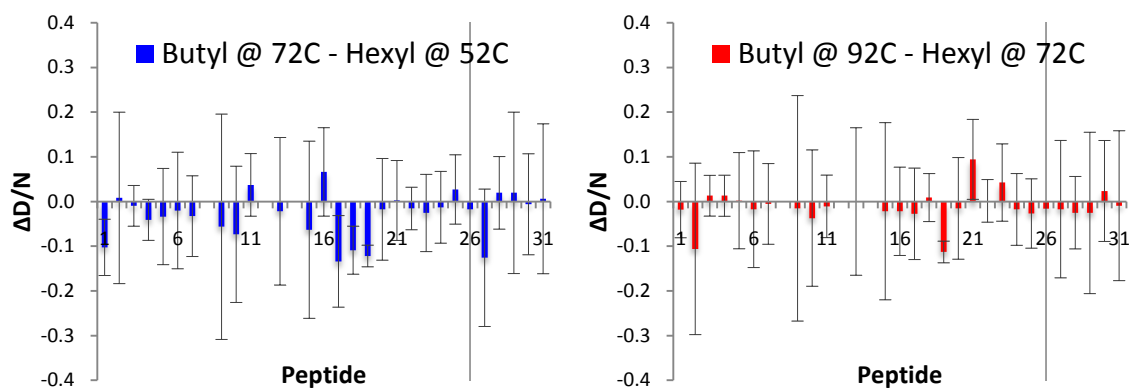
<b>Comparison</b>	<b><math>\Delta T</math> [°C]</b>	<b><math>\Delta D/N</math></b>	<b><math>\sigma</math></b>
<b>Phenyl-Butyl</b>	0	0.10	0.11
<b>Phenyl-Butyl</b>	10	0.10	0.087
<b>Phenyl-Butyl</b>	20	0.077	0.083
<b>Phenyl-Butyl</b>	30	0.077	0.089
<b>Phenyl-Butyl</b>	40	0.076	0.094
<b>Phenyl-Butyl</b>	50	0.085	0.10
<b>Phenyl-Butyl</b>	60	0.10	0.10
<b>Phenyl-Hexyl</b>	0	0.10	0.090
<b>Phenyl-Hexyl</b>	10	0.086	0.079
<b>Phenyl-Hexyl</b>	20	0.066	0.059
<b>Phenyl-Hexyl</b>	30	0.087	0.060
<b>Phenyl-Hexyl</b>	40	0.084	0.069
<b>Phenyl-Hexyl</b>	50	0.095	0.077
<b>Phenyl-Hexyl</b>	60	0.10	0.088
<b>Butyl-Hexyl</b>	0	0.083	0.093
<b>Butyl-Hexyl</b>	10	0.088	0.089
<b>Butyl-Hexyl</b>	20	0.077	0.086
<b>Butyl-Hexyl</b>	30	0.091	0.084
<b>Butyl-Hexyl</b>	40	0.10	0.079
<b>Butyl-Hexyl</b>	50	0.11	0.084
<b>Butyl-Hexyl</b>	60	0.11	0.089

Figure 2.26 shows the difference between deuterium labeling patterns for antitrypsin adsorbed on Phenyl 650M and Hexyl 650C at a 20°C difference. At this temperature, differences between deuterium labeling patterns for most regions are zero within uncertainty. For the comparison of Phenyl 650M at 52°C to Hexyl 650C at 32°C (Figure 2.26a), 27 of the 31 reporter peptides have differences of zero. Further, of the 4 that do not have  $\Delta D/N$  less than 0.1, the differences are small. For the comparison of Phenyl 650M at 72°C to Hexyl 650C at 52°C (Figure 2.26b), 27 of the 31 reporter peptides have differences of zero. Of the 4 that do not, 3 have  $\Delta D/N$  magnitudes greater than 0.1, indicating more differences for this temperature range comparison.



**Figure 2.26. Difference in  $D/N$ ,  $\Delta D/N$ , between antitrypsin adsorbed on Phenyl 650M at 52°C and on Hexyl 650C at 32°C (left), and adsorbed on Phenyl 650M at 72°C and on Hexyl 650C at 52°C (right).**

Comparisons of labeling between Butyl 650M and Hexyl 650C at a 20°C difference also show similar  $D/N$  values. Figure 2.27 shows the difference between deuterium labeling patterns for antitrypsin adsorbed on Butyl 650M and Hexyl 650C at a 20°C difference. At this  $\Delta T$ ,  $\Delta D/N$  for most regions is zero within margin of uncertainty. For the comparison of Butyl 650M at 72°C to Hexyl 650C at 52°C (Figure 2.27a), 27 of the 31 reporter peptides have zero differences. However of the 4 that do not, all have a  $\Delta D/N$  magnitude equal or greater than 0.1. For the comparison of Butyl 650M at 92°C to Hexyl 650C at 72°C (Figure 2.27b), 29 of the 31 reporter peptides have differences of zero. Of the 2 that do not, only 1 peptide (reporter Peptide 19) has a  $\Delta D/N$  magnitude equal or greater than 0.1.



**Figure 2.27.** Difference in  $D/N$ ,  $\Delta D/N$ , between antitrypsin adsorbed on Butyl 650M at 72°C and on Hexyl 650C at 52°C (left), and adsorbed on Butyl 650M at 92°C and on Hexyl 650C at 72°C (right).

As with transferrin, this superposition of labeling at different temperatures works best for surface-surface comparisons. The superposition works less well for comparisons of Butyl 650M to Phenyl 650M and comparisons of the three different surfaces to the solution (data not shown, but similar to Figure 2.25).

Superposition also works for the case of BSA aggregation on the HIC surfaces. Table 2.12 shows the average  $\Delta D/N$  and  $\sigma$  in  $D/N$  for solution-surface and surface-surface comparisons at different  $\Delta T$ . For solution-butyl and solution-hexyl comparisons the lowest average  $\Delta D/N$  and  $\sigma$  in  $D/N$  are observed at  $\Delta T$ s of 50°C. For the hexyl-butyl comparisons a minimum in average  $\Delta D/N$  and  $\sigma$  in  $D/N$  is observed at  $\Delta T$  of 20°C. As with transferrin and antitrypsin, other representations of  $\Delta D/N$  that account for superficial  $\Delta D/N$  values of  $\sim 0$  from fully labeled reporter peptides do not alter the outcome that surface-solution comparisons work better at  $\Delta T$ s  $> 0$ . A more detailed analysis for individual peptides for select solution-surface and surface-surface comparisons is presented below.

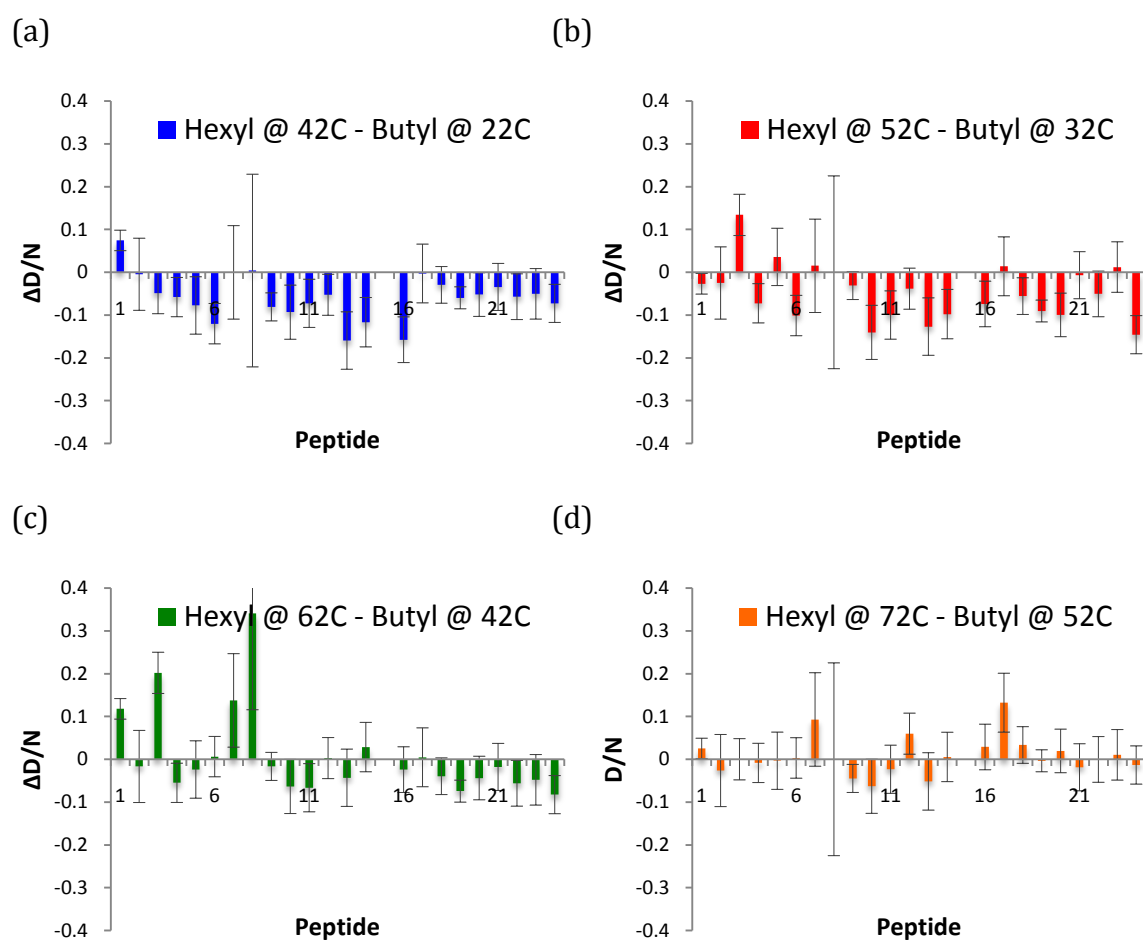
**Table 2.12.  $D/N$  differences for different solution-surface and surface-surface labeling comparisons at various temperature offsets.**

<b>Comparison</b>	<b><math>\Delta T</math> [°C]</b>	<b><math>\Delta D/N</math></b>	<b><math>\sigma</math></b>
<b>Solution-Butyl</b>	0	0.19	0.15
<b>Solution-Butyl</b>	10	0.17	0.13
<b>Solution-Butyl</b>	20	0.15	0.11
<b>Solution-Butyl</b>	30	0.12	0.098
<b>Solution-Butyl</b>	40	0.066	0.047
<b>Solution-Butyl</b>	50	0.047	0.025
<b>Solution-Hexyl</b>	0	0.15	0.12
<b>Solution-Hexyl</b>	10	0.13	0.094
<b>Solution-Hexyl</b>	20	0.10	0.090
<b>Solution-Hexyl</b>	30	0.11	0.076
<b>Solution-Hexyl</b>	40	0.12	0.072
<b>Solution-Hexyl</b>	50	0.095	0.062
<b>Hexyl-Butyl</b>	0	0.074	0.079
<b>Hexyl-Butyl</b>	10	0.073	0.074
<b>Hexyl-Butyl</b>	20	0.058	0.054
<b>Hexyl-Butyl</b>	30	0.076	0.057
<b>Hexyl-Butyl</b>	40	0.12	0.061
<b>Hexyl-Butyl</b>	50	0.13	0.064

Figure 2.28 shows the difference between deuterium labeling patterns for BSA adsorbed on Hexyl 650C and Butyl 650M at 20°C differences. Many regions in the 4 comparisons have differences of zero within uncertainty. As the absolute temperature increases, more reporter peptides have differences of zero. For example, 13, 15, 18, and 24



reporter peptides have  $\Delta D/N$  of zero as shown in panels (a), (b), (c), and (d), respectively. An explanation for why the comparisons become better as the absolute temperature increases, may be from aggregation not being initiated on the Hexyl 650C surface until a higher temperature than on Butyl 650M.



**Figure 2.28.** Difference in  $D/N$ ,  $\Delta D/N$ , between BSA adsorbed on Hexyl 650C at 42°C and on Butyl 650M at 22°C (a), adsorbed on Hexyl 650C at 52°C and Butyl 650M at 32°C (b), adsorbed on Hexyl 650C at 62°C and on Butyl 650M at 42°C (c), and adsorbed on Butyl 650M at 72°C and on Hexyl 650C at 52°C (d).

The superposition patterns presented here demonstrate that the regions that unfold or become more solvent-protected from surface interactions or from aggregation are generally retained on all the different surfaces. Superposition works best for comparisons

of the two alkyl ligands, suggesting that unfolding data on one surface type could be used to anticipate unfolding on a different others. However, the differences between alkyl and phenyl indicate that ligand chemistry is important to consider when anticipating the regions of proteins that may unfold when adsorbed. Additional thermal denaturation data collected on Ethyl 650M would provide more information on how far superposition can be extended to other surfaces.

An explanation for the lack of superposition between the alkyl ligands and the phenyl ligand may be from different binding orientations on the surface. Even if the same regions are unfolded, differences among those interacting with the surface would lead to different labeling patterns. HIC surfaces with alkyl ligands primarily interact with proteins through just hydrophobic interactions. However, HIC surfaces with aromatic ligands also show pi-pi interactions in addition to hydrophobic interactions. This could explain why superposition works well for Butyl 650M to Hexyl 650C, as binding orientation may be very similar on the alkyl ligands.

## ***2.5. Conclusions***

In this study, we showed that the unfolding of an adsorbed protein depends on its stability and on the chemistry of the HIC surface. We demonstrated that the regions of antitrypsin and transferrin that unfold in their solution intermediate are more prone to unfold when adsorbed on the HIC surfaces. Further, Butyl 650M and Hexyl 650C both affect similar regions of the proteins but at different magnitudes. When the unfolding patterns of the two are superimposed at different temperatures, these magnitudes become similar. Phenyl 650M unfolds the proteins to a lesser extent and affects different regions, which may be attributed to its different ligand chemistry (aromatic) and pi-pi interactions compared to Butyl 650M and Hexyl 650C (unbranched carbon chain).

We also demonstrated that the thermodynamic and kinetic effects of HIC surfaces on protein unfolding can be determined. Together with solution stability data, this information would aid in HIC media selection for the purification of therapeutic proteins. Media of higher hydrophobicity could be chosen up to the point that they would not cause unfolding during purification as indicated by surface  $\Delta\Delta G_u < \text{protein } \Delta G_u$ .

Finally, we showed that Butyl 650M and Hexyl 650C both unfold BSA at lower temperatures while facilitating its aggregation at higher temperatures. As with antitrypsin and transferrin, the two surfaces both unfold BSA, but to different extents. In this work, we presented the hypothesis that hydrophobic surfaces can also act as catalysts for aggregation for proteins that are aggregation-prone based on the formation of BSA aggregates on the surface at temperatures where none are observed in solution.

### **3. Protein Self-Association: Identification of Association Sites and Related Conformational Changes by Hydrogen Exchange**

#### ***3.1. Background and Introduction***

Protein self-association is an important process both physiologically and in the pharmaceutical industry. Physiologically, many proteins only function normally as homo-oligomers (Chan et al. 2010; Qi et al. 2010; Doyle 1998). In the pharmaceutical industry, self-association can pose several challenges in the purification development of therapeutic proteins. Furthermore, there is a growing interest in assessing how different steps in development can affect conformation including through self-association. There is thus a need to assess conformation and higher-order structure early in the product development cycle.

The focus of this work was the self-association of a therapeutic protein candidate, hereafter referred to as Protein X. Protein X consists of a modified region with therapeutic activity and an unmodified region. Purified Protein X exhibited multiple peaks on ion-exchange chromatograms, prompting an investigation to determine their compositions. Analysis by SEC revealed they contained dimers and oligomers, depending on the solution conditions. Additional studies via ultracentrifugation and SLS further confirmed the existence of self-associated species. Variations of protein concentration, ionic strength, and pH were all found to influence the extent of self-association (Dumetz et al., 2010; Zadnik et al., 2009).

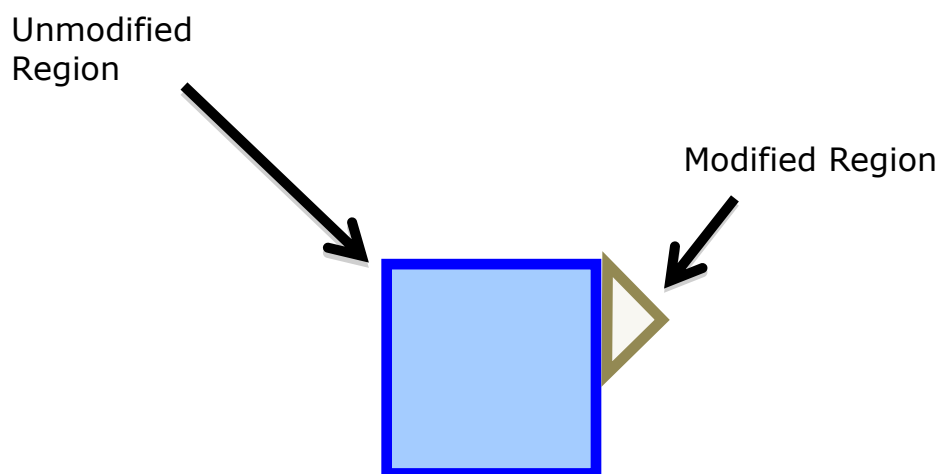
Previously, studies with hydrogen exchange mass spectrometry (HX-MS) have identified protein self-association and aggregation sites (Man et al., 2010; Qi et al., 2008; Tokihiro et al., 1996). By examining changes in hydrogen-deuterium exchange, changes in the protein's conformation and association sites can be determined. In particular, if an area of a protein unfolds, previously buried residues in that region will be exposed to the surrounding deuterated solvent and the hydrogen-deuterium exchange will increase, while if molecules associate, there can be reduced hydrogen-deuterium exchange, due to burial of residues within the interaction site and loss of solvent access.

The present study uses HX-MS to investigate the region(s) involved in Protein X association, including any local or global conformational changes that might occur. Hydrogen-deuterium exchange rates are compared for Protein X monomers, dimers, and oligomers. In addition, isothermal titration calorimetry (ITC) and size-exclusion chromatography (SEC) are used to reveal the thermodynamics of dimeric and oligomeric association from enthalpies of dissociation and dissociation equilibrium constants at different pH values.

## 3.2. *Materials and Methods*

### 3.2.1 Materials

Protein X with and without the modified region were generously provided by GlaxoSmithKline. Figure 3.1 shows a representation of Protein X and its two regions. Sodium succinate, potassium phosphate, guanidine hydrochloride (GndHCl), and ethylenediaminetetraacetic acid (EDTA) were purchased from Fisher Scientific (Houston, TX, USA) and were of HPLC-grade quality or better. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Thermo Scientific (Rockford, IL, USA).



**Figure 3.1.** Protein X representation. The unmodified region is represented by a blue square and the modified region is represented by a brown triangle.

### 3.2.2. Hydrogen Exchange Mass Spectrometry

Samples of Protein X were prepared in solution conditions associated with dimer and oligomer formation. Control samples of Protein X with and without the modified region were also prepared under conditions where only monomers were expected. All Protein X samples were prepared in deionized, distilled water (ddH<sub>2</sub>O) with 30 mM sodium succinate and adjusted to a pH of 6.4 or 5.2. Table 3.1 shows the different solutions used and the dominant species at each condition. Samples were allowed to equilibrate overnight before analysis. Deuterated buffers without protein were created at solution conditions identical to the ones listed in Table 3.1.

**Table 3.1. Sample Conditions for HX-MS Studies**

<b>Modified</b>	<b>Concentration</b>	<b>Sodium Succinate</b>	<b>pH</b>	<b>Dominant Species</b>
Yes	0.4 mg/mL	30 mM	6.4	Monomer
Yes	0.8 mg/mL	30 mM	6.4	Monomer
Yes	16.6 mg/mL	30 mM	6.4	Dimer
No	16.6 mg/mL	30 mM	6.4	Monomer
Yes	0.5 mg/mL	30 mM	5.2	Monomer
Yes	18.9 mg/mL	30 mM	5.2	Oligomer

Prior to hydrogen exchange experiments, digestion and sequence coverage were optimized for the molecule. It was found that, 80% of the whole molecule was covered by 35 reporter peptides, with 11 of these giving 60% coverage of the modified region.

Hydrogen exchange was initiated by mixing 5  $\mu$ L of sample with 45  $\mu$ L of deuterated buffer. Labeling times of 30 seconds, 2 minutes, 10 minutes, 30 minutes, and 120 minutes were investigated. After labeling, 4.25  $\mu$ L of pH 1.5, 150 mM  $\text{PO}_4^-$  buffer was added to lower the pH of the solution to 2.6, to bring the hydrogen exchange rate to a minimum. A volume of 180  $\mu$ L of 8 M GndHCl, 100 mM TCEP, and 27 mM EDTA was added and allowed to unfold the protein molecules for 2 minutes in preparation for proteolytic digestion. Finally, the sample was diluted with 265  $\mu$ L of 95% ddH<sub>2</sub>O, 5% acetonitrile, and 0.06% formic acid solution (solvent S) prior to injection into a 500  $\mu$ L stainless steel sample loop. In addition to the labeling times examined, non-deuterated and fully-deuterated controls were obtained.

An isocratic pump delivered solvent S and the injected sample to an immobilized pepsin column at a flow rate of 0.15 mL/min. Peptides eluting from the pepsin column were trapped and concentrated on a C<sub>8</sub>-desalting column (TR1/25109/02, MICHROM Bioresources, Inc, Auburn, CA). After 9 minutes, flow was switched from the sample pump to the Surveyor MS HPLC pump to elute the peptides off the C<sub>8</sub> column by using a 17-minute gradient of 70% solvent A (ddH<sub>2</sub>O, 0.1% formic acid, 0.01% TFA) and 5% solvent B (acetonitrile, 0.8% formic acid) to 40% solvent A. The eluted peptides from the C<sub>8</sub> column were further resolved with a C<sub>18</sub> column (18003563, Waters, Milford, MA) just before introduction to the mass spectrometer.

Eluted peptides were sent to a LTQ linear quadrupole ion trap mass spectrometer (Thermo Finnigan, San Jose, CA) where electrospray ionization (ESI) was used to spray the peptides. Data were collected in both positive ion and profile modes with an ESI voltage of 4.3 kV, a capillary temperature of 250°C, and a sheath gas flow rate of 40 units.



The fraction of a peptide labeled,  $D/N$ , was determined from

$$\frac{D}{N} = \frac{m_t - m_0}{m_{100} - m_0} \quad (3.1)$$

where  $D$  is the number of deuterated amides,  $N$  is the number of exchange-competent residues in the peptide,  $m_t$  is the mass of the peptide after a given labeling time,  $m_0$  is the non-deuterated mass of the peptide, and  $m_{100}$  is the fully-deuterated mass of the peptide.

### 3.2.3. Isothermal Titration Calorimetry

Samples of Protein X in 30 mM sodium succinate at 21.5 and 20.2 mg/mL at pH 6.4 and 5.2, respectively, were analyzed by ITC. Forty injections of 6  $\mu$ L of sample were added at a rate of 0.5  $\mu$ L/s every 300 seconds to a 1.8 mL sample cell of a VP-ITC (Microcal, North Hampton, MA) at 37°C. The sample concentration and temperature were roughly optimized to provide the best signal to noise ratio. Integrated heat pulse data were analyzed to yield dissociation constants,  $K_{diss}$ , and enthalpies of dissociation,  $\Delta H_{diss}$ , as described elsewhere (Czypionka et al. 2007). Free energies,  $\Delta G_{diss}$ , and entropies,  $\Delta S_{diss}$ , of dissociation were calculated from:

$$\Delta G_{diss} = -RT \ln(K_{diss}) \quad (3.2)$$

$$\Delta S_{diss} = (\Delta H_{diss} - \Delta G_{diss}) / T \quad (3.3)$$

where  $R$  is the universal gas constant, 8.314 J/mol-K, and  $T$  is the temperature of the experiments, 310K.

### 3.2.4. Size-Exclusion Chromatography

SEC experiments were performed on an AKTA Purifier with a G3000SW<sub>XL</sub> column (08541, Tosoh Bioscience, King of Prussia, PA) at room temperature. Injections of 0.9 mL at a flow rate of 1 mL/min were monitored with a UV detector signal at 214 and 280 nm. Injections consisted of 0.1 to 17.6 mg/mL protein in 10 mM sodium succinate at pH 5.8. Running buffer consisted of 10 mM sodium succinate at pH 5.8. Peak areas for monomers and dimers were converted to concentrations,  $C_{species}$ , with

$$C_{species} = \frac{(\% \text{ peak areas})(\text{total protein [mg/mL]})}{(MW_{species})} \quad (3.4)$$

where  $MW_{species}$  is the molecular weight of the monomer or the dimer. Equilibrium dissociation constants,  $K_{diss}$ , were calculated with

$$K_{diss} = \frac{C_{monomer}^2}{C_{dimer}} \quad (3.5)$$

and  $\Delta G_{diss}$  was calculated with Equation (3.2).

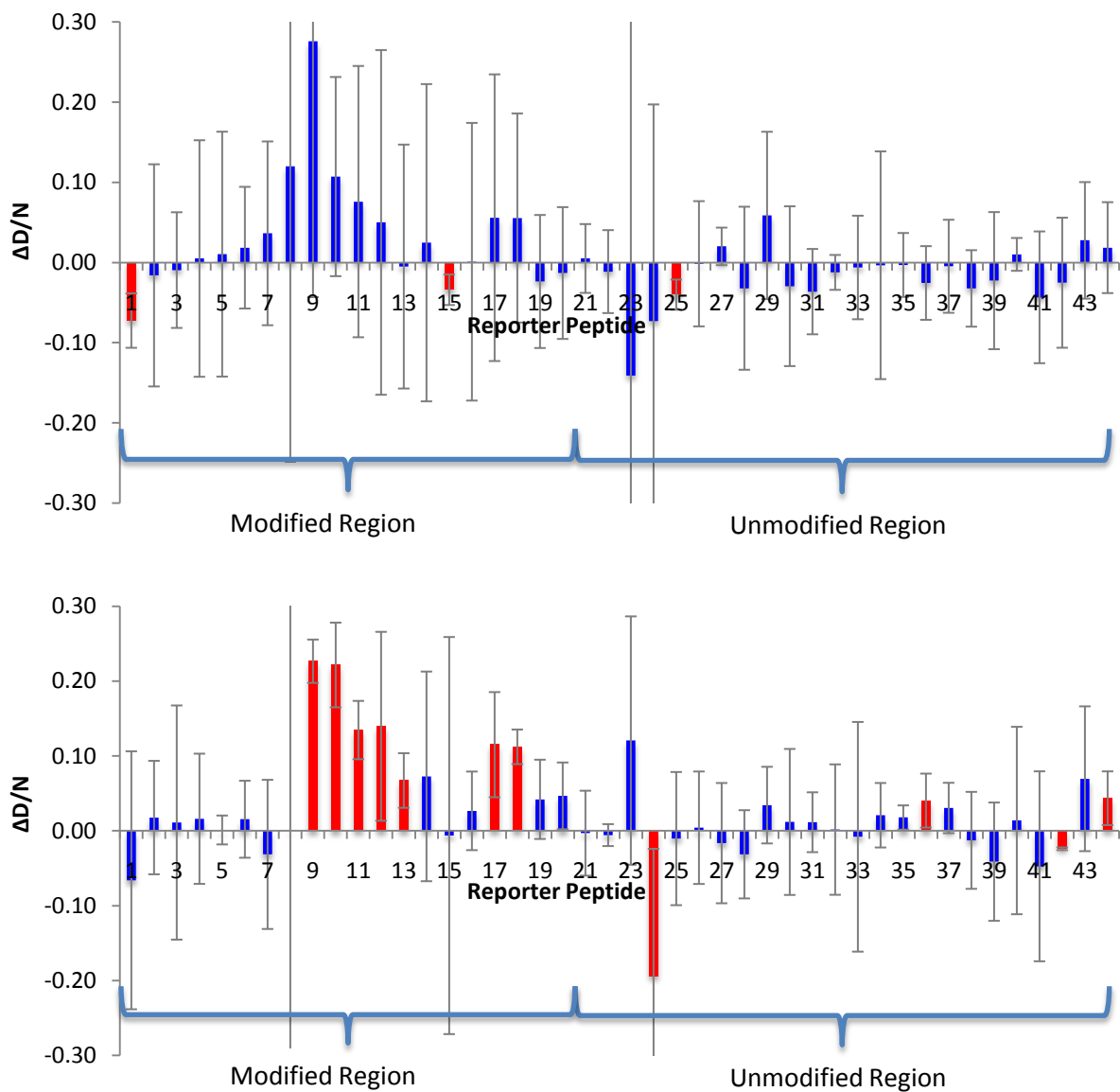
### **3.3. Results and Discussion**

#### **3.3.1 HX-MS Measurements of Conformational Changes**

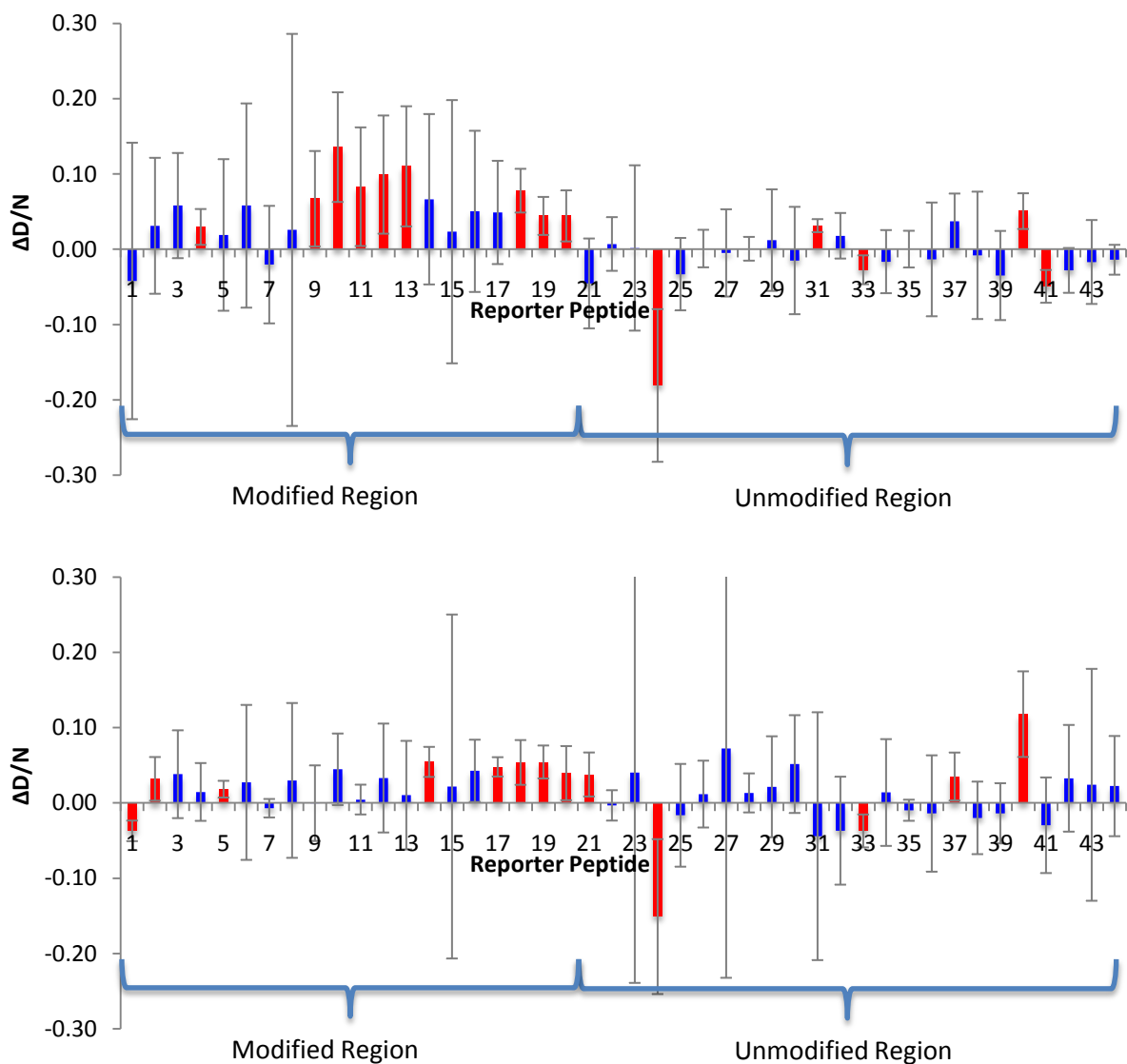
The difference in  $D/N$  values,  $\Delta D/N$ , at the different solution conditions for each of the HX-MS reporter peptides are plotted and compared. Figure 3.2 shows  $\Delta D/N$  values between concentrations of 0.8, monomers, and 16.6 mg/mL, dimers, at 2 and 10 minutes for the reporter peptides of Protein X at a pH of 6.4 in 30 mM sodium succinate. Figure 3.3 shows the  $\Delta D/N$  values at 30 and 120 minutes. Positive  $\Delta D/N$  values correspond to an increase in solvent protection at 16.6 mg/mL compared to 0.8 mg/mL. Negative  $\Delta D/N$  values correspond to a decrease in solvent protection at 16.6 mg/mL compared to 0.8 mg/mL. As expected, there is no significant difference in labeling between the two concentrations for the unmodified region (reporter peptides 21 to 44) with the exception of reporter peptide 24. This is expected, as Protein X without the modified region shows the same folding/association behavior at all concentrations.

The reporter peptides containing only the modified region (reporter peptides 1 to 20) generally show positive  $\Delta D/N$  values corresponding to more protection for this region at 16.6 mg/mL. Reductions in labeling with increased protein concentration suggest regions of the protein involved in the self-association interaction. Almost all cases of significant reduction in labeling with increased protein concentration occurred in the modified region-containing reporter peptides. At 10 and 30 minutes labeling, a statistically significant decrease in  $D/N$  at 16.6 mg/mL was observed for reporter peptides 9 to 13, 17, and 18 of the modified region. These decreases with increased protein concentration strongly suggest that the modified region is the main association site. The only other

significant decreases in labeling are in reporter peptides 19 and 20 at 30 minutes but these differences are smaller.



**Figure 3.2.**  $\Delta D/N$  values between concentrations of 0.8 and 16.6 mg/mL for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 2 minutes (top) and 10 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.



**Figure 3.3.**  $\Delta D/N$  values between concentrations of 0.8 and 16.6 mg/mL for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 30 minutes (top) and 120 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.

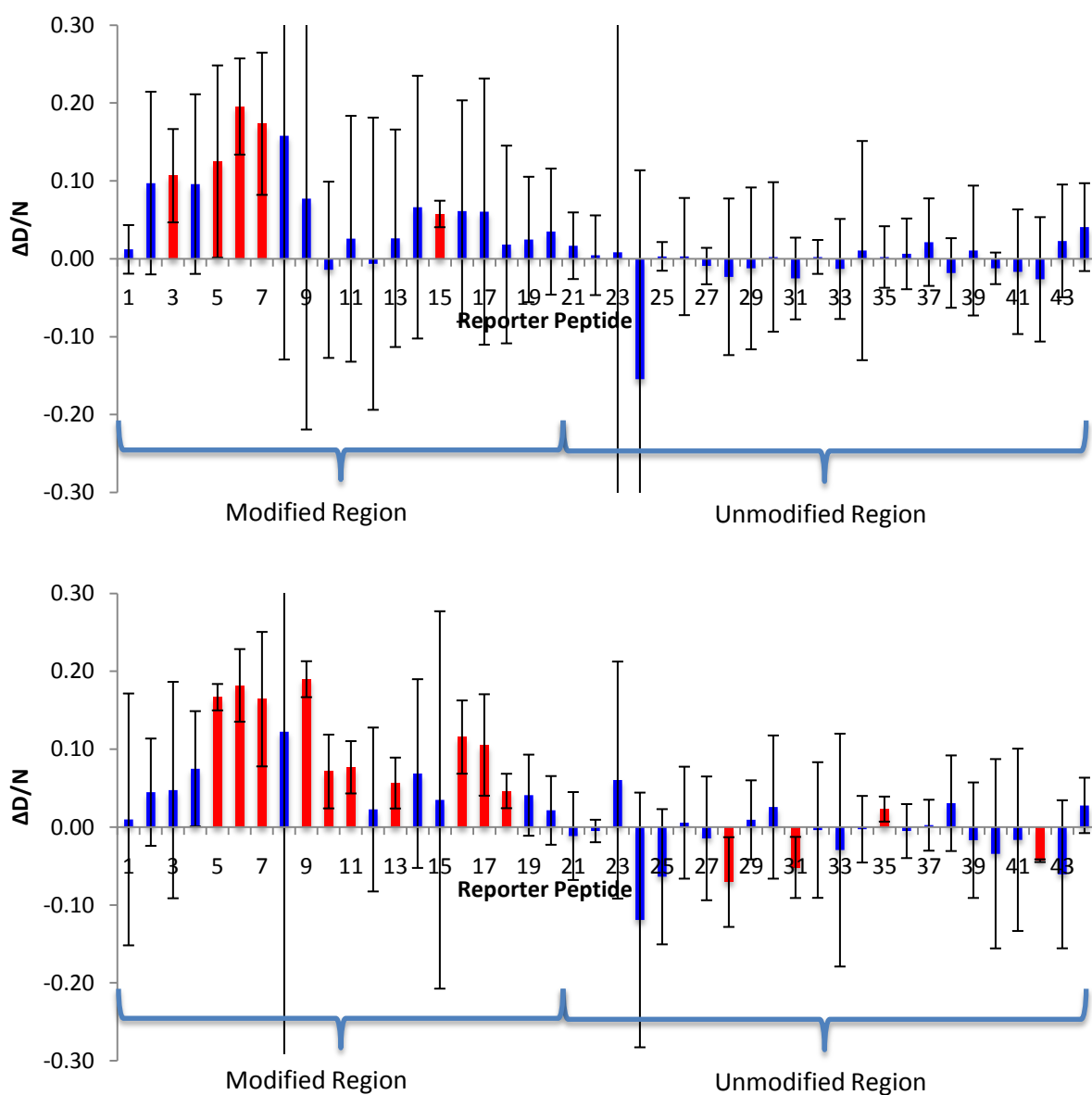
The unmodified region of Protein X shows no statistically significant differences in labeling greater than 0.1  $D/N$  between the two concentrations except for Peptide 24. In this case, there is a decrease in solvent protection at 16.6 mg/mL, suggesting that Protein X association involves a conformation change in this part of the molecule that gives an increase in solvent accessibility. Together these findings suggest that the main interaction between the dimers forming at pH 6.4 involves association within the modified region, but that there is also a conformation change in the unmodified region.

At pH 5.2, the patterns of solvent accessibility are qualitatively similar, shown in Figures 3.4 and 3.5, but show increases in degree of protection relative to pH 6.4. At the lower pH, a larger number of reporter peptides also show significantly decreased labeling with increased protein concentration. These reporter peptides are again largely restricted to the modified region-containing peptides (reporter peptides 3 to 7 and 9 to 20).

As with Protein X at pH 6.4, the unmodified region portion of the molecule does not show statistically significant differences in labeling greater than 0.1  $D/N$  between the two concentrations except for Peptide 24. Here again there is a decrease in solvent protection, suggesting that a change in conformation occurs in the unmodified region.

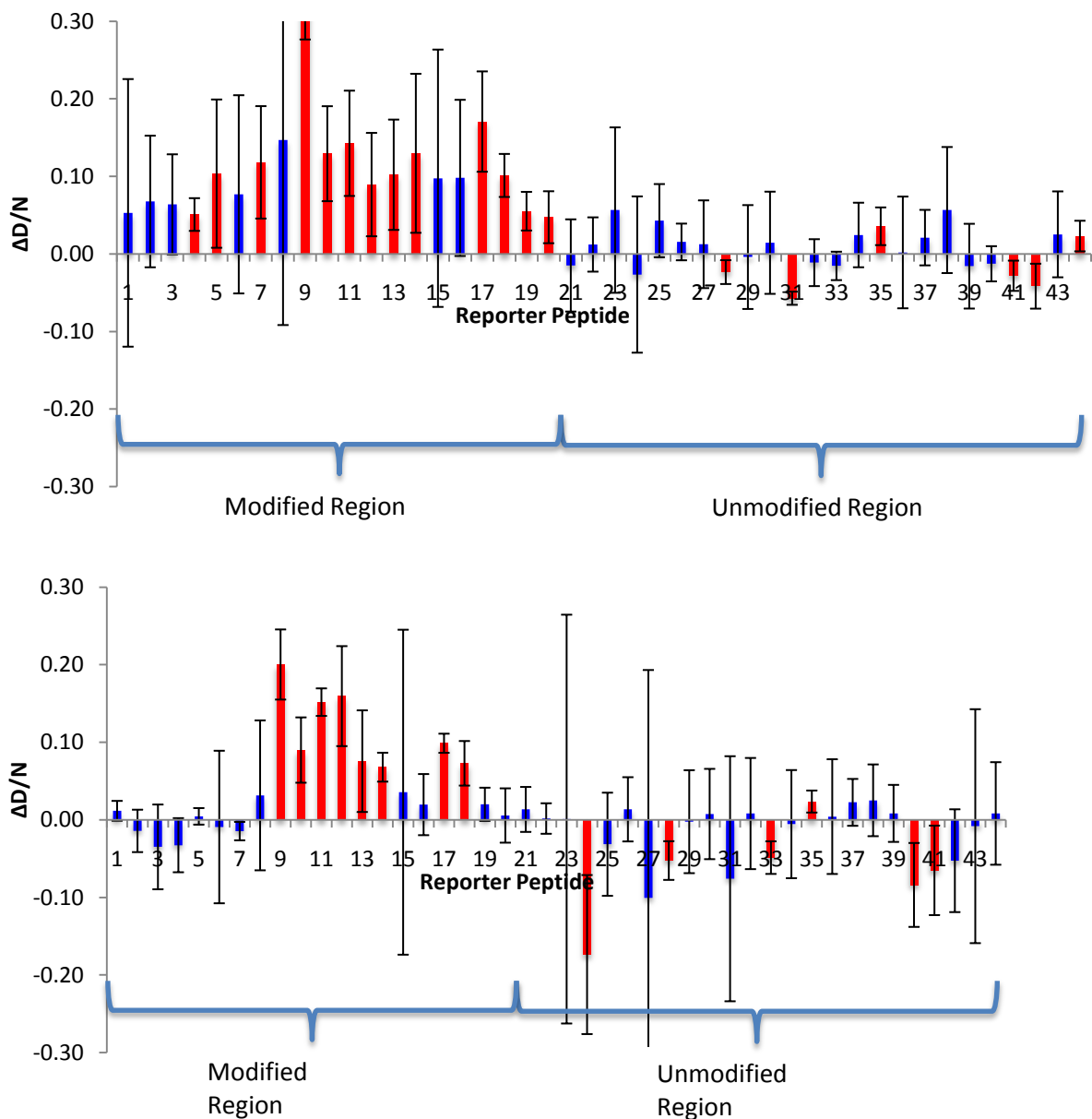
It is possible that other regions of the molecule self-associate but their effect on solvent protection could be too weak for detection by our HX-MS measurements. Our protocol involves dilution of the protein sample into deuterated buffer at a ratio of 1:9, bringing the concentration of the protein to 1/10<sup>th</sup> of its original concentration. For samples with dimers and oligomers this brings the concentration close to where monomers are expected to appear. Although such dissociation would not be instantaneous, certain

interactions between associated molecules may weaken to the point where the solvent protection is insufficient to be detected by HX-MS.



**Figure 3.4.**  $\Delta D/N$  values between concentrations of 0.5 and 18.9 mg/mL for Protein X reporter peptides in pH 5.2, 30 mM sodium succinate at labeling times of 2 minutes (top) and 10 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.



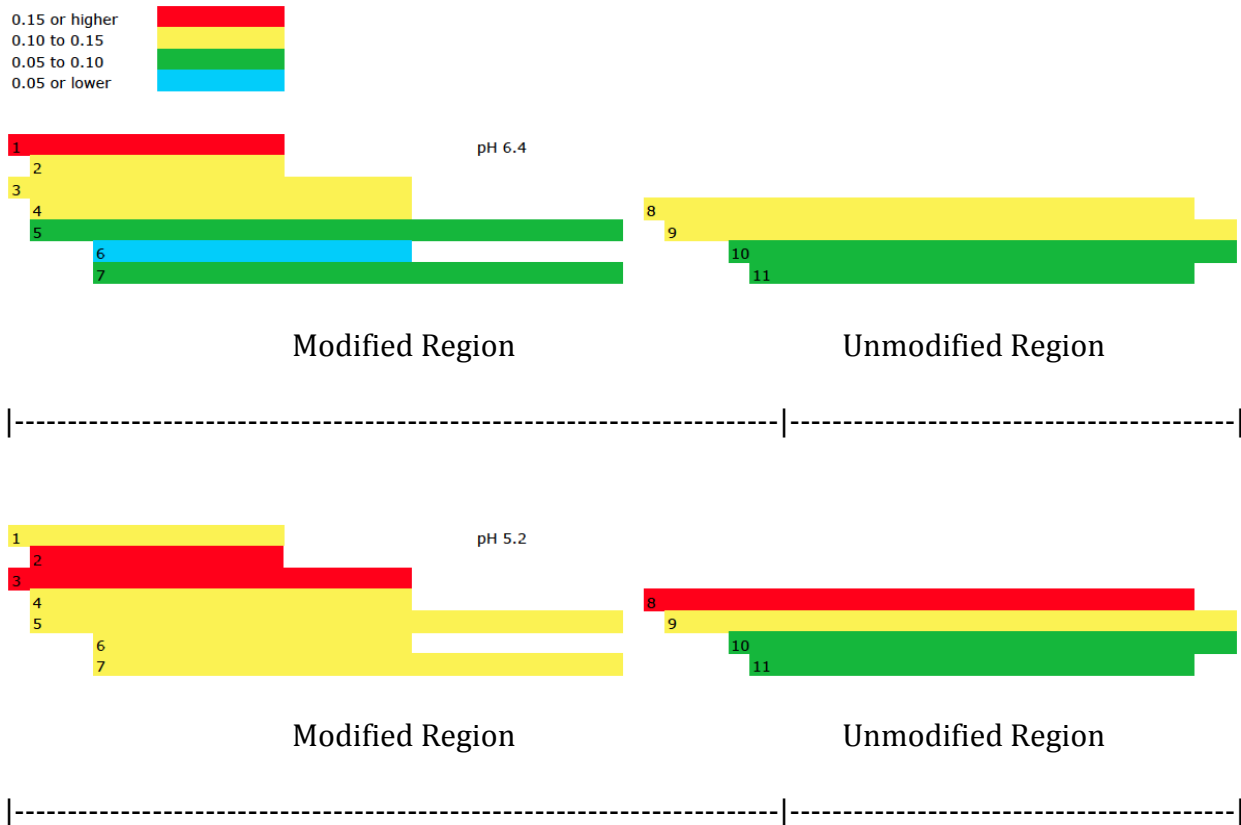


**Figure 3.5.**  $\Delta D/N$  values between concentrations of 0.5 and 18.9 mg/mL for Protein X reporter peptides in pH 5.2, 30 mM sodium succinate at labeling times of 30 minutes (top) and 120 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.

Certain locations in the modified region are more involved in association than others and these interactions are stronger at pH 5.2. Figure 3.6 shows a comparison of the labeling differences between the two pHs investigated for the residues of the modified region and the N-terminus of the unmodified region. The colored bars, showing the residues in the peptides numbered in the bars, represent the differences in labeling between the concentrations investigated at each pH. Red, yellow, green, or blue represent peptides that at high concentrations show a decrease in  $D/N$  in the range of 0.15 or greater, 0.10 to 0.15, 0.05 to 0.10, or 0 to 0.05, respectively. These values reflect the maximum difference observed at any labeling time. By comparing the colors of peptides with some overlapping residues, those that are involved in association may be inferred.

The following observations are apparent. First, the interactions in the modified region are stronger at a pH of 5.2 than at 6.4. All of the peptides, except for Peptide 1, show the same or higher increase in protection at pH 5.2 as protein concentration is increased. Second, certain locations in the modified region appear to be more involved in the association than others. Comparing Peptides 1 to 4 with Peptide 6 suggests that the N-terminus of the modified region is highly involved in the association. Comparing Peptide 6 with Peptides 5 and 7 also suggests that the residues of Peptides 5 and 7 that are not in Peptide 6 are also more involved in the association.

### Difference in $D/N$

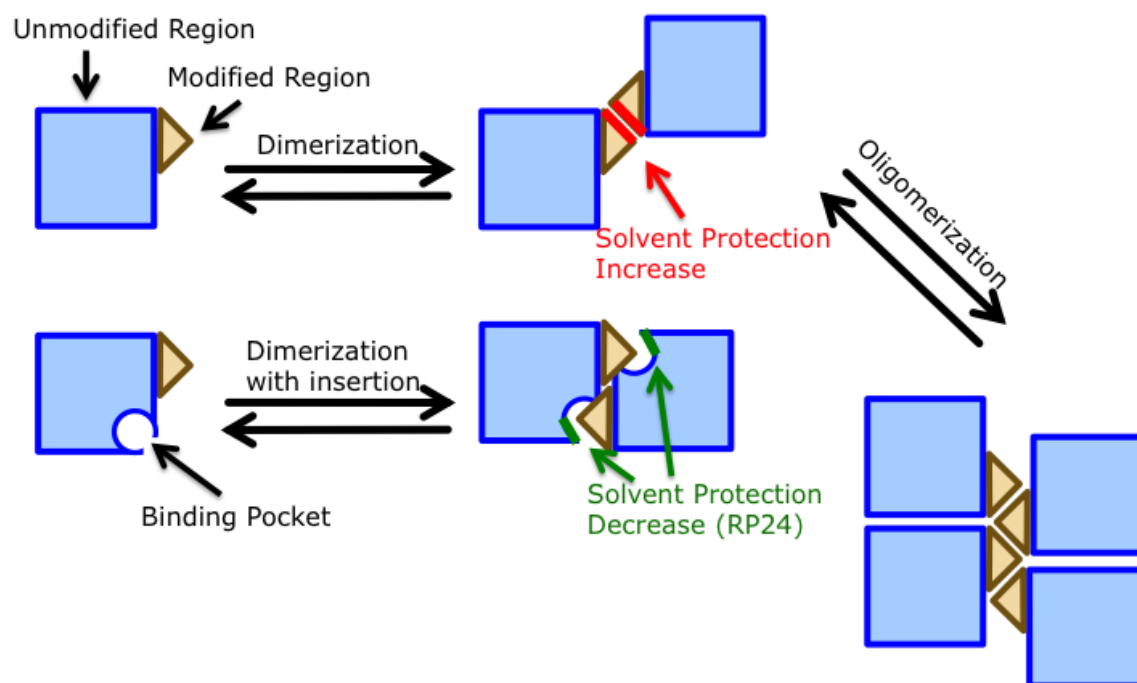


**Figure 3.6.**  $D/N$  differences between concentrations of 0.8 and 16.6 mg/mL at pH 6.4 (top) and between concentrations of 0.5 and 18.8 mg/mL at pH 5.2 (bottom) for reporter peptides in the modified region and beginning of unmodified region. Red, yellow, green, or blue represent peptides that at high concentrations show a decrease in  $D/N$  of 0.15 or greater, from 0.10 to 0.15, from 0.05 to 0.10, and from 0 to 0.05, respectively.

From comparison of Peptides 8 and 9 with Peptides 10 and 11 at both pHs, it is apparent the residues towards the C-terminus of the modified region are involved in the association, while the residues towards the N-terminus of the unmodified region are probably not involved. These findings suggest that even the N-terminus of the unmodified region is not involved in the association between Protein X molecules while the C-terminus of the modified region still factors into the association.

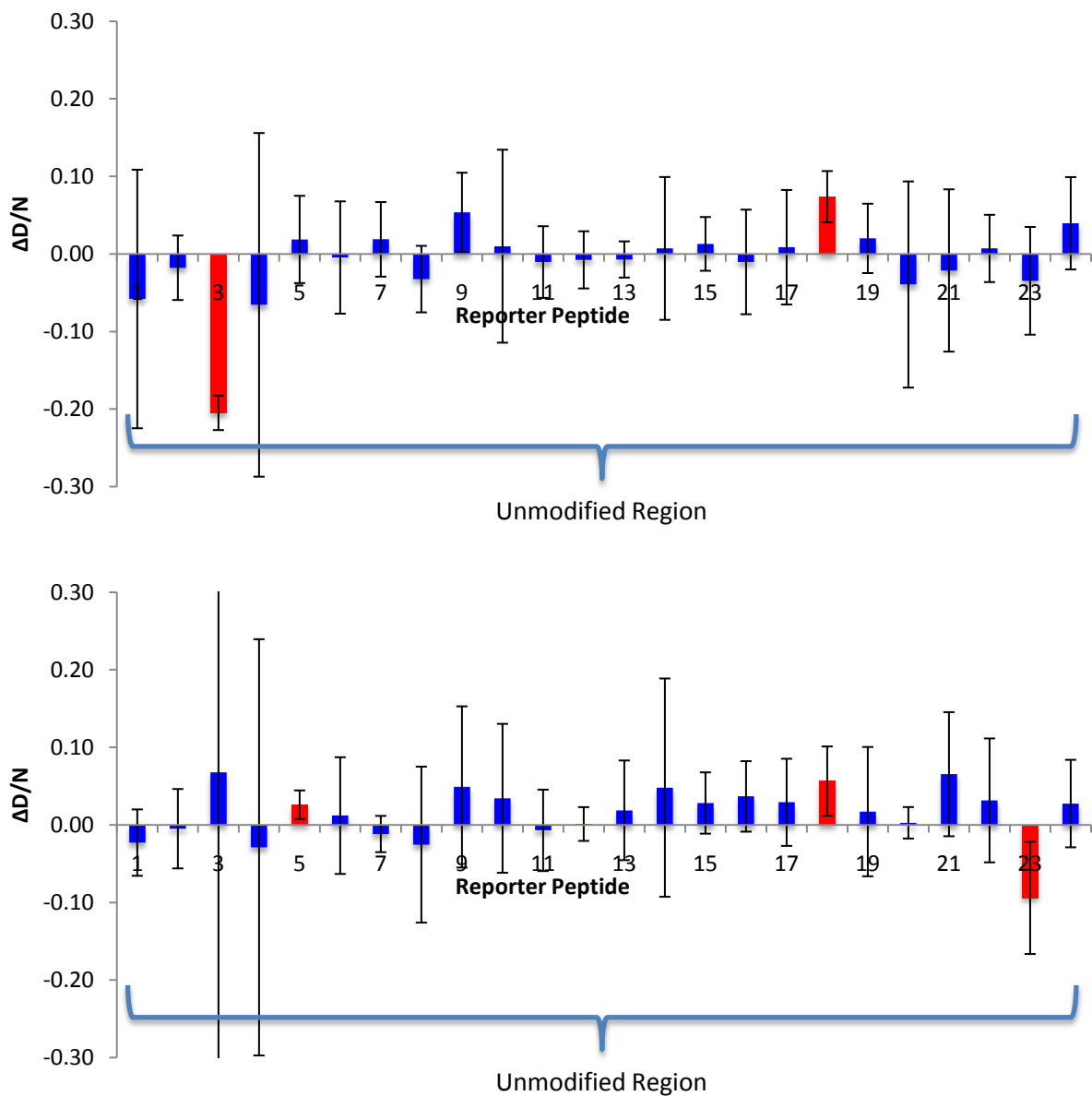
At both pHs, reporter Peptide 24 is the only region of Protein X that significantly increases in solvent exposure with increased protein concentration. The location of this peptide is in close spatial proximity to a known binding pocket of Protein X. The proximity of this peptide to the binding pocket suggests that Protein X association may not involve the direct association of only the modified region but instead, or in addition, may involve penetration of parts of the modified region into the known binding pocket.

Figure 3.7 shows schematic representations of proposals for Protein X molecules interacting based on patterns of change in solvent accessibility. For dimers, all significant increases in solvent protection occur in the modified region, suggesting the molecules associate through contact of their respective modified regions. However, the decrease in solvent protection of reporter Peptide 24 suggests that conformational change of a binding pocket with insertion of the modified region may be an alternative mechanism for dimerization. For oligomers, more and larger increases in solvent protection are observed in the modified region suggesting intermolecular interactions between multiple modified regions results in oligomerization. Although reporter Peptide 24 also partially unfolds under oligomer conditions, there is no obvious way to incorporate the unfolding of the binding pocket into a schematic of oligomerization.

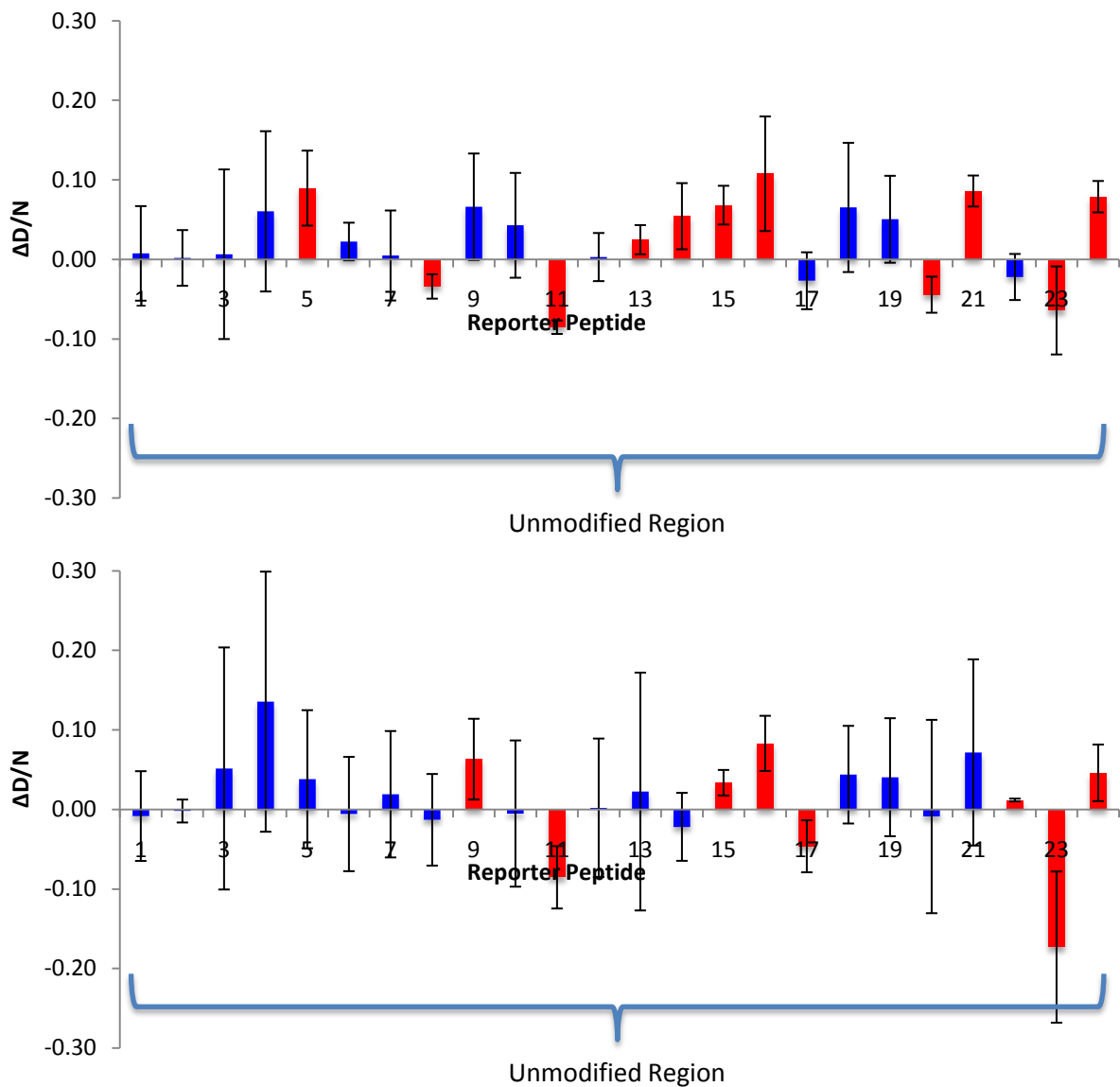


**Figure 3.7. Schematic representation of different association mechanisms based on changes in solvent protection. Shown are possible mechanisms of dimerization without and with insertion of the modified region into a binding pocket. Also shown is a possible structure for oligomers.**

Figures 3.8 and 3.9 show  $\Delta D/N$  values for reporter peptides of Protein X with and without the modified region at a pH of 6.4 in 30 mM sodium succinate at a concentration of 16.6 mg/mL. The small  $\Delta D/N$  values show that the addition of the modified region to Protein X does not affect the conformation of the unmodified region. Statistically significant differences do exist but the magnitudes of these differences are small. Thus, this suggests that the structure of the unmodified portion of the molecule has remained native like.



**Figure 3.8.**  $\Delta D/N$  values between Protein X at 16.6 mg/mL with and without modified region for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 30 seconds (top) and 2 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.



**Figure 3.9.**  $\Delta D/N$  values between Protein X at 16.6 mg/mL with and without modified region for Protein X reporter peptides in pH 6.4, 30 mM sodium succinate at labeling times of 10 minutes (top) and 30 minutes (bottom). Error bars represent 95% confidence intervals by Student t-test with the Bonferroni correction. Red bars represent statistically significant differences.

### 3.3.2. Isothermal Titration Calorimetry of Protein X Samples

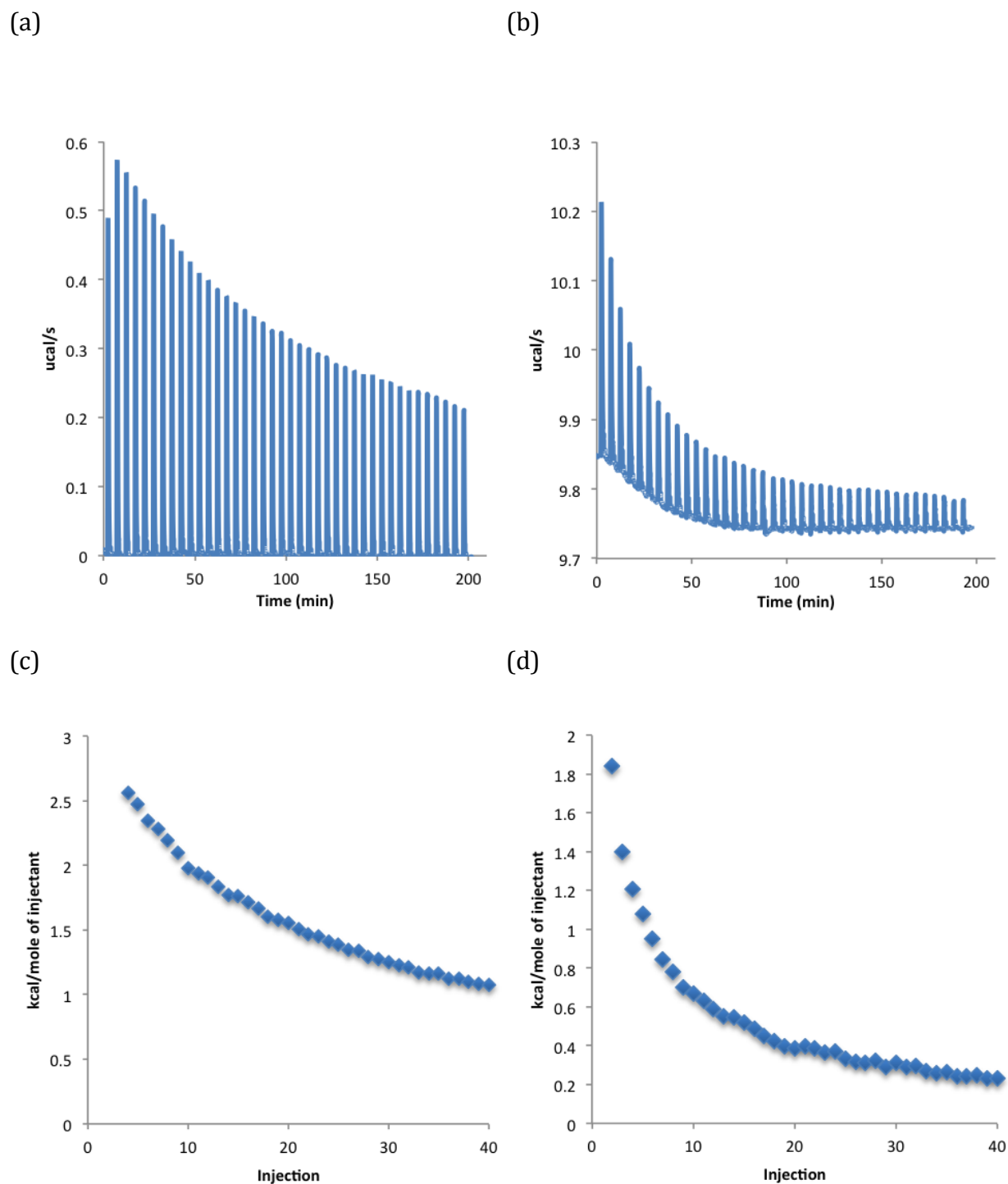
ITC can provide additional thermodynamic information about self-association. The enthalpy difference of injecting small amounts of protein into a solution can be integrated and fitted to a thermodynamic model to obtain equilibrium constants and standard-state Gibbs energies of the association.

Figure 3.10 shows the raw calorimetric data and integrated injection heats for Protein X at pH 6.4 and 5.2. The data were fitted to a dimer dissociation equilibrium model (Czypionka et al. 2007) to give the thermodynamic values in Table 3.2. Consistent with the SEC and HX-MS results, while association is seen at pH 6.4, it is strong association at pH 5.2 (lower  $K_{diss}$  and higher  $\Delta G_{diss}$ ). Notably, while  $\Delta H_{diss}$  at pH 5.2 is less favorable than at pH 6.4 for association, the change in  $\Delta S_{diss}$  is large enough to make association overall more favorable at pH 5.2 than 6.4.

**Table 3.2. Thermodynamic Parameters of Protein X Dissociation**

pH	Concentration [mg/mL]	$K_{diss}$ [mM]	$\Delta G_{diss}$ [kJ/mol]	$\Delta H_{diss}$ [kJ/mol]	$\Delta S_{diss}$ [J/mol-K]
<b>6.4</b>	21.5	0.156±0.020	33.0±0.3	127±1	304±5
<b>5.2</b>	20.2	0.0152±0056	39.1±0.9	79.1±6.6	129±21





**Figure 3.10.** Calorimetric data for the dissociation of Protein X. Raw data (a) and integrated injection heats (c) for injections of 21.5 mg/mL Protein X, 30 mM sodium succinate, pH 6.4 solution. Raw data (b) and integrated injection heats (d) for injections of 20.2 mg/mL Protein X, 30 mM sodium succinate, pH 5.2 solution.

### 3.3.3. Size-Exclusion Chromatography

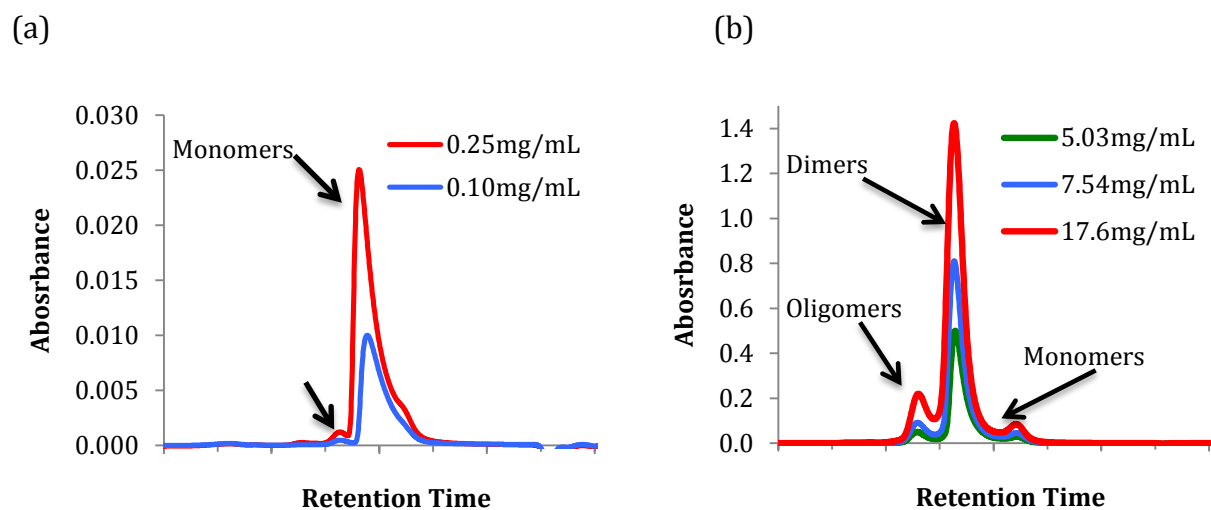
SEC chromatograms were analyzed at pH 5.8 for various Protein X concentrations in 10 mM sodium succinate. Figure 3.11 shows the chromatograms for two protein concentrations below 0.50 mg/mL and three above 0.50 mg/mL. The peak areas for the monomers and dimers were converted to concentrations to calculate  $K_{diss}$  and  $\Delta G_{diss}$  with Equations (3.4), (3.5), and (3.2).

At concentrations below 0.50 mg/mL (Figure 3.11a) monomers are the dominant species with only a small fraction of dimers present. At concentrations above 0.50 mg/mL (Figure 3.11b) monomers, dimers and oligomers are present with dimers now the dominant species. Figure 3.11 shows that there is a concentration threshold at 0.50 mg/mL for the formation of dimers and oligomers at pH 5.8. At all concentrations above 0.50 mg/mL (Figure 3.11a)  $\Delta G_{diss}$  is  $35.5 \pm 1.0$  kJ/mol. Together with the ITC results, a trend is observed that as pH decreases,  $\Delta G_{diss}$  increases making, association more favorable.

Hydrogen bonding, salt bridges, steric (excluded volume), electrostatic, van der Waals and hydrophobic interactions all contribute to protein-protein interactions (Larson 1999). The ITC and SEC data support the importance of electrostatics in Protein X self-association. As pH decreases, so the net positive charge of the protein increases, the association interaction becomes stronger, suggesting charge is one of the more important factors in Protein X self-association.

Crossing the pI, 5.7, seems to determine whether dimers or oligomers are the dominant species present. Specifically, the histidines in the modified region may play an important role in the formation of dimers or oligomers. The pKa of Histidine's side chain

can depend on its burial within the protein (Edgcomb & Murphy 2002; Antosiewicz et al. 1996). With an experimental side chain pKa of  $6.6 \pm 1$ , histidine will be mostly protonated and positively charged at pHs below 6.6 assuming no large effects on the pKa from the surrounding local environment. These additional charges may be important in the formation of oligomers as the interaction between the modified regions becomes stronger at lower pHs as seen with the HX-MS data.



**Figure 3.11. SEC chromatograms of Protein X at various concentrations below (a) and above (b) 0.50 mg/mL in 10 mM sodium succinate, pH 5.8 solution.**

### ***3.4. Conclusions***

HX-MS has been used to identify different exchange patterns for Protein X at high concentrations, where self-association is expected, relative to low concentrations dominated by monomeric species. At both pH 6.4 and 5.2, increased solvent protection is nearly completely in the modified region, identifying it as the primary association site. At pH 5.2, the association interaction is stronger and may involve different conformations and higher order oligomers. There are slight variations in association strength among the locations of the modified region involved in the aggregation process, as specific residues in the modified region show different increases in solvent protection with increased protein concentration. The only location of the unmodified region of Protein X that shows a change (decrease) in solvent protection is close to a known binding site. This suggests that association may involve the insertion of modified region residues into a binding cavity. ITC and SEC results confirm the findings of HX-MS that there is a stronger association at lower pHs, driven by a favorable decrease in the entropy of dissociation.

## **4. Unfolding of a Model Protein on Ion Exchange and Multi-Mode Chromatography Surfaces**

### ***4.1. Background and Introduction***

Ion-exchange chromatography (IEC) is a valuable tool for protein purification in the pharmaceutical industry. IEC can be used as a capture step for therapeutic proteins, as an orthogonal polishing step to HIC, or even in assisting the refolding of unfolded protein molecules. IEC has the added advantage that, in general, protein unfolding and denaturation is not observed. Almost all documented cases of unfolding in chromatography involve hydrophobic surfaces, but there is recent evidence that IEC surfaces can also cause unfolding and aggregation of proteins. Unusual elution profiles with multiple peaks and lower than expected yields had been previously observed, but role of surface unfolding was uncertain (Voitl et al. 2010; Azevedo et al. 2009). Even in cases without unusual elution profiles, unexpected increases in high molecular weight species have been observed after IEC steps, suggesting surface-induced unfolding and aggregation (Arakawa et al. 2007). In one study, the unfolding and aggregation of an IgG1 on a strong cation exchanger was demonstrated using size-exclusion chromatography (SEC) and hydrogen-exchange mass spectrometry (HXMS) (Gillespie et al. 2012). In that study, residence time, temperature, pH, and ionic strength all influenced protein unfolding and aggregation on the IEC surface.

Not only does unfolding impact yield losses, inactive product, and/or high levels of aggregate, it also affects retention. Protein conformation in IEC can affect chromatographic

separations for proteins unfolded with urea or dithiothreitol (Yamamoto et al. 2007; Hou et al. 2010). These effects on retention time may have positive or negative effects on selectivity. For example, if the unfolding is reversible, it could improve resolution as applied previously in hydrophobic interaction chromatography (HIC) (Lindahl & Vogel 1984; Wada et al. 1983; Valliere-Douglass et al. 2008; Deitcher et al. 2009).

There seem to be no studies on how different operational variables affect unfolding in IEC. Studies with protein unfolding in HIC have demonstrated that both mobile and stationary phase properties are important to consider (Gagnon et al. 1995), so it is likely that these are also important in IEC unfolding. Of particular interest with the mobile phase is how the variables that affect binding, such as pH and ionic strength, also affect unfolding. For stationary phases, it is unknown if a protein that unfolds on a cation exchanger will also unfold on an anion exchanger or vice versa. Further, it is uncertain if proteins unfold more or less on strong IEC media (charged over a wide pH range) or versus weak media (charged only over a narrow pH range).

Understanding unfolding on IEC surfaces is also important when considering multi-mode chromatography (MMC). MMC stationary phases have both IEC and HIC characteristics to enhance selectivity for removing certain impurities and providing operational advantages over single-mode steps such as IEC or HIC. IEC often requires dilution of samples to lower ionic strength to ensure protein binding, and HIC operation often requires the addition of lyotropic salts to promote protein binding to the surface. By incorporating multiple modes of adsorption (e.g. both hydrophobic and electrostatic), MMC often does not require mobile phase modification of protein samples for binding. This provides a cost benefit over HIC and IEC processes that require these modifications.

Recent attention in industry has shifted to MMC as a tool for purifying recombinant proteins. MMC is currently in use as an alternative polishing step to IEC or HIC (DePalma 2012; Eriksson et al. 2009; Kallberg et al. 2012). MMC can be an attractive alternative for capture steps of therapeutic proteins. Protein A resins are generally used for monoclonal antibody capture, but the high cost associated with them has driven interest towards finding alternative media. Further, certain classes of therapeutic proteins (i.e. fusion proteins) may not necessarily have the binding sites required for interaction with Protein A. MMC has shown promise in the capture of monoclonal antibodies and other therapeutic proteins, but more work needs to be done to improve process yield and purity (Bak & Thomas 2007; Gagnon et al. 2011).

MMC uses a variety of ligands and can be operated in several ways. The characteristics of the target protein often influence the choice of ligand. Most systems have both electrostatic and hydrophobic interactions for adsorption of the target protein. Typically, the choice of the ion-exchanging ligand depends on the operating pH, while choice of the hydrophobic ligand should not destabilize the protein (Zhao et al. 2009).

Although the ligands in MMC come in a variety of chemistries, all have functionalized groups with IEC and HIC characteristics. Recently, many MMC stationary phases have also included functionalized groups for hydrogen bonding to promote adsorption. For example, in addition to electrostatic interactions the strong anion-exchanger Canto™ adsorbs proteins through a phenyl group for hydrophobic interactions and a hydroxyl group for hydrogen bonding. The existence of multiple interactions makes understanding the adsorption mechanism a challenge. Rational screening of ligands, however, demands a thorough understanding of the adsorption mechanism for efficient

chromatographic separation (Zhao et al. 2009). At this time, it is unknown how each type of interaction contributes to protein destabilization.

In this study, the binding and unfolding behavior of a model multi-domain protein, bovine serum albumin (BSA), was examined on a series of IEC and MMC surfaces. It was hypothesized that by changing the degree of electrostatic interaction between the protein and IEC surface (through pH) the unfolding of BSA would vary. It was also hypothesized that changing the type of electrostatic interaction (by switching from cationic to anionic surfaces) would also affect BSA unfolding. Finally, it was hypothesized that by varying pH and ammonium sulfate concentration for BSA adsorbed on MMC surfaces, the individual effects of electrostatic and hydrophobic interactions on unfolding might be delineated.

Although it was uncertain whether BSA would unfold on these surfaces, the low stability of BSA, and detailed studies of its unfolding on HIC for comparison, made it a good candidate. To start, the effect of pH and ammonium sulfate concentration were examined with UV measurements and hydrogen-exchange mass spectrometry (HXMS) to identify condition(s) giving unfolding. Additional HXMS experiments were then done with proteolytic digestion to observe how unfolding in the different regions of BSA is affected by changing pH and type of IEC surface. Finally, this study showed that unfolding of BSA on Capto MMC may be more dependent on pH than on hydrophobic interactions.



## **4.2. Materials and Methods**

### **4.2.1. Materials**

BSA was purchased from Sigma-Aldrich (St. Louis, MO, USA). Citric acid and ethylenediaminetetraacetic acid (EDTA) were purchased from Sigma Aldrich (St. Louis, MO, USA). Monosodium citrate and disodium citrate were purchased from ACROS Organic (New Jersey, USA). Guanidine hydrochloride (GdnHCl) was purchased from MP Biomedicals (Solon, OH, USA). Disodium phosphate was purchased from Fisher Scientific (Fair Lawn, NJ, USA). Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Thermo Scientific (Rockford, IL, USA).

The IEX resins, SP Sepharose Fast Flow (FF) and Q Sepharose Fast Flow (FF), and MMC resins, Capto MMC and Capto Adhere, used in this study were purchased from GE Healthcare (Uppsala, Sweden). Ultrafree®-MC centrifugal filter units were purchased from Fisher Scientific (Houston, TX, USA) for the separation of supernatant liquid from resin particles.

### **4.2.2. pH Studies**

For no-surface control (solution) experiments, 5  $\mu$ L of 20 mg/mL protein solution were mixed with 45  $\mu$ L of deuterated buffer at room temperature. The protein solutions and labeling buffers were prepared at pH 3.0, 3.5, 4.0, and 4.5 in 50 mM citrate for SP Sepharose FF and pH 6.5, 7.0, 7.5, and 8.0 in 50 mM phosphate for Q Sepharose. Labeling times varied for each experiment, as outlined below. After labeling, 5  $\mu$ L of quench buffer

(150 mM potassium phosphate, pH 1.5), kept in an ice bath, were added, bringing the final solution pH to 2.6, near the pH minimum of the hydrogen-deuterium exchange reaction. Samples were kept at room temperature for 40 seconds before 145  $\mu$ L of desorption buffer were added to the solution. The desorption buffer was at pH 2.6, with 100 mM citric acid, 8M GdnHCl, 100 mM TCEP, and 27 mM EDTA in H<sub>2</sub>O. After addition of desorption buffer, the sample was placed in ice for 2 minutes before being kept at room temperature for 40 seconds. Then, 600  $\mu$ L of 95% H<sub>2</sub>O, 5% acetonitrile, 0.1% formic acid, and 0.01% trifluoroacetic acid were added to dilute the protein and GdnHCl concentration. Solution phase samples were placed at room temperature for 40 seconds to replicate the time between sample quenching and introduction into the MS for the adsorbed phase experiments with 2 additional centrifugation steps (40 seconds each).

For adsorbed phase experiments, 35  $\mu$ L of 20 mg/mL protein solution were added to 65  $\mu$ L of resin slurry (50:50 dry resin:working buffer) in an Ultrafree®-MC centrifugal filter unit within a 1.5 mL microcentrifuge tube. The samples were equilibrated overnight to ensure adsorption equilibrium. Prior to labeling, the sample was centrifuged at 7.4 rcf for 30 seconds to separate supernatant liquid. The concentration of the supernatant was measured via UV at 280 nm to obtain protein binding under the different conditions by material balance. To initiate labeling, 100  $\mu$ L of deuterated buffer were added to the filter unit at room temperature. Labeling times were the same as those for the solution experiments, as described below. After labeling, 10  $\mu$ L of quench buffer were added to the filter unit and the microcentrifuge tube was immediately centrifuged at 7.4 rcf for 30 seconds. The filter unit was transferred to a new microcentrifuge tube in ice and 145  $\mu$ L of desorption buffer were added. The sample was placed in ice for 2 minutes and then

centrifuged at 7.4 rcf for 30 seconds. Finally, 600  $\mu$ L of sample pump solution were added to dilute the protein and GdnHCl concentration.

Labeling times varied with surface and pH. For whole protein experiments on Q Sepharose and Capto Adhere, labeling was 10 minutes for all pHs. For whole protein experiments on SP Sepharose FF and Capto MMC a different labeling time for each pH was used: 100, 50, 10, and 5 minutes for pH 3.0, 3.5, 4.0, and 4.5, respectively. This tiered labeling time accounts for the logarithmic decay of the hydrogen-deuterium exchange rate with pH (Berger et al. 1959). This use of tiered labeling was intended to ensure an measurable level of deuterium uptake at the lower pHs. The tiered labeling times were also used for the no-surface (solution) control experiments.

BSA on SP Sepharose FF and Capto MMC were further investigated with proteolytic digestion at pH 4.0 and 4.5. The same protocols were used for these studies except additional labeling times of 5, 20, 40, 60, and 120 minutes were examined. Again, no-surface (solution) controls were treated with the same conditions.

For whole protein studies, 95% sample confidence intervals were obtained from triplicate data of BSA labeling under one condition: on SP Sepharose FF at pH 4.5. For proteolytic studies, 95% sample confidence intervals for each reporter peptide were obtained from a previous study for BSA labeling on Butyl 650M at 22°C and pH 7.0. Ideally, triplicate data would be collected for each experimental condition but the long processing time (~60 minutes) made this difficult. Because much of the variability in the measurements is associated with labeling for protein adsorbed on the surface, the confidence intervals obtained for BSA on Butyl 650M were used as estimates of confidence intervals under other conditions.

### 4.2.3. Salt effect studies

Twenty mg/mL BSA samples and labeling buffers were prepared at pH 5.5 in 50 mM phosphate with varying concentrations of ammonium sulfate (0 to 1.5M in 0.5M increments). For no-surface control experiments, 5  $\mu$ L of 20 mg/mL protein solution were mixed with 45  $\mu$ L of deuterated buffer at room temperature. For experiments with Capto MMC and Capto Adhere, 35  $\mu$ L of protein solution were added to 65  $\mu$ L of resin slurry in an Ultrafree®-MC centrifugal filter unit within a 1.5 mL microcentrifuge tube. The labeling time was 20 minutes. The remainder of the protocol for the solution and adsorbed phase experiments was the same as with the pH studies.

### 4.2.4. HPLC-MS

For the whole protein studies on both surfaces and in both solutions, samples were injected into a 200  $\mu$ L stainless steel sample loop using a 500  $\mu$ L glass syringe. A sample pump (LabAlliance™ Series I) pumped 95% ddH<sub>2</sub>O, 5% acetonitrile, 0.05% formic acid, and 0.01% trifluoroacetic acid solution and the injected sample at 100  $\mu$ L/min through the sample loop. Proteins were trapped, desalted, and concentrated on a C<sub>4</sub> column (TR1/25109/02, 1 mm inner diameter by 8 mm length, Michrom Bioresources, Inc, Auburn, CA). After this 5 minute desalting step, flow was switched from the sample pump to the Surveyor MS HPLC pump to elute the proteins off the C<sub>4</sub> column. Short gradients were used to minimize back-exchange during the desorption step. The whole proteins desorbed with a 16 minute linear gradient from 95% solvent A (ddH<sub>2</sub>O, 0.1% formic acid,

0.01% TFA) and 5% solvent B (acetonitrile, 0.8% formic acid) to 5% solvent A, followed by 1 minute at 5% solvent A, followed by a 1 minute gradient from 5% solvent A to 95% solvent A, and completed with 2 minutes at 95% solvent A.

For the digestion studies on both surfaces and in both solutions, the protein was injected into the sample loop in the same manner as the whole protein studies. Different columns were used to optimize peptide generation and resolution. The protein sample was first pumped into an immobilized pepsin column (2.1 mm inner diameter by 60 mm length) where proteolytic digestion took place. Pepsin preferentially cleaves at the C-terminal side of phenylalanine, leucine, tryptophan, tyrosine, alanine, glutamic Acid, and glutamine, allowing consistent peptide fragments to be generated for different runs. In this study, the large number of disulfide bridges on BSA probably prevented obtaining high sequence coverage since disulfide bridges hinder digestion of proteins and reporter peptide identification by increasing conformational stability of proteins (Pace et al. 1988). TCEP, a disulfide-reducing agent, was added in the experimental protocol but its contact time (~3-4 minutes) with the protein may not have been sufficient to reduce all disulfides since TCEP is listed as capable of doing most reductions at concentrations of 5 to 50mM within 5 minutes at room temperature (info courtesy of Thermo Scientific).

Peptides exiting the column were trapped, desalted, and concentrated on a C<sub>8</sub>-desalting column (TR1/25109/02, 1 mm inner diameter by 8 mm length, Michrom Bioresources, Inc, Auburn, CA). After this 6 minute desalting step, flow was switched from the sample pump to the Surveyor MS HPLC pump to elute the peptides off the C<sub>8</sub> column. Additionally, an XBridge C<sub>18</sub> column (186003563, 2.1 mm inner diameter by 50 mm length, 3.5  $\mu$  m pore size, Waters, Milford, MA) downstream of the C<sub>8</sub> column was used for

improved resolution of the large number of peptides. A short gradient run was employed to minimize back-exchange, while effectively resolving the peptides. The treatment for peptide desorption was a 17 minute gradient of 70% solvent A to 40% solvent A, followed by a 2 minute gradient from 40% solvent A to 10% solvent A, followed by 4 minutes at 10% solvent A.

Whole proteins and peptides were eluted directly to a LTQ linear electrospray ionization quadrupole ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA). Data were collected in a positive ion, profile mode with an ESI voltage of 4.3 kV, a capillary temperature of 250°C, and sheath gas flow rate of 15 units. The peptides identified in MS/MS experiments are shown in Table A3.

## 4.3. Theory

### 4.3.1. Measuring and Calculating Fractions of Labeled Peptide

The extent to which a peptide has been labeled with deuterium is determined by (Zhang & Smith 1993)

$$\frac{D}{N} = \frac{m_t - m_0}{m_{100} - m_0} \quad (4.1)$$

where  $D$  is the number of deuterated amides,  $N$  is the number of exchange-competent residues in a peptide,  $m_t$  is the mass of a peptide after a given labeling time,  $m_0$  is the non-deuterated mass of that peptide, and  $m_{100}$  is the fully-deuterated mass of that peptide.

In the present cases, exchange-competent refers to all peptide residues except proline, which does not have amide hydrogen, and the N-terminal residue of the peptide, which does not have a backbone amide. Back-exchange of deuterium for hydrogen occurs when the deuterated protein molecules are introduced back into the H<sub>2</sub>O-solvents used in high performance liquid chromatography (HPLC) for peptide resolution. Thus, the residue immediately after the N-terminal residue of a peptide is also not counted; the back-exchange of this residue is unusually high (Bai et al. 1993). Equation (4.1) also accounts for back-exchange experienced during the time between sample quenching and introduction into the MS (Zhang & Smith 1993).

### 4.3.2. Fast, medium, and slow exchangers

The hydrogen-deuterium exchange of a residue's backbone amide can depend on several factors including amino acid type, surrounding residues, pH, salt, and temperature (Bai et al. 1993). For simplicity, however, the residues of a protein or reporter peptide can

be grouped into three different classes: fast, medium, and slow exchangers. In this classification, fast exchangers are residues capable of exchanging backbone amide hydrogens for deuteriums at time constants smaller than the shortest labeling times used, i.e. on the order of seconds. Medium exchangers exchange at time constants on the order of minutes/hours, while slow exchangers can have time constants greater than the longest labeling times used here, i.e. on the order of days. The  $D/N$  value of a protein or reporter peptide as a function of labeling time,  $t$ , can then be modeled with multiple decayed functions

$$\frac{D}{N}(t) = 1 - A_f e^{-k_f t} - A_m e^{-k_m t} - A_s e^{-k_s t} \quad (4.2)$$

where  $A_f$ ,  $A_m$ ,  $A_s$ , are the fraction of fast, medium, and slow exchangers, respectively for a protein or reporter peptide, and  $k_f$ ,  $k_m$ , and  $k_s$ , are the corresponding exchange rate constants. Equation (4.2) is further simplified in the limit of  $k_s \rightarrow 0$  and  $k_f \rightarrow \infty$

$$\frac{D}{N}(t) = 1 - A_s - A_m e^{-k_m t} \quad (4.3)$$

In the case where  $D/N$  values are available at multiple labeling times:

$$A_f = \frac{D}{N}(t_{shortest}) \quad (4.4)$$

$$A_s = 1 - \frac{D}{N}(t_{longest}) \quad (4.5)$$

where  $D/N(t_{shortest})$  and  $D/N(t_{longest})$  are  $D/N$  values at the shortest and longest labeling time, respectively, in the data set. Then,  $A_m$  can be determined from mass balance

$$A_M = 1 - A_f - A_s \quad (4.6)$$



Then,  $K_m$  can then be determined from a regression of rearranged Equation (4.3)

$$\ln \left( \frac{1 - \frac{D}{N} - A_s}{A_m} \right) = -k_m t \quad (4.7)$$

A more detailed derivation and explanation of the simplifications is presented elsewhere (Tobler & Fernandez 2002).

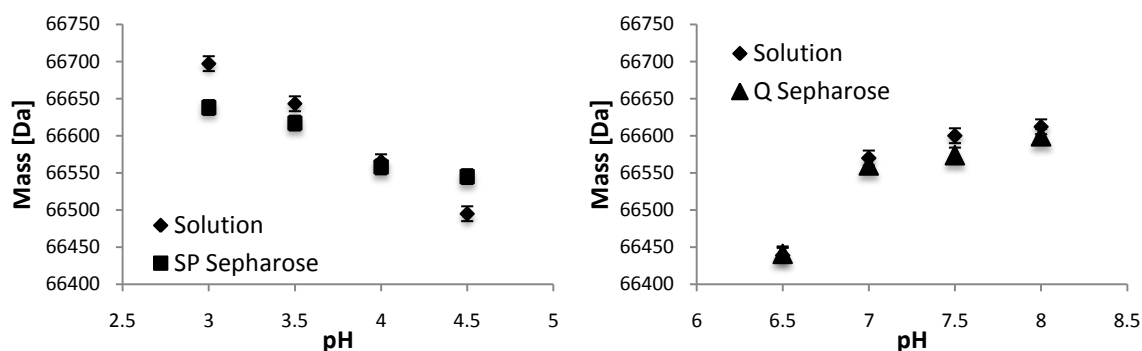
## ***4.4. Results and Discussion***

### **4.4.1. Whole protein studies**

MS spectra were collected and compared for BSA labeled in solution and while adsorbed on the two IEC and two MMC surfaces. Comparisons of labeled masses were done with Student t-test at  $p=0.05$  to infer whether differences in labeled masses were statistically significant. The whole protein studies provide qualitative trends on the effects of pH, ammonium sulfate, and surface type on the solvent exposure of BSA. Further, these studies identified conditions in which unfolding occurs on the surface that could be further examined with proteolytic hydrogen deuterium exchange studies. First, the effects of pH were examined on SP Sepharose FF, Q Sepharose FF, Capto MMC, and Capto Adhere to observe the effect on solvent exposure of electrostatic interaction strength. Finally on the MMC surfaces, ammonium sulfate was varied in concentration in an attempt to reduce electrostatic interactions relative to hydrophobic interactions for separating their effects on solvent exposure.

In general, the differences between labeled masses for BSA in solution and on the IEC surfaces were not statistically significant over most of the pH range studied. Figure 4.1 shows comparisons of the labeled masses between BSA in solution and on SP Sepharose FF (left) and between BSA in solution and on Q Sepharose FF (right) at various pHs. It should be noted that in order to obtain measurable levels of deuterium uptake progressively longer exchange times were used as pH was decreased. This complicates comparisons between different pH values, but because solution and adsorbed experiments were performed with the same exchange time, the effect of adsorption was detected. As pH

decreases from 4.5 to 3.0, Figure 4.1 shows labeling increases both in solution and on the surface. In solution, this is consistent with BSA denaturing over this pH range (Williams & Foster 1960; Lin & Koenig 1976). Given all of these data, BSA seems to unfold on SP Sepharose, but the unfolding is less strongly pH dependent than in solution. At pH 3.5 and 4.0, labeling is statistically the same between the solution and on the surface. At pH 4.5, however, 50 more residues are labeled on the surface compared to solution. The additional labeled residues suggest a conformation change on the surface at this pH.



**Figure 4.1.** Solvent exposure changes of BSA in solution (diamonds), adsorbed on SP Sepharose FF (left, squares), and adsorbed on Q Sepharose FF (right, triangles) in 50 mM citrate at pHs 3.0 to 4.5 and in 50 mM phosphate at pHs 6.5 to 8.0. In order to obtain measurable levels of deuterium uptake at each pH different labeling times were used for pH 3.0 (100 minutes), 3.5 (50 minutes), 4.0 (10 minutes), 4.5 (5 minutes), and 6.5 to 8.0 (5 minutes) as described in Methods (Section 4.2.2). Error bars represent sample 95% confidence intervals estimated based on triplicate data points collected on SP Sepharose FF at pH 4.5.

On the other hand, at pH 3.0, 59 more residues are labeled in solution when compared to SP Sepharose FF. Less labeling on the surface may be due to several reasons. Interactions with the surface could be limiting solvent exposure for several residues as has been observed for bovine  $\alpha$ -lactalbumin absorbed on a hydrophobic polystyrene surface

(Engel et al. 2004). However since exclusion of water from a hydrophobic surface is surprising, that would be even less likely on a polar charged surface.

Alternatively, the protein molecules may be aggregating or self-associating which could also limit solvent exposure for residues involved in intermolecular contacts. The effects of intermolecular interactions and aggregation on decreasing solvent exposure have been observed previously for diphtheria toxin and bovine insulin dimers formed at low pH and for amyloid beta aggregates (Man et al. 2010; Tokihiro et al. 1996; W. Qi et al. 2008). In addition, surfaced-induced aggregation for an IgG on a CEX surface leading to increased solvent protection has been observed previously as found by SEC (Gillespie et al. 2012). HXMS further characterized the aggregates as having lower solvent exposure relative to native molecules. With the present data alone, it is not possible to discern what causes decreased labeling. However, it would seem that an apparent stabilizing effect of the surface (at least relative to the solution), perhaps caused by trapping partially folded aggregated protein, could be worthwhile investigating.

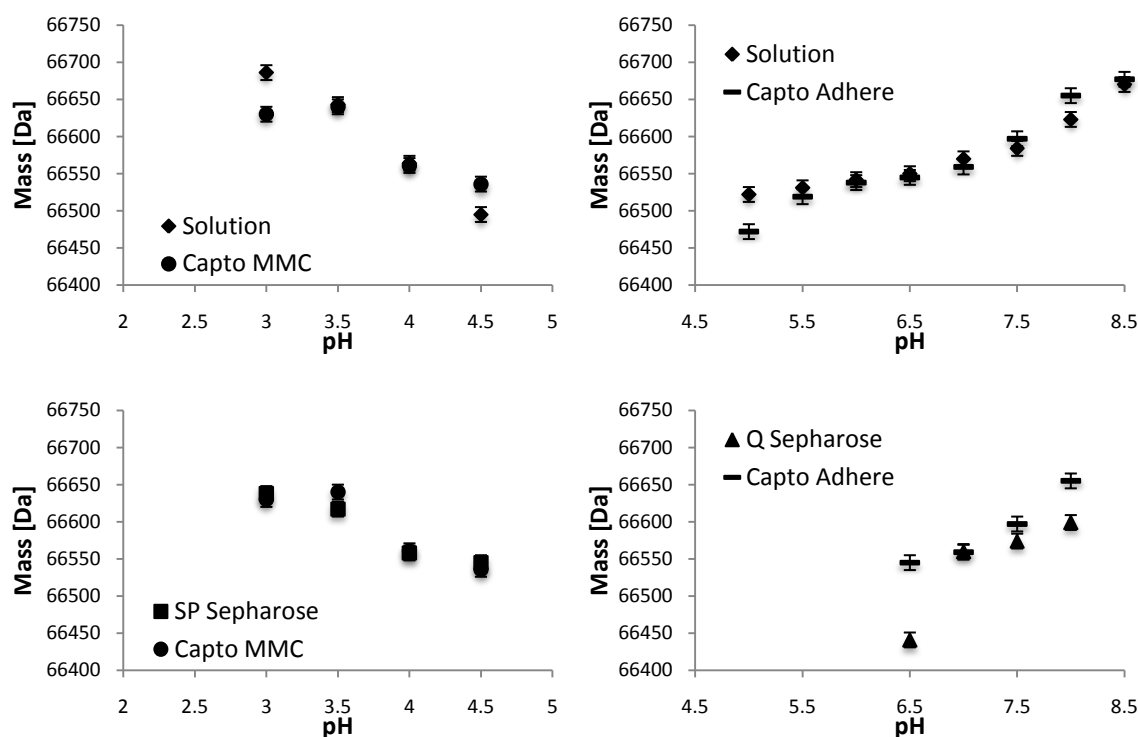
In contrast, adsorption on Q Sepharose in the pH range of 6.5 to 8.0 does not affect the labeling of BSA. Labeling both in solution on the surface is not significantly different over the pH range 6.5 to 8.0, as shown in the right of Figure 4.1. The observation that unfolding occurs on the cationic surface but not on the anionic surface suggests that the type of electrostatic interaction and/or the effect of low pH on protein stability is important for BSA unfolding. However, it should be noted that the unfolding observed on SP Sepharose FF occurs in a pH range where BSA undergoes acid-induced denaturation in solution: 4.5 to 3.0 (Williams & Foster 1960; Lin & Koenig 1976). In contrast, the alkaline pH range investigated in this study (7.0 to 8.0) is lower than the alkaline pH range (10+)

where major structural perturbations of BSA in solution have been observed (Ahmad et al. 2004; Sen et al. 2008). Additional labeling experiments at higher pH may reveal unfolding on the surface.

BSA adsorption and solvent exposure changes were also investigated on the MMC surfaces of Capto MMC and Adhere. Different solution conditions were chosen to separate electrostatic and hydrophobic interactions. To amplify electrostatic interactions and reduce hydrophobic interactions, HXMS was performed under low ionic strength (50 mM) and at pH values away from the isoelectric point of BSA. To minimize electrostatic effects, HXMS was performed at high ionic strength through the addition of ammonium sulfate (0.5 to 1.5 M), and at the isoelectric point of BSA (~5.5).

Figure 4.2 shows a comparison of labeling patterns of BSA with and without the MMC surfaces under solution conditions amplifying electrostatic interactions. The labeled masses for BSA in solution, adsorbed on Capto MMC, and adsorbed on Capto Adhere at various pHs are shown in Figure 4.2 (top). A comparison between the labeled masses of BSA on the MMC and IEC surfaces is also shown in Figure 4.2 (bottom). The comparison of BSA labeling in solution and on Capto MMC (Figure 4.2 top left) shows similar behavior as the earlier comparison between solution and SP Sepharose FF labeling (Figure 4.1 left). At pH 3.0 less labeling is observed on the surface compared to in solution while at pH 4.5 more labeling is observed. At pH 3.5 and 4.0, labeling is statistically the same in solution and on the surface. When the labeling of BSA on Capto MMC and on SP Sepharose FF are directly overlaid (Figure 4.2 bottom left), it is shown that the labeling is statistically the same. Although the two surfaces have different ligand chemistries, both have negatively charged functional groups for binding. These functional groups may be critical for BSA

unfolding on the surfaces at pH 4.5. Proteolytic HXMS studies (presented later) were done to determine how the surfaces might affect local regions of BSA.



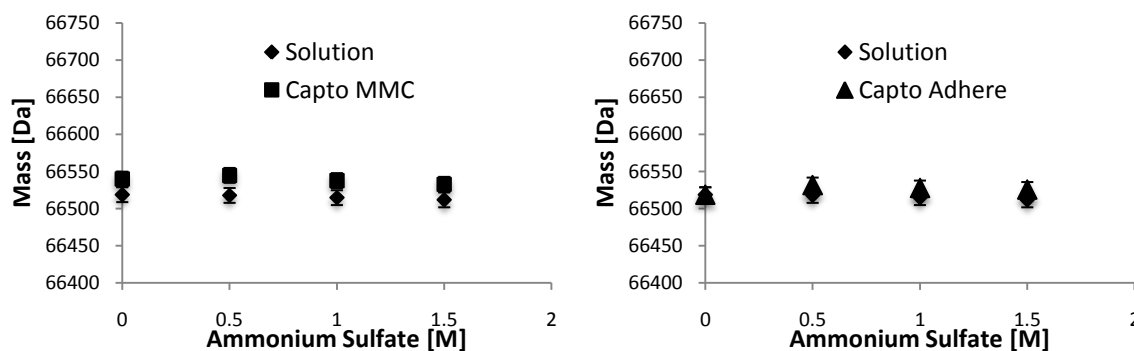
**Figure 4.2.** Solvent exposure changes of BSA under various conditions in 50 mM citrate at pHs 3.0 to 4.5 and in 50 mM phosphate at pHs 5.0 to 8.5. Top: BSA in solution (diamonds), adsorbed on Capto MMC (left, circles), and adsorbed on Capto Adhere (right, horizontal lines). Bottom: Solvent exposure changes of BSA adsorbed on Capto MMC (left, circles), adsorbed on SP Sepharose FF (left, squares), adsorbed on Capto Adhere (right, horizontal lines), and adsorbed on Q Sepharose FF (right, triangles). In order to obtain measurable levels of deuterium uptake at each pH different labeling times were used for pH 3.0 (100 minutes), 3.5 (50 minutes), 4.0 (10 minutes), 4.5 (5 minutes), and 6.5 to 8.0 (5 minutes) as described in Methods (Section 4.2.2). Error bars represent sample 95% confidence intervals estimated based on triplicate data points collected on SP Sepharose FF at pH 4.5.

Similar to BSA on Q Sepharose FF, adsorption on Capto Adhere does not have large effects on solvent exposure. For most pHs on Capto Adhere, no statistically significant differences are observed between labeling in solution and on the surface (Figure 4.2 top right). Two exceptions exist: pH 5.0 (less labeling on the surface) and pH 8.0 (more

labeling on the surface). At pH 5.0, 50 less residues are labeled on the surface compared to when in solution. At pH 8.0, 32 more residues are labeled on the surface than when in solution. Additional replicates and/or digestion studies, however, are recommended to confirm these apparent exceptions to the general trend of no difference in labeling between solution and surface.

In the case of reducing electrostatic interactions through the addition of ammonium sulfate, no effect on labeling in solution or on the two MMC surfaces was observed. The labeled masses for BSA in solution, adsorbed on Capto MMC, and adsorbed on Capto Adhere at pH 5.5 and various ammonium sulfate concentrations are shown in Figure 4.3. Although increasing ammonium sulfate concentration increases binding on both surfaces (data not shown) the concentration has no effect on labeling in solution or on the two surfaces.

It was expected that increasing ammonium sulfate concentration would decrease labeling in solution while increasing labeling on the surface, based on previous studies on the effects of ammonium sulfate on protein stability in solution and on HIC surfaces (Xiao et al. 2006; Fogle et al. 2006; Deitcher et al. 2009). At pH 5.5, near the BSA pI, it is expected that adsorption to the surface at high ionic strengths would be dominated by hydrophobic interactions as opposed to electrostatic ones. Ammonium sulfate is well known to have a stabilizing effect on proteins in solution. On hydrophobic surfaces, however, ammonium sulfate has been observed to destabilize proteins (Yunzhi Xiao et al. 2006; Fogle et al. 2006; Deitcher et al. 2009). It is not immediately obvious why this effect is not observed in this case. The protein may be very stable at this less acidic pH, resisting unfolding by increased salt concentration.



**Figure 4.3. Solvent exposure changes of BSA in solution (diamonds), adsorbed on Capto MMC (left, squares), and adsorbed on Capto Adhere (right, triangles) in 50 mM phosphate and varying concentrations of ammonium sulfate at pH 5.5. Samples were labeled with deuterated buffer for 20 minutes. Error bars represent sample 95% confidence intervals of triplicate data points collected on SP Sepharose FF at pH 4.5.**

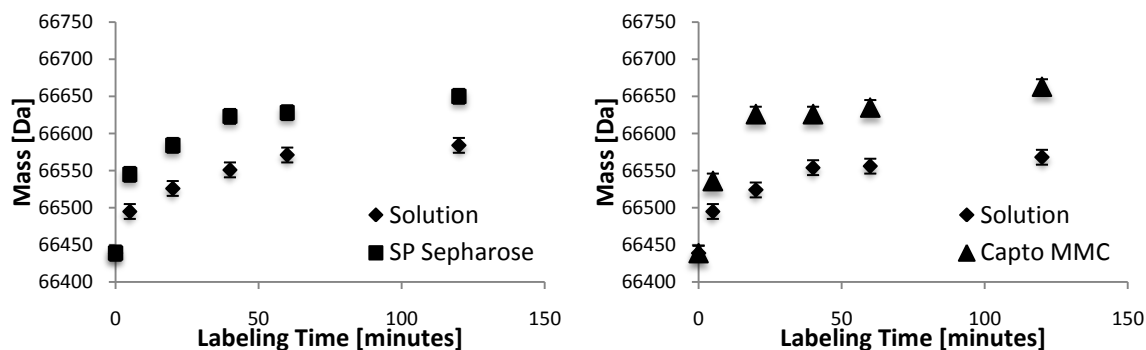
The unfolding of BSA on SP Sepharose FF and Capto MMC at pH 4.5 was further investigated with additional labeling times to determine if BSA unfolds at similar rates on both surfaces. The labeled masses for BSA in solution, adsorbed on SP Sepharose FF, and adsorbed on Capto MMC at pH 5.5 and various labeling times are shown in Figure 4.4. As expected, the mass of BSA increases with an increase in labeling time until it reaches a plateau at long labeling times.

The kinetic labeling profiles up to the plateau region of BSA on SP Sepharose FF and on Capto MMC are similar. This plateau corresponds to the “fast” and “medium” exchanging residues of BSA being deuterated. The residues of a protein can be grouped into fast, medium, and slow exchangers as described in 4.3.2. For the purposes of this study, the residues that have exchanged by the shortest labeling time, 5 minutes, were classified as fast exchangers. Conversely, those that have not exchanged by the longest



labeling time, 120 minutes, were classified as slow exchangers. The remaining residues from mass balance were then classified as medium exchangers.

On the two MMC surfaces, the mass of BSA plateaus at higher values than in solution: 66,650 Da for SP Sepharose FF, 66,663 Da for Capto MMC, and 66,568 in solution. This offset in mass on the two surfaces reflects a change in the number of fast and medium exchangers that populate BSA upon adsorption on the surface. The 82 Da offset in the plateau for BSA on SP Sepharose FF indicates that at least 82 residues that were slow exchangers in solution have become fast or medium exchangers when adsorbed on SP Sepharose FF. Likewise, the 95 Da offset of the plateau indicates that at least 95 residues that were slow exchangers in solution have become fast or medium exchangers when adsorbed on Capto MMC. The total distribution of fast, medium, and slow exchangers can theoretically be determined but a fully deuterated control (M100) mass is needed to account for back exchange. A M100 mass for intact BSA was not obtained due to unknown signal issues probably caused by the high concentrations of denaturant. However, M100 masses were successfully obtained for the digestion studies, so a more detailed analysis was possible for individual reporter peptides, as presented in Section 4.4.2.



**Figure 4.4.** Solvent exposure changes of BSA in solution (diamonds), adsorbed on SP Sepharose FF (left, squares), and adsorbed on Capto MMC (right, triangles) in 50 mM citrate at pH 4.5. Samples were labeled with deuterated buffer for 5 to 120 minutes. Error bars represent estimated 95% confidence intervals based on triplicate data points collected on SP Sepharose FF at pH 4.5 and 5 minutes labeling.

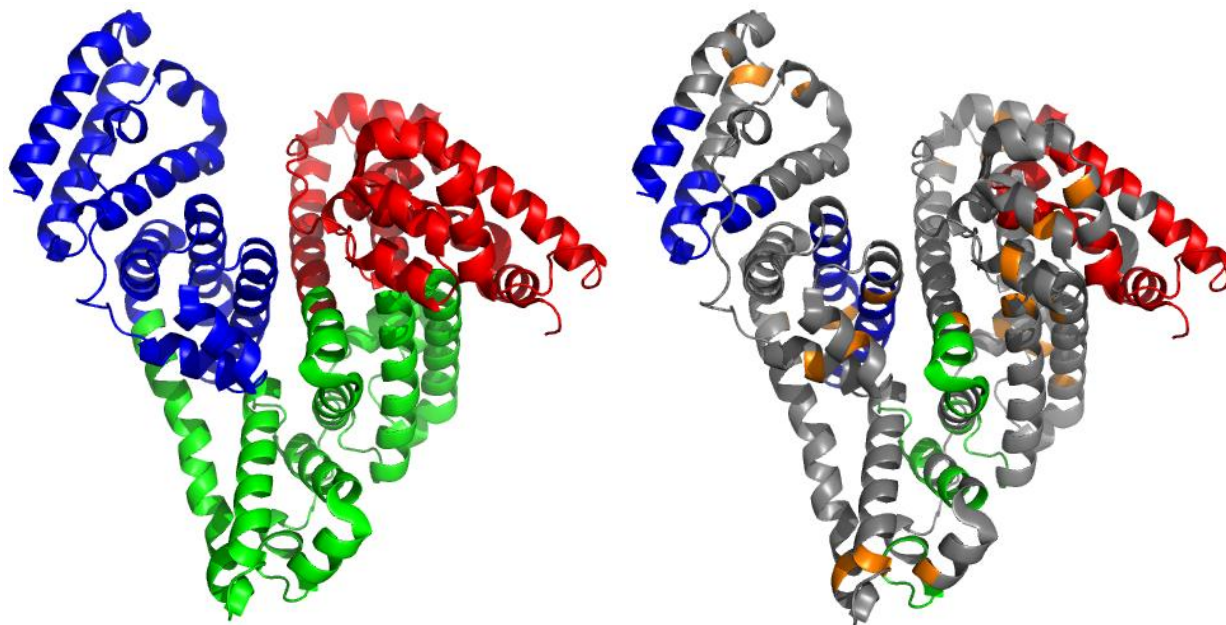
The largest increases in labeling occur before 5 minutes on both surfaces. BSA on Capto MMC has another large increase in labeling between 5 and 20 minutes, in effect reaching the plateau region by 20 minutes. The rate of increase, however, is smaller for BSA on SP Sepharose FF, as labeling still continues to increase after 20 minutes with the plateau region not reached until 40 minutes. Additional replicates, however, are needed to confirm this trend. However, the digestion studies presented in Section 4.4.2 support this trend. At 5 minutes labeling, BSA has a similar labeled mass on SP Sepharose FF, 66,545 Da, and on Capto Adhere, 66,536 Da. Presuming only fast exchangers get labeled within the first 5 minutes, the number of fast exchangers on both surfaces is the same within experimental uncertainty. The number of slow exchangers is also the same on both surfaces due to similarity in plateau mass. By mass balance, the number of medium exchangers is also the same. Therefore, what differs between the two surfaces is their effect on the unfolding rate of the medium exchangers with Capto MMC being faster. More

detailed analysis of the quantitative effect these surfaces have on unfolding rates is presented next.

#### **4.4.2. Digestion studies on SP Sepharose FF and Capto MMC at pH 4.5**

HXMS spectra were also collected with protease-digested samples to identify what regions of BSA are affected by adsorption on SP Sepharose and Capto MMC at pH 4.5. In addition, an analysis on how the surfaces affect the unfolding rate and distribution of fast, medium, and slow exchanging residues was performed. The mass spectra for reporter peptides under the different conditions were converted to  $D/N$  values using Equation (4.1) with experimentally determined  $M0$  and  $M100$ .

Although only 34% sequence coverage was obtained for the 585 residues of BSA, all three domains of BSA are represented. Figure 4.5 shows the native structure of BSA and its three domains (left). Also shown in Figure 4.5 are the location of the reporter peptides and cysteines of BSA (right). Many of the areas with missing sequence coverage are populated with cysteines involved in disulfide bridges which can impact proteolytic digestion as discussed in Section 4.2.4. Although reporter peptides are available for all three domains, the greatest coverage is obtained in Domain I, 40%, followed by Domain III, 37%. Domain II has the least coverage at 25%.



**Figure 4.5. Left: Native structure of BSA (PDB ID 3V03) highlighting Domain I (red), Domain II (green), and Domain III (blue). Right: Location of reporter peptides (highlighted according to domain) obtained in this study. Orange depicts cysteines of BSA. Grey depicts missing sequence coverage.**

The unfolding behavior of BSA reporter peptides on SP Sepharose FF and Capto MMC at pH 4.5 can be grouped into four classes. Class 1 consists of reporter peptides where the unfolding behavior mimics the behavior observed in the whole protein studies at pH 4.5: unfolding occurs to the same extent on both surfaces but occurs faster on Capto MMC. Class 2 consists of reporter peptides where unfolding occurs on both surfaces, but more on SP Sepharose FF. Class 3 consists of reporter peptides that unfold only on SP Sepharose FF. Finally, Class 4 consists of one reporter peptide where less labeling than in solution is observed on both surfaces. Table 4.1 lists the 4 classes, their characteristics, and the residues in each class. A more detailed description of the 4 classes is presented below.

**Table 4.1. Four classes of unfolding behavior of BSA on SP Sepharose FF and Capto MMC at pH 4.5**

	Unfolding?		Residues		
	<i>SP Sepharose</i>	<i>Capto MMC</i>	<i>Domain I</i>	<i>Domain II</i>	<i>Domain III</i>
<b>Class 1</b>	++	++	8-14 127-137	200-209 324-329 333-340	
<b>Class 2</b>	++	+	2-7 39-45	219-226 307-313	422-435 440-460 542-547
<b>Class 3</b>	+	/	18-32 49-70		529-541
<b>Class 4</b>	-	-			575-583

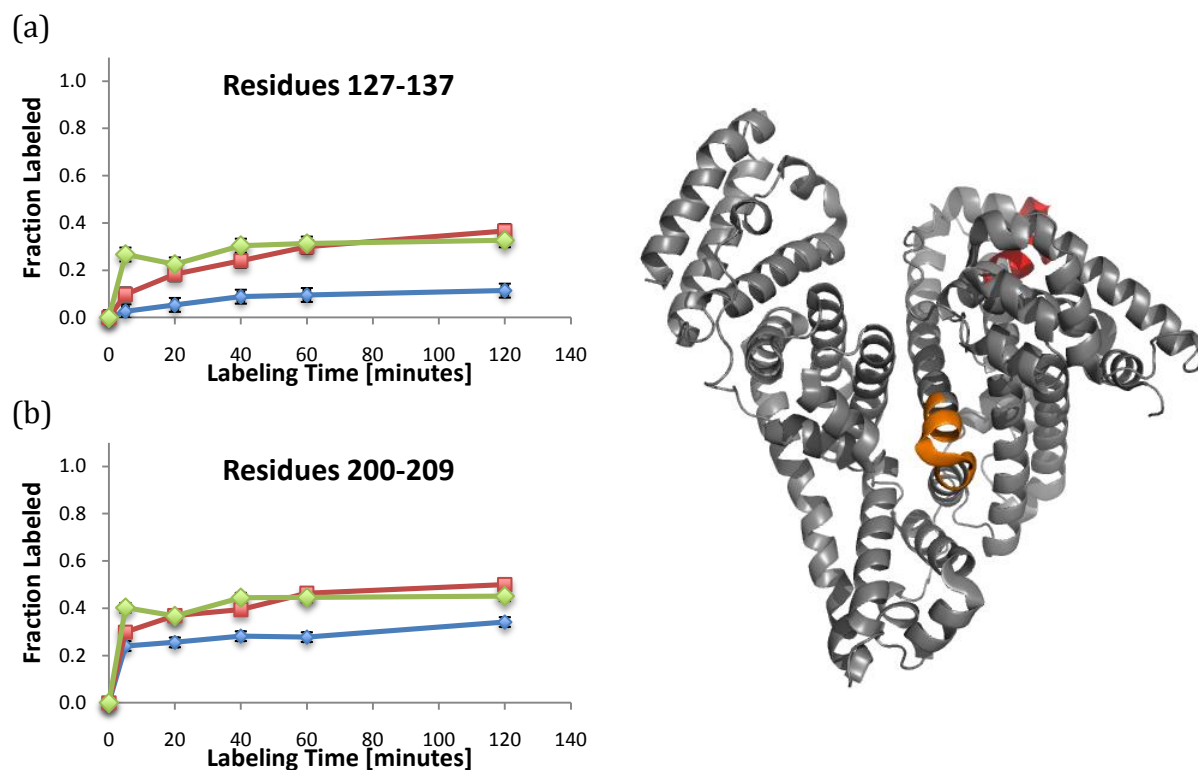
++ Unfolding

+ Some unfolding

/ No unfolding

- Solvent protection increase

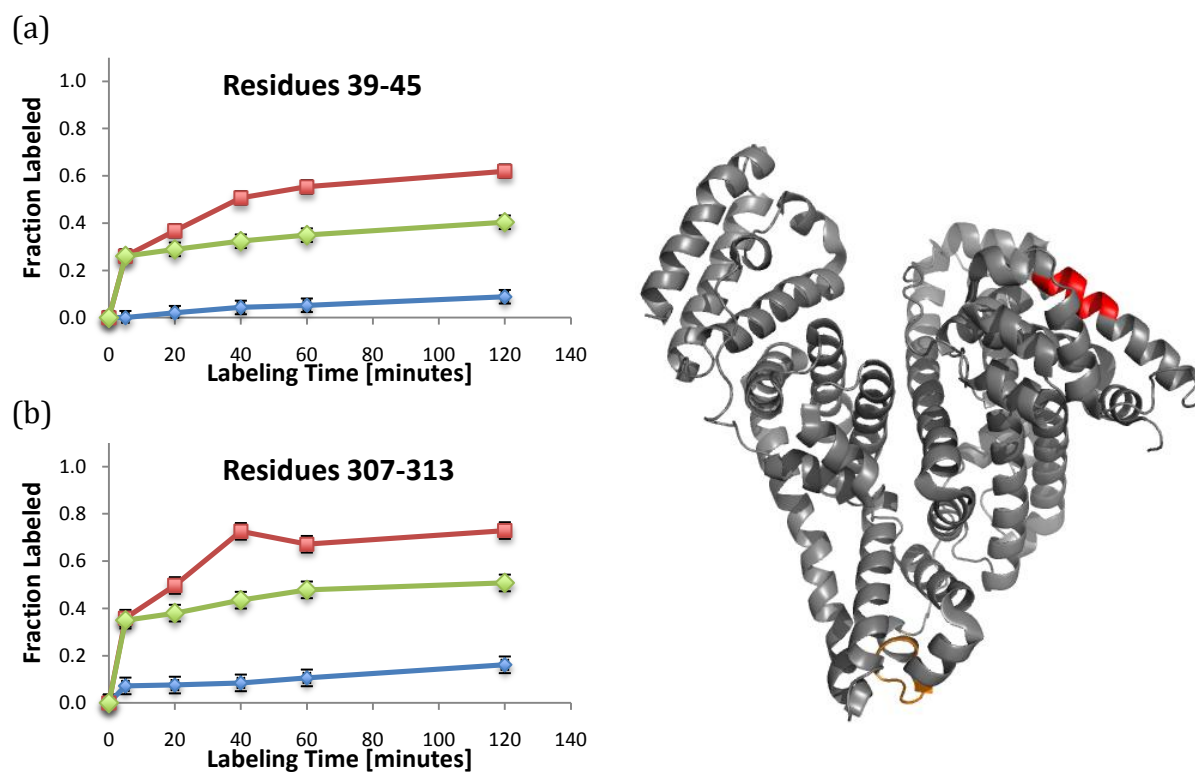
The unfolding behavior of Class 1 peptides is consistent with what is observed with the whole protein studies. The  $D/N$  values for two of these peptides in solution, on SP Sepharose FF, and on Capto MMC are shown in Figure 4.6a and 4.6b with their location on the native structure shown on the right. Residues 127-137 and 200-209 unfold on both surfaces with the increase in solvent exposure for each peptide plateauing at the same  $D/N$  value. However, with the whole protein studies, the rate at which the plateau is reached is not the same between the two surfaces. Unfolding occurs faster on Capto MMC for both peptides. Almost all of the labeling occurs within 5 minutes on Capto MMC whereas on SP Sepharose FF labeling continues after 5 minutes up to 120 minutes. In addition to these examples, Residues 8-14, 324-329, and 333-340 (data not shown) also unfold in this manner: equal extents of unfolding on both surfaces but faster on Capto MMC.



**Figure 4.6. Class 1 peptides in which BSA unfolds on SP Sepharose FF and Capto MMC to equal extents. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The  $D/N$  for Residues 127-137 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The  $D/N$  for Residues 200-209 is shown in panel (b) and their location on the native structure is shown in orange on the right.**

The unfolding behavior of Class 2 peptides deviates from the whole protein observations. In this case, unfolding still occurs on both surfaces, but more on SP Sepharose FF than Capto MMC. The  $D/N$  values for two of these peptides in solution, on SP Sepharose FF, and on Capto MMC are shown in Figure 4.7a and 4.7b with their location on the native structure shown on the right. Residues 39-45 and 307-313 unfold on both surfaces but to a greater extent on SP Sepharose FF. At 5 minutes labeling the  $D/N$  values are the same on both surfaces, but after 5 minutes further labeling occurs on SP Sepharose FF. This is in contrast to the Class 1 peptides and the whole protein studies. Therefore, not

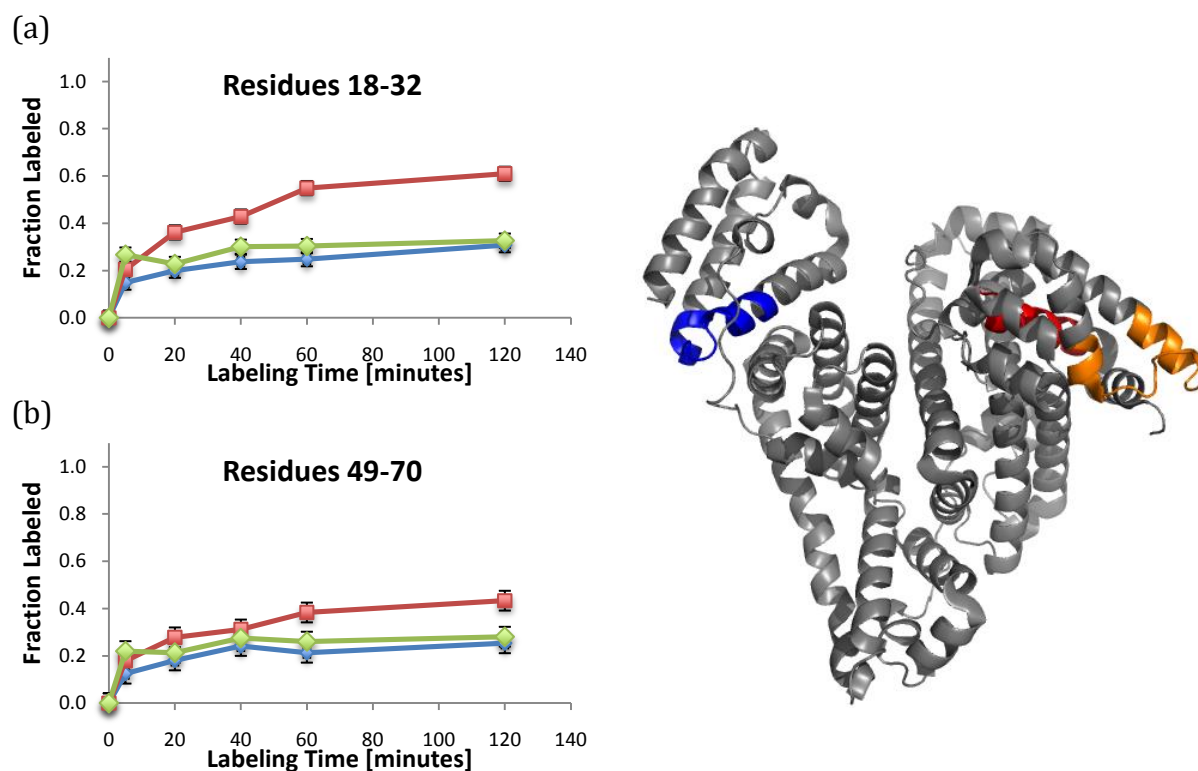
all regions of BSA unfold in the same manner on the surfaces. In addition to these examples, Residues 2-7, 219-226, 422-435, 440-460, and 542-547 (data not shown) also unfold in this manner.



**Figure 4.7.** Class 2 peptides in which unfolding occurs on SP Sepharose FF and Capto MMC but more on SP Sepharose FF. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The  $D/N$  for Residues 39-45 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The  $D/N$  for Residues 307-313 is shown in panel (b) and their location on the native structure is shown in orange on the right.

Class 3 peptides unfold on SP Sepharose FF but not on Capto MMC. The  $D/N$  values for two of these peptides in solution, on SP Sepharose FF, and on Capto MMC are shown in Figure 4.8a and 4.8b with their location on the native structure shown on the right. Residues 18-32 and 49-70 both are in this class. On SP Sepharose FF, the solvent exposure

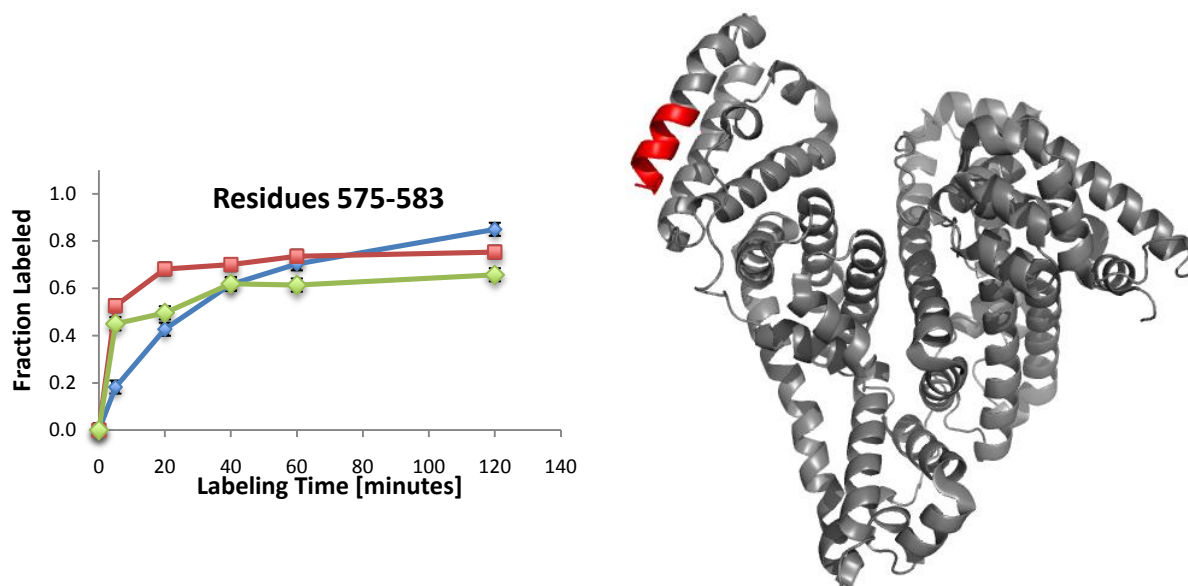
for both peptides increases at higher rates than it does in solution. On Capto MMC, more solvent exposure is observed in the first 5 minutes compared to solution, but after 5 minutes the solvent exposure in solution reaches the same  $D/N$  values as on Capto MMC. In these cases, a small amount of unfolding does occur at short times but not to the level seen in Class 1 and 2 peptides. Therefore, not all regions of BSA have their conformation affected by a large amount when adsorbed to Capto MMC. Residues 529-541 (data not shown) also show this behavior.



**Figure 4.8.** Class 3 peptides in which unfolding occurs only on SP Sepharose FF. Unfolding of different regions of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The  $D/N$  for Residues 18-32 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The  $D/N$  for Residues 49-70 is shown in panel (b) and their location on the native structure is shown in orange on the right. Also shown in blue on the native structure are the only other residues (529-541) that follow this behavior.



The 4<sup>th</sup> and final class of peptides consists of one peptide where at certain labeling times, more labeling is observed in solution than on the surfaces. The  $D/N$  values for Residues 575-583 in solution, on SP Sepharose FF, and on Capto MMC are shown in Figure 4.9 with their location on the native structure shown on the right. This region still unfolds on the surfaces with more solvent exposure than the solution at the shorter labeling times. Similar to the Class 2 peptides, more unfolding occurs on SP Sepharose FF than Capto MMC. At longer labeling times, however, the solvent exposure in solution surpasses that of both surfaces. This occurs on Capto MMC between 40 and 60 minutes labeling and on SP Sepharose FF between 60 and 120 minutes labeling. This decrease in solvent exposure on the surface relative to solution was also observed in the whole protein studies at pH 3.0 for both SP Sepharose FF and Capto MMC. As discussed previously, interactions between BSA molecules on the surface or the surface itself could prevent solvent access to residues. At present, it is not possible to discern the cause for the lower labeling.



**Figure 4.9.** Class 4 peptide in which less labeling is observed on the surfaces. Unfolding of Residues 575-583 of BSA in solution (blue, diamonds), on SP Sepharose FF (red, squares), and on Capto MMC (green, triangles) in 50mM Citrate, pH 4.5 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The  $D/N$  for Residues 575-583 is shown on the left and their location on the native structure (PDB ID 3V03) is shown in red on the right.

As first discussed in Section 4.4.3, the effect of the surfaces on the distribution of fast, medium, and slow exchanging residues can be determined from labeling data at different times. Table 4.2 lists the changes in the fractions of the three types of exchangers and the unfolding rate of the medium exchangers for the four classes of peptides. Values are averaged for all reporter peptides in each class that gave  $R^2$  values  $\geq 0.8$  when fitting the labeling data to Equation (4.7).

**Table 4.2. Change in fast, medium, and slow exchanging residues and unfolding rate from solution to surface for the 4 classes of peptides**

	SP Sepharose FF				Capto MMC			
	<i>Fast</i>	<i>Medium</i>	<i>Slow</i>	$k_{unf} [min^{-1}]$	<i>Fast</i>	<i>Medium</i>	<i>Slow</i>	$k_{unf} [min^{-1}]$
<b>Class 1</b>	+0.07	+0.14	-0.21	+0.008	+0.20	-0.04	-0.17	N/A
<b>Class 2</b>	+0.28	+0.27	-0.55	+0.030	+0.27	+0.06	-0.33	+0.010
<b>Class 3</b>	+0.06	+0.18	-0.24	+0.08	+0.11	-0.09	-0.03	N/A
<b>Class 4</b>	+0.35	-0.45	+0.10	+0.015	+0.27	-0.46	+0.19	+0.005

On both surfaces the fraction of slow exchangers is reduced for Class 1 peptides. The destabilizing effect of the two surfaces is sufficient that slow exchangers are converted to one of the faster exchanging groups. In the case of SP Sepharose FF, the largest increase in exchangers is in the medium category. On Capto MMC a large increase is observed in the fraction of fast exchangers while the fraction of medium exchangers is nearly constant. In the case of SP Sepharose FF, the unfolding rate of the medium exchangers is slightly increased relative to solution. On Capto MMC, a negligible fraction ( $<0.05$ ) of medium exchangers in Class 1 peptides are present therefore no unfolding rate was calculated.

For Class 2 peptides, more medium exchanging residues are present on SP Sepharose FF than on Capto MMC. Both surfaces increase fast exchangers by a similar amount but their medium and slow exchangers are affected differently. Adsorption on SP Sepharose FF reduces the number of slow exchangers more than it does on Capto MMC. As a result, there are more medium exchangers on SP Sepharose FF than on Capto MMC. For Class 2, the unfolding rate of the medium exchangers is affected by adsorption on SP Sepharose FF more than for the Class 1 peptides. With sufficient medium exchangers, the unfolding rate is found to increase for the Class 2 peptides on Capto MMC.

Class 3 peptides are characterized by how little Capto MMC affects the slow exchangers of BSA. Instead, Capto MMC only converts medium exchangers to fast exchangers, reflected by the greater increase in labeling within 5 minutes compared to solution (see Figure 4.8). Similar to the previous groups, SP Sepharose FF reduces the amount of slow exchangers and increases the other two exchangers, mostly medium exchangers. However, since Capto MMC does not affect the amount of slow exchangers, the  $D/N$  values plateau at the same value in solution and on Capto MMC. As with the Class 1 peptides, a negligible amount of medium exchangers are present on Capto MMC preventing an unfolding rate analysis.

Unlike the first three classes of peptides, the number of slow exchangers for the Class 4 peptide increases on the two surfaces. The trends of the two surfaces are similar: a large decrease in medium exchangers results in mostly an increase in fast exchangers, but also a lesser increase in slow exchangers. Although, the fraction of medium exchangers is reduced on the two surfaces, the unfolding rate of the remaining medium exchangers still increases.

In total, the behavior of most digested regions show unfolding behavior that deviates from the whole protein studies. This does not necessarily conflict with the results of the whole protein study where equal labeling was observed on both surfaces after 40 minutes (see Figure 4.4). The missing sequence coverage of 66% prevents making more complete estimates of the distribution of the 4 different groups. More regions may or may not fall into the 1<sup>st</sup> group than into the others. In addition, it is possible a 5<sup>th</sup> group of peptides exists where more unfolding occurs on Capto MMC than on SP Sepharose FF.

Peptides in this group could compensate for the lower labeling observed in Groups 2-4 on Capto MMC.

In general, the labeling patterns of the different classes are not domain specific. This suggests that residue/domain-specific unfolding patterns of BSA on IEC/MMC may not necessarily be inferred from solution unfolding patterns as was done successfully for the protein-HIC systems in Chapter 2. Some general patterns do exist between solution and surfaced unfolding at low pH when labeling data at pH 4.0 (Section 4.4.3 below) is considered. A more detailed discussion will be presented in Section 4.4.4.

#### **4.4.3. Digestion studies on SP Sepharose FF at pH 4.0**

HXMS spectra were also collected with protease-digested samples for BSA adsorbed on SP Sepharose FF at pH 4.0 to identify any regions that unfold due to adsorption. Although no unfolding from the whole protein study was detected at pH 4.0 for these surfaces, a digestion study was nevertheless done to determine local unfolding at this pH or whether more complex behavior may be occurring. The mass spectra for reporter peptides were converted to  $D/N$  values using Equation (4.1).

The unfolding behavior of BSA reporter peptides on SP Sepharose FF at pH 4.0 can be grouped into three categories. The 1<sup>st</sup> group consists of reporter peptides where the labeling patterns mimic the behavior observed in the whole protein studies at pH 4.0: no unfolding occurs on the surface. The 2<sup>nd</sup> group consists of reporter peptides where unfolding does occur on the surface. Finally, the 3<sup>rd</sup> group consists of reporter peptide where less labeling is observed for adsorbed protein.

**Table 4.3. Three classes of unfolding behavior of BSA on SP Sepharose FF at pH 4.0**

	Unfolding?		Residues	
	<i>SP Sepharose FF</i>	<i>Domain I</i>	<i>Domain II</i>	<i>Domain III</i>
<b>Class A</b>	/	18-32 127-137		529-547
<b>Class B</b>	+	2-7 39-45	219-226 307-313	440-460
<b>Class C</b>	-	8-14	200-209 333-340	575-583

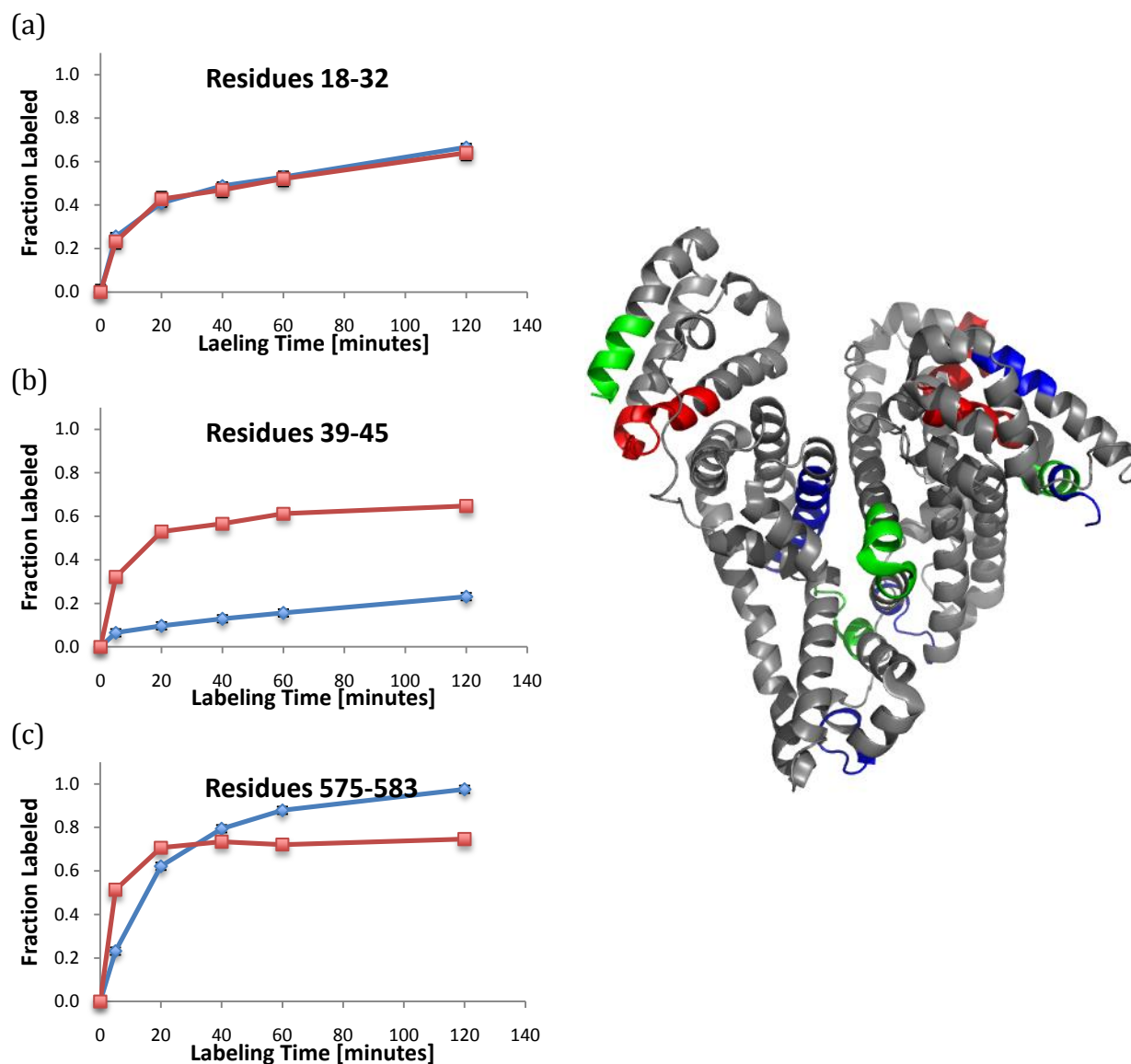
/ No unfolding

+ Unfolding

- Solvent protection increase

For brevity, only one example from each class is presented in detail. The *D/N* values for these three example peptides in solution, and on SP Sepharose FF are shown in Figure 4.10a, 4.10b, and 4.10c with the location of all the reporter peptides in each group on the native structure shown on the right. Residues 18-32 are representative of peptides from Class A where the labeling profile is the same between the solution and surface. The absence of unfolding on the surface for these residues along with Residues 127-137 and 529-543 is consistent with what was observed in the whole protein study. Residues 39-45 are representative of peptides from Class B where more labeling occurs on the surface compared to solution. These residues along with Residues 2-7, 219-226, 307-313, and 440-460 indicate that BSA does unfold at pH 4.0; an observation not detected in the whole protein studies. Finally, Residues 575-583 are representative of peptides from Class C where less labeling is observed on the surface. Initially at shorter labeling times (5 and 20 minutes), more labeling is observed on the surface but between 40 and 60 minutes labeling a crossover occurs, in which more labeling is observed in solution. Residues 8-14, 200-209,

and 333-340 are also observed to have less labeling on the surface but differ with respect to the labeling time in which that is first observed: after 10, 60, and 40xminutes, respectively. In addition, these residues differ from Residues 575-583 in that labeling is the same in solution and on the surface before those labeling times.



**Figure 4.10.** Labeling patterns of the three different classes of BSA peptides in solution (blue, diamonds) and on SP Sepharose FF (red, squares) in 50mM Citrate, pH 4.0 buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. The  $D/N$  for Residues 18-32 is shown in panel (a) and their location on the native structure (PDB ID 3V03) along with other Class A peptides are shown in red on the right. The  $D/N$  for Residues 39-45 is shown in panel (b) and their location on the native structure along with other Class B peptides are shown in blue on the right. The  $D/N$  for Residues 575-583 is shown in panel (c) and their location on the native structure along with other Class C peptides are shown in green on the right.

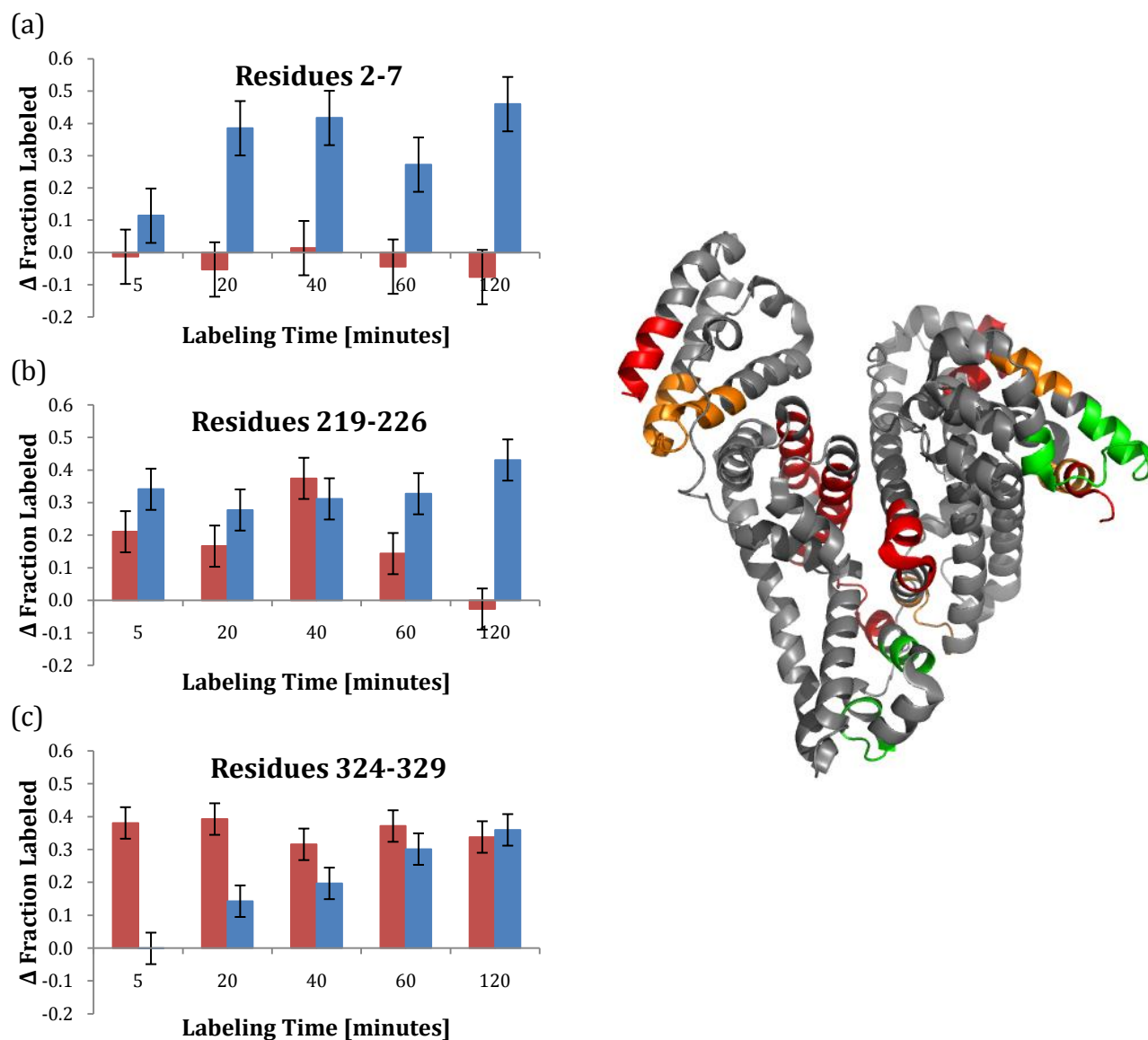
More variability in the labeling patterns is present at pH 4.0 rather than 4.5 for BSA adsorbed on SP Sepharose FF. At pH 4.5, all regions with sequence coverage (except 575-



583) are more labeled on the surface at labeling times greater than 5 minutes. At pH 4.0, however, this is only true for the 1<sup>st</sup> group of peptides. Two possible explanations exist for this observation. First, the destabilizing effect of the surface may not be as large at pH 4.0 as it is at pH 4.5. Therefore, fewer regions would be unfolded on the surface at the lower pH. The second possibility is that the destabilizing effect of the surface is the same at both pHs but in solution BSA unfolds more as pH is lowered from 4.5 to 4.0. A comparison was done of the extents of labeling at 4.5 and 4.0 to identify the more likely possibility.

In general, as pH is lowered from 4.5 to 4.0 labeling increases in solution but not on the surface, although two other trends also exist. Differences in  $D/N$  values between pH 4.0 and 4.5 for solution and surface labeling are shown in Figure 4.11. An example of each of the three observed trends is presented in panels (a-c). The location of each example is shown on the native structure at the right.

Residues 2-7, shown in panel (a), are representative of the regions of BSA where lowering pH increases labeling in solution but not on the surface. In addition, Residues 18-32, 127-137, 200-209, 330-340, 422-435, 440-460, and 575-583 also demonstrate this behavior. All these regions are partially unfolded on the surface at pH 4.5. If no conformational changes occur, lowering pH should, if anything, decrease labeling due to lower intrinsic exchange rates (Berger et al. 1959). The opposite, however, is observed for these regions in solution signifying a conformational change is occurring where these regions are more solvent-exposed. Again, this is consistent with BSA reportedly denaturing in this pH range (Williams & Foster 1960; Lin & Koenig 1976). This observation will be further discussed in Section 4.4.4.



**Figure 4.11.** Difference in  $D/N$  from pH 4.0 to 4.5 for different regions of BSA in solution (blue) and on SP Sepharose FF (red) in 50mM Citrate buffer. Samples were labeled with deuterated buffer for 10 to 120 minutes. Positive values indicate more labeling at pH 4.0 while negative values indicate less labeling. The difference in  $D/N$  for Residues 2-7 is shown in panel (a) and their location on the native structure (PDB ID 3V03) is shown in red on the right. The difference in  $D/N$  for Residues 219-226 is shown in panel (b) and their location on the native structure is shown in orange on the right. The difference in  $D/N$  for Residues 324-329 is shown in panel (c) and their location on the native structure is shown in green on the right. All other residues that follow 1 of the 3 patterns are also shown on the native structure in their representative color.

Labeling also increases on the surface as pH decreases from 4.5 to 4.0 in two other ways. Residues 219-226, shown in panel (b), are representative of regions where solvent exposure increases but at equal or less than in solution. Residues 324-329, shown in panel (c), are representative of regions where solvent exposure also increases as the pH is lowered but equal to or more than in solution. These two examples indicate that BSA does unfold more on the surface as pH is lowered, but fewer regions are involved. Eight regions were found to exhibit the behavior seen in panel (a) but only 5 (Residues 8-14, 39-45, 219-226, and 529-547) and 3 (Residues 49-70, 307-313, and 324-329) exhibit the behaviors seen in panels (b) and (c), respectively.

Similarities exist between the reporter peptides that populate the classes at pH 4.5 (Class 1 to 4) and at pH 4.0 (Classes A to C). The first similarity occurs between Class 1 and Class C reporter peptides. Three reporter peptides (Residues 8-14, 200-209, and 333-340) that unfold to equal extents on SP Sepharose FF and Capto MMC at pH 4.5 are more solvent-protected on SP Sepharose FF at pH 4.0. An even greater overlap is observed between Class 2 and Class B residues. Five reporter peptides (2-7, 39-45, 219-226, 307-313, and 440-460) that unfold more on SP Sepharose FF than on Capto MMC at pH 4.5 are at the only groups that unfold on SP Sepharose FF at pH 4.0. Further, 2 of the 3 reporter peptides from Class 3 (Residues 18-32 and 529-541) where unfolding only occurs on SP Sepharose FF fall into Class A at pH 4.0: no difference in labeling between the solution and surface. Finally, one reporter peptide, 575-583, from Class 4 and C, is the only one that is more solvent-protected at both pHs. These comparisons show that even though the labeling on the surface relative to solution may change as pH changes, the specific changes (e.g. more, the

same, or less labeling between surface and solution) are conserved amongst the different groups of peptides.

#### **4.4.4. Relating BSA acid-induced denaturation in solution and on the surface**

Proteins are known to unfold under acidic conditions. These proteins also do not necessarily fully unfold but instead can populate partially unfolded states commonly referred to as molten globules. These structures are characterized by highly dynamic states with native-like secondary structure and protein core that admits solvent but lacks close-packing (Ptitsyn 1995). BSA denatures under acidic conditions in the pH range of 4.5 to 3.0 with structure similarities to that of a molten globule (Sadler & Tucker 1993; Williams & Foster 1960; Lin & Koenig 1976). Sadler and Tucker used NMR to identify the “most dramatic changes” occurring over the pH range of 4.5 to 4.0. Consistent with the classical definition of a molten globule, more residues were observed to be solvent accessible with minimal loss of secondary structure (from 51 to 44% helical content as measured by optical rotation at 233nm) during the pH 4.5 to 4.0 transition (Sadler & Tucker 1993). Far-UV circular dichroism giving secondary structure and ultrasound giving adiabatic compressibility have confirmed this transition. These structural changes continue to evolve at least to pH 2.0 (El Kadi et al. 2006).

Although many studies have been done to characterize and understand the acid denaturation of BSA, no detailed structures of acid-denatured BSA are available. While studies have been done with bovine  $\alpha$ -lactalbumin and  $\beta$ -lactoglobulin to provide detailed structural characteristics of the respective proteins' molten globule (Alexandrescu et al.

1992; Baum et al. 1989; V Forge et al. 1999; Molinari et al. 1996), the BSA molten globule has not been so fully characterized. Relevant to the present work, it has been shown that the structures of bovine  $\alpha$ -lactalbumin and  $\beta$ -lactoglobulin adsorbed on HIC surfaces closely resemble that of the molten globule (Gospodarek et al. 2011).

In the case of BSA adsorbed on SP Sepharose FF, the transition of BSA into its acid-induced molten globule may happen at higher pH than in solution. At pH 4.5 all regions of BSA (with coverage) have an increase in solvent exposure on the surface relative to solution. As pH decreases to 4.0, BSA in solution begins its acid-induced unfolding where at this point the solvent exposure in solution reaches the levels observed for BSA on SP Sepharose FF at pH 4.5. BSA unfolding on SP Sepharose FF however, also continues as even more solvent exposure is observed for several regions at pH 4.0. As with BSA in solution continuing to denature as pH is lowered to 2.0, a similar phenomenon may be occurring on the surface where denaturation is not complete until a lower pH is reached. From the whole protein study (see Figure 4.1) it was observed that as pH decreases from 4.0 to 3.0, labeling still increases for both BSA in solution and on SP Sepharose FF. However the effect of pH is now greater in solution than on the surface in this range. In addition, the solution labeled mass becomes greater than the surface one over this range. Together, this suggests that acid-denaturation to the molten globule on the surface is becoming complete while in solution it is still proceeding. Another possibility is aggregation on the surface may begin at the lower pHs leading to a reduction in the solvent exposure increase. Additional HXMS and SEC experiments at lower pHs would provide more detailed information about the structural characteristics of both BSA in solution on the surface

allowing a better understanding of the relationship between acid-induced unfolding in solution and on the surface.

#### 4.4.5. Digestion studies on Capto MMC at pH 5.5 with varying ammonium sulfate concentrations

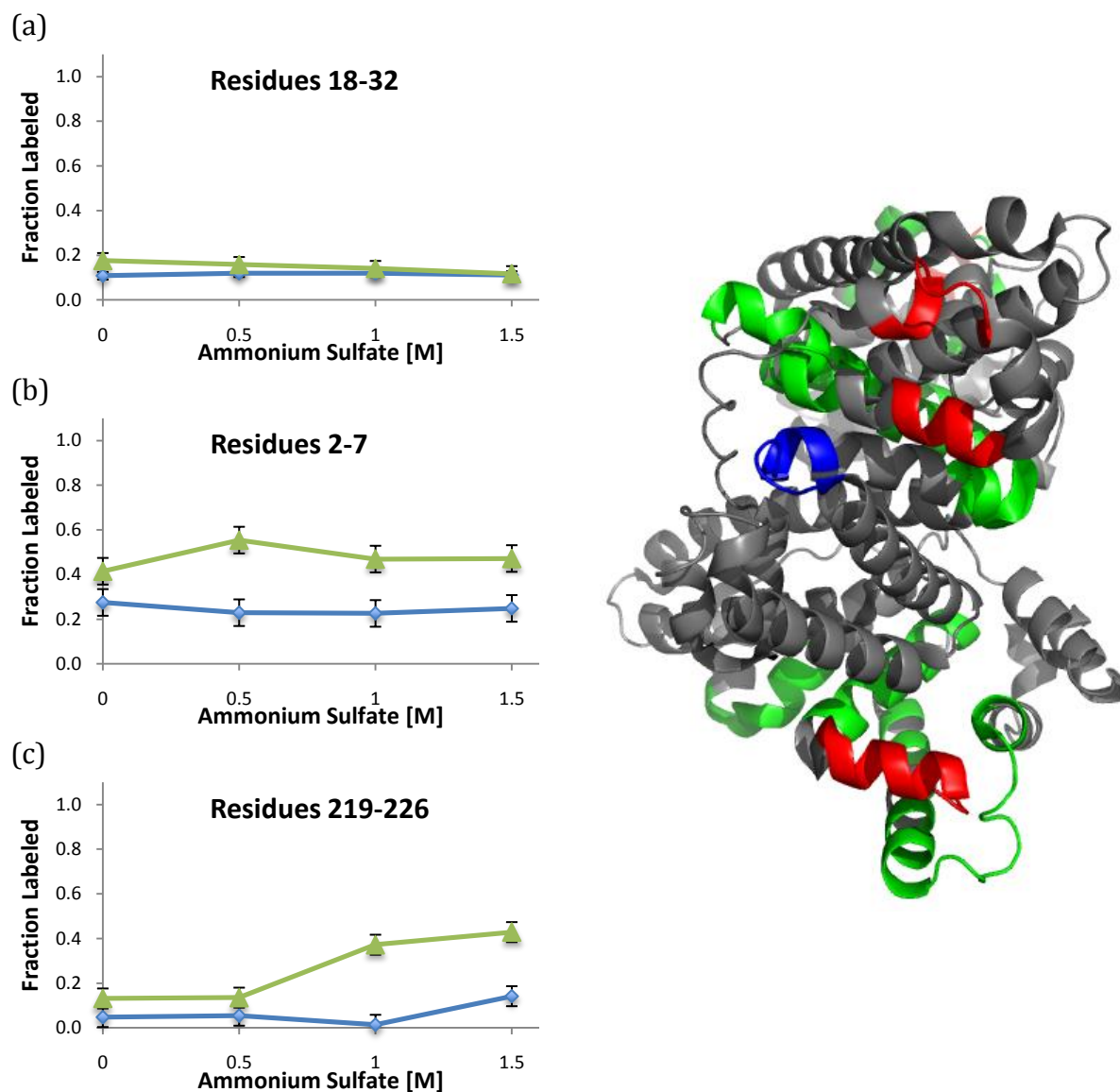
HXMS spectra were also collected with protease-digested samples for BSA adsorbed on Capto MMC at pH 5.5 under varying ammonium sulfate concentrations to identify if any regions unfold. Although no unfolding from the whole protein study was detected under these conditions, digestions studies were nevertheless done to confirm if any regions unfold at this pH or is there cancellation of regions of more exposure by regions of less exposure. The mass spectra for reporter peptides were converted to  $D/N$  values using Equation (4.1).

**Table 4.4. Three classes of unfolding behavior of BSA on Capto MMC at pH 5.5**

			<b>Residues</b>		
	<i>Unfolding?</i>	<i>Ammonium Sulfate effect?</i>	<i>Domain I</i>	<i>Domain II</i>	<i>Domain III</i>
<b>General</b>	<b>No</b>	No	18-32	200-209	422-435
			39-45		440-460
			49-70		529-547
			127-137		
<b>Exception 1</b>	<b>Yes</b>	No	2-14	307-313 323-329	575-583
<b>Exception 2</b>	<b>Yes</b>	Yes		219-226	

In general, no increase in labeling on the surface nor effect of ammonium sulfate concentration on solution or surface labeling was observed. Two exceptions to the general labeling pattern do exist where more labeling is observed on the surface and increasing ammonium sulfate concentration increases labeling. The residues that comprise each

labeling pattern are shown in Table 4.4. The  $D/N$  values in solution and on Capto MMC for the general case and the two exceptions are shown in Figure 4.12 (a-c). Their location on the native structure is shown on the right. Residues 18-32, shown in panel (a), are representative of what is observed for most regions of BSA: adsorption on Capto MMC does not affect the labeling compared to solution. No detectable unfolding on the surface relative to the solution is observed in this case. Residues 2-7, shown in panel (b), are representative of the few regions of BSA that are more labeled on the surface but labeling is still not affected by ammonium sulfate concentration. Along with Residues 2-7, Residues: 8-14, 307-313, 324-329, and 575-583 show similar behavior. Here, the solvent exposure is greater on the surface at 0.5, 1.0, and 1.5M ammonium sulfate indicating some local unfolding does occur on the surface. Residues 219-226, shown in panel (c), are the one exception where labeling on the surface increases as ammonium sulfate concentration increases. At 0 and 0.5M ammonium sulfate no statistically significant differences are observed between the solution and surface. At 1.0 and 1.5M ammonium sulfate, however, more labeling is observed on the surface. In solution only a small statistically significant increase is observed from 0.01 at 1.0M to 0.14 at 1.5M. In general, for most residues within the sequence coverage obtained, the similar labeling between solution and surface with no effect of ammonium sulfate is consistent with what was observed in the whole protein studies (see Figure 4.3).



**Figure 4.12.** Three labeling patterns observed for different regions of BSA in solution (blue, diamonds), and on Capto MMC (green, triangles) in 50mM phosphate, pH 5.5 buffer at varying ammonium sulfate concentrations. Samples were labeled with deuterated buffer for 20 minutes. Panel (a) shows the  $D/N$  for an example, Residues 18-32, of the general pattern where no unfolding on surface relative to solution and no effect of ammonium sulfate is observed. Panel (b) shows the  $D/N$  for an example, Residues 2-7, of the first exceptional pattern in which unfolding is observed on the surface but without any effect from ammonium sulfate. Panel (c) shows the  $D/N$  for the one group of residues, Residues 219-226, of the second exceptional pattern in which both unfolding on the surface and ammonium sulfate effects are observed. The location of the groups of residues in the general pattern, and the first and second exceptional labeling pattern are shown on the native structure (PDB ID 3V03) on the right in green, red, and blue, respectively.



An increase in unfolding on the surface from increasing ammonium sulfate concentration has been observed previously (Jungbauer et al. 2005; Xiao et al. 2006). It has been hypothesized that a relative difference in salt dependence on adsorption for native and unfolded molecules is responsible for this (Xiao et al. 2006). A similar phenomenon may be occurring here as molecules with this region unfolded may preferentially bind to the hydrophobic moieties of Cpto MMC. It is unclear why this is the only region that shows this type of behavior. As only 34% sequence coverage was obtained, it is possible other regions also show this behavior although the whole protein studies suggest this is not likely the case.

## ***4.5. Conclusions***

In this study, we showed that BSA unfolds on surfaces with cation exchange characteristics, SP Sepharose FF and Capto MMC, and that the unfolding is dependent on pH. This suggests that the type and magnitude of electrostatic interactions are important in unfolding. At pH 4.5, several regions of BSA unfold to an equal extent on both SP Sepharose FF and Capto MMC. There are, however, some regions that unfold to a lesser extent or not at all on Capto MMC, suggesting the type of cation exchanging ligand is important for unfolding.

Comparisons of the unfolding patterns in solution and on SP Sepharose FF at pH 4.5 and 4.0 suggest the acid-induced denaturation of BSA observed in solution initiates at milder acidic conditions when adsorbed on the surface. In solution, the solvent exposure levels observed for BSA on Sepharose at pH 4.5 are not reached until the pH is lowered to 4.0. Further, several regions of BSA continue to gain solvent exposure on the surface at pH 4.0 indicating the denaturation on the surface may not be complete. The whole protein studies reinforce this as solvent exposure continues to increase to pH 3.0 but now at a rate less than in solution. The structure of BSA on the surface may resemble the acid-induced molten globule but additional HXMS measurements are needed at lower pH to confirm this. Although domain-specific labeling patterns were not observed, the different reporter peptides could be grouped into four (pH 4.5) and three (pH 4.0) distinct classes of labeling behavior based on unfolding or solvent protection increase occurring on SP Sepharose FF and/or Capto MMC. It was observed that reporter peptides grouped into one class at pH 4.5 also tended to be grouped together again at pH 4.0. This suggests similarities exist

between how peptides unfold on BSA despite not being located in the same region or domain.

It was shown in Chapter 2 that BSA unfolds to a large extent ( $D/N$  values  $> 0.8$ ) on the HIC surfaces studied, more so than the other protein systems examined. BSA also underwent surface-induced aggregation at elevated temperatures. Generally, unfolding on IEC surfaces is not observed, as the denaturing capability of these surfaces is believed to be not as strong as those of HIC surfaces. This suggests that only proteins with very low folding stability and/or aggregation propensities may unfold in IEC although a larger array of proteins with varying folding stabilities would need to be examined to confirm this.

Finally, a few regions of BSA unfolded on Capto MMC under conditions where electrostatic interactions were minimized by added salt. Interestingly, the level of unfolding is lower than what was observed on Capto MMC at pH 4.5 where electrostatic interactions are expected to dominate. Nevertheless, to our knowledge this is the first report of the ability of MMC surfaces to unfold a protein. This study has provided further evidence that protein unfolding can occur on surfaces with ion-exchanging characteristics. These findings should help in understanding unusual peak behavior in IEC and MMC processes.

## 5. Conclusions and Future Work

### 5.1. Conclusions

The objective of this thesis was to help develop models/relationships to predict protein unfolding on chromatographic surfaces. This first required identifying the effect adsorption has on the structure of different regions of a protein. Although sequence coverage was limited to less than 50% for all proteins in this thesis, enough detailed information about the protein structure was obtained to elucidate information about unfolded structures in solution. Second, HXMS was demonstrated to be effective in characterizing unexpected behavior such as self-association in the solution phase. Thus, this work reinforces the value of detailed spectroscopic techniques such as HXMS for maximum characterization of the variety of protein behaviors in different situations. Finally, more evidence and information on protein unfolding on IEC and MMC surfaces was gathered which should be useful in evaluating unusual peak behavior in these processes.

In Chapter 2, patterns between the structures of proteins partially unfolded in solution and on hydrophobic surfaces were observed. The regions that are unfolded in solution intermediates are more prone to unfolding on HIC surfaces while the more stable regions in solution are more likely to retain native structure on a surface. Further, it was found that ligand type is probably important in determining the extent and mechanism of protein unfolding. Superposition of labeling patterns at different temperatures works quite well for comparing HIC surfaces with alkyl ligands but significant differences are found in some protein regions when comparing alkyl and aromatic ligands. Finally, it was demonstrated that unfolding rates and free energies of unfolding can be determined for

proteins adsorbed on a surface and they are different than in solution. This information should be helpful in quantifying the specific effects an HIC surface has on protein stability.

In Chapter 3, HXMS increased understanding of the self-association of a therapeutic protein by identifying the residues involved in the association site and a conformational change related to the association. Self-association was shown to primarily involve the modified region of the protein with all large solvent protection increases occurring there. A conformational change, however, was identified in the unmodified region in an area that has a known binding pocket. Self-association, therefore, may involve a more complicated mechanism in which insertion of the modified region into this binding pocket drives the association. Further, it was also shown that under oligomer conditions more residues become involved in the association site and that the interactions between the residues on different molecules are stronger. This information along with ITC and SEC measurements provided a comprehensive description of the self-association which should be useful for future investigations of this type of behavior.

In Chapter 4, potential protein unfolding in IEC and MMC was investigated with HXMS. BSA unfolding on SP Sepharose FF and Capto MMC had similar trends with pH but more subtle differences existed and were identified with digestion studies. It was shown that BSA unfolds more on SP Sepharose FF than on Capto MMC with the effect varying for different regions of BSA. It was also shown that solvent exposure levels within BSA's acid denaturation range are reached at milder pH conditions on SP Sepharose FF than in solution but solution labeling ultimately exceeds surface labeling by pH 3.0. This indicates adsorption to the surface may initiate acid-induced denaturation outside the normal range required in solution but that the final unfolded solvent exposure levels are less than in

solution. Finally, it was shown that BSA unfolds on Capto MMC under high salt conditions where electrostatic effects are minimized and hydrophobic interactions can become important, although the unfolding is much less than what was observed under low salt conditions where electrostatic interactions dominate.

## ***5.2. Future Work***

The results of the present studies should have moved the field of protein unfolding in chromatography closer to being able to predict if, from a thermodynamic perspective, and how, in terms of which domains/residues, a protein will unfold on specific surfaces. However, much work remains to achieve reliable predictions. The results suggest potentially productive directions for the next steps.

The major focus of this work was to determine protein structure changes as system conditions change its interactions with a surface, such as an array of surfaces that of varying hydrophobicity for HIC/MMC media and of varying charge for IEC/MMC media. In Chapter 2 three Tosoh Inc. resins with ligands (Phenyl, Butyl, and Hexyl) of significant hydrophobicities were used. Examining other available ligand types, e.g., PPG and Ether, would provide additional information about protein unfolding over a greater range of denaturing conditions.

It is also recommended from the findings of Chapter 2 to explore other measures of hydrophobicity to accurately quantify the degree of hydrophobic interaction between the protein and the surface. In this study, a relatively simple measure of hydrophobicity, based on retention times and dynamic binding capacities, was used. This analysis may not accurately reflect the denaturing effect since here more unfolding was observed on Butyl 650M than on Hexyl 650C even though the latter was characterized as more hydrophobic. Studies have been done by the Steve Cramer (Rensselaer Polytechnic Institute, Troy, New York) and Alois Jungbauer (University of Natural Resources and Life Sciences, Vienna, Austria) groups to provide more quantitative and accurate measurements of

hydrophobicity (Machold et al. 2002; Xia et al. 2004; Ladiwala et al. 2006). Using the metrics may provide more reliable measures of hydrophobicity.

Although the guanidine denaturation studies in Chapter 2 demonstrated that the effect of a surface on unfolding rates and free energies of unfolding could be determined, a larger database is needed to provide more complete information from this technique. Most importantly, more measurements and replicate experiments in the transition region of the denaturation curves are needed to reduce uncertainties obtained in this study. The lengthy analysis time (60-70 minutes) of the HPLC-MS system needed for the present studies prevented obtaining more data. More advanced UPLC and MS systems are available that should significantly reduce analysis time.

An important direction arising from the guanidine denaturation studies is to determine values of  $\Delta\Delta G_{unf}$  and  $\Delta k_{unf}$  from Equations (2.7) to (2.15) for a larger set of proteins on a single resin type. The goal would be to determine if  $\Delta\Delta G_{unf}$  and  $\Delta k_{unf}$  depends on resin but independent of the protein. If so, then  $\Delta\Delta G_{unf}$  and  $\Delta k_{unf}$  can be determined for any resin used in HIC and together with solution stability data, allowing prediction of conditions for protein unfolding upon adsorption.

The studies from Chapters 2 and 4 also do not distinguish between reversible and irreversible unfolding either on the surface or upon desorption since labeling was done while the protein was adsorbed. The state of the protein is captured at a particular moment and the method cannot distinguish if refolding occurs. For practical purposes, the degree and mechanism of unfolding may not matter if unfolding is reversible. But full knowledge of the behavior should still be useful.



The data collected with HXMS from Chapters 2 and 4 can be used for comparisons with protein stability calculations obtained via computer simulation. Presently very few models exist that include surfaces or ligands in the molecular modeling. The observed deuterium uptake data can be readily converted to protection factors, which can be straightforwardly compared with computer simulation protection factors. Such comparisons would guide validation of new molecular interaction models.

There is also a need to continue the investigation of protein unfolding on IEC and MMC surfaces from Chapter 4. The protein chosen in Chapter 4, BSA, is not very stable in solution and on surfaces. Since it has recently been discovered that antibodies may unfold and aggregate on IEC surfaces, the BSA behavior may be relevant. New materials for chromatography are being continuously developed in order to maximize protein binding to reach high loading. It is therefore imperative to continue collecting data over a wide variety of conditions to learn how different surfaces and operational variables affect unfolding. Proteins with known acid-induced denaturation pathways may be of particular value to compare protein denaturation in solution and on IEC surfaces, especially as pH changes.

Initially, it is recommended to rapidly screen for unusual chromatograms and elution profiles suggestive of unfolding. High-throughput robotic systems with liquid handling devices and RoboColumns™ are ideal for such screening since many data can be generated quickly. It is recommended to first focus on varying pH and ionic strength followed by incorporating commonly used mobile phase modifiers such as arginine or ethylene glycol.

Then, cases with unusual elution profiles could be analyzed with techniques such as SEC and stability assays to determine if unfolding and aggregation may occur. The case of the therapeutic protein that self-associates but does not unfold or aggregate from Chapter 3 suggests that the self-association properties of a protein should also be thoroughly investigated so as no false red flags are raised with unexpected elution behavior is observed. Finally, HXMS could be incorporated to determine the structural characteristics of apparently unfolded or aggregated proteins on HIC, IEC, and MMC resins and compared with structures in solution and on HIC surfaces.

## 6. References

1. Ahmad, B. & Khan, R.H., 2005. Urea induced unfolding of F isomer of human serum albumin: A case study using multiple probes. *Archives of Biochemistry and Biophysics*, 437(2), pp.159–167.
2. Alexandrescu, A.T. et al., 1992. <sup>1</sup>H-NMR assignments and local environments of aromatic residues in bovine, human and guinea pig variants of alpha-lactalbumin. *European Journal of Biochemistry*, 210(3), pp.699–709.
3. Antosiewicz, J., McCammon, J. a & Gilson, M.K., 1996. The determinants of pKas in proteins. *Biochemistry*, 35(24), pp.7819–33.
4. Arakawa, T. et al., 2007. The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography. *Protein expression and purification*, 54(1), pp.110–6.
5. Aune, K.C. & Tanford, C., 1968. Thermodynamics of the Denaturation of Lysozyme by Guanidine Hydrochloride II: Dependence on Denaturant Concentration at 25C. *Biochemistry*, 8(11), pp.4586–4590.
6. Azevedo, A.M. et al., 2009. Downstream processing of human antibodies integrating an extraction capture step and cation exchange chromatography. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 877(1-2), pp.50–8.
7. Bai, Y. et al., 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins*, 17(1), pp.75–86.
8. Bai, Y. et al., 1995. Thermodynamic Parameters from Hydrogen Exchange Measurements. *Methods in Enzymology*, 259, pp.344–356.
9. Bak, H. & Thomas, O.R.T., 2007. Evaluation of commercial chromatographic adsorbents for the direct capture of polyclonal rabbit antibodies from clarified antiserum. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 848(1), pp.116–30.
10. Baum, J. et al., 1989. Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig alpha-lactalbumin. *Biochemistry*, 28(1), pp.7–13.
11. Benedek, K., 1988. Thermodynamics of alpha-Lactalbumin Denaturation in Hydrophobic-Interaction Chromatography and Stationary Phase Comparison. , 458, pp.93–104.
12. Berger, A., Loewenstein, A. & Meiboom, S., 1959. Nuclear Magnetic Resonance Study of the Protolysis and Ionization of N-Methylacetamide. , 81, pp.62–67.
13. Chan, W.-H. et al., 2010. Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *Journal of virology*, 84(14), pp.7337–45.
14. Conchillo-Solé, O. et al., 2007. AGGREGSCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC bioinformatics*, 8(i), p.65.
15. Czypionka, A. et al., 2007. The isolated C-terminal domain of Ring1B is a dimer made of stable, well-structured monomers. *Biochemistry*, 46(44), pp.12764–76.

16. Deitcher, R W et al., 2009. Protein instability during HIC: evidence of unfolding reversibility, and apparent adsorption strength of disulfide bond-reduced alpha-lactalbumin variants. *Biotechnology and bioengineering*, 102(5), pp.1416–27.
17. Deitcher, R.W., O'Connell, J.P. & Fernandez, E.J., 2010. Changes in solvent exposure reveal the kinetics and equilibria of adsorbed protein unfolding in hydrophobic interaction chromatography. *Journal of Chromatography A*, 1217, pp.5571–5583.
18. DePalma, A., 2012. Leading-Edge Separation Techniques. *Genetic Engineering & Biotechnology News*, 32(9), pp.1–39.
19. Dismer, F. & Hubbuch, J., 2007. A novel approach to characterize the binding orientation of lysozyme on ion-exchange resins. *Journal of chromatography. A*, 1149(2), pp.312–20.
20. Doyle, D. a., 1998. The Structure of the Potassium Channel: Molecular Basis of K<sup>+</sup> Conduction and Selectivity. *Science*, 280(5360), pp.69–77.
21. Edgcomb, S.P. & Murphy, K.P., 2002. Variability in the pKa of histidine side-chains correlates with burial within proteins. *Proteins*, 49(1), pp.1–6.
22. Engel, M.F.M., Visser, A.J.W.G. & Van Mierlo, C.P.M., 2004. Conformation and orientation of a protein folding intermediate trapped by adsorption. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31), pp.11316–21.
23. Eriksson, K., Rodrigo, G. & Brekkan, E., 2009. MAb Contaminant Removal with a Multimodal Anion Exchanger. *Bioprocess International*.
24. Fausnaugh, J.L., Kennedy, L.A. & Regnier, F.E., 1984. Comparison of Hydrophobic-Interaction and Reversed-Phase Chromatography of Proteins. *Journal of Chromatography*, 311, pp.141–155.
25. Fogle, J.L., O'Connell, John P & Fernandez, Erik J, 2006. Loading, stationary phase, and salt effects during hydrophobic interaction chromatography: alpha-lactalbumin is stabilized at high loadings. *Journal of chromatography. A*, 1121(2), pp.209–18.
26. Forge, V et al., 1999. Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *Journal of Molecular Biology*, 288(4), pp.673–88.
27. Gagnon, P. et al., 2011. Minibodies and Multimodal Chromatography. , 8(2), pp.26–35.
28. Gagnon, P., Grund, E. & Lindbäck, T., 1995. Large Scale Process Development for Hydrophobic Interaction Chromatography , Part 1 : Gel Selection and Development of Binding Conditions. *Journal of the Honours Biology & Pharmacology Program*, 8, pp.1–9.
29. Ghaemmaghami, S., Fitzgerald, M C & Oas, T.G., 2000. A quantitative, high-throughput screen for protein stability. *Proceedings of the National Academy of Sciences of the United States of America*, 97(15), pp.8296–301.
30. Gillespie, R. et al., 2012. Cation exchange surface-mediated denaturation of an aglycosylated immunoglobulin (IgG1). *Journal of chromatography. A*, 1251, pp.101–10.
31. Gospodarek, A.M. et al., 2011. Protein stability and structure in HIC: hydrogen exchange experiments and COREX calculations. *Langmuir: the ACS journal of surfaces and colloids*, 27(1), pp.286–95.

32. Greene, F. & Pace, C N, 1974. Urea and Guanidine Hydrochloride Denaturation of Ribonuclease, Lysozyme, alpha-Chymotrypsin, and beta-Lactoglobulin. *The Journal of biological chemistry*, 249(17), pp.5388–5393.
33. Hjerten, S., 1973. Some General Aspects of Hydrophobic Interaction Chromatography. *Journal of Chromatography*, 87, pp.325–331.
34. Horinek, D. & Netz, R.R., 2011. Can simulations quantitatively predict peptide transfer free energies to urea solutions? Thermodynamic concepts and force field limitations. *The journal of physical chemistry. A*, 115(23), pp.6125–36.
35. Hou, Y. et al., 2010. Effects of urea induced protein conformational changes on ion exchange chromatographic behavior. *Journal of chromatography. A*, 1217(47), pp.7393–400.
36. Ingraham, R. et al., 1985. Denaturation and the Effects of Temperature on Hydrophobic-Interaction and Reversed-Phase High-Performance Liquid Chromatography of Proteins. *Journal of Chromatography*, 327, pp.77–92.
37. Jungbauer, A., Machold, C. & Hahn, R., 2005. Hydrophobic interaction chromatography of proteins. *Journal of Chromatography A*, 1079(1-2), pp.221–228.
38. El Kadi, N. et al., 2006. Unfolding and refolding of bovine serum albumin at acid pH: ultrasound and structural studies. *Biophysical journal*, 91(9), pp.3397–404.
39. Kallberg, K., Johansson, H.-O. & Bulow, L., 2012. Multimodal chromatography: an efficient tool in downstream processing of proteins. *Biotechnology journal*, 7(12), pp.1485–95.
40. Kato, Y., Kitamura, T. & Hashimoto, T., 1984. Operational Variables in High-Performance Hydrophobic Interaction Chromatography of Proteins on TSKGEL Phenyl-5PW. *Journal of Chromatography*, 298, pp.407–418.
41. Krishnan, B. & Gierasch, L.M., 2011. Dynamic local unfolding in the serpin  $\alpha$ -1 antitrypsin provides a mechanism for loop insertion and polymerization. *Nature structural & molecular biology*, 18(2), pp.222–6.
42. Ladiwala, A. et al., 2006. Investigation of protein retention and selectivity in HIC systems using quantitative structure retention relationship models. *Biotechnology and bioengineering*, 93(5), pp.836–50.
43. Larson, R.G., 1999. *The Structure and Rheology of Complex Fluids*, New York City: Oxford University Press.
44. Lin, V.J. & Koenig, J.L., 1976. Raman studies of bovine serum albumin. *Biopolymers*, 15(1), pp.203–18.
45. Lindahl, L. & Vogel, H., 1984. Metal-ion-dependent hydrophobic-interaction chromatography of  $\gamma$ -lactalbumins. *Analytical Biochemistry*, 140(2), pp.394–402.
46. Machold, C. et al., 2002. Hydrophobic interaction chromatography of proteins. I. Comparison of selectivity. *Journal of chromatography. A*, 972(1), pp.3–19.
47. Man, P. et al., 2010. Accessibility changes within diphtheria toxin T domain when in the functional molten globule state, as determined using hydrogen/deuterium exchange measurements. *The FEBS journal*, 277(3), pp.653–62.
48. Mccue, J.T., 2009. Theory and Use of Hydrophobic Interaction Chromatography in Protein Purification Applications. *Methods in Enzymology*, 463(09), pp.405–414.
49. Mcnay, J. & Fernandez, E., 1999. How does a protein unfold on a reversed-phase liquid chromatography surface? *Journal of Chromatography A*, 849(1), pp.135–148.

50. Molinari, H. et al., 1996. Partially folded structure of monomeric bovine beta-lactoglobulin. *FEBS letters*, 381(3), pp.237–43.
51. Myers, J.K., Pace, C Nick & Scholtz, J.M., 1995. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Science*, 4, p.98.
52. Norde, W & Giacomelli, C E, 2000. BSA structural changes during homomolecular exchange between the adsorbed and the dissolved states. *Journal of biotechnology*, 79(3), pp.259–68.
53. Pace, C N et al., 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *The Journal of Biological Chemistry*, 263(24), pp.11820–11825.
54. Powell, K.D. & Fitzgerald, Michael C, 2003. Accuracy and precision of a new H/D exchange- and mass spectrometry-based technique for measuring the thermodynamic properties of protein-peptide complexes. *Biochemistry*, 42(17), pp.4962–70.
55. Ptitsyn, O.B., 1995. Structures of folding intermediates. *Current opinion in structural biology*, 5(1), pp.74–8.
56. Qi, S. et al., 2010. Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4. *Cell*, 141(3), pp.446–57.
57. Qi, W. et al., 2008. Simultaneous monitoring of peptide aggregate distributions, structure, and kinetics using amide hydrogen exchange: application to Aβ(1-40) fibrillogenesis. *Biotechnology and Bioengineering*, 100(6), pp.1214–27.
58. Queiroz, J., 2001. Hydrophobic interaction chromatography of proteins. *Journal of Biotechnology*, 87(2), pp.143–159.
59. Sadler, P.J. & Tucker, a, 1993. pH-induced structural transitions of bovine serum albumin. Histidine pKa values and unfolding of the N-terminus during the N to F transition. *European journal of biochemistry / FEBS*, 212(3), pp.811–7.
60. Santra, M.K. et al., 2005. Unfolding pathways of human serum albumin: evidence for sequential unfolding and folding of its three domains. *International Journal of Biological Macromolecules*, 37(4), pp.200–4.
61. Sen, P., Ahmad, B. & Khan, R.H., 2008. Formation of a molten globule like state in bovine serum albumin at alkaline pH. *European biophysics journal: EBJ*, 37(8), pp.1303–8.
62. Sethuraman, A. & Belfort, G., 2005. Protein structural perturbation and aggregation on homogeneous surfaces. *Biophysical journal*, 88(2), pp.1322–33.
63. Shen, Z.M. et al., 1992. Conformational stability of porcine serum transferrin. *Protein science: a publication of the Protein Society*, 1(11), pp.1477–84.
64. Sluzky, V. et al., 1991. Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 88(21), pp.9377–81.
65. Suda, E.J. et al., 2009. Comparison of agarose and dextran-grafted agarose strong ion exchangers for the separation of protein aggregates. *Journal of chromatography. A*, 1216(27), pp.5256–64.
66. Tang, L., Roulhac, P.L. & Fitzgerald, Michael C, 2007. H/D exchange and mass spectrometry-based method for biophysical analysis of multidomain proteins at the domain level. *Analytical chemistry*, 79(22), pp.8728–39.

67. Tayyab, S., Sharma, N. & Mushahid Khan, M., 2000. Use of domain specific ligands to study urea-induced unfolding of bovine serum albumin. *Biochemical and biophysical research communications*, 277(1), pp.83–8.
68. Tobler, S. & Fernandez, E., 2002. Structural features of interferon aggregation revealed by hydrogen exchange. *Protein Science*, 11, pp.1340–1352.
69. Togashi, D.M., Ryder, A.G. & O'Shaughnessy, D., 2010. Monitoring local unfolding of bovine serum albumin during denaturation using steady-state and time-resolved fluorescence spectroscopy. *Journal of fluorescence*, 20(2), pp.441–52.
70. Tokihiro, K. et al., 1996. Mass spectroscopic evidence on inhibiting effect of matosyl-a-cyclodextrin on insulin self-association. *Pharmaceutical Sciences*, 2, pp.519–522.
71. Trovato, A., Seno, F. & Tosatto, S.C.E., 2007. The PASTA server for protein aggregation prediction. *Protein engineering, design & selection: PEDS*, 20(10), pp.521–3.
72. Valliere-Douglass, J., Wallace, A. & Balland, A., 2008. Separation of populations of antibody variants by fine tuning of hydrophobic-interaction chromatography operating conditions. *Journal of chromatography. A*, 1214(1-2), pp.81–9.
73. Vermonden, T., Giacomelli, Carla E. & Norde, Willem, 2001. Reversibility of Structural Rearrangements in Bovine Serum Albumin during Homomolecular Exchange from AgI Particles. *Langmuir*, 17(12), pp.3734–3740.
74. Vesely, F. et al., 2012. Simple colorimetric methods for determination of sub-milligram amounts of ultra-high molecular weight polyethylene wear particles. *Acta biomaterialia*, 8(5), pp.1935–8.
75. Voitl, A., Butté, A. & Morbidelli, M., 2010. Behavior of human serum albumin on strong cation exchange resins: I. experimental analysis. *Journal of chromatography. A*, 1217(34), pp.5484–91.
76. Wada, A., Saito, Y. & Ohogushi, M., 1983. Multiphasic conformation transition of globular proteins under denaturing perturbation. *Biopolymers*, 22(1), pp.93–9.
77. West, G.M., Tang, L. & Fitzgerald, Michael C, 2008. Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass spectrometry-based strategy. *Analytical chemistry*, 80(11), pp.4175–85.
78. Williams, E. & Foster, J., 1960. The Aggregation of Bovine Plasma Albumin at Low pH. *Journal of the American Chemical Society*, 82(14), pp.3741–3745.
79. Wu, H. et al., 2011. Site-selective probe for investigating the asynchronous unfolding of domains in bovine serum albumin. *Talanta*, 84(3), pp.881–6.
80. Wu, S.-L.W., Figueroa, A. & Karger, B.L., 1986. Protein Conformational Effects in Hydrophobic Interaction Chromatography. , 371, pp.3–27.
81. Xia, F. et al., 2004. Evaluation of selectivity changes in HIC systems using a preferential interaction based analysis. *Biotechnology and bioengineering*, 87(3), pp.354–63.
82. Xiao, Yunzhi et al., 2006. Protein instability during HIC: describing the effects of mobile phase conditions on instability and chromatographic retention. *Biotechnology and bioengineering*, 93(6), pp.1177–89.
83. Yamamoto, S. et al., 2007. Effects of protein conformational changes on separation performance in electrostatic interaction chromatography: unfolded proteins and PEGylated proteins. *Journal of biotechnology*, 132(2), pp.196–201.

84. Yamasaki, M. et al., 2010. Loop-sheet mechanism of serpin polymerization tested by reactive center loop mutations. *The Journal of biological chemistry*, 285(40), pp.30752–8.
85. Zhang, Z. & Smith, D.L., 1993. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *October*, 2, pp.522–531.
86. Zhao, G., Dong, X.-Y. & Sun, Y., 2009. Ligands for mixed-mode protein chromatography: Principles, characteristics and design. *Journal of biotechnology*, 144(1), pp.3–11.



## Appendix A. Supplementary Data

**Table A1. Transferrin Reporter Peptides**

Peptide	Start	End	Sequence
1	64	81	VYDAYLAPNNLKPVVAEF
2	67	81	AYLAPNNLKPVVAEF
3	125	131	WNIPIGL
4	180	188	NQYFGYSGA
5	202	209	VKHSTIFE
6	202	211	VKHSTIFENL
7	202	249	VKHSTIFENLANKADRDQYELLCLDNTRKPVDEYKDCHLAQVPSHTVV
8	209	222	ENLANKADRDQYEL
9	210	249	NLANKADRDQYELLCLDNTRKPVDEYKDCHLAQVPSHTVV
10	212	249	ANKADRDQYELLCLDNTRKPVDEYKDCHLAQVPSHTVV
11	223	259	LCLDNTRKPVDEYKDCHLAQVPSHTVVVARSMGGKEDL
12	250	259	ARSMGGKEDL
13	251	259	RSMGGKEDL
14	260	278	IWELLNQAQEHFGKDKSKE
15	260	281	IWELLNQAQEHFGKDKSKEFQL
16	263	278	LLNQAQEHFGKDKSKE
17	263	281	LLNQAQEHFGKDKSKEFQL
18	264	281	LNQAQEHFGKDKSKEFQL
19	264	291	LNQAQEHFGKDKSKEFQLFSSPHGKDLL
20	279	291	FQLFSSPHGKDLL
21	292	306	FKDSAHGFLKVPPRM
22	292	308	FKDSAHGFLKVPPRMDA
23	292	310	FKDSAHGFLKVPPRMDAKM
24	386	392	MSLDGGF
25	424	435	FAVAVVKKSASD
26	529	555	FVKHQTPQNTGGKNPDPWAKNLNEKD
27	530	553	VKHQTPQNTGGKNPDPWAKNLNE
28	530	555	VKHQTPQNTGGKNPDPWAKNLNEKD
29	530	557	VKHQTPQNTGGKNPDPWAKNLNEKDYE
30	636	651	AKLHDRNTYEKYLGEE
31	652	659	YVKAVGNL
32	652	667	YVKAVGNLRKCSTSSL

**Table A2. Antitrypsin Reporter Peptides**

<b>Peptide</b>	<b>Start</b>	<b>End</b>	<b>Sequence</b>
<b>1</b>	24	32	NKITPNLAE
<b>2</b>	78	89	EGLNFnLTEIPE
<b>3</b>	86	92	EIPEAQI
<b>4</b>	98	108	ELLRTLNLQPDS
<b>5</b>	100	109	LRTLNLQPDSQ
<b>6</b>	110	119	LQLTTGNGLF
<b>7</b>	111	119	QLTTGNGLF
<b>8</b>	131	142	LEDVKKLYHSEA
<b>9</b>	151	166	EEAKKQINDYVEKGTQ
<b>10</b>	183	189	ALVNYIF
<b>11</b>	197	207	PFEVKDTEED
<b>12</b>	201	207	KDTEED
<b>13</b>	238	241	WVLL
<b>14</b>	252	266	FFLPDEGKLQHLNE
<b>15</b>	252	270	FFLPDEGKLQHLENELTHD
<b>16</b>	265	279	NELTHDIITKFLNE
<b>17</b>	266	273	ELTHDIIT
<b>18</b>	278	288	NEDRRSASLHL
<b>19</b>	304	317	GQLGITKVFSNGAD
<b>20</b>	306	314	LGITKVFSN
<b>21</b>	318	338	LSGVTEEAPLKLSKAVHKAVL
<b>22</b>	326	343	PLKLSKAVHKAVLTIDEK
<b>23</b>	353	372	LEAIPMSIPPEVKFNKPFVF
<b>24</b>	354	372	EAIPMSIPPEVKFNKPFVF
<b>25</b>	356	372	IPMSIPPEVKFNKPFVF
<b>26</b>	357	373	PMSIPPEVKFNKPFVFL
<b>27</b>	367	385	NKPFVFLMIEQNTKSPLFM
<b>28</b>	376	384	EQNTKSPLF

**Table A3. BSA Reporter Peptides**

<b>Peptide</b>	<b>Start</b>	<b>End</b>	<b>Sequence</b>
<b>1</b>	1	14	DTHKSEIAHRFKDL
<b>2</b>	2	7	THKSEI
<b>3</b>	20	31	KGLVLIAFSQYL
<b>4</b>	39	45	HVKLVNE
<b>5</b>	49	70	FAKTCVADESHAGCEKSLHTLF
<b>6</b>	127	137	KADEKKFWGKY
<b>7</b>	154	164	LYYANKYNGVF
<b>8</b>	169	182	QAEDKGACLLPKIE
<b>9</b>	200	209	ASIQKFGERA
<b>10</b>	219	226	SQKFPKAE
<b>11</b>	307	313	DFAEDKD
<b>12</b>	324	329	AFLGSF
<b>13</b>	331	340	YEYSRRHPEY
<b>14</b>	333	340	YSRRHPEY
<b>15</b>	393	398	QFEKLG
<b>16</b>	422	435	LVEVSRSLGKVGTR
<b>17</b>	423	436	VEVSRSLGKVGTRC
<b>18</b>	440	460	PESERMPCTEDYLSLILNRLC
<b>19</b>	529	543	VELLKHKPKATEEQL
<b>20</b>	530	541	ELLKHKPKATEE
<b>21</b>	531	545	LLKHKPKATEEQLKT
<b>22</b>	531	547	LLKHKPKATEEQLKTVM
<b>23</b>	559	573	AADDKEACFAVEGPK
<b>24</b>	575	583	VVSTQTALA

Table A4. Unfolding Rates and Free Energies of Transferrin on Hexyl 650C

	N-terminal <sup>a</sup>	C-terminal <sup>b</sup>
$\Delta G_{u,sol}$ [kcal/mol] <sup>1</sup>	9.3	13.6
$\Delta G_{u,sol}$ [kcal/mol] <sup>2</sup>	11.1	14.4
$\Delta G_{u,ads}$ [kcal/mol] <sup>1</sup>	6.5	7.2
$\Delta \Delta G_u$ [kcal/mol] <sup>3</sup>	2.8	6.4
$m_{sol}$ [kcal/mol•M] <sup>1</sup>	3.1	5.5
$m_{sol}$ [kcal/mol•M] <sup>2</sup>	2.5	2.7
$m_{ads}$ [kcal/mol•M] <sup>1</sup>	1.6	3.0
$k_{u,sol}$ [min <sup>-1</sup> ] <sup>4</sup>	0.046	0.028
$k_{u,ads}$ [min <sup>-1</sup> ] <sup>4</sup>	0.14	0.071
$\Delta k_u$ [min <sup>-1</sup> ] <sup>5</sup>	0.094	0.043

<sup>a</sup> Data from GdnHCl denaturation curve of Residues 223-259 used

<sup>b</sup> Data from GdnHCl denaturation curve of Residues 554-589 used

<sup>1</sup> Calculated with Equations 2.11, 2.12, and 2.13

<sup>2</sup> Values taken from Tang et al. 2007

<sup>3</sup> Calculated with Equation 2.14

<sup>4</sup> Calculated with Equation 2.11

<sup>5</sup> Calculated with Equation 2.15