PROTEIN STRUCTURAL INVESTIGATION USING HYDROXYL RADICAL PROTEIN FOOTPRINTING

by

CAROLINE MCCULLOUGH WATSON

(Under the Direction of Joshua S. Sharp and Ron Orlando)

ABSTRACT

The focus of my research is using hydroxyl radical protein footprinting (HRF) coupled with mass spectrometry (MS) to study protein structure, protein – protein, and protein – ligand interactions in solution. Even with high-resolution techniques such as NMR and X-ray crystallography, not all protein structures or protein interactions can be studied using these techniques due to the nature of the protein. Hydroxyl radicals have become a popular labeling technique because they provide a fast, relatively nonspecific, covalent label that probes a variety of solvent accessible amino acid residues with one experiment. The reaction rate of hydroxyl radicals with each amino acid is based on two factors, the chemical nature of the residue and the average accessibility of the residue to the hydroxyl radical. By monitoring the change in the amount of oxidation under different experimental conditions, changes in protein structure can be determined by how the structural changes affect the solvent accessibility of the different regions of the protein. The work presented describes protein pharmaceuticals and a biologically relevant protein complex that were investigated by HRF coupled with mass spectrometry, as well as a new method for producing hydroxyl radicals for protein labeling without the use of hydrogen peroxide.

INDEX WORDS: Mass Spectrometry, Hydroxyl Radicals, Protein Footprinting, CCL5/RANTES, β-lactoglobulin, Protein Structure, Fast Photochemical Oxidation of Proteins (FPOP), Water Radiolysis, Biosimilars, Granulocyte Colony Stimulating Faction, Interferon α-2A, Erythropoietin

PROTEIN STRUCTURAL INVESTIGATION USING HYDROXYL RADICAL PROTEIN FOOTPRINTING

by

CAROLINE MCCULLOUGH WATSON

B.S., College of Charleston, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2012

PROTEIN STRUCTURAL INVESTIGATION USING HYDROXYL RADICAL PROTEIN FOOTPRINTING

by

CAROLINE MCCULLOUGH WATSON

Major Professors:	Joshua S. Sharp Ron Orlando
Committee:	I. Jonathan Amster Lance Wells

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2012

DEDICATION

Mom, this is for you!

ACKNOWLEDGEMENTS

I would like to acknowledge the support for this work was provided by grants from the National Center for Research Resources (5P41RR005351-23) and the National Institute of General Medical Sciences (8 P41 GM103390-23) from the National Institutes of Health. Thanks to Dr. Prestegard, Dr. Xu Wang and Younghee Park at the Complex Carbohydrate Research Center at the University of Georgia for your NMR knowledge and assistance, and thanks to Tracy Handel at Skaggs School of Pharmacy and Pharmaceutical Sciences at the University California San Diego for the work with CCL5. Thanks to Dr. Ireneusz Janik at the University of Notre Dame Radiation Laboratory for his time and instrumentation. Thanks to Dr. Jeffery Urbauer and Ramona Urbauer for their circular dichroism instrumentation and assistance. I would also like to thank Dr. Elizabeth R. Wright and Hong Yi at the Robert P. Apkarian Integrated Electron Microscopy Core Facility at Emory University for obtaining the electron microscopy images. I would like to thank my committee members for their time and suggestions throughout my graduate career.

I extend my thanks to my advisors, Joshua Sharp and Ron Orlando, for, well, everything. Josh, I cannot thank you enough for your patience and guidance; you have provided me with an invaluable education and I hope to make you proud in my post graduate career. To Jesse Hines, I thank you for the calming lava lamps and your irreplaceable assistance. To all my past and present group members, I thank you for the ideas, support, and laughs. To the next “group mom”, good luck.

A special thanks to my friends and family for your unending love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Protein tertiary and quaternary structure are fundamental to determining mechanisms of protein function. Understanding the details of protein structure and protein interactions in macromolecule assemblies are of biological importance. Protein interactions and assemblies control almost all biological processes, like replication, transcription, translation, metastasis and apoptosis. The methods available for studying proteins provide different types of structural information. Two commonly used high resolution methods for structural analysis are X-ray crystallography and Nuclear Magnetic Resonance Spectroscopy (NMR).

X-ray crystallography is a powerful tool for obtaining structural information [1], and accounts for almost 90% of the structures entered into the protein databank (www.pdb.org) [2]; however, it comes with significant limitations. Proteins must be examined in crystalline form; therefore, the behavior of molecules in solutions cannot be studied and only one conformation of the protein can be observed [1]. Even with milligrams of pure sample, adequate crystals of proteins cannot always be obtained; for instance, conformational flexibility of the protein and/or ligand make crystallization nearly impossible [3].

NMR is another dominant tool for obtaining high resolution protein structure, dynamics, and interaction information by studying molecules in solutions using a variety of experiments [1]. Even though NMR doesn't require crystals and studies proteins under more physiological conditions, disadvantages to NMR include protein size limitation of ~35 kDa [4], difficulty in manual data interpretation [3], and requirement of milligrams of isotopically labeled proteins for

comprehensive studies. Both NMR and X-ray crystallography are high resolution techniques used to examine protein structures; however, both methods are low throughput, labor intensive and are not suitable for all proteins or protein complexes.

Low resolution techniques, such as absorption, fluorescence, and circular dichroism (CD) spectroscopy, can be used to provide a quick look at the structural integrity of the protein analyte. These techniques do not provide atomic coordinates like high resolution techniques, but are commonly used to examine structural reproducibility, overall secondary structure, and real-time folding.

Mass spectrometry-based protein footprinting has become a widely used tool for obtaining an intermediate level of structural analysis of proteins. Protein footprinting techniques determine the solvent accessibility of the backbone or side chain structures of macromolecules by chemical modification [5]. The standard approach to protein footprinting is to determine which sites or areas of the protein reveal a change in solvent accessibility between two or more states of the protein. Acetylation [6], crosslinking [7, 8], deuterium exchange [9-11], and hydroxyl radicals [12] are examples of chemical modifications used to map macromolecules. The drawback of acetylation and crosslinking is both usually only react with lysine, limiting the number of sites to probe for investigation. Deuterium reversibly exchanges with backbone amide hydrogens of all twenty amino acids [13-18], but its reversibility complicates data analysis. Hydroxyl radical methods readily label fourteen of the twenty amino acids scaring the modification to the protein unlike the reversible deuterium exchange [5, 12]; an irreversible modification to the protein gives snapshots of native protein structure under different conditions.

Hydrogen-deuterium exchange (HDX) in conjunction with MS is commonly used to study protein structure, dynamics, and protein-ligand interactions. As more deuteriums are

incorporated, the protein mass will increase making it possible to study protein dynamics such as rates of folding and unfolding, and determine binding constants of a protein-ligand interaction [19], The location of deuteriums can be roughly determined by digesting the deuterated protein with an acid protease such as pepsin to decrease the rate of back-exchange, and fast HPLC separation at 0ºC prior to MS. Disadvantages of HDX include the use of pepsin for digestion which introduces a complex mixture of overlapping partial digestion products, and complex data analysis because tandem MS is not possible due to back exchange.

Hydroxyl radicals have been used in protein footprinting experiments for years. Initially, hydroxyl radicals were used to nonspecifically cleave the protein backbone [20-23] followed by SDS-PAGE to separate cleavage products. Hydrogen abstraction from the Cα carbon and subsequent radical reaction with oxygen leads to backbone cleavage [24, 25]. However, the side chain reactions with hydroxyl radicals occur 10 to 1000 times faster than hydrogen abstraction from the Cα carbon [24, 26, 27]. Therefore, amino acid side chains serve as better probes for protein structural studies.

Hydroxyl radical footprinting (HRF) coupled with MS has become increasingly popular as a labeling technique to probe intact protein structure [28-34], protein-protein interactions [35-37], protein folding [38, 39], and protein-small molecule ligand interactions [40] in solution. HRF uses diffusing hydroxyl radicals to oxidize the protein of interest and the amount of oxidation of each amino acid is measure by MS. Hydroxyl radicals are attractive for labeling for several reasons. First, the van der Waals area and solvent properties are similar to water molecules, making hydroxyl radicals ideal as solvent accessibility probes. Second, they have high reactivity with well understood chemical selectivity. Third, they can be generated safely and conveniently under multiple solution conditions. Finally, their non-specificity allows many

residues to be covalently probed with one experiment. Covalently labeling the protein using HRF overcomes the shortcomings of HDX and facilitates sample analysis making it easy to denature, digest and analyze the peptides by MS.

Generally, in an HRF experiment the protein solution is exposed to hydroxyl radicals and the radicals covalently react with the solvent accessible side chains of residues [26]. Following irradiation the modified protein can be frozen for later analysis or immediately subjected to specific proteases to maximize sequence coverage and simplify data analysis. The stable modifications to the protein side chains [26] leads to straightforward and relatively easy sample analysis. Reverse-phase high performance liquid chromatography (HPLC) separates both modified and unmodified peptides. The chromatographic peak area or mass spectral intensity of modified and unmodified peptides is used to calculate the faction oxidized.

Hydroxyl radicals for labeling proteins have been generated by Fenton chemistry [34, 41], hydrogen peroxide (H_2O_2) photolysis [29, 32, 33], and water radiolysis by gamma (γ) rays [42-46], exposure to an X-ray synchrotron beam [12, 36, 37, 39, 47, 48], or using a Van de Graaff electron accelerator [49]. The time scale for large scale protein motions, like helix coiling/uncoiling, ranges from long microseconds to milliseconds [50]; however, most hydroxyl radical generation techniques, including water radiolysis by gamma (γ) rays and X-rays, are performed on a millisecond to minute time scale. A time scale longer than large scale protein motion increases the likelihood that the protein of interest will experience unfolding due to oxidative modifications [51, 52]. Previous strategies to determine native protein probing include circular dichroism to monitor conformational changes as oxidation is increased [28], limiting the amount of oxidation to the point that a majority of protein molecules had one or no oxidations [33, 34], or monitoring the reaction kinetics of the protein [53] or constituent peptides [35, 36,

46] to be oxidation follows the expected pseudo-first order rate law. These strategies limit the overall amount of oxidation to prevent oxidatively unfolded protein, which limits the amount of oxidation sites resulting in lower structural resolution.

Successful footprinting methods achieve high levels of protein oxidation on a sub-microsecond time scale, shorter than large scale protein motions. Fast photochemical oxidation of proteins (FPOP) [29, 30] and pulsed electron beam water radiolysis [49] produce hydroxyl radicals in less than a microsecond. Addition of a scavenger controls radical exposure time to the protein; this is important because a single oxidation event can alter the native protein structure and lead to labeling of the oxidatively unfolded conformation [51-53]. The hydroxyl radical extinction coefficient with Beer-Lambert Law was used to calculate the expected lifetime of hydroxyl radicals in solution [54]. Addition of twenty millimolar glutamine is sufficient to consume the hydroxyl radicals in around one microsecond under the described conditions [55]; calculations show consumption at around one microsecond but does not consider hydroxyl radical reactions with buffer components or protein analyte, both of which will further decrease the half-life of the radical. Immediately following irradiation, methionine amide and catalase are added to devour any secondary oxidation products formed in the sample. By using glutamine to control radical lifetime and minimizing exposure to peroxide post FPOP, proteins can be labeled without excessively perturbing their conformation [55].

The rate of oxidation in hydroxyl radical protein footprinting is primarily a function of two factors: solvent accessibility and the reactivity of the exposed residues [12, 34, 35, 56]; this observation allows us to directly relate the relative rate of oxidation with the relative solvent accessibility of the residue in the structure [38, 56]. The rate constants for the reaction of each free amino acid with hydroxyl radical are found in **Table 1.1**; the relative reactivity of the side

chains with hydroxyl radicals is as follows: Cys > Met > Trp > Tyr > Phe > Cysteine > His > Leu , Ile > Arg, Lys, Val > Ser, Thr, Pro > Gln, Glu > Asp, Asn > Ala > Gly [27]. Protein oxidation leads to a variety of products with each amino acid side chain often having multiple competing reaction mechanisms [27, 57]. Hydroxyl radical reactions with each amino acid have been widely studied and the major oxidation products have been identified [26, 27, 57-62] (**Table 1.1**) and the structure of the reaction products are shown in **Figure 1.1**. Most commonly, all reactive residues experience a net addition of oxygen, i.e. +16, +32, +48 Da, etc. [26, 27, 57-63]; the aliphatic residues have additional +14 mass shifts due to the formation of carbonyl groups at the aliphatic hydrocarbon side chains. Some side chains give rise to characteristic oxidation products. Arginine has a characteristic mass shift of -43 Da, acidic residues have a -30 Da mass shift due to the decarboxylation of the C-terminal carboxyl, and oxidation of histidine residues leads to very complex products including +16, -22, -23, +5, and -10 Da. Multiple oxidation products complicate the spectrum analysis; however, characteristic oxidation products can facilitate residue level quantitation. **Table 1.1** gives a more comprehensive list of the common modifications observed in hydroxyl radical mediated protein footprinting experiments.

While hydroxyl radicals are considered non-specific reactants, they do not label side chains with equal efficiency. The second order rate constants listed in **Table 1.1** pertain to amino acids free in solution and not as residues in a protein sequence. The rates reflect the inherent residue reactivity in a fully exposed protein structural content [27].

Table 1.1. Initial ·OH-amino acid side chain reaction rates and common mass spectrometry observed products of ·OH-mediated protein footprinting

amino acid	$k_{\cdot OH}$ (M^{-1} sec^{-1}) [64]	common modifications (Da) [5]				
Cys	3.5×10^{10}	-15.9772	+31.9898	+47.9847		
Trp	1.3×10^{10}	+3.9949	+15.9949	+31.9898	+47.9847	
Tyr	1.3×10^{10}	+15.9949	+31.9898	+47.9847		
Met	8.5×10^{9}	-32.0085	+15.9949	+31.9898		
Phe	6.9×10^{9}	+15.9949	+31.9898	+47.9847		
His	4.8×10^{9}	-23.0160	-22.0320	-10.0320	+4.9789	+15.9949
Arg	3.5×10^{9}	-43.0534	13.9793	+15.9949		
Ile	1.8×10^{9}	+13.9793	+15.9949			
Leu	1.7×10^{9}	+13.9793	+15.9949			
Val	8.5×10^{8}	+13.9793	+15.9949			
Pro	6.5×10^{8}	+13.9793	+15.9949			
Gln	5.4×10^{8}	+13.9793	+15.9949			
Thr	5.1×10^{8}	-2.0157	+15.9949			
Lys	3.5×10^{8}	+13.9793	+15.9949			
Ser	3.2×10^{8}	-2.0157	+15.9949			
Glu	2.3×10^{8}	-30.0106	+13.9793	+15.9949		
Ala	7.7×10^{7}	+15.9949				
Asp	7.5×10^{7}	-30.0106	+15.9949			
Asn	4.9×10^{7}	+15.9949				
Gly	1.7×10^{7}	n.d.				

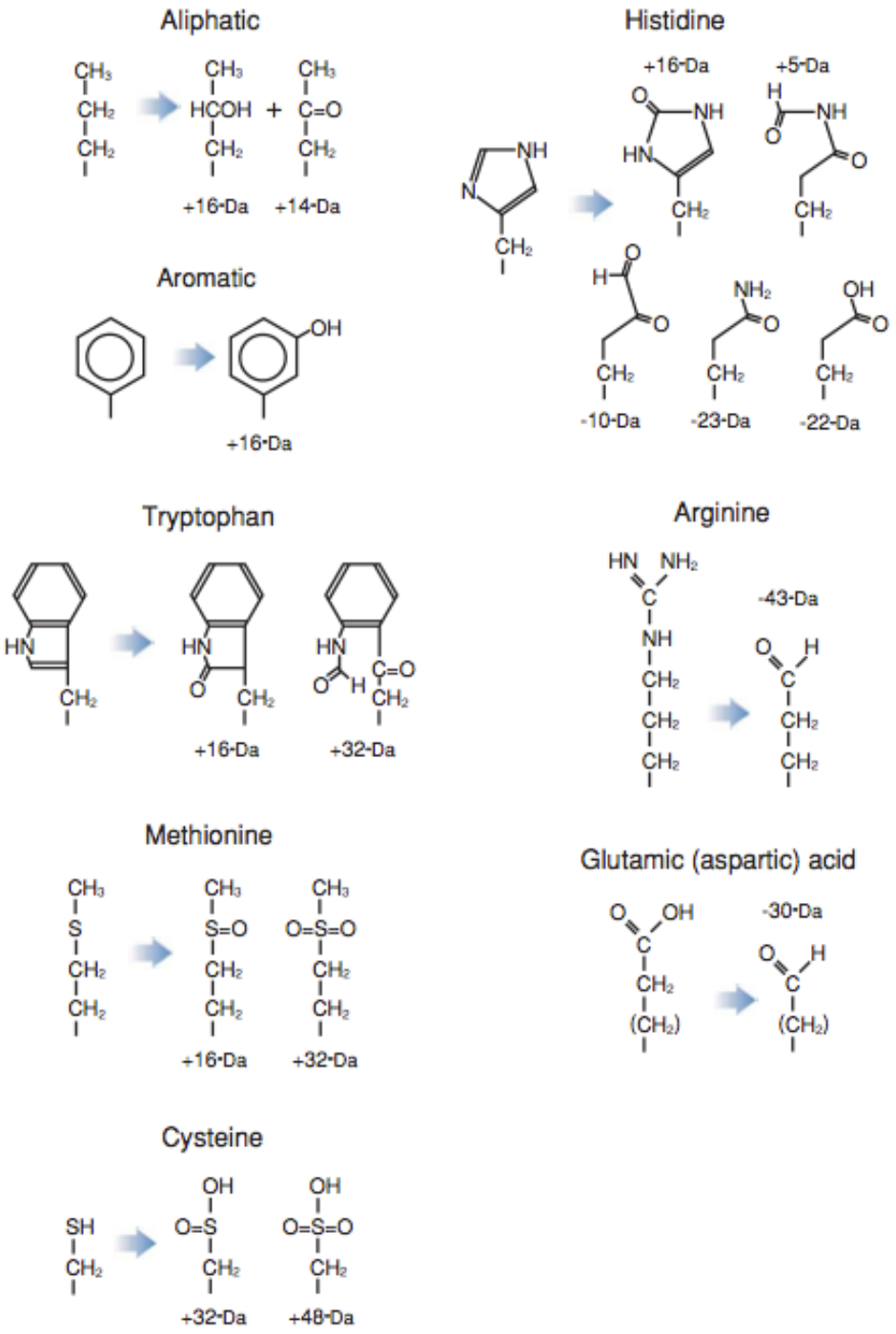


Figure 1.1. Major Oxidation Products. Major reaction products of the most reactive amino acid side chains [12].

Electron beam water radiolysis method for producing hydroxyl radicals is described in **Chapter 4**. The FPOP method for producing hydroxyl radicals is employed by using a 248 nm pulsed laser beam to photodissociate H_2O_2 into two hydroxyl radicals (·OH). The diffusing radical is dispersed through the protein solution reacting with the analyte as well as buffer components and glutamine scavenger. A flow system is designed such that the laser produces a small window of high flux light designed to maximize radical exposure to a small volume of protein solution (**Figure 1.2**). The syringe pump flow rate and the frequency of the laser are set to allow for 90% of the sample to be oxidized; this prevents any portion of the sample from being irradiated more than once and accounts for laminar flow. Following irradiation the sample is collected in a tube containing methionine amide and catalase to immediately quench secondary oxidants.

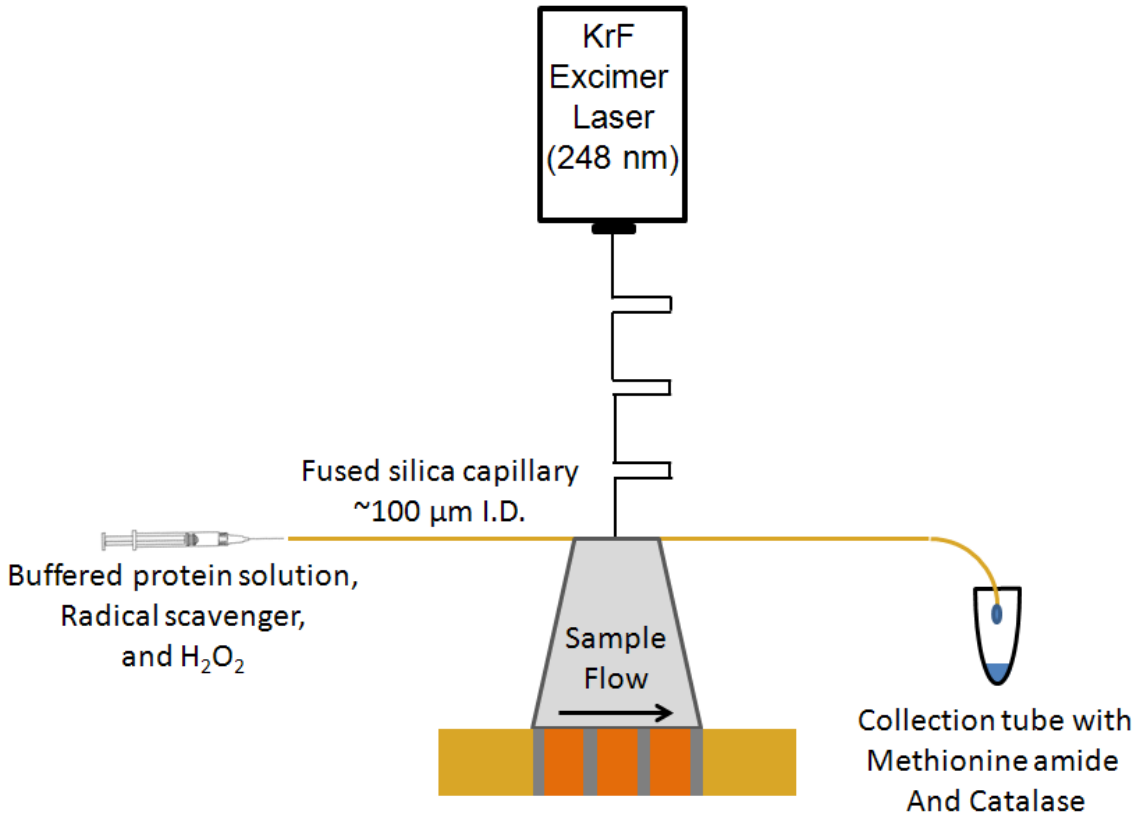


Figure 1.2. FPOP Schematic. Schematic of typical FPOP setup and conditions.

Hydroxyl radical footprinting is often used to determine the degree of difference from one sample to another; being able to quantify the difference in the amount of oxidation leads to conclusions about solvent accessibility in various samples. Quantitation of peptide oxidation is done using liquid chromatography (LC) – mass spectrometry (MS). The digested protein is typically subjected to reverse-phase HPLC, which separates the digested peptides and radiolytically modified peptides from their unmodified parent peptides. The modified peptides generally elute earlier than the unmodified peptide for simple alcohol or keto additions. The abundance of unmodified and modified peptides can be determined using the selected ion chromatogram peak area or the intensity of the peaks in the mass spectra. The sum of the modified peptide peak intensities (areas) divided by the sum of the unmodified and modified peptide peak intensities (areas) provides the fraction oxidized for that peptide. Tandem MS is used to confirm the peptide identity and the site(s) of oxidation on the peptide. Peptide fragmentation is classically done using collision-induced dissociation (CID). By investigating CID fragmentation spectra corresponding to a peptide oxidation product, the b- and y- ion series can be used to map the site(s) of oxidation to specific amino acids. **Figure 1.3** depicts an example of identifying the site of oxidation on a peptide.

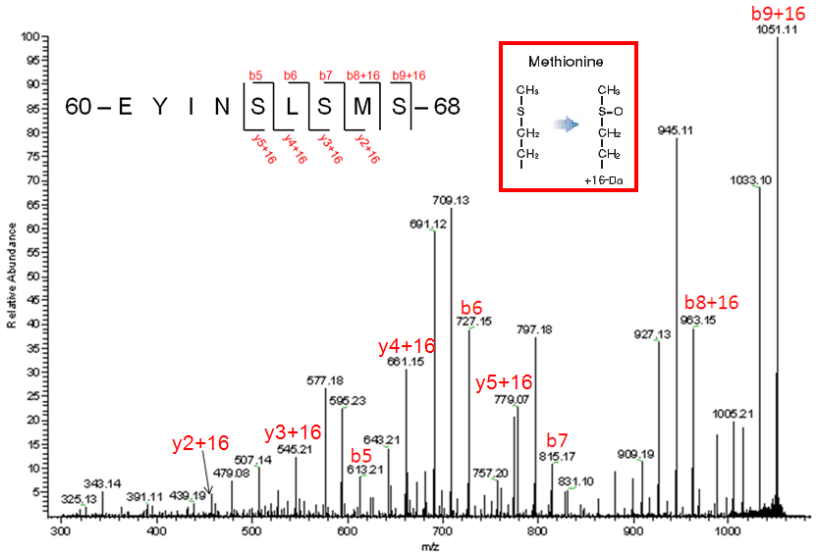


Figure 1.3. MS/MS Spectrum of an Oxidized Peptide. Sample spectrum mapping the oxidation site to methionine 67 on peptide 60-EYINSLSMS-68.

There are two peptide quantitation techniques used to quantify hydroxyl radical footprinting results, isotopic labeling and product monitoring. At lower radical doses, both quantitation methods are comparable but at higher radical doses, product monitoring can underestimates the amount oxidized due to difficulty of identifying all oxidation products. The advantage to isotopic labeling is that by monitoring the change in the unoxidized isotopically labeled peptide, the amount of oxidation can be accurately determined even if the oxidation products cannot be observed. Although isotopic labeling, especially ^{15}N labeling, is valuable it is not always feasible due to commercial availability; therefore, product monitoring must be used and label-free quantification does a reasonable job. Product monitoring is the commonly used quantitation method because isotopic labeling is expensive and only special cases warrant an extra protein labeling step. Isotopic labeling was employed in the CCL5 projects found in **Chapters 4** and **5** because the protein of interested could be expressed in-house; however the

other projects found in **Chapters 2** and **3** were quantitated using oxidation products of each peptide within the protein.

In **Chapter 2**, we demonstrate a pulsed electron beam water radiolysis method for performing sub-microsecond hydroxyl radical protein footprinting. Unlike other sub-microsecond radical labeling techniques, this method produces hydroxyl radicals without exposing the sample to hydrogen peroxide. Although hydrogen peroxide photolysis is sufficient in most cases, cases exist where the presence of hydrogen peroxide can harm your sample.

Therapeutic protein pharmaceuticals have become widespread and due to increased patent expiry, many follow-on formulations of the therapeutic proteins are being produced. To avoid extensive clinical trials for the biosimilars, the follow-on formulations must exhibit equivalency of both primary structure and protein conformation. **Chapter 3** describes an abbreviated hydroxyl radical protein footprinting for rapid conformational comparison of therapeutic protein formulations using samples of recombinant erythropoietin, interferon α-2A, and granulocyte colony stimulating factor.

CCL5 is chemokine that plays a role in recruiting leukocytes to sites of inflammation. Recently, studies have shown CCL5 oligomerization [65-68] and CCL5 – glycosaminoglycan (GAG) interaction [67-69] is essential for cell migration *in vivo*. The crystal structure for CCL5 dimer exists but, recent studies have emphasized the biological importance of CCL5 oligomerization. Using an integrated approach including NMR, small angle X-ray scattering, hydroxyl radical footprinting, and electron microscopy a novel structural model for CCL5 oligomerization is investigated and presented in **Chapter 4**. The structure of the oligomer explains the disaggregating effects of two CCL5 mutants and provides mechanisms for GAG promoted oligomerization. **Chapter 5** focuses on exploring CCL5 – GAG interactions with

techniques such as dynamic light scattering, electron microscopy, and hydroxyl radical protein footprinting. Three homogeneous sulfated GAGs are used as ligands to determine if the sulfation pattern or the amount of sulfation affects binding.

CHAPTER 2

PULSED ELECTRON BEAM WATER RADIOLYSIS FOR SUBMICROSECOND HYDROXYL RADICAL PROTEIN FOOPTRINTING[1]

ABSTRACT

Hydroxyl radical footprinting is a valuable technique for studying protein structure, but care must be taken to ensure that the protein does not unfold during the labeling process due to oxidative damage. Footprinting methods based on sub-microsecond laser photolysis of peroxide that complete the labeling process faster than the protein can unfold have been recently described; however, the mere presence of large amounts of hydrogen peroxide can also cause uncontrolled oxidation and minor conformational changes. We have developed a novel method for sub-microsecond hydroxyl radical protein footprinting using a pulsed electron beam from a 2 MeV Van de Graaff electron accelerator to generate a high concentration of hydroxyl radicals by radiolysis of water. The amount of oxidation can be controlled by buffer composition, pulse width, pulse current, and dissolved nitrous oxide gas in the sample. Our results with ubiquitin and β-lactoglobulin A demonstrate that one sub-microsecond electron beam pulse produces extensive protein surface modifications. Highly reactive residues that are buried within the protein structure are not oxidized, indicating that the protein retains its folded structure during the labeling process. Time-resolved spectroscopy indicates that oxidation of the protein is complete in a time scale shorter than that of large scale protein motions.

[1] Watson, C.; Janik, I.; Zhuang, T.; Charvatova , O.; Woods, R.; Sharp, J. 2009. *Journal of Analytical Chemistry*. *81* (7): 2496-2505. Reprinted here with permission of publisher.

INTRODUCTION

Mass spectrometry is becoming a widely used technique for structural analysis of proteins. Radical-based surface mapping techniques, such as hydroxyl radical protein footprinting, followed by mass spectrometry analysis has become increasingly popular for studying protein structure, protein-protein and protein-ligand interaction interfaces [12]. Hydroxyl radicals have gained popularity as a labeling technique because they provide a fast, relatively nonspecific, covalent label that probes a variety of solvent accessible amino acid residues with one experiment [57]. Proteins experience structural distortion due to conformational changes, multimerization, ligand binding, and aggregation, and these structural changes expose different solvent accessible surfaces of the protein. By labeling a protein with hydroxyl radicals before and after inducing a structural change, a mass spectrometric comparison can be used to determine protection or exposure of specific residues of each protein conformation [12, 47]. However, common methods of production of hydroxyl radicals produce undesired side reactions that complicate analysis. Here a superior method with few uncontrolled side reactions is presented.

Hydroxyl radicals for labeling of proteins have been successfully produced by Fenton chemistry [34, 41], hydrogen peroxide (H_2O_2) photolysis [29, 32, 33], and water radiolysis by either gamma (γ) rays [42-46] or exposure to an X-ray synchrotron beam [12, 36, 37, 39, 47, 48]. Protein oxidation is a labeling process that is dependent on the solvent-accessible surface and the reactivity of the exposed residues [12, 34, 35, 70], although a small dependence on the local sequence has been noted [33]. Upon oxidation the native protein structure can be altered due to the oxidative modifications; even a single oxidation event can induce protein unfolding which causes a rapid increase in oxidation as compared to the folded conformation [51-53]. In order for

hydroxyl radical protein footprinting to be a reliable method for protein structural determination, hydroxyl radicals must react with the protein exclusively in its native conformation. Most hydroxyl radical-generation techniques, including water radiolysis by gamma (γ) rays and X-rays, are performed on a millisecond to minute time scale, making it more likely that the protein will experience unfolding due to modifications [51, 52]. Most previous strategies for ensuring that the native conformation is exclusively probed include the use of circular dichroism to determine when conformational changes occur in order to label the protein before structural changes can be detected [28], limiting the amount of oxidation to where the vast majority of protein molecules have one or no oxidations [33, 34], or monitoring of the reaction kinetics to ensure the oxidation of the protein [53] or its constituent peptides [35, 36, 46] follows an expected pseudo-first order rate law. However, these strategies all require that the overall amount of oxidation be strictly limited to prevent probing oxidatively-unfolded structures, and this limited amount of oxidation results in fewer oxidation sites and lower structural resolution.

A relatively new strategy for ensuring that the native conformation of the protein is probed that still allows for high levels of oxidation is to ensure that all oxidation is completed on a faster time scale than the protein can unfold by using a very short UV laser pulse to photolyse H_2O_2. This can complete the labeling process on a sub-microsecond time scale [29, 30]. The time scale for large scale protein motions, such as helix coiling/uncoiling, is in the long microseconds to milliseconds range [50]; therefore, oxidatively labeling a protein at or below this time frame without the use of a precursor oxidant will prevent labeling of the oxidatively unfolded conformation. However, hydrogen peroxide-based methods of oxidation, including Fenton chemistry [34, 41, 71] and UV photolysis [28-30, 33], all have several problems stemming from the presence of H_2O_2. The presence of H_2O_2 in a protein solution can induce uncontrolled

Fenton-like chemistry in redox-active metal-binding proteins [72, 73], and cysteine and methionine two-electron oxidation can proceed spontaneously [60]. Furthermore, H_2O_2 has different physical properties than water and may induce protein conformational changes independent of oxidation.

Here we demonstrate the importance of limiting the exposure to hydrogen peroxide using an NMR based assessment of its effects on a model cysteine containing protein, Galectin-3. We then introduce a novel pulsed electron beam water radiolysis technique for hydroxyl radical protein footprinting that does not require hydrogen peroxide. In this method, hydroxyl radicals are generated in less than 1 μs using a 2 MeV Van de Graaff electron accelerator. Using a single electron beam pulse to oxidize the protein sample and subsequently scavenging secondary oxidants using methionine amide, protein oxidation is complete before the protein conformation can change. Ubiquitin, a small protein that ionizes efficiently, was used as a model to determine if controllable protein oxidation could be achieved by this water radiolysis method. Furthermore, β-lactoglobulin A was used as a model protein for method development because of its sensitivity to oxidation-induced conformational changes [53] and the availability of a high-resolution X-ray crystal structure [74]. In order to account for the motion of the side chains, molecular dynamics (MD) simulations of the protein were performed. These simulations enable the average solvent accessibilities to be computed for each side chain, and thus in principle provide a more complete model for the protein surface, under the conditions of the experiment, than possible from a single static crystal structure [70, 75]. We also determined the half-life of the hydroxyl radical label in anoxic solution with the protein by time-resolved UV spectroscopy to ensure that the label is consumed on a time scale consistent with large scale protein motions. The detected sites of

oxidation are compared with residues known to be present on the surface of the natively-folded protein to ensure that we are not labeling an oxidatively-unfolded protein.

MATERIALS AND METHODS

Reagents

All solvents were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise noted, at the highest purity available and used as supplied without further purification. Methionine amide and ammonium phosphate were purchased from Bachem (Torrance, CA) and J.T. Baker (Phillipsburg, NJ), respectively. Deionized water (18 MΩ) was prepared in-house with a Millipore Mill-Q water purification system (Millipore Bedford, MA). Ubiquitin (Sigma-Aldrich) was prepared at 15 μM in 20 mM sodium phosphate buffer. B-lactoglobulin A (Sigma-Aldrich) was prepared at 20 μM in 20 mM sodium phosphate, ammonium bicarbonate, and ammonium phosphate (J.T. Baker, Phillipsburg, NJ) buffers.

NMR Spectroscopy

Heteronuclear single quantum coherence (HSQC) spectra of the ^{15}N-labeled galectin-3 samples were acquired at 600 MHz using a gradient enhanced version of a standard HSQC experiment (gNHSQC from the Varian BioPack). A small number of t1 points (32) and a minimal number of scans were used to reduce acquisition times to three minutes and still allow observation at a reasonable S/N radio. Data were processed using nmrPipe and plotted using nmrDraw [76]. Unambiguous resonance assignments were taken from BMRB deposition # 4909 [77].

Electron Pulse Irradiation

20 μM buffered (pH=7) protein solutions (ubiquitin or β-lactoglobulin A) were subjected to electron pulses from Radiation Laboratory 2 MeV Van de Graaff accelerator by Ireneusz

Janik. Solutions were saturated prior irradiation either with air or a gas mixture of nitrous oxide/oxygen (N_2O/O_2; 4/1 v/v Mittler Supply Inc., ultrahigh purity, South Bend, IN). For initial studies with ubiquitin the dose was changed either by varying the beam current, via changes in heater settings of the dispenser cathode, at fixed electron pulsewidth (660 ns) or by varying the electron pulsewidth (480-660 ns) at fixed beam current. Typically to attain high levels of labeled protein, the dose had to be around 300 Gy achieved by a heater setting of 15 and we routinely used a dose in this range. A 250 μl syringe (Hamilton, Reno, NV, custom grinded to minimize the loss of electrons due to scattering in the glass wall) was placed in the front of the electron beam exit window. The beam was focused to a 2 mm diameter and was spatially adjusted to hit the very front volume of the syringe (**Figure 2.1**). From the glass coloration after irradiation we could determine the spread of the beam in the sample cell and estimate the total irradiated volume to be about 5 μl. After each pulse of electrons the piston was moved to release irradiated solution plus a 0.5 μL volume of unirradiated solution and refill the irradiation volume with the fresh protein solution. The released solution was directed to an empty 1 mL centrifuge vial or a 1 mL centrifuge vial prior filed with 50 μl of buffered (pH=7) 20 mM methionine amide (Bachem) to stop any chain oxidation processes. After 10 subsequent electron pulses the vial containing 100 μl of 10 μM oxidized protein solution was stopped, frozen and stored awaiting labeling analysis. In initial studies the frequency of pulses and corresponding flow in the irradiated volume was varied to establish contribution of secondary irradiation before contact with the quencher. Electron pulse frequency of 1 Hz was used for results presented in this report. The doses applied to the protein solutions were estimated with the Fricke dosimeter [78]. Oxygen saturated solutions of Fricke dosimeter were irradiated at the same conditions like the protein samples (beam current, electron pulsewidth, cell position) after protein oxidation process was

completed. Over twelve months of studies, the dose varied in the range of 30% at initially established accelerator settings i.e. the beam current and the electron pulsewidth.

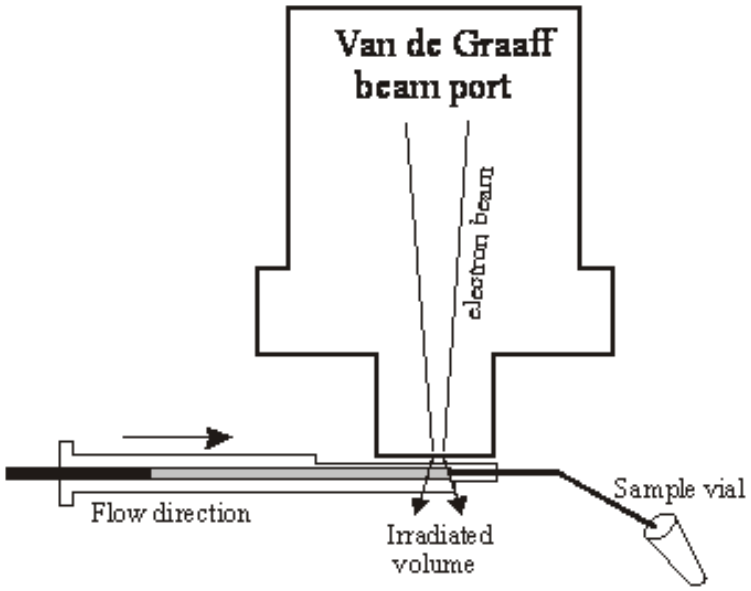


Figure 2.1. Schematic Representation of the Electron Pulse Protein Oxidation Setup.

Time-Resolved UV Spectroscopy

Pulse radiolysis experiments were performed using 100-1500 ns pulses from the Radiation Laboratory 8 MeV electron LINAC by Ireneusz Janik to obtain the range of doses comparable with doses applied in other protein oxidation experiments. Analyzing light from a pulsed 75 W xenon lamp (Photon Technology International, Birmingham, NJ) was selected using monochromator SPEX-270M. UV kinetics were measured at 250 nm where hydroxyl radicals extinction coefficient is ε=535 $M^{-1}cm^{-1}$ [79]. Samples of protein were prepared by dissolving β-lactoglobulin A to a concentration of 20 μM in ammonium phosphate buffer (pH=7). Buffer solutions were prepared in deionized water (18.2 MΩ-cm, Barnstead Nanopure System, Dubuque, IA). Prior to experiment, samples were bubbled with N_2O (Mittler Supply Inc., Ultrahigh Purity) at ambient conditions resulting in a final concentration of N_2O of 25 mM. To

improve the signal-to-noise ratio, UV signals were averaged over 15 consecutive traces. Due to transient absorption of quartz in the flow sample cell after the pulse of electrons (pronounced mostly at the highest doses) the blank traces for a given dose were collected in the empty cell purged with argon gas (Mittler Supply Inc., Ultrahigh Purity). These blank traces were subtracted from the corresponding sample traces before kinetic analysis. The UV transient absorption traces were fit using a system of first order kinetic differential equations model (fitting code written in Igor Pro 5.00, WaveMetrics Inc., Lake Oswego, OR) that incorporates most important reactions expected after pulse radiolysis of β-lactoglobulin A in N_2O saturated solutions. The initial concentration of hydroxyl radicals after the radiation pulse was taken as a sum of initial yields of hydroxyl radicals and hydrated electrons given that hydrated electrons convert to hydroxyl radicals within the duration of the pulse in N_2O saturated solutions (see reaction (7)). Extinction coefficients of hydrated electrons and hydroxyl radicals at 250 nm are comparable within the experimental error [80]. Thus, their rapid inter-conversion occurring in the first couple hundreds of nanoseconds does not obstruct the overall kinetic analysis extending up to 40 μs after the irradiation pulse. A fitting program was setup to fit all collected UV traces at once fitting the reaction rate of hydroxyl radicals with protein and the resulting product extinction coefficient, while keeping most other (known) parameters of the model fixed. For simplicity the rate constant of the protein radical decay was assumed to be first order since this reaction is many orders of magnitude slower than the initial decay of hydroxyl radicals in the monitored transient absorption time window.

LC-MS of Oxidized Protein

Intact ubiquitin samples were analyzed using a hand pulled fused silica (Technologies, Phoenix, AZ) spray column (75 μm x 10 cm; tip 15 ± 1μm) that was prepared by packing silica

C_{18} resins (Rainin Microsorb MV, 5 μm, 300 Å pore size) from a 50% isopropanol and 50% methanol slurry into the pulled fused silica capillary using a pressurized stainless steel bomb. Prior to reverse-phase HPLC, the column was equilibrated with 0.1% formic acid in water and the intact samples were loaded onto the column using a stainless steel bomb pressurized with nitrogen gas at 1,000 psi for 45 minutes. Liquid chromatography was initiated at a primary flow rate of 4 μL through the Agilent 1100 Series reversed-phase HPLC system (Agilent Technologies, Waldbronn, Germany) that ran through a splitter and resulted in a flow rate of 400 nL/min over the column with a 10 min rinse in 95% Buffer A (0.1% formic acid in water) followed by a 20-min linear gradient of 5 to 95% Buffer B (0.1% formic acid in acetonitrile). The spectra were acquired by nano-electrospray ionization on a Thermo Finnigan LTQ-FT mass spectrometer (San Jose, CA). The capillary temperature was 250°C and the spray voltage was 2.2 kV for ubiquitin.

For all intact β-lactoglobulin A samples, 8 μL were injected using an Agilent autosampler module (Agilent Technologies) over an Agilent ZORBAX 300SB C_{18} (150 x 0.3 mm, 5 μM particles) reverse-phase column (Agilent Technologies). The HPLC, directly coupled to the LTQ-FT for mass spectrometry, was run using a linear gradient of 95% Buffer A to 95% Buffer B over 20 minutes at a flow rate of 2 μL/min, followed by a 10 minute wash with 95% Buffer B and an 80 minute wash with 95% Buffer A. The capillary temperature was 250°C and the spray voltage was 2.1 kV for β-lactoglobulin A.

Tryptic Digestion of Oxidized β-Lactoglobulin A

One β-lactoglobulin A irradiated with 480 ns electron pulses was selected for tryptic digestion to determine the sites and amounts of oxidation. Ammonium bicarbonate (50 μL, 50 mM) and DTT (25 μL, 25 mM) were added to the β-lactoglobulin A sample (50 μL) and the

samples were incubated at 50°C for 3 hours to denature and reduce the protein. Iodoacetamide (25 μL, 90 mM) was added to the solution to carbamidomethylate the reduced disulfide bonds, and incubated in the dark at room temperature for 45 minutes. Sequencing grade modified trypsin (Promega, Madison, WI) was added (25 μL, 0.1 μg/μL) and incubated at 37ºC overnight while rotating to digest the protein samples. The samples were analyzed as is in triplicate and the remaining sample was stored at 0°C.

LC-MS/MS of Digested Oxidized β-Lactoglobulin A

8 μL of the oxidized β-lactoglobulin A tryptic digest were injected over an Agilent ZORBAX 300SB C_{18} (150 x 0.3 mm, 5 μM particles) reverse-phase column (Agilent Technologies). The HPLC, interfaced to an LTQ-FT mass spectrometer, was run using a linear gradient of 95% Buffer A to 60% Buffer B over 47 minutes then to 90% Buffer B over 15 minutes at a flow rate of 2 μL/min, followed by a 5 minute wash with 90% Buffer B and an 55 minute wash with 95% Buffer A. The capillary temperature was 250°C, and the spray voltage was 2.3 kV for β-lactoglobulin A tryptic digest. After the LC-MS/MS experiment, the measured peptides were screened computationally for different modifications using MASCOT in conjunction with ProteoIQ [81], and ByOnic [82]. Semi-tryptic peptides were included in the search and analysis to obtain better sequence coverage as well as to determine other sites of oxidation. All tandem mass spectra assignments and sites of oxidation verified manually due to sample and search space complexity.

MD simulations and solvent accessibility calculations

A crystal structure of β-lactoglobulin (PDB ID code 1BSY) [74] was retrieved from the Research Collaboratory for Structural Biology (RCSB) database [83]. In order to calculate the time-averaged solvent accessible surface area (<SASA>) of each side chain in β-lactoglobulin,

Olga Charvátová performed a 10-ns MD simulation of the protein was performed using the all-atom AMBER 8 [84] force field with the PARM99 protein parameters [85]. Prior to the MD simulation, all histidine residues were considered as neutral and were protonated at the Nε position. Nine Na^+ ions were added to neutralize charge and the protein was solvated by 5,417 TIP3P [86] water molecules, using the Protein Builder component of the GLYCAM-Web tool (http://www.glycam.com) [87], which employs the tLEaP [84] module of AMBER. The energy was minimized using the SANDER module [84] of AMBER by performing 5000 steps of steepest descent followed by 5000 steps of conjugate gradient minimization, again using GLYCAM-Web. The water molecules were then subjected to a simulated annealing protocol in which they were heated from 5 K to 300 K over a period of 50 ps, held at 300 K for 100 ps, before being cooled to 5 K over a final 50 ps. Following solvent annealing, the entire system was heated from 5 K and to 300 K over 100 ps with no restraints applied to the coordinates. A production MD data set was then collected for 10 ns with the temperature held constant at 300 K. All simulations were performed using the NPT ensemble, at 1 atm using a 2 fs integration time step, with the SHAKE algorithm [88] treatment of all hydrogen-containing bonds, and a unit dielectric constant. The atomic co-ordinates were stored every 10 ps for analysis, for a total of 1000 snapshots. SASA values were calculated for individual snapshots employing the NACCESS program [89] and averaged over the 10 ns simulation by in-house program to obtain <SASA> values and standard deviations for each of the 162 residues.

RESULTS AND DISCUSSION

Rapid NMR Spectroscopy of Proteins in Solution with H_2O_2

H_2O_2 photolysis is one of the fastest methods for producing hydroxyl radicals, but there is concern about sensitivity of proteins to the physical and chemical effects that H_2O_2 can impart.

Before dismissing the use of methods dependent on H_2O_2 for transition metal-free protein preparations, we felt it is important to assess the extent to which its presence affected conformation. Using time-resolved NMR spectroscopy we were able to observe changes in the conformation of a model protein, Galectin-3, in the presence of H_2O_2 versus H_2O. Galectin-3 is a small (15 kDa) stable protein that binds galactose terminated oligosaccharides. It is not suspected to be usually sensitive to hydrogen peroxide. However, it does have a cysteine residue that is not particularly solvent exposed. A three minute HSQC spectrum of galectin-3 was recorded in the presence and absence of H_2O_2 (**Figure 2.2**). The small shifts labeled (**Figure 2.2**) in the HSQC spectrum can be interpreted as a conformational change. Many of the shifted peaks are assigned to resonances near the lactose binding site. Due to the very short time scale (3 minutes) and known conformational flexibility of the binding site, direct oxidation by hydrogen peroxide is not thought to be the driving force behind this noted conformation change. The subtle shifts are more likely due to physical association of H_2O_2 versus water. Some of the perturbed resonances do, however, cluster near C173, the partially buried cysteine that is expected to be most sensitive to oxidative chemical events. At longer times, perturbations of additional resonances more indicative of rapid two-electron oxidation events are seen (data not shown). An LC-MS analysis of Galectin-3 after a three minute H_2O_2 exposure determined the protein is not unusually sensitive to hydrogen peroxide, as no noticeable oxidation products were detected (data not shown). Regardless of the origin of the shifts, the data illustrate that proteins are not necessarily probed in their native conformation when using methods that involve H_2O_2 to produce hydroxyl radicals even in the absence of peroxide-mediated oxidation events, and that these perturbations can occur on short time scales.

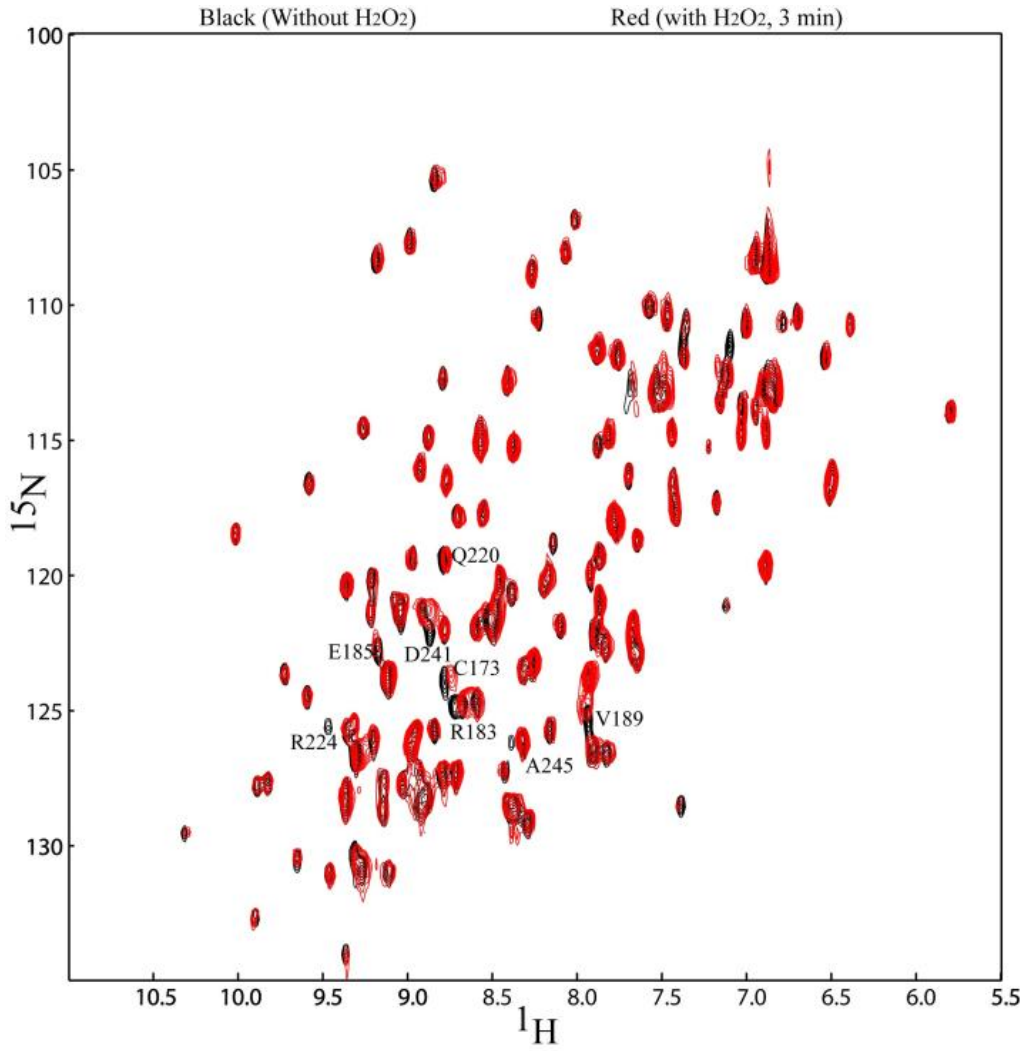


Figure 2.2. HSQC Spectrum of Galectin-3. A three minute HSQC spectrum of galectin-3 in the absence (red) and presence (black) of 1% hydrogen peroxide (v/v).

Protein Oxidation

High energy electrons passing through the aqueous solution ionize/excite water molecules forming a number of transient species and stable products according to equation (1).

$$H_2O \xrightarrow{irr} e_{aq}^-, {}^{\bullet}OH, H^{\bullet}, H^+, OH^-, H_2O_2, H_2 \qquad (1)$$

Among all species generated during water radiolysis, hydroxyl radicals ($^{\bullet}OH$) and hydrated electrons (e_{aq}^-) are the most reactive with peptides and proteins [57]. When the radiolysis is performed in the presence of oxygen, superoxide radicals and its acidic form of hydroperoxyl radicals are generated in reaction (2) suppressing the effect of hydrated electrons on the proteins.

$$e_{aq}\text{-}\ (H^{\bullet}) + O_2 \rightarrow O_2^{\bullet-}(HO^{\bullet}_2) \quad (2)$$

$$HO^{\bullet}_2 \rightleftarrows O_2^{\bullet-} + H^+ \quad (3)$$

Since the pKa of $HO_2^{\bullet}$ radical is around 4.8 (reaction (3)) and oxidation experiments are performed at buffered pH 7, most of the hydroperoxyl radicals are in the form of a superoxide radical ($HO_2^{\bullet}/O_2^{\bullet-}$). $HO_2^{\bullet}/O_2^{\bullet-}$ radicals are known to react with amino acids at very low rate constants ranging from 10-10^2 dm^3mol^{-1} [90, 91] therefore one can expect very little effect of $HO_2^{\bullet}/O_2^{\bullet-}$ radicals with the metal free proteins that are subject of the current studies. $HO_2^{\bullet}/O_2^{\bullet-}$ radicals undergo comparatively rapid disproportionation to H_2O_2 and O_2 in reactions (4) and (5) via a pH dependent mechanism that also involves reaction (3).

$$HO^{\bullet}_2 + HO^{\bullet}_2 \rightarrow H_2O_2 + O_2 \quad (4)$$

$$HO^{\bullet}_2 + O_2^{\bullet-} \rightarrow H_2O_2 + O_2 + OH^- \quad (5)$$

Based on the kinetic equation for this decay mechanism [92, 93], one can estimate the rate constant for disproportionation at pH=7 to be about $2k_{obs}=1.2\times10^6$ dm^3mol^{-1}. This indicates that if left alone $HO^{\bullet}_2/O_2^{\bullet-}$ radicals can last milliseconds after the irradiation pulse; however, it is not long enough to cause any significant oxidation of the available protein amino acids. Converting $^{\bullet}OH$ radicals leading to water and recovering initially consumed oxygen in reaction (6) is an additional reaction that consumes $HO_2^{\bullet}/O_2^{\bullet-}$ radicals before they can react with the protein.

$$HO_2^{\bullet}/O_2^{\bullet-} + {}^{\bullet}OH \rightarrow H_2O + O_2 \quad (6)$$

Since the rate constant of reaction (6) is as fast as $1\times10^{10}dm^3mol^{-1}$, it inconveniently consumes some of important $^{\bullet}OH$ radicals for oxidation.

In the initial microseconds following pulse irradiation, the balance of the above reactions leads mostly to the very reactive $^{\bullet}OH$ radicals and relatively nonreactive $HO_2^{\bullet}/O_2^{\bullet-}$ radicals. $^{\bullet}OH$ radical reaction with amino acids is well established and proceeds with reaction rate constants [64] varying from 10^{10} to 10^7 dm^3mol^{-1}. From a kinetic point of view the majority of free $^{\bullet}OH$ radicals that happen to be close to the protein surface will interact with the fastest reacting amino acids; those include amino acids containing sulfur as well as unsaturated/aromatic side groups. Amino acids with the aliphatic side groups react with $^{\bullet}OH$ radicals relatively slower. In oxygenated solutions protein surface radicals add oxygen forming the corresponding peroxyl radicals. The mechanism of initial sulfur centered radicals is not completely explained yet but leads to well characterized and recognizable products [94, 95]. The fate of the relatively fast surface located peroxyl radicals depends entirely on their nature and surroundings. Some of the aromatic peroxyl radicals will eliminate $HO^{\bullet}_2$ radicals leaving behind a stable hydroxyl group [94] or a carbonyl group [96]. C-centered aliphatic peroxyl radicals can undergo a bimolecular recombination leading to the creation of hydroxyl and/or carbonyl group on the parent carbon atom, or transfer the radical to form a hydroperoxyl group that propagates oxidation to neighboring amino acids [96]. Those reactions are well understood and summarized in the number of articles and textbooks [94, 96, 97]. In either case in diluted protein solution, sites of oxidation should be anchored to the place where the initial $^{\bullet}OH$ radical attack occurred. Chances are that some of the initially created peroxyl radicals can survive on the surface and do not undergo any of the mentioned transformations. In this case they can react with lengthy present $HO^{\bullet}_2/O_2^{\bullet-}$ radicals producing hydroxyl group, oxygen and hydrogen peroxide [98]. The presence

of oxygen is quite essential in the $^{\bullet}OH$ radical induced protein oxidation even though the fastest surface labeling occurs by way of hydroxylation with $^{\bullet}OH$ radicals [97].

Our initial experiments were performed with aerated buffered ubiquitin solutions in order to establish optimal doses and concentrations for maximum oxidation. We varied the pulsewidth and current of the electron beam, as well as, the dissolved gas in the sample. In order to double the concentration of hydroxyl radicals in the pulse irradiation we saturated our solutions with nitrous oxide/oxygen combination (N_2O/O_2 4/1 v/v). In this case the oxygen concentration remained on the same level as in the air saturated solutions, but additionally we benefited from the well-known reaction (7) of fast conversion of hydrated electrons to hydroxyl radicals [64]. Since the solubility of N_2O in water is about 20 times higher than oxygen, most hydrated electrons convert to hydroxyl radicals within the electron pulse, doubling the hydroxyl radical concentration.

$$e_{aq}^{-} + N_2O \xrightarrow{H_2O} {}^{\bullet}OH + OH^{-} + N_2 \qquad (7)$$

The influence of gas composition on ubiquitin oxidation is presented in **Figure 2.3**. Due to the large dose applied, we see very substantial oxidation in both air and N_2O/O_2 saturated solutions. The solutions containing N_2O show higher abundance of heavily oxidized ubiquitin. The initial experiments showed that we can easily control any excess of labeled protein by varying the dose and the gas composition. Additionally, we noted that the increase of the applied radiation dose does not increase the level of oxidation proportionally. It is mostly related to the oxygen uptake in the solution after the pulsed irradiation. In aerated or N_2O/O_2 saturated solutions, oxygen concentrations are on the order of 2.7×10^{-4} mol/dm^3 that corresponds to the concentration of water radiolysis transient species ($^{\bullet}OH$, $H^{\bullet}$, e_{aq}^{-}) achievable at roughly 500 Gy of absorbed dose. Considerably increasing the dose above 500 Gy decreases the amount of available oxygen and

leads to undesired reactions of hydrated electrons with the protein surface amino acids. Lack of oxygen can distort the formation process of carbon centered peroxyl radicals as precursors of hydroperoxides [98] and other oxidation products on the protein surface. The quantitative studies of the dose dependence on the pulse radiolysis $^{\bullet}OH$ protein footprinting were not the subject of the current investigation but will be addressed in the future reports.

LC-FTMS Analysis of Ubiquitin and β-lactoglobulin A

Sodium phosphate-buffered ubiquitin was used as a model to determine if protein oxidation could be achieved by this method. LC-FTMS of ubiquitin (**Figure 2.3**) shows the range of intact protein oxidation obtained with and without N_2O, and with and without methionine amide, a scavenger of secondary oxidants. Ubiquitin was also irradiated with various pulsewidths (480 or 660 ns) and electron beam amperages (doses) (**Figure 2.4**). In the absence of methionine amide (**Figure 2.3b**) no unoxidized ubiquitin is detected even at relatively low radiation doses despite the fact that 10% of the solution was unirradiated. However, in the presence of methionine amide (**Figure 2.3c and d**), unoxidized ubiquitin from the unirradiated fraction of the solution was still detected despite heavy amounts of oxidation. These results indicate that the methionine amide is sufficient to prevent oxidation of the unirradiated ubiquitin even at high radiation dosages, while secondary oxidation sufficient to consume all of the unirradiated ubiquitin occurs in the absence of the methionine amide even at relatively low radiation dosages (**Figure 2.4**). Quenching is necessary to ensure the vast majority of labeling occurs on a microsecond time scale and before the protein can unfold, rather than by secondary oxidants (e.g. peroxides in the presence of metal ion traces of UV light) on the second to minute time scale. With electron beam pulsewidths under 700 ns, we detected an extensive amount of

oxidation (**Figure 2.3d**) using a sub-microsecond electron pulse. By adjusting various parameters of the irradiation (pulsewidth, dissolve N_2O and dose), we are able to control the amount of oxidation from very little oxidation to extensive oxidation (**Figure 2.4**).

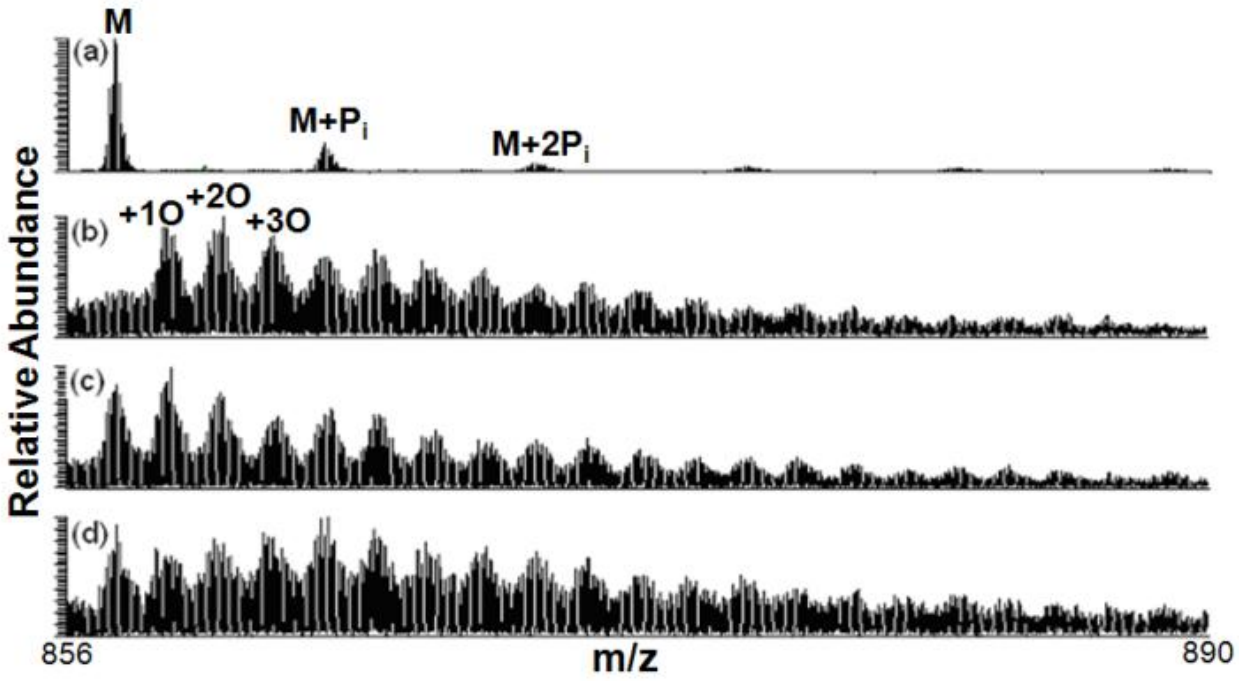
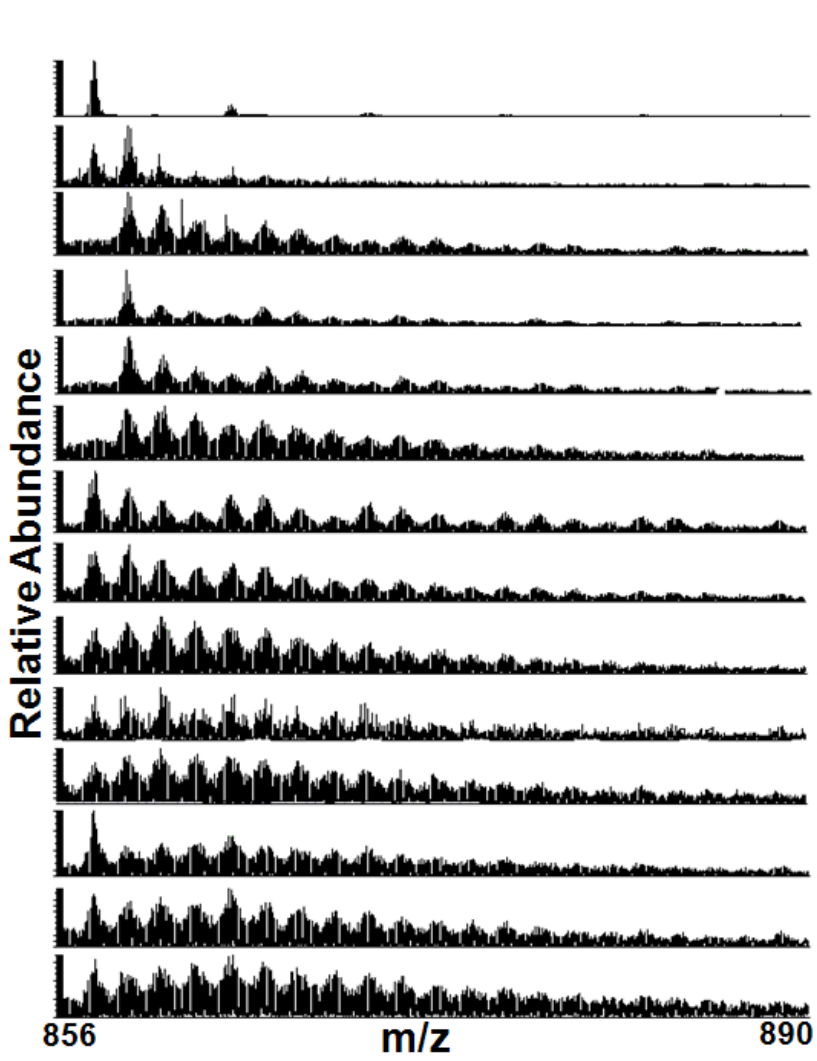


Figure 2.3. LC-MS of the 10+ Charge State of Intact Ubiquitin. Phosphate adducts were seen from sodium phosphate buffer. (a) Unirradiated sodium phosphate buffered ubiquitin. Ubiquitin irradiated with 800 ns pulsewidth in (b) air, (c) air + methionine amide, and (d) N_2O/O_2 + methionine amide.



Conditions	Pulsewidth	Heater Setting
intact	n/a	n/a
Air	400ns	13
Air	400ns	14
Air	400ns	15
Air	800ns	13
Air	800ns	14
Air + amide	800ns	13
Air + amide	800ns	14
Air + amide	800ns	15
N2O/O2 + amide	800ns	13
N2O/O2 + amide	800ns	14
N2O/O2 + amide	800ns	15
N2O/O2 + amide	400ns	14
N2O/O2 + amide	400ns	15

Figure 2.4. LC-MS of the 10+ Charge State of All Intact Ubiquitin Investigated. Phosphate adducts were seen from sodium phosphate buffer. The conditions of irradiation are listed to the right of each spectrum.

In initial experiments on β-lactoglobulin A, a sodium phosphate buffer was irradiated and we detected sodium adduction (M + n22) in each sample despite multiple attempts to desalt the solution. This interfered with our data analysis (data not shown). Ammonium bicarbonate and ammonium phosphate buffers were used for subsequent experiments. Upon irradiation, the amount of oxidation in the ammonium bicarbonate-buffered samples was significantly lower,

probably due to radical scavenging by the bicarbonate buffer. Also, an adduct or modification of unknown origin (observed mass shift of 174) is present in all ammonium bicarbonate-buffered β-lactoglobulin A samples (data not shown). Ammonium phosphate buffer gave the best results for intact β-lactoglobulin A, as it did not scavenge the radical, nor did it result in sodium or other adducts (**Figure 2.5**).

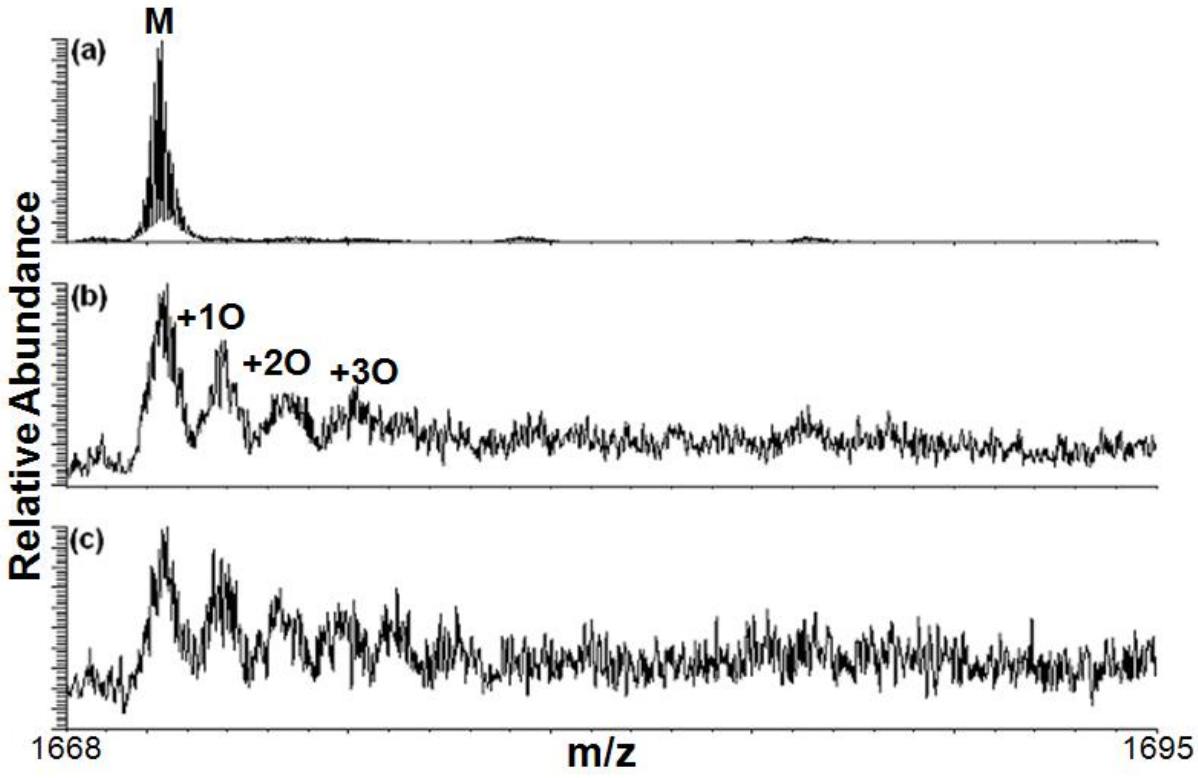


Figure 2.5. LC-MS of the 13+ Charge State of Intact β-Lactoglobulin. (a) Unirradiated ammonium phosphate buffered β-Lactoglobulin. β-Lactoglobulin was irradiated for 200 ns (b) and 400 ns (c) in air + methionine amide.

Time-Resolved UV Absorbance Spectra Reveal the Hydroxyl Radical Lifetime

Since proteins can unfold significantly on the order of a few microseconds [99], labeling chemistry should be completed faster than that to prevent labeling of oxidatively unfolded protein. One way to observe the progress of labeling chemistry is to monitor the decay of hydroxyl radicals in the microsecond time scale after their formation during the pulse of electrons. The lifetime of hydroxyl radicals in aqueous solutions after each pulse of electrons can be estimated based on the known reaction rate constants of hydroxyl radicals with solutes [64]. In pure buffered water hydroxyl radicals recombine with each other since reactions with buffer components are usually much slower. The second order hydroxyl radical recombination (8) competes with slower first order reactions especially at higher doses when higher concentrations of hydroxyl radicals are produced.

$$^{\bullet}OH + {}^{\bullet}OH \rightarrow H_2O_2 \quad (8)$$

We performed time resolved pulse radiolysis experiments with transient absorption detection at 250 nm to look at the upper limit of hydroxyl radical life time in buffered N_2O saturated (no oxygen present) solutions without proteins. The traces from these experiments are provided in **Figure 2.6**. The initial increase of signal corresponds to formation of hydroxyl radicals and is related to their concentration via an extinction coefficient. Analyzing the transient absorption one can see that the apparent signal decays very quickly and within 8 μs reaches the plateau of H_2O_2 absorption. Reaction (9) of hydroxyl radicals with H_2O_2 is relatively slow (2.7×10^7 $mol^{-1}dm^3$) [64] and consumes only about 1% of the formed H_2O_2 40 microseconds after the irradiation pulse.

$$^{\bullet}OH + H_2O_2 \rightarrow HO_2^{\bullet}/O_2^{\bullet-} + H_2O \quad (9)$$

However, this minimal concentration of superoxide radical contributes some 20% to the final absorption, since the extinction coefficient of $HO_2^{\bullet}/O_2^{\bullet-}$ radicals greatly exceeds the extinction coefficient of hydrogen peroxide. Contribution of $HO_2^{\bullet}/O_2^{\bullet-}$ radical absorption increases, obviously, with the increase of the applied dose since reaction (9) proceeds more efficiently but, the final concentrations of formed $HO_2^{\bullet}/O_2^{\bullet-}$ are always at a lower fraction than the initial $^{\bullet}OH$ radical concentration. Extension of the pulsewidth from 400 ns to 1500 ns increases the dose almost 4 times but not the peak transient absorption of hydroxyl radicals (**Figure 2.6**). Reaction (8) effectively decreases the concentration of hydroxyl radicals during the 1500 ns pulse and the rate of hydroxyl radical formation approaches the rate of hydroxyl radical decay giving steady state concentration of hydroxyl radicals just only 50% higher than the peak concentration of hydroxyl radicals after 400 ns pulse. It is important to note that the final absorption of H_2O_2 and $HO_2^{\bullet}/O_2^{\bullet-}$ after the completion of reaction (8) is more than 3 times higher for 1500 ns pulse than 400ns (**Figure 2.7**), which confirms that reaction (8) with reaction (9) is the main channel of hydroxyl radicals decay in the protein free N_2O saturated buffered water. Based on the extinction coefficients at 250 nm of species present, we determined that the time for hydroxyl radicals to decay to 0.2 μM (1% of protein in oxidation experiments) is about 20.6 μs + pulsewidth of electrons. It is very fast but still comparable with the duration of some protein unfolding events [100]. Obviously, in the presence of protein, hydroxyl radicals will not stay free for as long because they will react with the protein surface groups. For most amino acids side changes, the reactions are very fast [64] and should contribute very effectively to the decrease of hydroxyl radical concentration. To confirm the effect of protein, we performed an experiment in which β-lactoglobulin A was added to the buffered solution and the decay of hydroxyl radicals was monitored. The UV traces in **Figure 2.6** show that after the addition of 4 μM of protein the

apparent UV signal reaches a higher plateau than in buffer alone suggesting that some other absorbing product is being formed in addition to H_2O_2 and $HO_2^{\bullet}/O_2^{\bullet-}$. In fact, a further increase of the protein concentration results in an even higher increase in the amplitude of the formed product. The apparent UV signal results from a sum of contributions from various transient species produced upon hydroxyl radical reaction with the protein surface and can be represented in symbolic reaction (10). The initial step of hydroxyl radical reaction with protein is formation of a protein radical:

$$OH + \text{protein} \rightarrow \text{protein radical} \qquad (10)$$

The protein radical in oxygen-free, N_2O saturated solution undergoes further reactions leading to the formation of the final oxidized protein (in case of radicals formed on the sulfur containing or unsaturated/aromatic residues) or decays in the bimolecular processes of disproportionation/dimerization [57, 94]. For the time being we are interested in the fate of hydroxyl radicals, and based on the several UV transient signals collected for different protein concentrations, we constructed a simple model to extract the portion of the UV signal that is related to the decay of hydroxyl radicals. The applied global fitting model allowed us to estimate the rate constant of reaction (9) and the overall extinction coefficient of the protein radical for the studied system. All other parameters used to obtain fitting results are tabulated in **Table 2.1**. The time profiles of most contributing species absorbing at 250 nm resulting from global fitting are shown on **Figure 2.7**. This basic experiment proved that the lifetime of hydroxyl radicals decreases considerably after addition of protein. In addition to the lifetime of hydroxyl radicals, the transient absorption experiment has also shown the importance of properly choosing the concentration of the protein in the pulse irradiation experiment. Lower concentrations of protein may be necessary for some experiments and can increase the extent of protein labeling [99], but

at the same time the lower protein concentration prolongs the hydroxyl radical's lifetime, increasing the chance of oxidizing a previously unfolded residue and necessitating the use of an exogenous quencher to limit the hydroxyl radical lifetime [29].

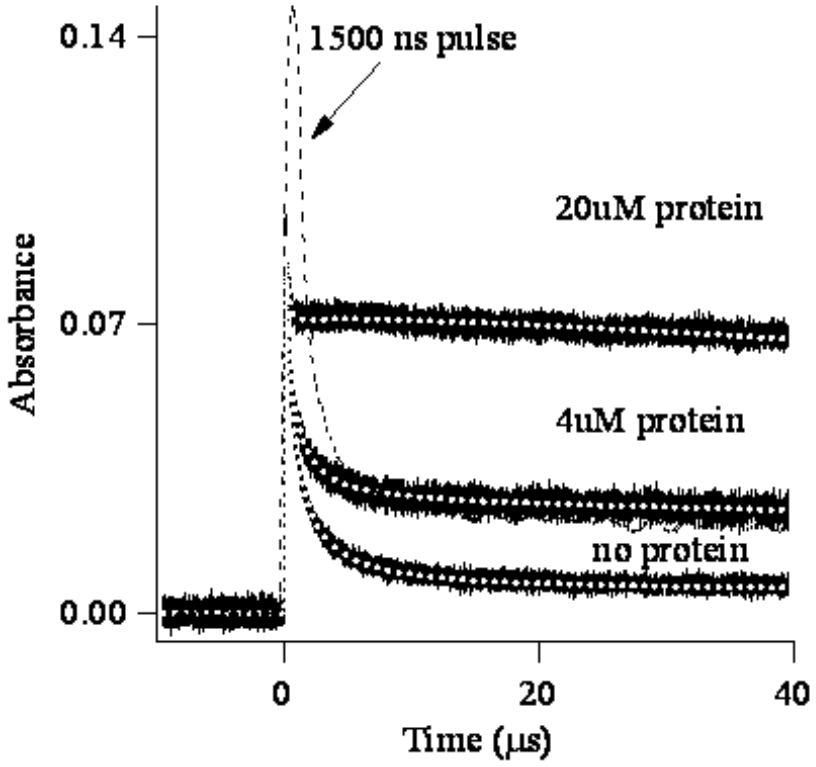


Figure 2.6. Transient Absorption Signals. Signals monitored after the pulse of electrons at 250nm. Solid lines-signals recorded after 400ns pulse; dashed line-signal recorded after 1500ns pulse; white dotted line-represents global fit of the transient absorption signals. All solutions were fixed at pH=7 with ammonium phosphate buffer and contained 25mM N_2O, the concentration of β-lactoglobulin was varied like indicated on the figure.

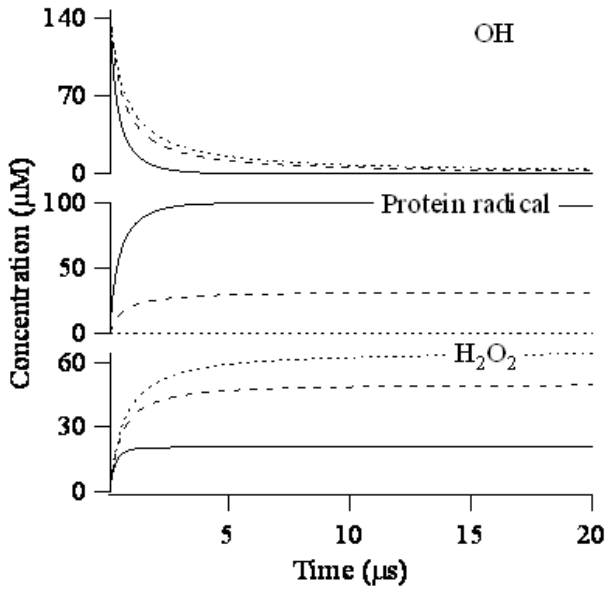


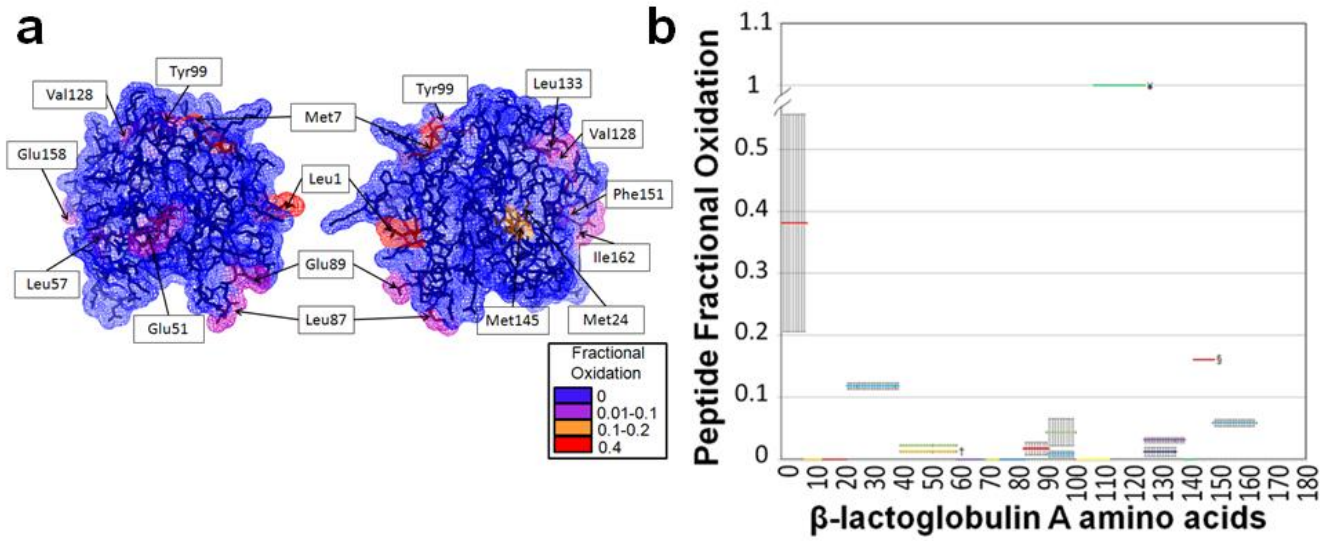
Figure 2.7. Species Concentration Profiles. Concentration profiles of species contributing most to the transient absorption signals at 250nm based on the global fitting results. Dotted lines - solutions with no protein; dashed lines-solutions with 4μM of β-lactogloglobulin; solid lines- 20μM of β-lactogloglobulin.

Table 2.1. Reactions and Parameters Used in the Transient Absorption Global Fitting

Reaction	**Product extinction coefficient (M^{-1} cm^{-1})**	**Rate constant (M^{-1} dm^{-3})**
$OH + OH \rightarrow H_2O_2$	26	5.0×10^{9}
$OH + H_2O_2 \rightarrow O_2^- + H_2O + H^+$	1890	3.7×10^{7}
$OH + H \rightarrow H_2O$	non absorbing	7.0×10^{9}
$H + H \rightarrow H_2$	non absorbing	5.0×10^{9}
$OH + PO_4^{3-} \rightarrow$ products	non absorbing	1.0×10^{5}
OH + protein → protein radical	691	1.24×10^{10}
H + protein → protein radical	691	8.7×10^{8}
protein radical → products	non absorbing	2.9×10^{3}

LC-MS/MS Analysis of Oxidized β-lactoglobulin A Tryptic and Semi-Tryptic Fragments

The β-lactoglobulin A sample irradiated with the dose of 260 Gy using a pulse length of 480 ns, was denatured, reduced, carbamidomethylated, and digested using trypsin. The peptide mixture was loaded on a C_{18} capillary column for LC-MS/MS analysis to locate sites of oxidation. The overall sequence coverage, including all oxidized and unoxidized peptides identified in the LC-MS/MS run, was 100%, with manual verification of all tandem mass spectra assignments. Fourteen oxidation sites were identified for β-lactoglobulin A in these LC-MS/MS experiments; however, one semi-tryptic peptide was detected as completely oxidized and although oxidation could not be mapped to a specific site(s), peptide 108-ENSAEPEQSLACQCLVR-124 contains many reactive residues with large <SASA> values. The identified oxidation sites are shown on the X-ray crystal structure of β-lactoglobulin A [74] in **Figure 2.8**.



[†]Mass shift of -30 which is characteristic of oxidized glutamic acid residue
[¥]Mass shift of +28 (two +14 shift from aliphatic residues)
[§]Only seen in one replicate

Figure 2.8. Hydroxyl Radical Footprinting β–Lactoglobulin Results. (a) Two views of X-ray crystal structure of β-lactoglobulin A with waters removed for clarity created using PyMOL (PDB identifier 1BSY). Oxidized residues confirmed by LC-MS/MS are colored with red. (b) Quantitation of peptide fractional oxidation. All oxidation mass shifts are +16 and seen in at least twice in the triplicate runs unless otherwise noted. Error bars represent the standard deviation of identical triplicate runs.

The surface average solvent accessibility (<SASA>) value of each amino acid residue for all 162 residues was obtained from 1000 snapshots over a 10 ns MD simulation (**Figure 2.9**). While 10 ns is too short a period to simulate large-scale protein backbone motions, it is adequate to capture much of the motion of the side chains, and provides a useful approach for achieving statistically significant average SASA values (<SASA>) on a given time scale. In addition, the use of MD simulations on this time scale provides the opportunity to probe the extent to which a

model based essentially on the fully folded protein is able to describe the data from an experiment performed on the millisecond time scale. The <SASA> values for each mapped oxidation site on β-lactoglobulin A are listed in **Table 2.2**. Using this method we are able to determine that, with the exception of M24, all amino acids determined to be oxidized under our conditions were predicted to be suitably solvent accessible by MD simulations.

The sulfur-containing residues (Met and Cys) are the most susceptible to oxidative modifications [60]. No cysteine residues were conclusively identified as sites of oxidation; even though they are chemically reactive, they are buried within the protein structure, and therefore should react only slowly with a hydroxyl radical if the protein retains its folded form during the labeling process. Three of the four methionine residues found in β-lactoglobulin A, Met7, Met24 and Met145, were identified as sites of oxidation with a mass shift of +16. Met24 has a very small <SASA> value of 0.21 $Å^2$, but Met7 and Met145 have small to moderate solvent accessibility (<SASA> value = 35.88 and 6.7 $Å^2$ respectively). Met107 also has a large <SASA> value and was not detected as oxidized by our measurements. However, it is present on a large peptide that was determined to be the site of multiple semi-tryptic cleavages and unusual oxidation events; it is quite possible that Met107 was oxidized quite readily, but that we are unable to resolve and identify the oxidized form of the peptide from the MS/MS spectra. All other sites of oxidation identified had large <SASA> values (**Table 2.2**), and can be clearly seen on the surface of the protein (**Figure 2.8**); importantly, all of the oxidized sites were predicted by MD simulation to be among the most highly accessible residues of each type within the protein (e.g. out of twenty-two total leucine residues, the four leucine residues oxidized were all among the five most solvent accessible according to MD simulations), with the exception of Met24 (**Figure 2.9**).

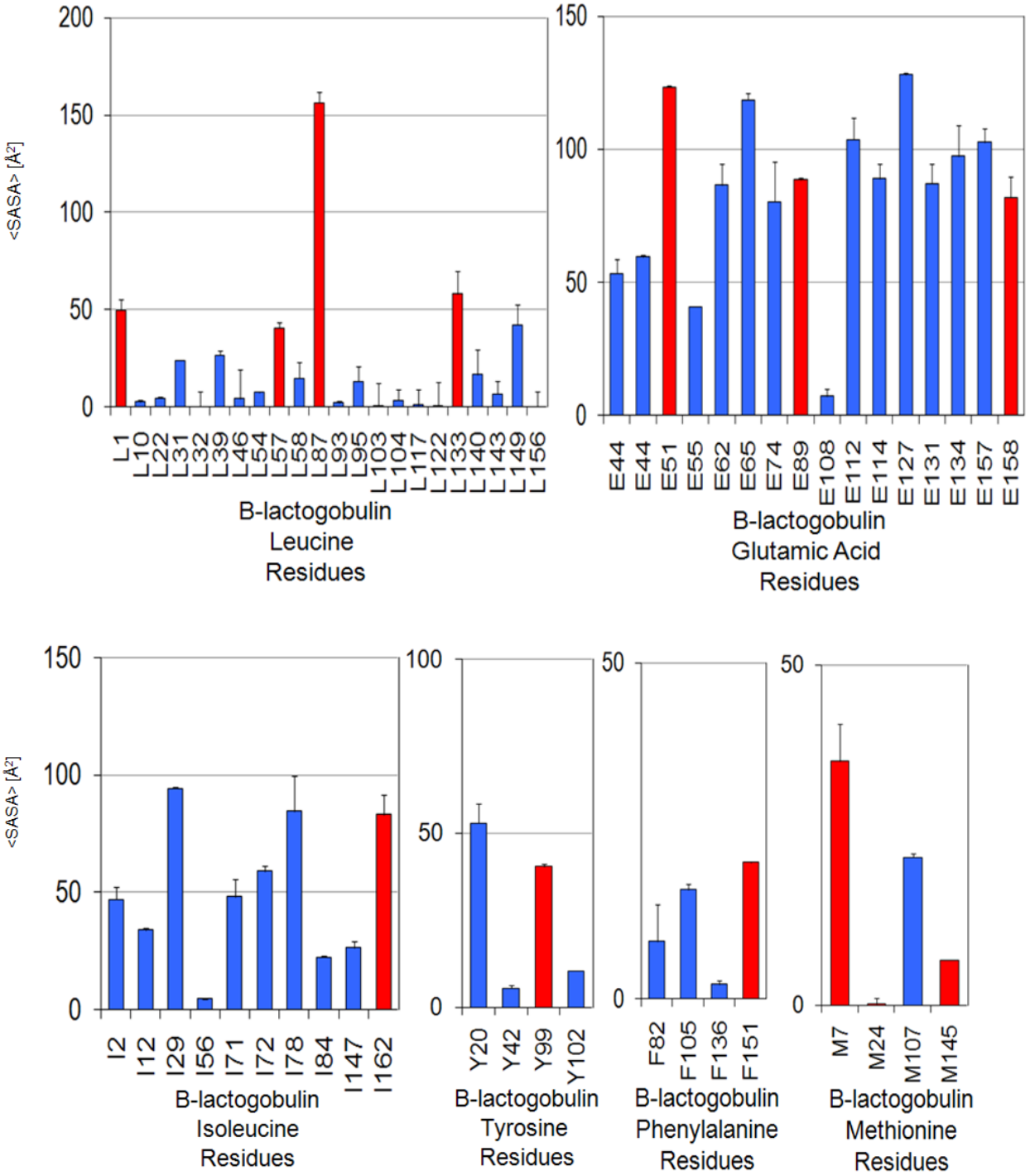


Figure 2.9. Amino Acid Surface Average Solvent Accessibility. <SASA> value for the amino acid groups of β-lactoglobulin that contained at least one oxidized residue. The oxidized residues are colored red.

Table 2.2. Average Solvent Accessibility of Oxidized Residues in β-Lactoglobulin.

Peptide (residues)	Peptide Fractional Oxidation	Site(s) of Oxidation	<SASA> ($Å^2$)
1-8	0.3861±0.1686	L1	49.78
		M7	35.88
23-40	0.1184±0.0052	M24	0.21
41-60	0.0127±0.0003[†]	E51	123.54
	0.0227±0.0009	L57	40.63
84-91	0.0174±0.0103	L87	156.27
		E89	88.65
92-100	0.0091±0.0043	Y99	40.51
92-101	0.0437±0.0218		
125-135	0.0124±0.0072	V128	28.86
125-138	0.0310±0.0040	L133	58.05
142-148	0.1607[§]	M145	6.7
149-162	0.0590±0.0057	F151	20.35
		E158	81.89
		I162	83.35

[†]Mass shift of -30 which is characteristic of oxidized glutamic acid residue
[§]Only seen in one replicate

This leaves an unresolved question: why was Met24 oxidized under these conditions? One potential cause is incomplete scavenging of secondary oxidants, which have been noted to lead to uncontrolled oxidation of methionine [60]. However, these same secondary oxidants would also be capable of oxidizing cysteine, which is not detected. A more likely possibility is that the protein undergoes larger backbone motions, in the region of M24, on the time scale of the experiment, than are captured in the 10 ns MD simulation. In that case, the oxidation of M24 would be larger than predicted by the <SASA> values. It is important to note here that by using

average <SASA> values, rather than those from a single protein structure, the possibility that this discrepancy is related to an anecdotal orientation of the side chain of M24 is greatly reduced.

Because most of the oxidized peptides contain more than one oxidation site, accurate quantitation of the amount of oxidation at individual sites is currently not possible. We were able to quantitate the amount of oxidized versus unoxidized peptide using the triplicate LC-MS/MS runs to determine the peak area and the standard deviation of the MS total ion chromatogram of a given m/z for each peptide and its oxidized form(s). The peak areas were used to calculate peptide fractional oxidation by dividing the peak area of the oxidized peptide by the sum of the peak areas of the oxidized and unoxidized forms of that peptide. The peptide fractional oxidation is shown for each β-lactoglobulin A peptide in **Figure 2.8b** and values are listed in **Table 2.2**. The sites of oxidation that could be determined by MS/MS analysis are also listed in **Table 2.2**; however, while these identified sites are major sites of oxidation for the peptide, they may not be the only substantial sites of oxidation. It is important to note that, as is all LC/MS experiments of complex mixtures, it is quite possible that major oxidation products exist that were not detected or that did not fragment sufficiently well to determine the site(s) of oxidation.

CONCLUSIONS

The overall purpose of this study was to demonstrate a method for hydroxyl radical footprinting of proteins in solution that is completed faster than large scale oxidation-induced conformational changes can occur, without requiring a precursor oxidant. Time-resolved NMR spectroscopy illustrates that the presence of hydrogen peroxide can result in uncontrolled oxidation or peroxide-induced conformational changes, causing the labeling of a non-native protein structure. Additionally, the presence of redox-active transition metals in solution with the hydrogen peroxide can lead to uncontrolled metal catalyzed oxidation [41], and in many cases it

is not possible to completely purify redox-active transition metals from the protein-ligand preparation. The development of a hydrogen peroxide-free method of hydroxyl radical footprinting that is capable of completing the labeling reaction faster than the protein can unfold increases the applicability of hydroxyl radical footprinting technology to peroxide-sensitive proteins, while still allowing for extensive labeling of the native conformation. As shown from the results presented, the electron accelerator pulsed water radiolysis method is suitable for heavy, controlled oxidation of proteins and oxidation of solvent accessible residues. Using an electron beam pulse, controllable protein oxidation can be obtained by adjusting variables such as pulsewidth, dose, and dissolved gas. Time–resolved UV spectroscopy indicates that the most reactive radical species are consumed in less than 2 μs without a quencher, and this rapid time scale allows for extensive oxidation before the protein can unfold due to the modifications. Our ability to heavily oxidize the protein without concern for oxidation-induced unfolding allows us to detect a much greater amount of oxidation than previously-reported results on the same protein [33], allowing for higher resolution hydroxyl radical protein footprinting data. MS/MS analysis and MD simulations indicate that all oxidation sites except one (Met24) identified by this method are moderately to highly solvent accessible. Additionally, as a rapid oxidation method, electron pulse water radiolysis has broad applicability in time-resolved structural studies, most notably UV-based spectroscopy methods where the high concentrations of hydrogen peroxide necessary for flash photolysis of peroxide can confound spectroscopy. Such a pulse labeling method will prove to be extremely useful in designing time-resolved spectroscopic studies of oxidation-induced protein unfolding, which is important for understanding the biophysical basis of oxidative stress-induced protein inactivation.

ACKNOWLEDGEMENTS

This research was supported by the National Center for Research Resources of the National Institute of Health (P41RR005351) and the Department of Energy, Office of Basic Sciences.

CHAPTER 3

CONFORMATIONAL ANALYSIS OF THERAPEUTIC PROTEINS BY HYDROXYL RADICAL PROTEIN FOOTPRINTING[2]

ABSTRACT

Unlike small molecule drugs, therapeutic protein pharmaceuticals must not only have the correct amino acid sequence and modifications, but also the correct conformation to ensure safety and efficacy. Here, we describe a method for comparison of therapeutic protein conformations by hydroxyl radical protein footprinting using liquid chromatography-mass spectrometry (LC-MS) as an analytical platform. Hydroxyl radical protein footprinting allows for rapid analysis of the conformation of therapeutic proteins based on the apparent rate of oxidation of various amino acids by hydroxyl radicals generated in situ. Conformations of Neupogen®, a patented granulocyte colony-stimulating factor (GCSF), were compared to several expired samples of recombinant GCSF, as well as heat-treated Neupogen®. Conformations of different samples of the therapeutic proteins interferon α-2A and erythropoietin were also compared. Differences in the hydroxyl radical footprint were measured between Neupogen® and the expired or mishandled GCSF samples, and confirmed by circular dichroism spectroscopy. Samples that had identical circular dichroism spectra were also found to be indistinguishable by hydroxyl radical footprinting. The method is applicable to a wide variety of therapeutic proteins and formulations through the use of separations techniques to clean up the protein samples after radical oxidation. The reaction products are stable, allowing for flexibility in sample handling, as

[2] Watson, C., Sharp, J.S. Accepted to *American Association of Pharmaceutical Scientists Journal.* Reprinted here with permission of the publisher.

well as archiving and reanalysis of samples. Initial screening can be performed on small amounts of therapeutic protein with minimal training in LC-MS, but samples with structural differences from the reference can be more carefully analyzed by LC-MS/MS to attain higher spatial resolution, which can aid in engineering and troubleshooting.

INTRODUCTION

Since the introduction of recombinant human insulin in 1982, therapeutic protein pharmaceuticals have grown into an estimated $102.4 billion industry in 2011, with more than 120 drugs on the market. Assays for protein pharmaceuticals have not kept pace because of the inherent complexity and variability of the proteins and their biological manufacturing processes. Unlike small molecules, therapeutic protein pharmaceuticals must have the proper primary structure of amino acid sequence, post-translational modification(s), disulfide bond formation, and also maintain the proper three-dimensional conformation. The conformation of a therapeutic protein is often sensitive to small changes of many factors in its production, formulation and handling, including temperature, pH, and buffer composition. Therapeutic proteins with improper conformation can lose efficacy, or even become highly toxic. For example, subcutaneously injected Eprex®, a recombinant human erythropoietin pharmaceutical manufactured by Johnson and Johnson, was linked to an increase in incidents of pure red-cell aplasia. A change in Eprex® formulation resulted in a decrease in the conformational stability of the protein, making the protein more sensitive to disruptions in the cold chain [101, 102].

Analysis of protein conformation is necessary during development of expression and purification methods for the therapeutic protein, development of deliverable formulations of the therapeutic protein, tests of drug stability and shelf-life, and quality control during the manufacturing process. Given the number of therapeutic protein patents that are either expired

or will be expiring shortly, follow-on formulations of the therapeutic protein must demonstrate equivalency of both the primary protein structure and the protein conformation to ensure equivalency with the therapeutic protein approved by regulatory agencies to avoid separate clinical trials. Conformational analysis is often cumbersome, requiring large amounts of sample, labor, and expertise, making routine conformational screening of drug lots difficult to achieve. High-resolution analysis of most protein therapeutics is possible by nuclear magnetic resonance (NMR) spectroscopy and/or X-ray crystallography, but these analyses are very laborious, require large amounts of sample, and have strict limitations on the formulation components that can be tolerated. Of particular note is the fact that neither X-ray crystallography nor NMR spectroscopy can readily analyze protein aggregation or highly heterogeneous mixtures of protein conformations, which are often problems with therapeutic protein formulations and can play a large role in the inducement of immune responses [103]. Spectroscopic techniques are also available for conformational analysis, and are often quite useful for detecting differences in conformation and stability. However, they typically yield little information about the regions of the therapeutic proteins that have been altered in conformation or stability, making troubleshooting more difficult, and usually require relatively large amounts of protein.

Mass spectrometry is the method of choice for analysis of the primary structure of therapeutic proteins due to its versatility and sensitivity [104]. Development of mass spectrometry (MS)-based methods for conformation analysis would allow for a rapid and sensitive analysis of both protein primary structure and conformation in one set of experiments using a single analytical platform. Several methods for probing the solution-phase conformation of proteins have been developed previously, such as native spray charge state distribution analysis [105]. Perhaps the most widely used method is amide hydrogen-deuterium exchange

(HDX) analysis coupled with mass spectrometry MS [106]. In HDX-MS, the protein of interest is diluted into deuterium oxide. The amide proton found in the backbone of every amino acid except proline undergoes exchange with water on a time scale that is amenable to analysis. In HDX-MS, the protein is allowed to exchange for differing periods of time, after which the rate of exchange is drastically reduced by lowering the pH and cooling the sample to near freezing. The protein is usually rapidly digested with a nonspecific protease, and the peptides are analyzed by rapid LC-MS to determine the shift in mass caused by exchange of a hydrogen for a deuterium in order to determine the kinetics of exchange between the amide hydrogen and the deuterated solvent. The rate of exchange is a function of the amino acid sequence, as well as the stability of any hydrogen bonds that the amide hydrogen is involved with (the solvent accessibility of the amide hydrogen may also play a role). By comparing two or more samples of the same amino acid sequence, changes in the kinetics of HDX can be interpreted as changes in the conformation of the protein [107]. HDX is a powerful technique for examining changes in protein secondary structure and dynamics in therapeutic proteins [108]; however, it is technically a very challenging technique. Challenges that arise are largely due to the reversible nature of the HDX process and difficulties with reproducibility due to small changes in pH, temperature, analysis time, protease digestion efficiency, or chromatography. Even after the exchange is "quenched" by acidification of the solution, exchange still occurs and the continued exchange results in loss of information because incorporated deuteriums back exchange with water. The back-exchange problem also limits the post-analysis sample handling that can be done, posing problems for therapeutic protein formulations that contain compounds that make mass spectrometric analysis more difficult.

To avoid the back-exchange problems of HDX, a group of methods uses differences in the apparent rate of covalent modification of a protein in two or more different conformations to probe the structure of the protein. In these methods, a protein of interest is mixed with a reactive agent and allowed to react under controlled conditions. This process is then repeated for the same amino acid sequence thought to be in a different conformation. The protein is then enzymatically digested into peptides, and the peptides are subjected to LC to clean up samples and separate peptides, coupled to MS to measure the mass and relative abundance of each modified and unmodified peptide. Tandem mass spectrometry (MS/MS) is used to identify the site(s) of modification of each peptide. Differences in the apparent rate of reaction as measured by changes in the relative abundance(s) of the modified form(s) of the peptide compared to the relative abundance of the unmodified form are then interpreted as differences in either the accessibility of the reactive protein functional group to the reagent, or differences in the local environment of the functional group that alter its reactivity [6]. Covalent modification has fewer technical challenges than HDX due to the usually irreversible nature of the modification; however, most covalent modification techniques have their own set of difficulties. Most notable is the fact that covalent modification itself is known to alter the conformation of the protein. This caveat is a central problem in covalent modification techniques. Usually, other biophysical techniques or functional analyses must be carried out to ensure that the modification is controlled to an extent that does not compromise the conformational integrity of the protein analyte. In addition, most covalent labeling techniques target specific functional groups of protein side chains (e.g. primary amines, free thiols, *etc*.). This specificity limits the amount of information that can be gathered for the protein, because only a strictly limited subset of amino acids is probed.

One covalent modification technique has been quite successful at overcoming the aforementioned limitations. Hydroxyl radical protein footprinting is a technique wherein freely diffusing hydroxyl radicals are generated *in situ* with the protein of interest [109]. Hydroxyl radicals are highly reactive, and will oxidize any amino acid side chain, albeit with different inherent rate constants. The rate of reaction depends primarily on two factors: the inherent reactivity of the amino acids—with aromatic and sulfur-containing amino acids being the most reactive, followed by aliphatic amino acids and arginine, followed by the other charged and hydrophilic amino acids [27, 64] —and the time-averaged solvent accessible surface area of the side chain [70]. Like other covalent modification technique, hydroxyl radical protein footprinting has been shown to alter the conformation of the protein if not sufficiently controlled. However, recent work has demonstrated that techniques that complete the oxidation process in less than one microsecond are capable of heavily oxidizing protein analytes faster than they can change their conformation due to the modifications. While the protein still changes conformation after the labeling process, the information from the original conformation is frozen in a chemical "snapshot" of the oxidation rates of the amino acid side chains, completed before the protein can change conformation [55]. Sub-microsecond oxidation has been reported to be performed by either radiolysis of water by brief electron pulses from a Van de Graaff accelerator [49] or by nanosecond laser photolysis of hydrogen peroxide by a KrF excimer laser [29]. Either method is capable of generating short bursts of very high concentrations of hydroxyl radicals, and through the proper use of scavengers and quenchers, the labeling chemistry can be controlled to the sub-microsecond regime [55]. After the oxidation chemistry is completed, the protein is digested with a protease and subjected to LC-MS/MS to identify oxidized and unoxidized peptides, quantitate the amount of oxidation of each peptide based on the relative abundance of

the oxidized form(s) of the peptide to the relative abundance of the unoxidized form, and finally to identify the major sites of oxidation of the peptide. The "snapshot" nature of hydroxyl radical footprinting (HRF) also decouples the conformational probe from the electrospray ionization process, allowing for post-oxidation clean-up of samples containing components that interfere with the electrospray ionization of proteins and peptides and giving the technique more flexibility than native spray charge distribution analysis [105]. While hydroxyl radical protein footprinting manages to overcome many of the problems with traditional covalent labeling techniques, data interpretation remains a daunting challenge. Due to the promiscuity of the hydroxyl radical, the MS/MS spectra to determine the major site(s) of oxidation of each peptide is challenging, and often requires manual interpretation from a mass spectrometrist experienced in analyzing oxidized peptides.

Here, we describe the application of an abbreviated hydroxyl radical protein footprinting protocol towards the conformational analysis of therapeutic protein formulations. We take samples of recombinant erythropoietin (EPO), interferon α-2A (IFN), and granulocyte colony-stimulating factor (GCSF), including samples of a Food and Drug Administration-approved formulation of GCSF (Neupogen®) and subject them to hydroxyl radical protein footprinting using fast photolysis of peroxide (FPOP) to generate the radical *in situ*. The samples are then digested and subjected to LC-MS in order to quantitate the amount of oxidation of each peptide. As MS/MS analysis is the most laborious step in the process, and the only step requiring a substantial amount of highly specialized expertise, we do not perform MS/MS analysis of oxidized peptides. Even without the MS/MS analysis, we demonstrate that we can successfully compare different samples of therapeutic proteins. We confirm all of our analyses with circular dichroism spectroscopy in order to validate the data from the abbreviated hydroxyl radical

footprinting technique. We propose abbreviated hydroxyl radical protein footprinting as a rapid, flexible, and sensitive technique for conformational comparison of therapeutic protein samples (consuming ~300 picomoles of sample per triplicate analysis) that can be performed by any laboratory with typical expertise in protein LC-MS. The technique is amenable to any typical protein LC-MS platform, and the only specialized instrumentation required is a relatively inexpensive KrF excimer laser. The data generated are sufficient to identify any substantial changes in conformation, and if substantial changes are detected, the samples can be archived and probed more deeply at a later time by more sophisticated separations and MS/MS techniques in order to develop models of conformational changes for troubleshooting.

MATERIALS AND METHODS

Reagents

Neupogen® Filgrastim (Lot 1025877, expiration date 08/13) was purchased from Amgen® (Thousand Oaks, CA, USA) and stored under manufacturer's recommended conditions until analysis. Acetonitrile, catalase from bovine liver, ammonium bicarbonate, and L-glutamine were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO, USA). Thirty percent hydrogen peroxide and formic acid were purchased from J.T. Baker (Phillipsburg, NJ, USA). Methionine amide was purchased from Bachem (Torrace, CA, USA). Dithiothreitol (DTT) was purchased from Fisher Biotech (Fair Lawn, NJ, USA). Sequencing-grade modified trypsin was purchased from Promega Corporation (Madison, WI, USA). Purified water (18 MΩ) was obtained from a Milli-Q Synthesis system (Millipore, Billerica, MA, USA). Recombinant samples of the therapeutic proteins IFN (20 mM potassium phosphate buffer, pH 5.7), EPO (50 mM sodium acetate, pH 5.8), and GCSF (100 mM sodium acetate, 0.005% Tween 80, 5%

sorbitol, pH 4) that are beyond their expected shelf life were generous gifts from a donor who wishes to remain anonymous.

Circular Dichroism Spectroscopy

The circular dichroism (CD) spectra were acquired at 25°C using a Jasco 710/810 spectropolarimeter. Samples concentrations were determined by ultraviolet (UV) absorbance on a Thermo NanoDrop 2000c spectrophotometer [110] (Thermo Scientific, Bremen, Germany). Data points were collected from 190 to 260 nm using a 0.1 cm path length cuvette. Data were analyzed and plotted using Microsoft Excel.

FPOP labeling

FPOP labeling of the protein pharmaceuticals were performed using the proteins as received, without further purification or buffer exchange. All proteins were combined with glutamine and nanopure water to give a final protein concentration of 10 µM in 20 µL final volume. The final glutamine and hydrogen peroxide concentrations were 20 and 100 mM, respectively. Concentrated hydrogen peroxide was diluted to 1 M solution and was added to each replicate to give a final concentration of 100 mM just prior to infusion into the tubing for oxidative modification by FPOP. FPOP was performed as previously described [29, 55]. The protein samples flowed through the beam path of an EX100 KrF excimer laser at 248 nm, with a laser power of 50 mJ/pulse (GAM Laser Inc., Orlando FL, USA). The flow rate of the syringe pump and the pulse frequency of the laser were set such that each segment of protein sample was irradiated with a single ~20 ns UV pulse with a 10% calculated buffer region between irradiated segments to help account for sample diffusion and laminar flow [55]. Methionine amide (0.5 µg/µL) and catalase (0.5 µg/µL) were in a collection tube with ammonium bicarbonate (50 mM) to immediately quench oxidation by destroying excess hydrogen peroxide.

Proteolysis

Hot DTT (64 mM, 95 °C) was added to each irradiated sample to give a final concentration of 5 mM and incubated at 37 °C for 30 minutes to denature and reduce the protein. Sequencing-grade modified trypsin (0.2 µg) was added and incubated at 37°C for 30 minutes while rotating to digest the protein samples. A 30-minute digestion was sufficient to digest all three protein pharmaceuticals with no detectable intact protein remaining when examined by LC-MS.

Mass Spectrometry

Samples were analyzed by electrospray time of flight mass spectrometry on a Synapt G2 HDMS in sensitivity mode controlled by Mass Lynx 4.1 software (Waters Corporation, Milford, MA, USA). Using the nanoAcquity UPLC system (Waters Corporation) the samples are cleaned by a Symmetry C18 trap column (5 µm, 180 µm x 20 mm) prior to injection on to the nanoAcquity UPLC BEH130 C18 column (1.7 µm, 100 µm x 100 mm, Waters Corporation). The gradient was pumped at 600 nL/min from 3-40% buffer B (99.9% acetonitrile, 0.1% formic acid) for 9 minutes, 40-95% buffer B over 1 minute, held at 95% buffer B for 4 minutes, and re-equilibrated to 97% buffer A (99.9% water, 0.1% formic acid) for 15 minutes. MS was performed with the following settings: capillary voltage of 3.0 kV, cone voltage of 45 V, m/z range of 100-2,000, desolvation temperature of 150 °C, and desolvation gas flow of 100 L/h.

Abbreviated FPOP Data Analysis

In silico digestion of GCSF, IFN, and EPO was performed using PROWL (http://prowl.rockefeller.edu). Only MS was performed so all data analysis was done manually using MassLynx 4.1 and quantitation was done on the peptide-level only. Using accurate masses, the tryptic peptides and corresponding oxidation products were identified from the LC-MS runs

manually to calculate the average oxidation events per peptide in each of the protein sample. Average oxidation events per peptide is calculated by summing the ion intensities of all the oxidized peptide masses multiplied by the number of oxidation events required for the mass shift (e.g., one event for +16, two events for +32) divided by the sum of the ion intensities of all unoxidized and oxidized peptide masses. Two-tailed independent Student's *t* tests were used for statistical analysis, with an α of 0.01 selected for statistical significance.

RESULTS AND DISSCUSSION

For each therapeutic protein, the protein sample was aliquoted into three replicates, and each sample was oxidized by FPOP. As the protein is exposed to the short burst of hydroxyl radicals, the radicals react with the amino acid side chains in predictable ways, with an apparent rate that is a function of the amino acid sequence and the exposure of the side chain to the solvent. As all of the therapeutic proteins had the same protein sequence, changes in the apparent rate of oxidation can be attributed to changes in the accessibility of the side chain, which changes as the population of protein conformations in the sample changes. After FPOP oxidation, the samples are quenched to eliminate less-reactive oxidants like hydrogen peroxide and superoxide, digested with trypsin for 30 minutes, and analyzed by a short LC-MS run. Control samples of unoxidized protein were also analyzed intact by LC-MS, as well as digested and analyzed to detect alterations in primary structure that occurred outside of the FPOP oxidation process. In order to ensure that changes in the HRF detected were due to changes in the conformation of the protein rather than the primary structure, one GCSF sample with an unexpected three-amino acid truncation of the N-terminus in ~70% of the protein population was identified during the screening and excluded from analysis (data not shown). All other samples showed near-identical

primary structure, with the only differences arising from minor changes in the amount of native methionine oxidation in the sample.

Granulocyte Colony-Stimulating Factor

One of the most pressing needs for rapid conformational analysis in therapeutic proteins is the comparison of patented pharmaceuticals with biosimilars, which are follow-on therapeutic protein formulations that seek to establish equivalency with the patented drug to avoid the thorough clinical trials usually necessary for regulatory approval of a drug. It is essential for patient safety to ensure that any follow-on therapeutic recombinant protein has not only a primary structural equivalence (amino acid sequence, disulfide bonds, glycosylations, etc.) but also a conformational equivalence. In order to test the sensitivity of HRF to misfolding, aggregation, or other conformational differences between a potential biosimilar and a patented recombinant therapeutic protein, we compared properly handled Neupogen®, the Amgen brand name pharmaceutical for GCSF, with Neupogen® that had been heated well beyond manufacturer's specifications. We also compared Neupogen® with recombinant GCSF samples that were manufactured at different dates, and were all well past their expected shelf life. These expired recombinant GCSF samples should give us a variety of different populations of non-native conformations, which should differ by varying extents from the patented therapeutic protein. After FPOP oxidation and tryptic digestion, GCSF is resolved into four peptides that cover >99% of the protein sequence, all of which yield a measurable hydroxyl radical protein footprint. To determine the ability of the HRF method to detect known changes in conformation that occur upon mishandling of the therapeutic protein, a portion of the Neupogen® sample was heated to 95 °C for one hour, cooled to room temperature and analyzed by HRF. We also irradiated and analyzed five expired recombinant GCSF preparations of unknown conformation,

expressed at different times. The fraction oxidized was calculated for each peptide of each sample and the average fraction oxidized is graphed in **Figure 3.1**. The gross fingerprints of the seven samples are all similar, with the two internal peptides being more heavily oxidized than the N- or C-terminus. However, comparison of each peptide in the Neupogen® sample with the corresponding peptide in each other preparation reveals measurable differences in the conformation of the samples. Student's *t* tests were used to determine the statistical difference in the amount of peptide oxidation between each sample and Neupogen® **(Table 3.1)**. Differences were considered significant if the calculated *p* value was less than or equal to 0.01. The heat-treated Neupogen® and all expired recombinant GCSF preparations saw a statistical difference in at least one of the four peptides by HRF, indicating by our technique that the heat-treated sample and each of the expired recombinant GCSF samples had different conformations than Neupogen®. Circular dichroism spectra were also taken of each unoxidized protein formulation in order to verify any differences identified by HRF with a biophysical technique that is established and in widespread use for therapeutic proteins **(Figure 3.2)**. GCSF is 61% helical, consisting of five helices [111]. The heat-treated Neupogen® and all expired recombinant GCSF samples saw a notable difference by circular dichroism, verifying the differences noted by HRF.

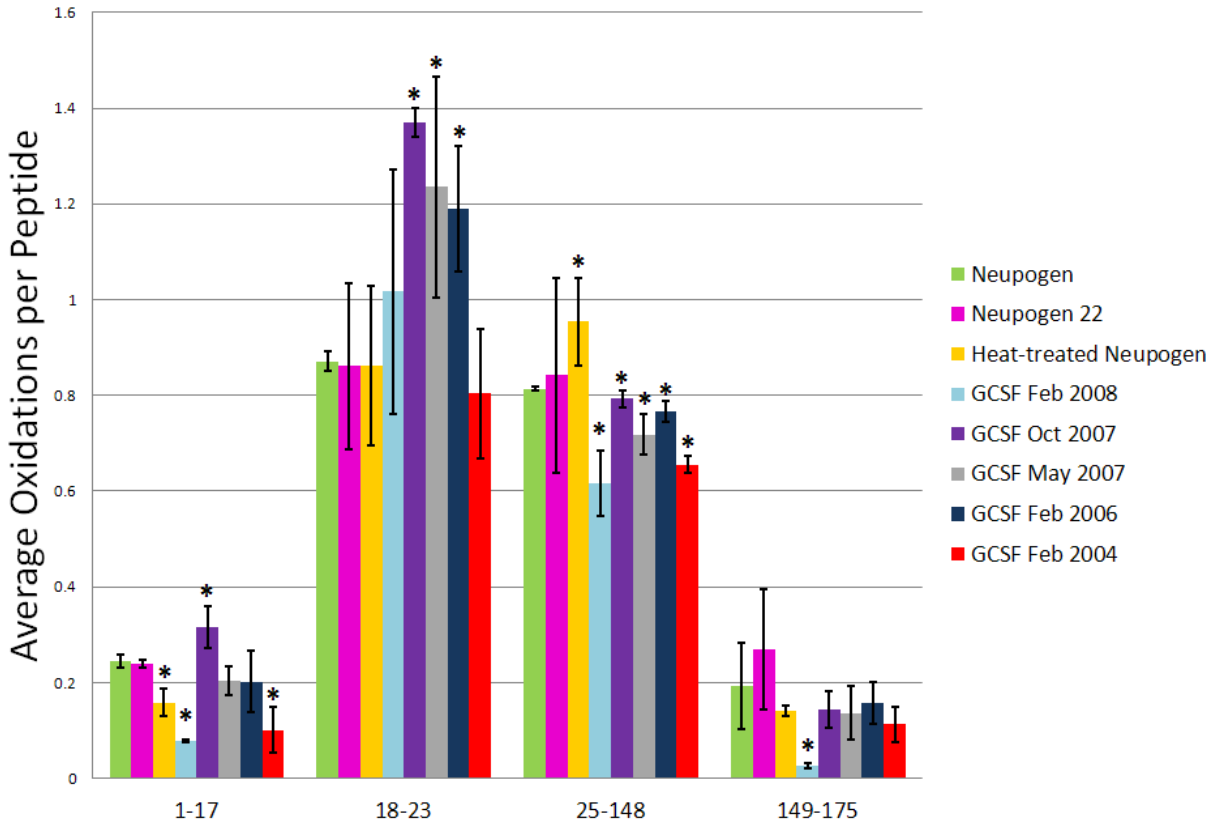


Figure 3.1. Hydroxyl Radical Footprinting of GCSF Samples. Each set of bars represents one peptide from Neupogen® (*green*), heat-treated Neupogen® (*yellow*), or recombinant GCSF samples generated in Feb 2008 (*light blue*), Oct 2007 (*purple*), May 2007 (*gray*), Feb 2006 (*navy*), or Feb 2004 (*red*). The *y*-axis represents the average number of oxidation events per peptide in the sample. *Error bars* 2 SD from a triplicate set of FPOP oxidations and analyses. *Asterisks* peptides with oxidation levels that significantly different than Neupogen® ($p \leq 0.01$).

Table 3.1. Independent Two-Tailed Student's T-Test of HRF between GCSF Samples

vs. Neupogen®				
Sample	1-17	18-23	25-148	149-175
Heat-treated Neupogen®	**$p < 0.001$**	$p = 0.850$	**$p = 0.006$**	$p = 0.114$
GCSF Feb 2008	**$p < 0.001$**	$p = 0.122$	**$p < 0.001$**	**$p = 0.003$**
GCSF Oct 2007	**$p = 0.005$**	**$p < 0.001$**	**$p = 0.004$**	$p = 0.894$
GCSF May 2007	$p = 0.012$	**$p = 0.006$**	**$p = 0.002$**	$p = 0.136$
GCSF Feb 2006	$p = 0.088$	**$p = 0.001$**	**$p = 0.002$**	$p = 0.282$
GCSF Feb 2004	**$p < 0.001$**	$p = 0.167$	**$p < 0.001$**	$p = 0.046$
vs. GCSF Feb 2006				
GCSF May 2007	$p = 0.953$	$p = 0.592$	$p = 0.027$	$p = 0.357$

Bold values are significant at $p \leq 0.010$

Neupogen® vs. Neupogen22. According to the manufacturer, Neupogen® is still potent when left at room temperature for up to 24 hours; therefore, our sample stored at room temperature for 22 hours should be conformationally identical to the Neupogen® sample stored at 4 °C. Looking at the HRF results in **Figure 3.1**, no statistical differences are detected on the peptide level between the Neupogen® and Neupogen22 samples (**Table 3.1**); the CD spectra are also identical (**Figure 3.2**). This positive control comparison establishes the capability of the HRF method to detect conformational equivalency between samples and adds a powerful aspect to the technique.

Neupogen® vs. heat-treated Neupogen®. As shown in **Figure 3.1**, two peptides exhibit significant changes in the hydroxyl radical footprint after heating of Neupogen® to 95° C and cooling back to room temperature. Peptide 25-148 contains 124 amino acids and represents almost 71% of the protein. The native protein structure shows 80% alpha helical content in peptide 25-148 by NMR. The heat treated Neupogen® sample is the only sample that had a relative increase in the amount of oxidation on peptide 25-148, and this increase was substantial.

The increase in hydroxyl radical oxidation of this peptide is probably due to the loss of structural stability here upon heating. A small decrease in oxidation at peptide 1-17 is possibly due to non-native structuring of the normally disordered N-terminus. CD data in **Figure 3.2** illustrates a loss of most of the original secondary structure in the heat treated Neupogen® (yellow) when compared to Neupogen® control (green), confirming the differences detected by HRF.

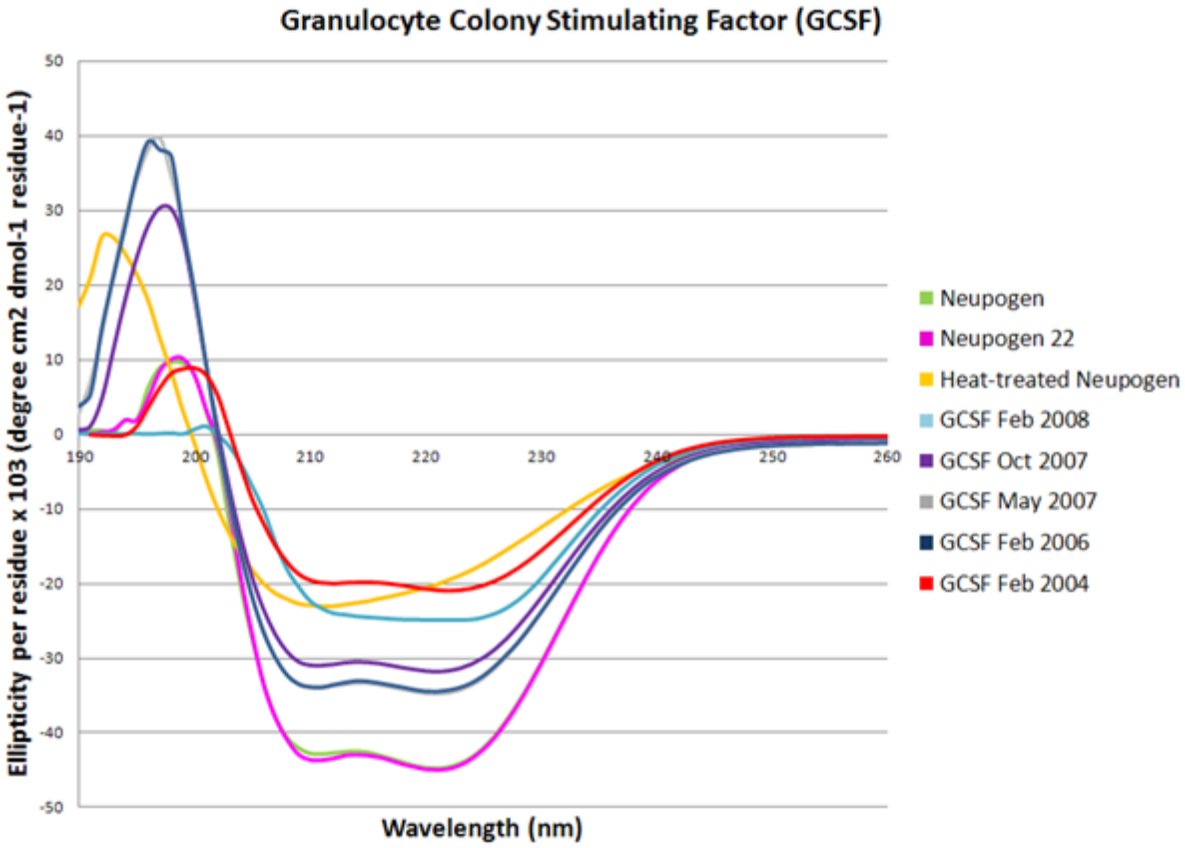


Figure 3.2. Circular Dichroism Analysis of GCSF Samples. The near-UV CD spectra are presented from Neupogen® (*green*), heat-treated Neupogen® (*yellow*), or recombinant GCSF samples generated in Feb 2008 (*light blue*), Oct 2007 (*purple*), May 2007 (*gray*), Feb 2006 (*navy*), or Feb 2004 (*red*).

Neupogen® vs. GCSF February 2008 sample. Three of the four GCSF peptides (all except for 18-23) representing 96% of the whole protein are less oxidized in the February 2008 sample. A reduced amount of oxidation indicates a reduced solvent accessibility, which could be due to aggregation or oligomerization, or collapse of flexible loops into a non-native secondary structure. The CD spectrum for GCSF Feb 2008 sample suggests a shift in alpha helical to more beta sheet and random coil secondary structure. The reduced amount of oxidation and the appearance of beta-sheet secondary structure, which is often associated with protein aggregation [112], may indicate that the GCSF Feb 2008 is aggregating.

Neupogen® vs. GCSF October 2007 sample. Two GCSF peptides representing 13% of the protein are reported to have increased oxidation in the October 2007 sample compared to the Neupogen® sample; a slight increase in oxidation for peptide 1-17, and a much larger increase in oxidation for peptide 18-23. Peptide 1-17 consists of a highly mobile N-terminus and the first part of helix 1, while peptide 18-23 is highly structured; an increase in oxidation suggests a loss of structural stability in the N-terminal and helix 1 region of the protein. GCSF Oct 2007 CD spectrum corresponds to a modest decrease in alpha helical structure that dominates peptides 1-17 and 18-23, without the appearance of substantial beta-sheet structure observed in the GCSF Feb 2008 sample.

Neupogen® vs. GCSF May 2007 and Feb 2006 sample. May 2007 and Feb 2006 samples have identical CD spectra to each other in the near-UV region (minor differences measured in the UV region around 197 nm are probably due to increased noise stemming from buffer interferences in this region), indicating that the two proteins should have identical conformations to each other **(Figure 3.2)**. Both expired recombinant GCSF samples are similar to Neupogen®, but show measurably less helical structure. If HRF is a reliable technique, these two separate

expired GCSF recombinant samples should show no differences in hydroxyl radical footprint when compared to each other, but both should show significant and similar differences when compared to Neupogen®. Using a Student's *t* test (*p* values ≤ 0.01) to determine the statistical difference in the amount of oxidation from the HRF experiments, the 2007 and 2005 samples are not statistically different from one another on the peptide level **(Table 3.1)**. When comparing the two expired recombinant GCSF samples to Neupogen®, two peptides show small but statistically significant differences from Neupogen® **(Figure 3.1, Table 3.1)**. Peptide 25-148 is slightly less oxidized than Neupogen®; this reduced amount of oxidation may represent a subpopulation that is beginning to oligomerize in solution, or it may represent a non-native conformation that is more compact in this region. Peptide 18-23 is more oxidized, suggesting a loss of structural stability in this region and confirming the results found by comparison of the CD spectra. These samples demonstrate clearly that HRF is capable of successfully identifying both identical conformations and small differences in conformation with sensitivity and robustness.

Neupogen® vs. GCSF February 2004 sample. The 2004 sample differs from Neupogen® by decreased oxidation on two peptides representing 81% of the protein. Both these peptides show a substantial decreases in oxidation similar to the Feb 2008 recombinant GCSF sample, suggesting the GCSF Feb 2004 sample has undergone significant conformational changes that protect surfaces that should be solvent accessible. Of all of the expired recombinant GCSF samples, the Feb 2004 sample has the CD spectrum that has the greatest loss of helical content and appearance of beta-sheet content, along with an increase in the random coil content. The HRF data would be consistent with a beta-sheet-mediated oligomer or aggregate of the recombinant GCSF sample.

Overall, the expired recombinant GCSF samples are all different from Neupogen®, although the extent of the differences varies. Hydroxyl radical footprinting was able to identify all of the differences in the recombinant GCSF samples as confirmed by CD. HRF was also able to confirm the identical conformations shared between the Feb 2006 and May 2007 GCSF samples, which were confirmed by identical CD spectra. Unlike the CD data, the HRF data identified the general regions of the protein that changed conformationally, and the samples could be reanalyzed more thoroughly to identify the regions of conformational change with higher spatial resolution. Our data support the use of HRF for comparison of therapeutic recombinant proteins and potential biosimilars to their patented counterparts to test for conformational equivalence, as well as describe differences detected. HRF is well-suited to the analysis of conformational changes that occur due to sample aging and mishandling, as it is sensitive to relatively small conformational changes, rapid, and able to analyze conformational differences using a single analytical platform that can simultaneously analyze differences in primary structure.

While the analysis of GCSF was in excellent agreement with the data generated by CD, in order to test the robustness and sensitivity of HRF across the range of therapeutic proteins, we need to test more than different preparations of one protein. In order to do this, we examined different preparations of two other therapeutic proteins, interferon α-2A and erythropoietin, that were generated at different times during the last twelve years in order to identify conformational differences between the samples generated at different times.

Interferon

IFN has 165 residues with seven helices constituting 61% helical secondary structure [113]. Ten tryptic peptides were analyzed giving 91% sequence coverage. The IFN 2005 sample was used as our "reference" conformation, and all other samples were compared to this conformation by HRF **(Figure 3.3)**, with observed differences verified by CD **(Figure 3.4)**. Student's *t* test was used to calculate significance of differences in the hydroxyl radical footprint between the IFN 2005 reference and the other recombinant IFN protein sample (p value ≤ 0.01) (**Table 3.2)**.

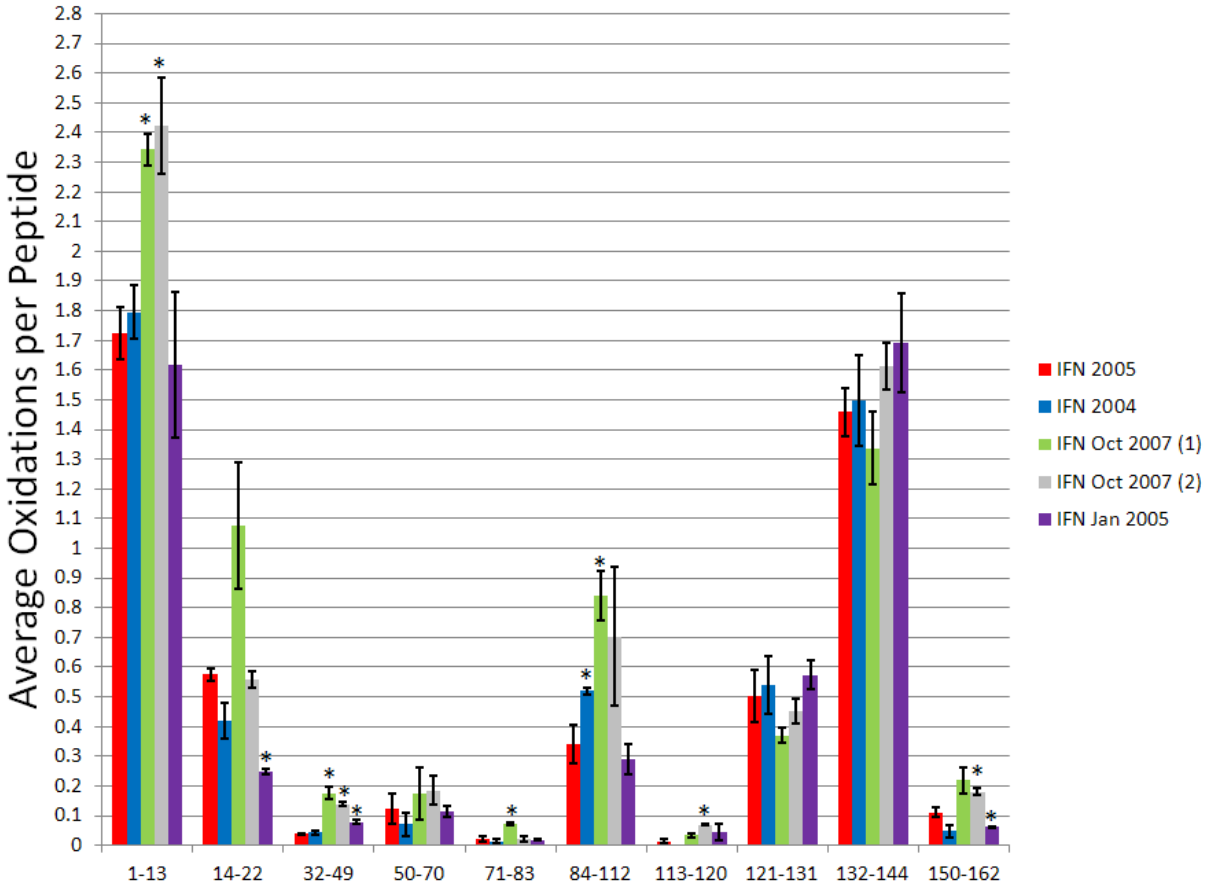


Figure 3.3. Hydroxyl Radical Footprinting of IFN Samples. Each set of bars represents one peptide from recombinant IFN samples generated in 2005 (*red*), 2004 (*blue*), two samples from Oct 2007 (*green* and *gray*), and Jan 2005 (*purple*). The *y*-axis represents the average number of oxidation events per peptide in the sample. *Error bars* 2 SD from a triplicate set of FPOP oxidations and analyses. *Asterisks* peptides with oxidation levels that significantly different than IFN 2005 ($p \leq 0.01$).

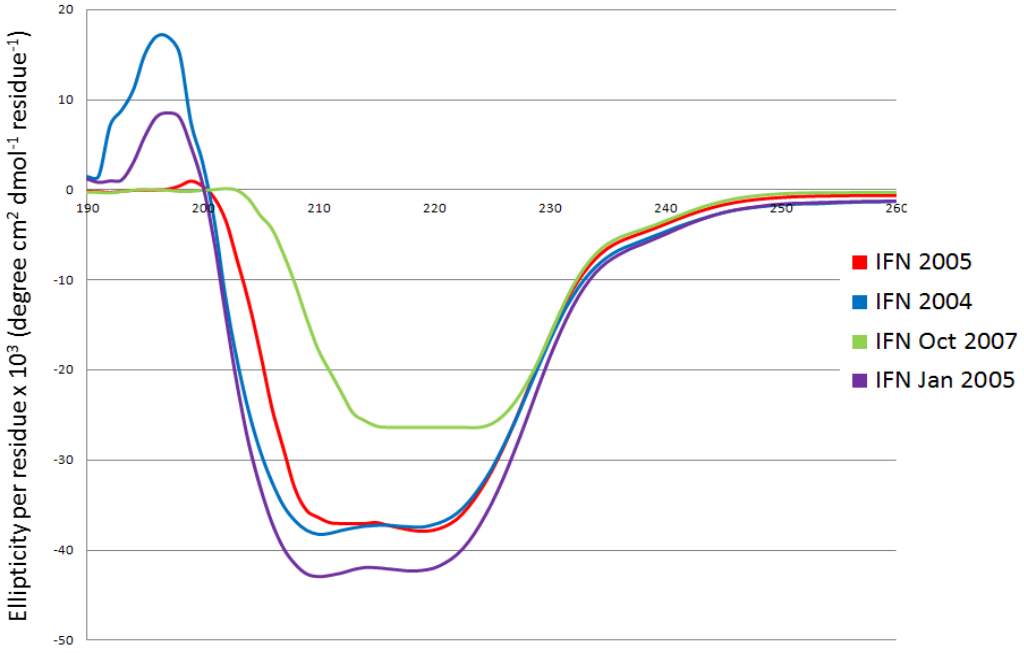


Figure 3.4. Circular Dichroism Analysis of IFN Samples. The near-UV CD spectra are presented from recombinant IFN samples generated in 2005 (*red*), 2004 (*blue*), Oct 2007 (*green*), and Jan 2005 (*purple*).

Table 3.2. Independent Two-Tailed Student's T-Test of HRF between IFN Samples and IFN 2005

Sample	1-13	14-22	32-49	50-70	71-83	84-112	113-120	121-131	132-144	150-162
IFN 2004	$p = 0.394$	$p = 0.014$	$p = 0.271$	$p = 0.234$	$p = 0.338$	$\mathbf{p = 0.010}$	$p = 0.018$	$p = 0.658$	$p = 0.712$	$p = 0.013$
IFN Oct 2007 (1)	$\mathbf{p < 0.001}$	$p = 0.015$	$\mathbf{p < 0.001}$	$p = 0.434$	$\mathbf{p = 0.001}$	$\mathbf{p = 0.001}$	$p = 0.026$	$p = 0.064$	$p = 0.219$	$p = 0.015$
IFN Oct 2007 (2)	$\mathbf{p = 0.003}$	$p = 0.416$	$\mathbf{p < 0.001}$	$p = 0.194$	$p = 0.789$	$p = 0.062$	$\mathbf{p < 0.001}$	$p = 0.423$	$p = 0.303$	$\mathbf{p = 0.004}$
IFN Jan 2005	$p = 0.516$	$\mathbf{p < 0.001}$	$\mathbf{p < 0.001}$	$p = 0.803$	$p = 0.765$	$p = 0.341$	$p = 0.115$	$p = 0.287$	$p = 0.092$	$\mathbf{p = 0.007}$

Bold values are significant at $p \leq 0.010$

IFN 2005 vs. IFN 2004. IFN 2004 shows a very similar hydroxyl radical footprint to IFN 2005, with very few quantitative differences in the amount of oxidation of any of the ten peptides. The only peptide that shows a statistically significant change in the amount of oxidation is peptide 84-112, which consists of a helix-loop-turn structure. This region exhibits a modest increase in oxidation in the 2004 sample, suggestive of either a minor change in conformation in this region; probably in the loop region as the helices that pack against this helix do not exhibit changes in solvent accessibility. Comparison of the CD spectra of IFN 2005 and IFN 2004 reveal a small but measurable change in the secondary structure content indicated by a small shift in the local minima around 207 and 222 nm, and a change in the maxima and amplitude around 197 nm (although the extent of these farther UV changes are often obscured by buffer interferences). Even though the conformational change was small as measured by CD, we were able to clearly detect it using hydroxyl radical footprinting.

IFN 2005 vs. IFN Oct 2007 (1). The IFN Oct 2007 (1) sample has four peptides that are more heavily oxidized (and therefore more solvent accessible) to a statistically significant extent than the IFN 2005 reference sample, as well as one additional peptide that is much more oxidized on average, but on which the replicate statistics are relatively poor. A significant increase in oxidation for the IFN Oct 2007 (1) samples occurs on peptides 1-13, 32-49, 71-83 and 84-112, along with an apparently large (but statistically insignificant) increase in oxidation of peptide 14-22. Peptide 1-22 represents the flexible N-terminus and the first helix of the IFN native structure, and the other two peptides, 71-112, stretch from the start of helix 4 through helix 5, turn 2 and 2 residues in helix 6. The substantial increase in solvent accessibility in these regions rich in secondary structure probably indicates destabilization of the structure of the protein, and should be reflected by a substantial loss of helical structure in the CD spectrum. The CD spectra for the IFN Oct 2007 (1) sample is much different than the other recombinant IFN samples. There is a large decrease in the amount of alpha helical secondary structure in the IFN Oct 2007 (1), consistent with the data from the HRF indicating a loss of structural stability in regions known to contain a large amount of helical content.

IFN 2005 vs. IFN Oct 2007 (2). Due to lack of sample we were unable to obtain an accurate CD spectrum of the IFN Oct 2007 (2) sample. IFN 2007 (2) gave an almost identical HRF as IFN Oct 2007 (1), with the sole differences being reduced oxidation and improved statistics for peptide 14-22, yielding results similar to IFN 2005, and improved statistics for the C-terminal peptide 150-162, indicating a slight increase in oxidation compared to IFN 2005. All other differences were essentially identical to IFN Oct 2007 (1) described above.

IFN 2005 vs. IFN Jan 2005. In the HRF analysis, two peptides of the IFN Jan 2005 sample showed statistically significant changes in the amount of oxidation compared to the IFN

2005 sample, peptide 14-22 (representing the latter half of helix 1) and 150-162 (representing the latter half of the helix 7 and a post-helix turn). Unlike the previous samples, the IFN Jan 2005 sample shows a decrease in the solvent accessibility from the IFN 2005 sample, indicating either an increased stability of the structure, and alternate, more compact conformation, or an oligomerization/aggregation event. The IFN Jan 2005 sample has a similar CD spectrum to the IFN 2005 reference, but exhibits an increase in alpha helical structure. An increase in secondary structure content caused by stabilization and/or extension of helixes 1 and 7 of the protein would explain both the CD spectra and the HRF data.

Erythropoietin

EPO is a therapeutic protein with 166 residues, with 56% of the protein forming five alpha helices [114]. Trypsin digestion of EPO yields eight detectable peptides resulting in 53% coverage. The relatively low sequence coverage was due to the primary sequence of EPO, which has a large number of lysines and arginines resulting in much of the protein being cleaved by trypsin into very small peptides that are very difficult to detect by LC-MS. Of the eight peptides, one had no oxidation modifications after HRF. We compared the seven peptides that were modified for each of the five recombinant EPO samples using HRF **(Figure 3.5)**. EPO June 2009 sample is the most recently made sample and will be used as a reference to compare the difference in conformation among the other EPO samples. A difference in the hydroxyl radical footprint was considered significant if the p value by Student's t test was ≤ 0.01 **(Table 3.3)**. Differences in conformation identified by HRF were confirmed by CD **(Figure 3.6).**

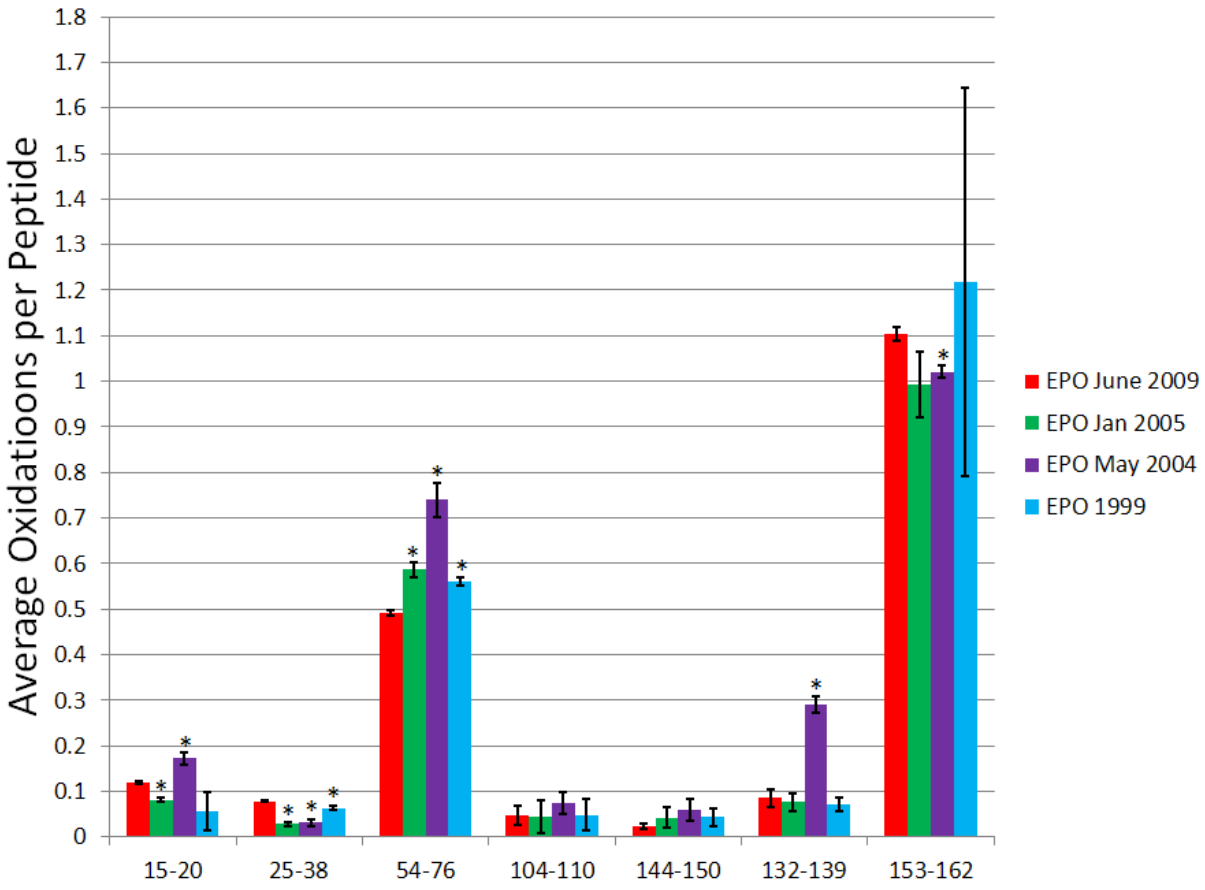


Figure 3.5. Hydroxyl Radical Footprinting of EPO Samples. Each set of bars represents one peptide from recombinant EPO samples generated in 2009 (*red*), Jan 2005 (*green*), May 2004 (*purple*), and 1999 (*blue*). The *y*-axis represents the average number of oxidation events per peptide in the sample. *Error bars* 2 SD from a triplicate set of FPOP oxidations and analyses. *Asterisks* peptides with oxidation levels that significantly different than EPO 2009 ($p \leq 0.01$).

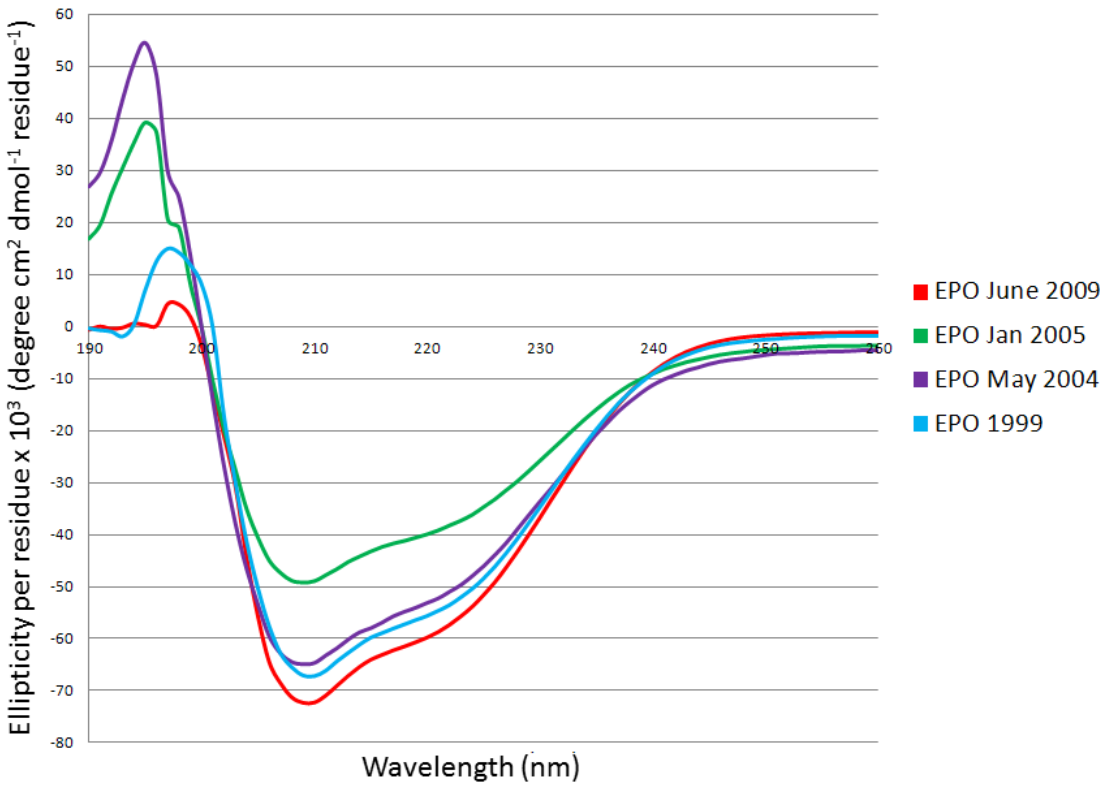


Figure 3.6. Circular Dichroism Analysis of EPO Samples. The near-UV CD spectra are presented from recombinant EPO samples generated in 2009 (*red*), Jan 2005 (*green*), May 2004 (*purple*), and 1999 (*blue*).

Table 3.3. Independent Two-Tailed Student's T-Test of HRF between EPO Samples and EPO June 2009

Sample	15-20	25-38	54-76	104-110	132-139	144-150	153-162
EPO Jan 2005	**p = 0.001**	**p < 0.001**	**p = 0.001**	p = 0.867	p = 0.208	p = 0.592	p = 0.061
EPO May 2004	**p = 0.003**	**p < 0.001**	**p < 0.001**	p = 0.226	p = 0.062	**p < 0.001**	**p = 0.003**
EPO 1999	p = 0.061	**p = 0.004**	**p < 0.001**	p = 0.962	p = 0.177	p = 0.365	p = 0.670

Bold values are significant at $p \leq 0.010$

EPO June 2009 vs. EPO Jan 2005. EPO Jan 2005 is more oxidized at peptide 54-76 compared to the 2009 reference sample. Peptide 54-76 is essentially helix three of native EPO. An increase in oxidation probably signifies a loss of structural stability, causing the peptide to become more solvent exposed and increasing peptide oxidation. The change in the amount of oxidation is minimal, but statistically significant. The sample also exhibits a small decrease in the solvent accessibility of peptide 15-20 and 25-38. All other peptides are not oxidized differently between the EPO 2009 and EPO 2005 samples to statistical significance. The CD spectra for EPO Jan 2005 appears to have the similar shape of EPO 2009, but with a slight shift of the minimum and maximum to shorter wavelengths, a decrease in ellipticity amplitude in the 200-240 nm range, and an increase in ellipticity amplitude in the 190-200 nm range.

EPO June 2009 vs. EPO May 2004. EPO May 2004 shows statistically significant increases in solvent accessibility in three peptides; peptide 15-20 exhibits a small increase in solvent accessibility, while peptides 54-76 and 132-139 show fairly large increases in solvent accessibility. Additionally, EPO May 2004 exhibits small but statistically significant decreases in solvent accessibility in two other peptides, 25-38 and 153-162. The increased protection of some regions and increased exposure of other regions argues against a global structural destabilization, and rather suggests a different but relatively stable conformation for EPO May 2004 compared to EPO June 2009. These results are confirmed by the CD spectra, which show the EPO May 2004 sample with a clearly structured, but different CD spectrum than the EPO June 2009 reference. The minimum is shifted to lower wavelengths, with a small decrease in the amplitude of ellipticity, while the maximum is shifted to lower wavelengths with a substantial increase in ellipticity amplitude.

EPO June 2009 vs. EPO 1999. The hydroxyl radical footprint of the EPO 1999 sample was almost identical to the 2009 sample, with the only statistically significant difference in the hydroxyl radical footprint appearing as a very small decrease in oxidation on peptide 25-38, which has a random coil structure in native EPO, and a small increase in oxidation in peptide 54-76, which contains the majority of helix 3. These data correlate very well with the CD data, where the 1999 sample, by far, has the most similar CD spectrum to the 2009 reference, with almost identical minima and maxima. However, the CD spectrum does indicate small differences in the secondary structure, which are confirmed by small differences in the hydroxyl radical footprint. The data indicate quite clearly that even small differences in conformation as measured by CD can be clearly observed in the hydroxyl radical footprint.

All of the EPO samples were different from the 2009 reference sample by HRF, although the differences varied in magnitude from the almost-identical 1999 sample to the significantly different 2004 sample. The differences were also confirmed in the CD spectrum, with the 1999 sample being nearly identical and the 2004 and 2005 samples being substantially more varied.

DISCUSSION

Hydroxyl radical footprinting was capable of identifying differences in the conformation between twelve different pairings of samples across three different therapeutic proteins. HRF proved to be able to detect differences due to loss of solvent accessibility (probably due in some cases to aggregation or oligomerization) or due to increases in solvent accessibility (probably due to structural destabilization). HRF was also able to successfully detect two pairs of samples that had no differences in conformation. All analyses were carried out in technical triplicate (i.e. the therapeutic protein was aliquoted into three samples which were independently oxidized by FPOP and independently analyzed by LC-MS). The overall time for a single analysis is

approximately 2 hours (~10 minutes for FPOP and quenching, 1 hour for denaturation and digestion, 30 minutes for LC-MS), with all of the FPOP steps being capable of automation with currently available commercial technology. Sample requirements are light, with only ~90 picomoles of sample required for the triplicate analysis. HRF is capable of accommodating almost any buffer or formulation component and is largely insensitive to pH [64], although some components like carrier proteins can complicate the resulting LC-MS spectra and the inclusion of high concentrations of radical scavengers will require the generation of higher burst concentrations of hydroxyl radicals to generate sufficient hydroxyl radical footprinting data to differentiate conformations. The method is also capable of handling heterogeneous protein conformations and oligomerization/aggregation states, as potentially demonstrated by the GCSF samples with non-native beta sheet content and substantially more protected surface areas, and also as demonstrated in other studies on confirmed polydispersed protein oligomers [115]. The reproducibility of the oxidation chemistry by FPOP and the quantitation of oxidation by LC-MS were more than sufficient for the purposes of routine analysis. The use of only MS data from short LC gradients, as opposed to MS/MS and long, careful separations to isolate oxidation isomers, limits the spatial resolution of the footprinting information obtained. However, interpretation of MS/MS spectra of peptides oxidized by hydroxyl radicals is difficult and time consuming, requiring at the very least manual verification of all identified oxidized peptides. Conversely, quantification at the peptide level is fast and can be performed without special expertise, as the products of the complicated hydroxyl radical protein chemistry resolve into a limited number of mass shifts, most of which represent net additions of one or more oxygen atoms [27]. Once a sample has been initially characterized, the data analysis can also be easily automated, as the hydroxyl radical footprint will boil down to a measurement of the abundance

of specific peak abundances eluting at specific times from the LC. Similarly, the quantification by LC-MS does not require high-end LC or MS instrumentation, as the experiment is quite compatible with almost any electrospray instrument that is suitable for peptide LC-MS.

These useful analytical qualities make HRF an attractive technique for the characterization of therapeutic proteins, both during the development and production stage and for quality control afterwards. The ability to use a single analytical platform to characterize both primary structure and three-dimensional conformation in a single analysis is a powerful option for screening of therapeutic proteins. The small sample quantities, the potential for automation of analysis, and the relatively short analysis times makes HRF suitable for insertion into the production process as a quality control measure, as well as for spot checks of protein shelf life during the storage and transportation chain. Similarly, HRF is a very useful analytical technique for the development and validation of biosimilars to patented therapeutic proteins; as abbreviated, HRF is capable of confirming conformational identity as well as capable of not only identifying conformational differences, whether they be in regions with stable secondary structure or in more dynamic loops, but also roughly localizing these conformational differences for either more thorough HRF analysis with UPLC-MS/MS or for thorough analysis by more demanding high-resolution structural techniques. The abbreviated HRF technique demonstrated here can fill an important niche as an intermediate resolution and rapid structural technique for the conformational analysis of therapeutic proteins.

While HRF performed well in all cases examined here, there are instances that should be handled with care when using HRF for comparing therapeutic proteins. One important factor that must be carefully considered in designing an HRF experiment for the analysis of therapeutic protein formulations is composition of the formulation other than the therapeutic protein.

Different formulation components (buffers, carrier proteins, etc.) will scavenge radicals to different extents, which can result in global reductions in the hydroxyl radical footprint of a protein even without any conformational change. This reduction is due to competition between the therapeutic protein and buffer components for the hydroxyl radical during the radical burst, resulting in less diffusing hydroxyl radical available to oxidize the therapeutic protein and a lower apparent rate of oxidation. However, it is quite possible with careful planning to apply HRF to different formulations with widely differing radical scavenging properties. In order to correct for different radical scavenging properties of the buffer, mixtures matching the formulations to be tested, without the therapeutic protein, should be prepared and spiked with a reporter. This reporter can either be a radical-sensitive chromophore, or a small unstructured peptide that can be monitored for oxidation by LC-MS. The radical dose (as controlled by laser pulse energy and hydrogen peroxide concentration) is then adjusted until the amount of oxidation of the reporter is equal between the two samples. These adjusted conditions are then compared to compensate for the radical scavenging properties of the different formulations.

CONCLUSION

Here, we describe an abbreviated hydroxyl radical footprinting technique for the conformational comparison of therapeutic protein samples using hydroxyl radical protein chemistry and LC-MS analysis. Using this method, we are able to identify conformational differences between Neupogen® and expired recombinant GCSF samples, as well as conformational differences between different samples of the therapeutic proteins erythropoietin and interferon α-2A. We are also able to confirm conformational equivalence between two different samples of GCSF produced nine months apart and two Neupogen® samples branded to have the same conformation. All conformational HRF analyses were successfully confirmed by

circular dichroism analysis, demonstrating the accuracy and robustness of this technique for the comparison of conformations of therapeutic proteins.

ACKNOWLEDGEMENTS

We acknowledge the National Center for Research Resources (a center of the NIH) for financial support of the Research Resource for Integrated Glycotechnology and the hydroxyl radical footprinting project (P41-RR005351). The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We would also like to thank Dr. Jeffrey Urbauer of The University of Georgia for use of his CD instrumentation. JSS would also like to thank Dr. Marshall Bern for numerous productive conversations on the topics explored here.

CHAPTER 4

OLIGOMERIC STRUCTURE OF THE CHEMOKINE CCL5/RANTES FROM NMR, MS, AND SAXS DATA

ABSTRACT

CCL5 is a pro-inflammatory chemokine known to activate leukocytes through its receptor, CCR5. Monomeric CCL5 is adequate to cause cell migration *in vitro*; however, CCL5 aggregation is essential for *in vivo* migration, T cell activation and apoptosis, and HIV entry into cells. CCL5's propensity for aggregation is biologically important yet no structural information is available on the CCL5 oligomer larger than the canonical CC chemokine dimer. Therefore, we investigated the solution structure of CCL5 oligomer using an integrated approach, including NMR, SAXS and hydroxyl radical protein footprinting. The proposed model not only explains the disaggregating effect of the E66S mutant, but also provides mechanisms whereby the oligomer can be stabilized by glycosaminoglycans and bind CCR5 without depolymerizing.

INTRODUCTION

Chemokines are an important class of immunoactive-signaling molecules that are widely expressed by many cells in different contexts, including immune surveillance and inflammation. Establishment of chemokine concentration gradients on endothelial layers and in the surrounding tissue provides directional cues to guide cell movement. The chemokines' ability to oligomerize likely plays a vital role in the establishment and maintenance of the gradients [65, 116]. Although monomeric mutants of chemokines are capable of causing cell migration *in vitro* [67, 117-119], several studies using non-oligomerizing variants have shown that induction of

migration *in vivo* is dependent on oligomerization [65-68]. Furthermore, chemokines that are unable to oligomerize have shown therapeutic benefit in a number of animal models of inflammatory disease [68, 120], hence the importance of understanding oligomeric interactions.

Glycosaminoglycans (GAGs) are sulfated polysaccharides that reside on cell surfaces and in the extracellular matrix. GAGs can promote chemokine oligomerization and the chemokine-GAG interaction was demonstrated experimentally when chemokine mutants incapable of binding GAGs were unable to activate cell migration *in vivo* [67-69]. Chemokine-GAG interactions have shown importance in processes separate from migration, such as T cell activation and induction of apoptosis [67-69].

Structural knowledge of chemokine oligomers could provide insight into molecular mechanisms by which chemokines facilitate the recruitment and activation of leukocytes *in vivo*. Currently, little information exists on the mechanism of oligomerization. Models of quaternary assemblies of chemokines larger than dimers do exist but they have been deduced from crystal structures [121-127]. Recently, Ren et al. (2010) solved the crystal structure of MIP-1α/β oligomers and used SAXS data to support the presence of similar structures in solution; this is one of the few cases where any experimental data exists [126]. Relatively little crystal or solution data exists for oligomeric structures of chemokines.

CCL5 is an aggregation-prone member of the chemokine family that is secreted by both endothelial cells and activated leukocytes. Several studies have shown CCL5's importance in different pathways [65, 69], suggesting that oligomeric CCL5 activates different signaling pathways than those induced my monomeric CCL5. Monomeric and oligomeric CCL5 appear to act as "functionally selective" or "biased" ligands of CCR5, terms usually reserved for different ligands of the same receptor [128].

CCR5 plays a crucial role in mediating HIV entry into T cells and monocytes. In the context of HIV, differences in the oligomeric state of CCL5 have functional consequences. High concentrations of aggregation competent wild-type (WT) CCL5 promote HIV entry, however, low concentrations of CCL5 and mutants incapable of forming aggregates suppress HIV infection [116]. Binding of monomeric CCL5 to CCR5 physically blocks access to the receptor, inhibiting HIV entry. On the contrary, high concentrations of aggregating CCL5 seem to enhance HIV entry by physically crosslinking HIV particles to target cells or by activating the kinase-signaling pathways [129]. Hence, understanding oligomeric tendencies of CCL5 is potentially an important role in both control of immune response and intervention of HIV infection.

Physical characterization of the oligomerization behavior of CCL5 has previously revealed that CCL5 forms large aggregates at micromolar concentrations in a neutral pH environment but smaller aggregates at reduced pH. Dissociation into dimers at pH 3-4 and mutants that limit oligomerization [116] have allowed dimer characterization by NMR and X-ray crystallography [130, 131]. Other structural studies were done under conditions that favor dimers [123, 132, 133], but none of the current studies provide insight into the quaternary conformation of CCL5 higher-order oligomers.

Utilizing CCL5's relatively high solubility at an intermediate pH (4.5), we now present CCL5's higher oligomeric structure using solution NMR, SAXS and hydroxyl radical footprinting mass spectrometry (MS). NMR and SAXS data revealed that the size distribution of the CCL5 oligomer at pH 4.5 is primarily tetramer. Residual dipole coupling (RDC) data was used to determine relative orientations of the dimeric units within the tetramer, which significantly restricts the number of possible orientations the tetramer can adopt. SAXS shape information was used to find the best translational placement of the dimers within the tetramer.

NMR cross-saturation experiments and hydroxyl radical footprinting MS were used to determine the interaction interfaces between the dimers. Hydroxyl radical footprinting was able to investigate the mutant and WT CCL5 at pH 7, biologically relevant pH for oligomer formation. Using an integrated approach, the data provides information on the assembly of the tetramer and the formation of linear higher-order aggregates. The resulting model shows that both GAG-binding sites and residues known to interact with CCR5 are exposed in a linear array, possibly facilitating movement of leukocytes along cell surfaces and aiding pathogen entry *in vivo*. Electron microscopy was also utilized to visualize the oligomeric fibers formed by WT CCL5 at pH 7.

MATERIALS AND METHODS

Expression of WT and E66S CCL5

Two versions of CCL5 were expressed in *E. coli* using a pET23a vector with conventional $^{15}NH_4Cl$ or $^{14}NH_4Cl$ containing M9 media for labeled and unlabeled CCL5. They were purified according to Czaplewski et al. [134]. Briefly, inclusion bodies of CCL5 were solubilized in 6 M guanidine hydrochloride and purified on a Superdex 75 column. Purified protein was then subjected to fast dilution refolding with refolding buffer consisting of 100 mM Tris, pH 8.0, 100 μM reduced glutathione and 10 μM oxidized dimeric glutathione. The refolded protein was then dialyzed against 0.1% TFA and purified further using strong anion exchange chromatography.

NMR Spectroscopy

All NMR data was acquired and analyzed by Xu Wang. NMR samples of CCL5 contained 1 – 2 mM of protein in 50 mM acetate buffer at pH 4.4. RDCs were measured on a 1 mM CCL5 sample in stretched 5% positively charged [135] or neutral polyacrylamide gels using

the two-stage gel tube as described in Lui and Prestegard [136]. RDC values were measured using a J-modulated $^{1}H,^{15}N$-HSQC experiment [136]. Samples for cross-saturation measurements were conducted on a sample consisting of 0.75 mM $^{15}N,^{2}H$-labeled protein mixed with 2.25 mM unlabeled protein in 50% D_2O, 50 mM acetate (pH 4.4) buffer. Cross-saturation experiments were carried out in a manner similar to that published by Takahashi et al. [137]. Protons were either saturated at 1 or 23 ppm as a control.

SAXS Measurements

SAXS data was acquired by Brian Jones at Bruker and SAXS measurements were carried out as described by Wang et al. [138]. Specifically, samples containing 1 – 1.25 mM of WT CCL5 at pH 4.5 were irradiated for 1 hour with a 1.54 Å X-ray source. The scattering pattern was collected using a Bruker Nanostar U X-ray system and the 2D scattering curves using Bruker's SAXS software. The scattering profile was determined as the difference between samples with and without protein. Values for maximum pair distances were extracted using the program GNOM [139].

Model Screening using Grid Search, SAXS, and Residue-Paring score

The grid was conducted in a manner similar to Wang et al. [140]. Specifically, the crystal structure of the CCL5 dimer (PDB accession code 1U4L) was placed in the alignment tensor principal axes frame such that the presumed symmetry axis became the x axis of the PDB frame. Another dimer of CCL5, obtained by rotating the first dimer by 180° around the x axis, was then translated over a grid on a plane perpendicular to the x axis, was then translated over a grid on a plane perpendicular to the x axis. Theoretical SAXS curves of the each tetramer/hexamer model obtained by placing the second dimer at a different grid point was calculated using CRYSOL [141] and a hydration shell contrast value of 0.07 e/Å^3. The curves were then used in

OLIGOMER [142] to find the best fit to a tetramer/hexamer mixture. The interface of each model was evaluated using residue-pairing score [143], and all models were evaluated using a combined score, which is obtained by scaling the residue-pairing score to the same numeric range as the SAXS-fitting χ^2 values for models with a χ^2 value of 2.50 or less.

Hydroxyl radical footprinting MS

Samples of ^{15}N-labeled CCL5 (E66S mutant or wild type) at 20 μM in acetate buffer (pH 4.5) containing 50 μM H_2O_2 with 20 μM glutamine to limit radical half-life was flowed through the beam path of a KrF laser at 248 nm, and pulsed so that each segment of sample was irradiated with a single ~20 ns UV pulse with a small buffer region between irradiated segments to help account for sample diffusion and laminar flow effects. The oxidation was immediately quenched using methionine amide (0.5 μg/μL) and catalase (0.5 μg/μL). Prior to trypsin digestion, unoxidized and unlabeled mutant or WT CCL5 (20 μM) was added to the corresponding irradiated samples to serve as an internal standard for quantitation. Ammonium bicarbonate (50 mM) and DTT (5 mM) were added to the CCL5 samples (50 μL) and the samples were incubated at 55°C for 3 hours to denature and reduce the protein, respectively. Sequencing grade modified trypsin was added (0.2 μg/μL) and incubated at 37ºC for 48 hours while rotating to digest the protein samples. The resulting peptides were analyzed by LC-MS to quantify the change in oxidation using the percent fractional oxidation. FT-MS was used to quantitate the unmodified, unlabeled peptide peak to unmodified, ^{15}N labeled peptide peak for each tryptic peptide, and the fractional oxidation was calculated from the reduction of signal for the unmodified protein in the oxidized sample relative to that for the identical light isotopomer in the internal standard. Tandem mass spectrometry was used in conjunction with LC-MS to

identify sites of oxidation, with sites of oxidation determined by manual annotation of MS/MS spectra.

Electron Microscopy

Images were obtained using electron microscopy by Hong Yi and Elizabeth Wright at the Robert P. Apkarian Integrated Electron Microscopy Core using a JEOL JEM-1400 Transmission Electron Microscope. WT CCL5 at pH 7 samples were prepared at 10 and 100 μM concentrations. In each case, a carbon grid was placed face down onto 5 μL of sample on a piece of laboratory film for 1 – 3 minutes, the grid is blotted with filter paper, placed on a drop of 1% phosphotungstic acid (PTA) negative staining dye for 30 seconds, and blotted again.

RESULTS AND DISCUSSION

Symmetry of the WT CCL5 Tetramer

The oligomeric state of ~1 mM WT CCL5 at pH 4.5 was investigated by NMR, dynamic light-scattering, and SAXS data. NMR data demonstrates the diffusional properties consistent with a tetramer. Light scattering and SAXS data indicate the presence of some higher order oligomers in addition to tetramer. The presence of only one set of cross-peaks in the two-dimensional (2D) ^{1}H-^{15}N HSQC (heteronuclear single quantum coherence) spectrum of an ^{15}N-labeled CCL5 sample suggests that the oligomers formed are symmetric structures, at least when averaged over the NMR time scale.

Residual dipolar coupling (RDC) [144, 145] measurements were employed to determine the structures of symmetric complexes, because RDCs provide orientational restraints rather than distance restraints like Nuclear Overhauser Effect (NOE)-based experiments [146]. RDCs are sensitive to orientations of N-H bond vectors relative to the external magnetic field giving information about the relative orientations of parts of the molecule that are far apart in the

structure; RDCs provide a test of preservation of the monomer structure as observed by X-ray crystallography [147, 148], and information on the orientation of monomers in higher-order structures [144]. Relative orientations of the dimeric units within the tetramer were determined by comparing the experimental RDCs with those back calculated from the monomer. RDCs for E66S mutant fit very well with calculated (**Figure 4.1(a)**) because it appears the monomer is preserved in the dimer of the crystal structure. More unexpectedly, the RDC data of the WT CCL5 tetramer fit the crystal structure nearly as well (**Figure 4.1(b)**). The results indicate dimer preservation within the tetramer structure.

The axis of rotational symmetry can also be determined from the RDC data. The axis of rotational symmetry was determined to be very closely aligned with C_{2V} symmetry. **Figure 4.1(c)** shows the orientation of the alignment tensor relative to the CCL5 dimer. The alignment tensor orientation and C_{2V} symmetry for WT CCL5 tetramer were determined independently. The fact that both the dimer and tetramer have the same alignment tensor and C_{2V} symmetry, and the dimer structure is preserved in the tetramer places substantial restrictions on the tetramer structure.

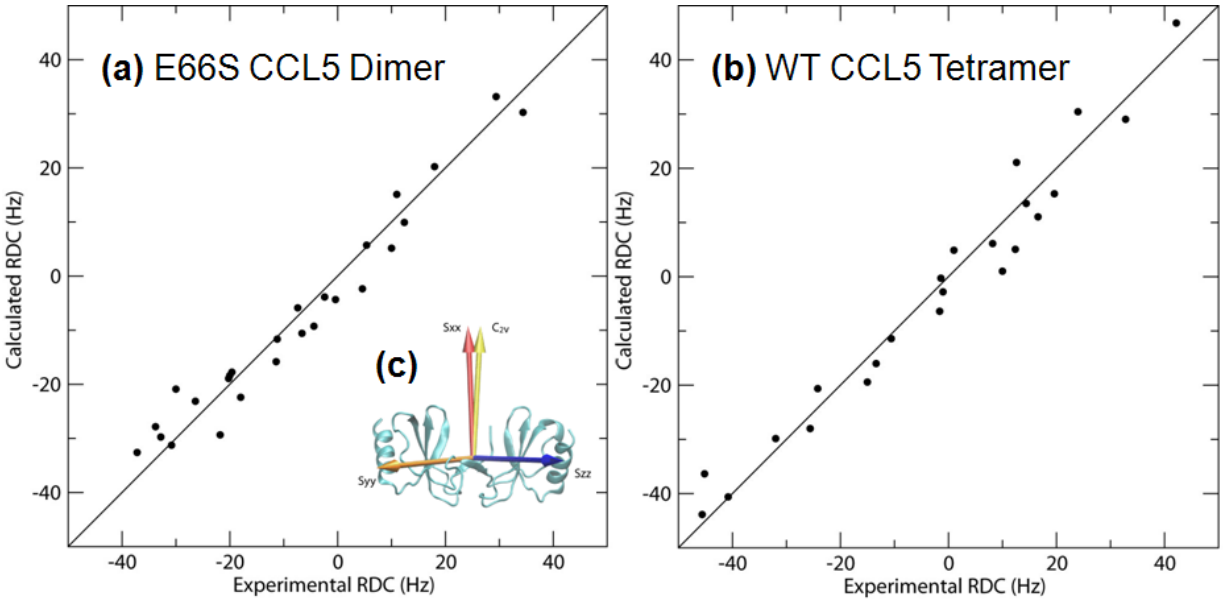


Figure 4.1. Experimental RDCs for E66S and WT CCL5. Experimental RDCs for E66S CCL5 dimer **(a)** and tetramer **(b)** compared to back calculated RDCs for the monomer. **(c)** Orientation of alignment tensor relative to the CCL5 dimer. The x axis of the principal axes of the alignment tensor is red, the y axis is blue, and the z axis is orange. The golden arrow indicated the orientation of the symmetry axis from the dimer crystal structure.

Tetramer Model

Small angle X-ray scattering is a scattering technique that measures the elastic scattering of X-rays by a sample. SAXS data are particularly sensitive to the spatial distribution of the electron density of the dominate species in the solution allowing shape restrictions to be placed on the tetramer model. Data were collected on a ~1.25 mM WT CCL5 sample at pH 4.5. Using the symmetry axis and preservation of the dimer restrictions determined by RDCs, the SAXS data can be used to find the best relative placement of the dimers within the tetramer.

Due to the restrictions on the symmetry and shape of the tetramer, a simple grid search [140] was used to generate all possible models satisfying the restrictions previously determine.

Two dimer subunits were produced by duplicating the single dimer oriented in the alignment tensor frame indicated by the RDC data and rotating the new copy by 180°. Tetramer models were produced by fixing one of the dimer subunits and translating the second dimer subunit in the plane perpendicular to the symmetry axis in 1 Å steps along either one of the two axes perpendicular to the symmetry axis. Owing to the symmetry restrictions, only a grid search around half of a plane was needed as models generated by placing the second dimer unit on positions in the other half of the plane produced the same models. The models were subsequently screened for agreement with the SAXS profile and favorable binding surface using a residue-pairing score [143].

The tetramer model in **Figure 4.2** is based on an interface formed by contacts between residues 25 and 30 in the second β strand from one monomer of one dimer and several residues of the C-terminal helix from one monomer of the other dimer. The tetramer model was confirmed using NMR cross-saturation and hydroxyl radical footprinting.

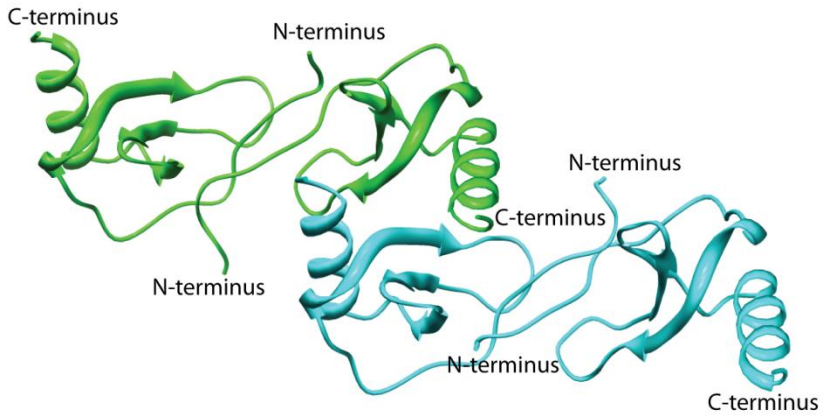


Figure 4.2. Tetramer Model from Grid-Search and SAXS Data. Dimer subunits are represented using two colors. The N- and C-termini of each monomer are label respectively.

Validating the Tetramer Binding Interface with NMR Cross-Saturation Data

A solution based NMR cross-saturation [137] experiment probed the inter-dimer interface to validate the tetramer model shown in **Figure 4.2**. In this method, $^{15}N,^{2}H$-labeled CCL5 was mixed at a 1:3 ratio with unlabeled CCL5, allowing mixed dimers and tetramers to form. By applying RF pulses in the methyl/methylene region of the protein spectrum, sites in the unlabeled molecules become saturated, and spin diffusion transfers this saturation across the interface to the nearest ^{1}H-^{15}N pairs of the labeled molecules. The proximity of the various ^{1}H-^{15}N pairs can be measure using the ratio of HSQC cross-peak intensities. This method provides information about both the dimer and tetramer interfaces; therefore, it is important to compare experiments for the E66S mutant and WT to distinguish between the monomer-monomer interface of the dimer versus the dimer-dimer interface of the tetramer. **Figure 4.3(a)** plots the HSQC saturated/unsaturated signal intensity ratios for WT and E66S CCL5. The data is not quantitatively comparable but serves as a good qualitative guide for identifying the tetramer interface. **Figure 4.3(b)** highlights the residues perturbed in the WT sample. The intensity decrease on the N-terminal residues (6-10 colored red in **Figure 4.3(b)**) is obviously due to the monomer-monomer interaction because the same residues are perturbed in the spectra of the E66S mutant. Residues 26-29, 33-34, and 64-68 were perturbed in the WT data only (orange region of **Figure 4.3(b)**), and represent the residues in the dimer-dimer interface and clearly support the model.

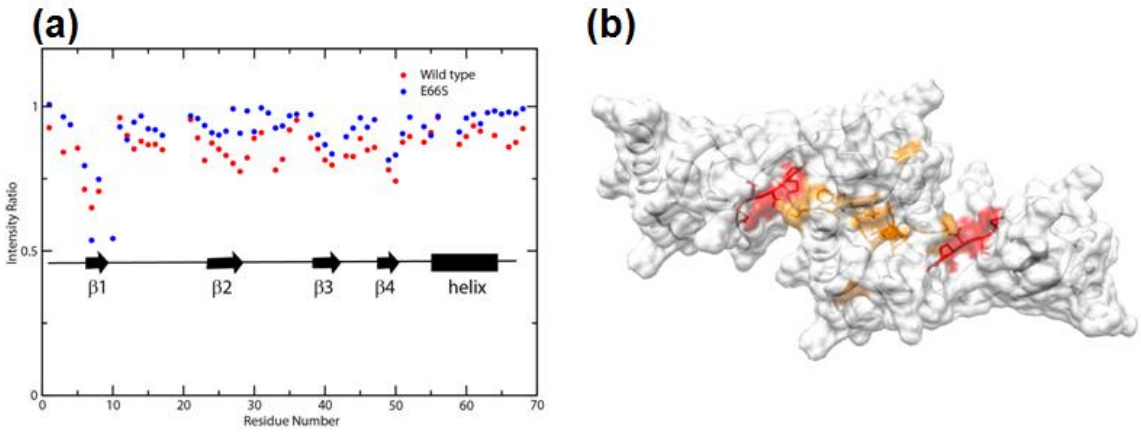


Figure 4.3. Identification of Dimer-Dimer Contacts from Cross-Saturation Data. (a) Plot of residue-specific cross-saturation-induced amide proton signal intensity changes for WT and E66S CCL5. **(b)** Surface plot of CCL5 tetramer model with residues identified as being specifically perturbed in the WT (residues 26-29, 33, 34, 66, and 67) colored orange and the dimer interface (residues 6-10) colored red.

Validating the Tetramer Binding Interface through Hydroxyl Radical Footprinting MS

The binding interface of the CCL5 tetramer was also validated through hydroxyl radical footprinting MS. UV radiation-induced hydrogen peroxide photolysis was used to produce hydroxyl radicals that rapidly oxidize exposed protein side chains, and the degree of oxidation was interpreted as an indicator of relative solvent exposure. By comparing the solvent exposure of the E66S mutant, which is a mixture of monomer and dimer at the concentrations studied in this experiment (20 μM, pH 4.5), with WT under the same conditions, regions that are protected in the dimer-dimer interface of the tetramer can be identified by a decrease in the rate of oxidation. In this experimental setup the radical was produced and consumed on a sub-microsecond time scale to ensure that modifications stemming from nonnative side chain exposures, i.e., as a result of conformation changes caused by oxidation of the protein, were kept

to a minimum [55]. For quantitation the amount of oxidation of each peptide within CCL5 was determined by measuring the amount of unmodified peptide compared to an isotopically labeled internal control, then MS/MS was used to determine the major residues oxidized for each peptide, providing amino acid level resolution. **Figure 4.4(a)** compares the peptides and residues protected in WT CCL5 to those protected in the E66S mutant. H23, E26, Y27, F28, Y29, V40, F41, and C50 were found to be highly protected from oxidation in the WT structures compared to the E66S dimer, whereas oxidation of M0, P2, Y3, and C11 on the N-terminal peptide and E61, Y62, M67, and S68 on the C-terminal peptide are only slightly more protected. M0, P2, Y3, C11, H23, V40, and C50 appear to be protected in the monomer-monomer interface of the dimer, and their increases in protection may reflect a depletion of monomer concentration or some additional stabilization of the dimer on the oligomerization. **Figure 4.4(b)** shows the locations of the remaining protected residues on the surface of the CCL5 tetramer. It is evident that the binding surface identified by protection of these residues from hydroxyl radical footprinting agrees very well with that of the cross-saturation data. The hydroxyl radical footprinting data was able to probe the heart of the dimer-dimer interface as shown in **Figure 4.4(b)**.

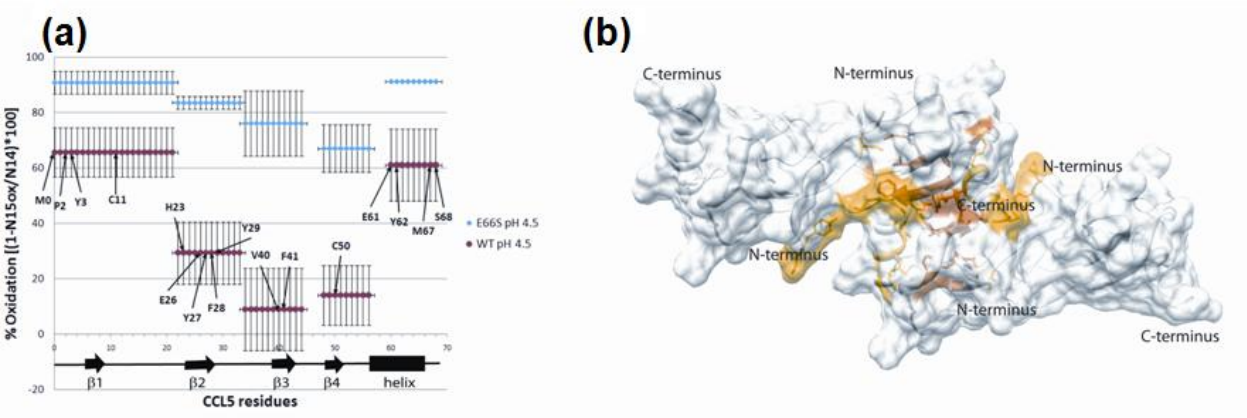


Figure 4.4. Identification of Dimer-Dimer Contacts from Hydroxyl Radical Footprinting. **(a)** Plot of residue-specific hydroxyl radical modification percentage for WT and E66S CCL5. The degree of modification is analyzed at the peptide level and the major sites of oxidation are identified at the residue level for each peptide. Residues identified as being protected from modification are indicated. **(b)** Surface plot of the CCL5 tetramer with residues identified as being in the tetrameric interface (residues 26-29, 41, 61, 62, 67, and 68) colored orange.

Formation of Higher-Order Oligomers

The tetramer model in **Figures 4.2-4.4** is one that leads naturally to the formation of higher order oligomers. Use of the residues of one monomer in the dimer leaves those residues of in the second monomer of the dimer open to interactions that propagate the structure to linear polymers of even order (hexamers, octamers, etc.). Native spray MS on WT CCL5 at pH 4.5 indicates the presence of higher oligomers (**Figure 4.5**). MS requires much lower concentration than NMR. The ability to use lower concentrations (~10 μM) causes less aggregation of CCL5 in solution; however, the loss of solvent from electrospray droplets is known to promote and allow detection of oligomers that normally form at higher concentrations. In addition to large amounts of monomer and dimer, significant amounts of tetramer, hexamer, and octamer are detected, but

no trimer, pentamer, or heptamer. Monomer is abundant because the concentration used is close to that of the Kd of the dimer [131, 149]. Our results indicate the higher-order oligomers are a result of a concatenation of the dimer structure, rather than the oligomerization of the monomer structure (resulting in odd-numbered oligomers), or oligomerization of the tetramer (which would not form hexamer). Hydroxyl radical footprint of both WT and E66S mutant CCL5 at pH 7 (**Figure 4.6**) showed further support for the oligomerization through the dimer subunit. An overall increase in protection is observed, possibly due to the decrease in monomer content and in the percentage of exposed dimer ends as the oligomers grow larger. More importantly, the highest percentage increase in protection is observed on peptide 22-33 that contains several residues suggested to comprise the dimer-dimer interface of the tetramer. The residues involved in the tetramer interface (26, 27, 28, 67, and 68) are contained in the two peptides that experience the largest change in the amount of oxidation, when comparing mutant and WT at pH 7. If you compare the same peptides in the WT sample, the tetramer interface residues experience approximately a 34% (26, 27, 28) and 57% (67, 68) decrease in oxidation at pH 7 versus pH 4.5 (**Figure 4.4**).

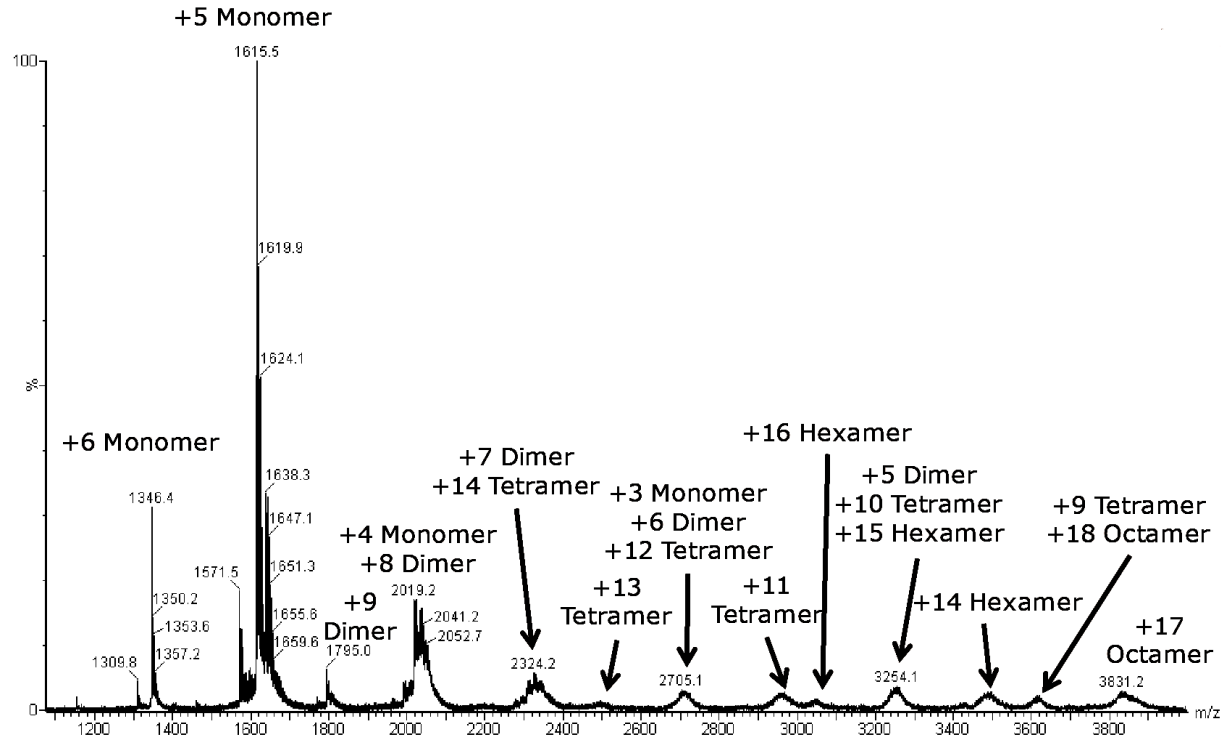


Figure 4.5. WT pH 4.5 Native Spray Mass Spectrometry Data. Native spray mass spectrum of WT CCL5 (10 μM) at pH 4.5. Even-numbered oligomers from dimer to octamer are observed, indicating that the oligomer is built from a concatenation of dimer subunits.

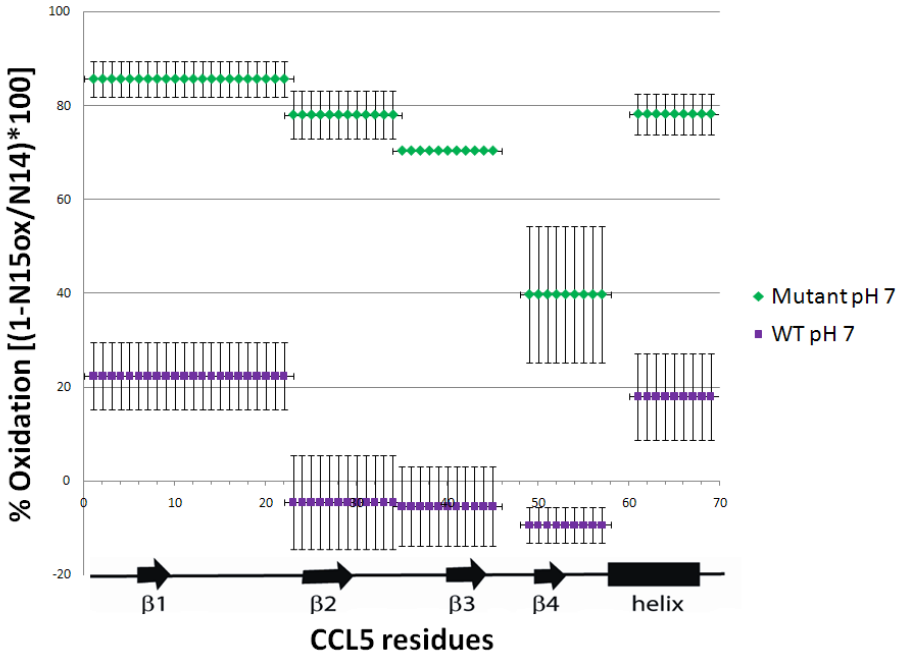


Figure 4.6. Hydroxyl Radical Footprinting Data at pH 7. Plot of residue-specific hydroxyl radical modification percentage for WT and E66S CCL5 at pH 7. The degree of modification is analyzed at the peptide level and the major sites of oxidation are identified at the residue level for each peptide.

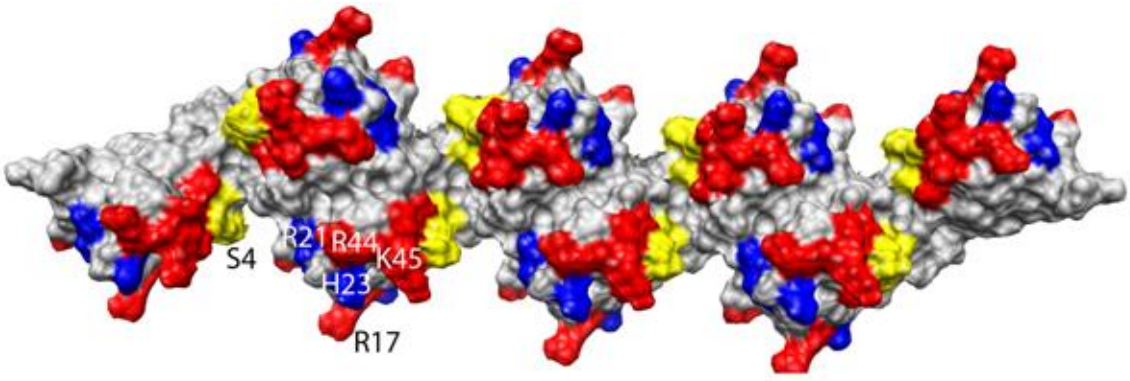


Figure 4.7. Extended Octamer Model. Surface plot of the CCL5 octamer model with residues pertured by GAGs (residues 44-48, 55, and 56) colored red and CCR5 contact residues at pH 6 (residues 16, 17, 21, and 23) colored blue. The N terminus of CCL5 (colored yellow) is perturbed by GAGs and binds to CCR5 in the CCL5 monomer.

The octamer model in **Figure 4.7** was constructed by joining four dimers using the same interface contacts identified in the tetramer. The type of interaction employed should normally propagate indefinitely. Native spray MS of WT CCL5 at pH 7 reveals a very broad, unresolved peak consistent with the presence of very large, highly charged and heterogeneous oligomers (**Figure 4.8**), supporting the idea that the dimers propagate indefinitely to point where MS cannot resolve the individual oligomers. Furthermore, an image of the large oligomers could be visualized using electron microscopy.

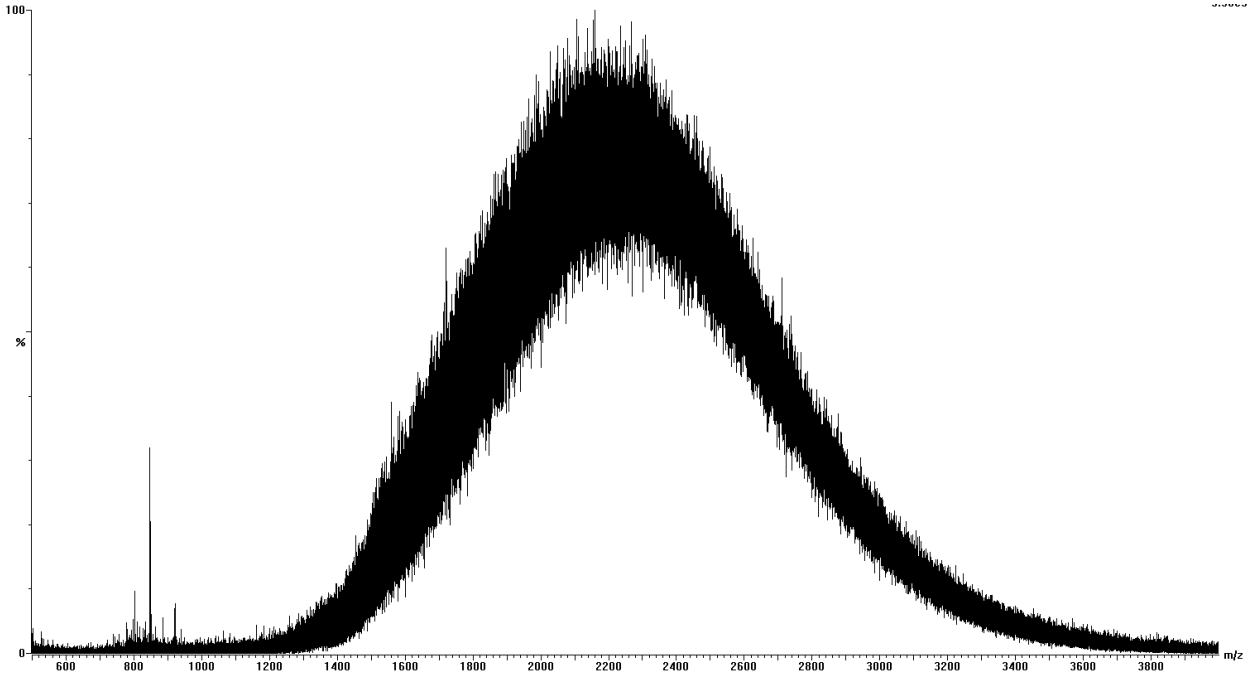
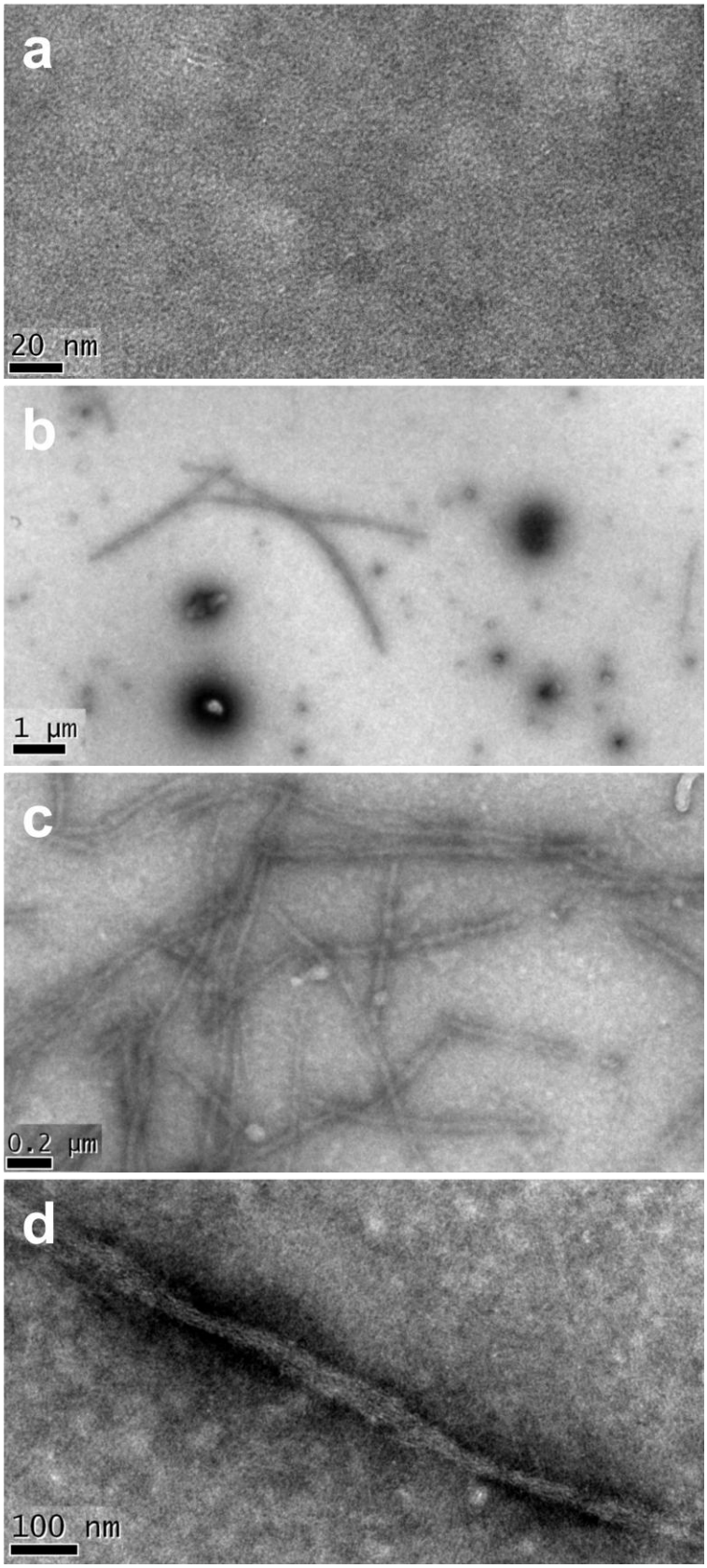


Figure 4.8. WT pH 7 Native Spray Mass Spectrometry Data. Native spray mass spectrum of WT CCL5 (10 μM) at pH 7.

Electron Microscopy

Electron microscopy (EM) was employ in an attempt to visualize the predicted pH-induced oligomers. The samples were stained using a tungstate based dye to scatter imaging electrons. Most biological materials are nearly "transparent" to the electron beam; staining the

sample adds electron density which results in more interactions between the electrons in the primary beam and those of the sample, which provides contrast in the resulting image. Initially the 100 μM WT CCL5 at pH 7 was prepared for viewing by EM. It was apparent that analyte was on the sample grid; however, the EM image appeared to contain analyte that was too densely packed to resolve (**Figure 5.3c**). A 10-fold dilution of the WT pH 7 sample solution allowed much better resolution and representative EM images are found in **Figure 4.9**. Imaging of the pH 7 buffer resulted in no fibrous material (**Figure 4.9(a)**). A gallery of images of WT CCL5 at pH7 under identical conditions from a single grid is presented in **Figure 4.9(b-g)**. **Figure 4.9(b)** shows the presence of resolved fibers in the 10 μM WT CCL5 sample at pH 7. **Figure 4.9 c-g** represent progressively smaller scales; therefore, zooming in on the detected fibers. The linear bundle of fibers in the images is actually multiple oligomers parallel to each other and can be clearly visualized in **Figure 4.10**. The presence of the bundles is perplexing; using the techniques employed in the study of CCL5 oligomerization there would be no way to predict what we see in the EM images. This is an exciting find; however, further investigation will be needed to determine the significance of the bundle of fibers formed by CCL5 under these conditions. The biological concentration of CCL5 is likely higher than 10 μM leaving us unable to conclude exactly how large the CCL5 oligomers can grow or predict the size of the bundles formed.


a
20 nm
b
1 µm
c
0.2 µm
d
100 nm

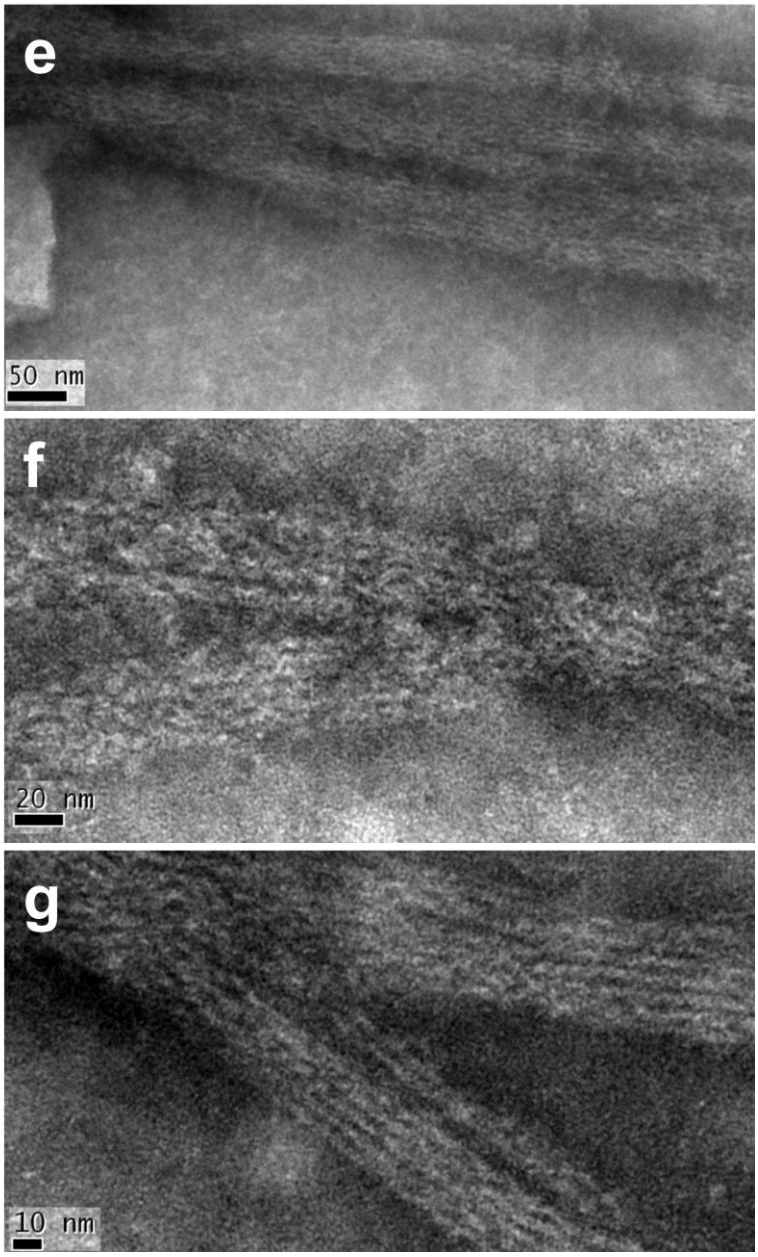


Figure 4.9. Electron Microscopy Images of Buffer and WT CCL5 at pH 7. (a) Buffer at pH 7, bar = 20 nm. The remaining electron microscope images are taken of 10 μM WT CCL5 at pH 7, bars = (b) 1 μm, (c) 0.2 μm, (d) 100 nm, (e) 50 nm, (f) 20 nm, and (g) 10 nm.

Using **Figure 4.9(g)** to measure the width of the oligomers, one oligomer is measured to be roughly 3 nm across and fibrous bundles are more than 1 μm throughout the sample. The length of the oligomers are much larger than initially expected and suggests the reason behind the unresolved mass spectrum of WT CCL5 at pH 7 (**Figure 4.8**). The EM images largely support our data that hypothesizes long fibrous assemblies of CCL5 dimers at pH 7.

CONCLUSIONS

The compelling oligomerization model presented was generated by integrating complementary data from multiple techniques, including existing crystal structure dimers, solution NMR, SAXS data, native spray MS, and MS hydroxyl radical footprinting. The results conclude that CCL5 is capable of forming oligomers without changing its dimeric structure or orientation of a C_{2V} symmetry axis. The resulting model is elongated, providing a mechanism for propagating dimer-dimer interactions indefinitely resulting in long, linear polymeric chains. Electron microscopy confirmed the elongated assembly of CCL5 into fiber-like structures.

Molecular Basis for the Dimer-Dimer Interaction

The residues in the dimer-dimer interface were identified through a combination of hydroxyl radical footprinting and NMR cross-saturation data. The interface primarily involves contacts between residues 25 and 30 in the second β stand and the residues at the C-terminal helix of one monomer of one dimer and similar residues in the neighboring dimer. The interface contact surface involves both hydrophobic and electrostatic interactions. Aromatic ring-stacking interactions in the middle of the interface occur between Y27s on separate dimer units. F28 provides further hydrophobic interaction surface by making cross-unit contacts with L65 and I62 on the C-terminus (**Figure 4.10(a)**). An intermolecular salt bridge between K25 on one dimer unit and E26 from neighboring unit stabilizes the interface. Additional salt bridges could be form

between K25 and E66 and R44, which are positioned closely to K25 as well (**Figure 4.11(b)**). Furthermore, the involvement of E66 and E26 in the oligomerization interface provides a rationalization for the mutants, the change these residues to serine or alanine, to have less propensity for oligomerization [134].

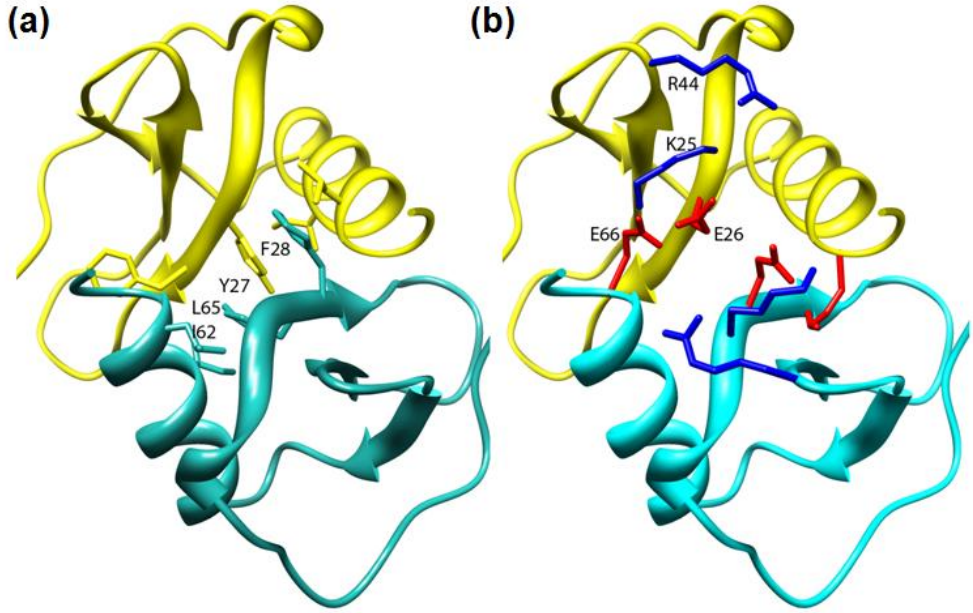


Figure 4.10. CCL5 Dimer-Dimer Interface Details. (a) Details of inter-dimer hydrophobic interactions formed by Y27, F28, I62 and L62. (b) Electrostatic interactions at the dimer-dimer interface. K25, E26, E66 and R44 form pairs of electrostatic bonds.

Predicted Interactions with GAGs

GAG binding is important in leukocyte migration; therefore, it is appropriate to investigate how the proposed oligomer model will interact with GAGs. Residues K44, K45, N46, R47, and to a lesser extent K55 and K56, highlighted red in **Figure 4.7**, are the most prominent residues involved in GAG binding. Notably, **Figure 4.7** shows these residues are all solvent accessible and form a linear arrangement along the CCL5 oligomer to where the GAG could

bind. Additionally, the linear arrangement offers explanation for GAG-promoted chemokine oligomerization [150, 151]. In the absence of GAG the electrostatic repulsive force would keep the CCL5 oligomer size small; however, the negatively charged sulfated GAGs would bind and neutralize the repulsive forces between dimers, increasing aggregation propensity. This may account for the ability of even low molecular weight GAGs to cause aggregation [151, 152]. GAG polymers have the ability to bridge binding sites on multiple chemokine subunits stabilizing the oligomers through avidity, somewhat akin to molecular Velcro.

Simultaneous Interaction with Chemokine Receptors and GAGs

CCL5 residues involved in binging to CCR5 have been identified by chemical shift perturbation from titrations with a CCR5 N-terminal peptide. CCR5 peptide has a stronger interaction with CCL5 monomer at low pH (3.8). The residues involved in the monomer-monomer interface of the dimer are among the most strongly perturbed [149], which is consistent with the monomer involvement in activation of CCR5 at low concentrations. Conversely, at pH 6 oligomers can interact with CCR5 at high concentrations and the perturbed residues shift to residues 16, 17, and 23. In **Figure 4.7** these residues are colored blue and are fully accessible for receptor interaction in our proposed oligomer.

Although never experimentally demonstrated, our model suggests mechanisms for chemokines immobilized on the cell surface to simultaneously interact with receptors and GAGs. **Figure 4.7** shows the GAG-binding residues and CCR5-interacting residues occupying separate regions of the CCL5 oligomer model, suggesting simultaneous interactions via the same area of the oligomer or on different blocks of the polymer. The linear arrangement of the binding sites could promote diffusion of leukocytes along the CCL5 polymer. One-dimensional diffusion offers eminent advantages over multi-dimensional diffusion and is exploited by many other

cellular contexts [153, 154]. Thus, facilitated diffusion on GAG chains could be effective in cell migration.

ACKNOWLEDGEMENTS

This chapter was adapted from "Oligomeric Structure of the Chemokine CCL5/RANTES from NMR, MS, and SAXS Data". Wang, X.; Watson, C.; Sharp, J. S.; Handel, T. M.; Prestegard, J. H. *Structure* **2011,** *19* (8), 1138-1148. We thank Bruker AXS for use of their NANOSTAR system and Brian Jones for collecting and evaluating the bioSAXS data on WT CCL5. We thank Dr. Hsiau-Wei Lee for assistance with RDC collection. Thanks to Dr. Elizabeth R. Wright and Hong Yi at the Robert P. Apkarian Integrated Electron Microscopy Core Facility at Emory University for obtaining the electron microscopy images. We also acknowledge the National Center for Research Resources (a part of the NIH) for financial support of the Resource for Integrated Glycotechnology and the CCL5 project (P41-RR005351) and the National Institute of General Medical Sciences K99 program for support of X.W. (K99GM088483)), and the National Institute of Allergy and Infectious Disease for support of T.M.H. (RO1AI37113). The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

CHAPTER 5

STRUCTURAL INVESTIGATION OF CCL5/RANTES INTERACTION WITH GLYCOSAMINOGLYCANS

ABSTRACT

CCL5 is a chemokine expressed by T-cells that is important for targeting sites of tissue injury, infection, and certain types of tumor metastasis. By interacting with sulfated glycosaminoglycans (GAGs) on cell surfaces, CCL5 controls cell migration by chemotactic signaling. The binding site of CCL5 for GAGs establishes a concentration gradient used for chemotactic migration. Biological importance and potential therapeutic targets of chemokine-GAG interactions has motivated studies to define and characterize the structure, affinity, and interactions. By combining dynamic light scattering, electron microscopy imaging, and hydroxyl radical footprinting, we plan to investigate CCL5 dimer (E66S CCL5 exists mostly as a dimer) both in the presence and absence of the three GAGs to characterize the interactions.

INTRODUCTION

In addition to CCL5 oligomerization discussed in the previous chapter it is important to investigate CCL5-GAG interactions. It is becoming apparent that the complete function of the chemokine system requires dynamically changing structures and interactions including chemokine oligomer association and dissociation, binding to GAGs and receptors, and possibly receptor oligomerization. In addition to interactions with chemokine receptors, some if not all chemokines have essential interactions with GAGs.

Structurally, it is not surprising that chemokines bind to GAGs. Most chemokines are highly basic proteins with many arginine, lysine, and histidine residues on their surfaces, resulting in high positive charge at physiological pH. The positively charged protein, therefore, favorably interacts with negatively charged GAGs rich in sulfate groups. The charge interaction implies a non-specific electrostatic chemokine-GAG interaction; however, previous research suggests that there is specificity [155, 156]. For example, while CCL5 does not have the highest affinity for cation exchange resin, compared other CC chemokines it binds the tightest to heparin sepharose beads [157], evidence for selectivity.

The importance of chemokine-GAG interactions were initially supported by studies that demonstrated binding of chemokines to purified GAGs *in vitro* [158], and to endothelial cell surface GAGs *in vitro* [159] and *in vivo* [160]. The biological relevance and selectivity of chemokine-GAG interaction was compellingly illustrated with chemokine variants (CCL5, CCL2, and CCL4) that had mutations to inhibit their heparin binding ability, but only minimally perturbed their ability to bind to their respective receptors and promote cell migration *in vitro*. However, the mutants were unable to induce cell migration *in vivo* like their WT counterparts, demonstrating GAG interactions are essential for full *in vivo* activity [67]. Most of the chemokine-GAG systems show that minimally the interaction provides a mechanism for chemokine localization and concentration on the cell surfaces, which provides guidance cues for migrating cells. Without the interactions a rapid diffusion of chemokines would occur, dissolving directional gradients, and resulting in chemokine concentrations lower than needed to activate receptors.

In an attempt to determine GAG binding sites on the surface of chemokines, basic residues with linear sequence motifs such as BBXB (where B is a basic residue) were mutated,

and heparin binding effects *in vitro* were evaluated. Proudfoot et al. determined residues R44, K45, and R47 of CCL5/RANTES are involved in the heparin binding site [157]. However, many chemokines oligomerize on GAGs, suggesting the GAG binding epitopes could be significantly larger and more distributed in higher order oligomeric structures [124]. Several tetrameric structures of chemokines have been crystalized [121, 122, 124, 127], and once the GAG binding site were mapped onto the tetramer structures, it could be suggested that oligomerization might increase the chemokine affinity for GAGs by providing more surface area for binding. Another notion suggests different types of GAGs, which are extremely heterogeneous, may induce different chemokine structures, which could contribute to the specificity [161]. Oligomerization could also cause or be required of simultaneous receptor and GAG binding; simultaneous binding is not currently known, but the extended CCL5 oligomer forms in such a way that residues for both receptor and GAG binding are aligned linearly on the surface of the oligomer [162].

CCL5 has shown great selectivity for different GAG families, because the affinities for heparin and chondroitin sulfate differ by three orders of magnitude [159]. The differing affinities introduce a level of selectivity that is only beginning to be explored by *in vitro* binding and activation assays. Since chondroitin sulfate (CS) is known to bind to CCL5, and the capture and purification of homogeneous CS hexamers was previously describe [163, 164], we sought to structurally investigate CCL5's interaction with three homogeneous GAGs, an ARIXTRA-like heptamer (**Figure 5.1a**), and 4,4,4- and 6,4,4-sulfated chondroitin sulfate (**Figure 5.1b and c**). Exploring the difference in binding using the homogeneous ligands will allow us to determine if the sulfation pattern or amount of sulfation affects binding.

E66S mutant of CCL5 does not form higher order aggregates at micromolar concentrations, existing mostly as a dimer at pH 4.5. Exploring GAG interactions with E66S mutant should allow determination of changes involved with ligand binding alone and not GAG induced oligomerization. By combining light scattering data, electron microscopy imaging, and mass spectrometry techniques, we seek to characterize GAG interactions with CCL5.

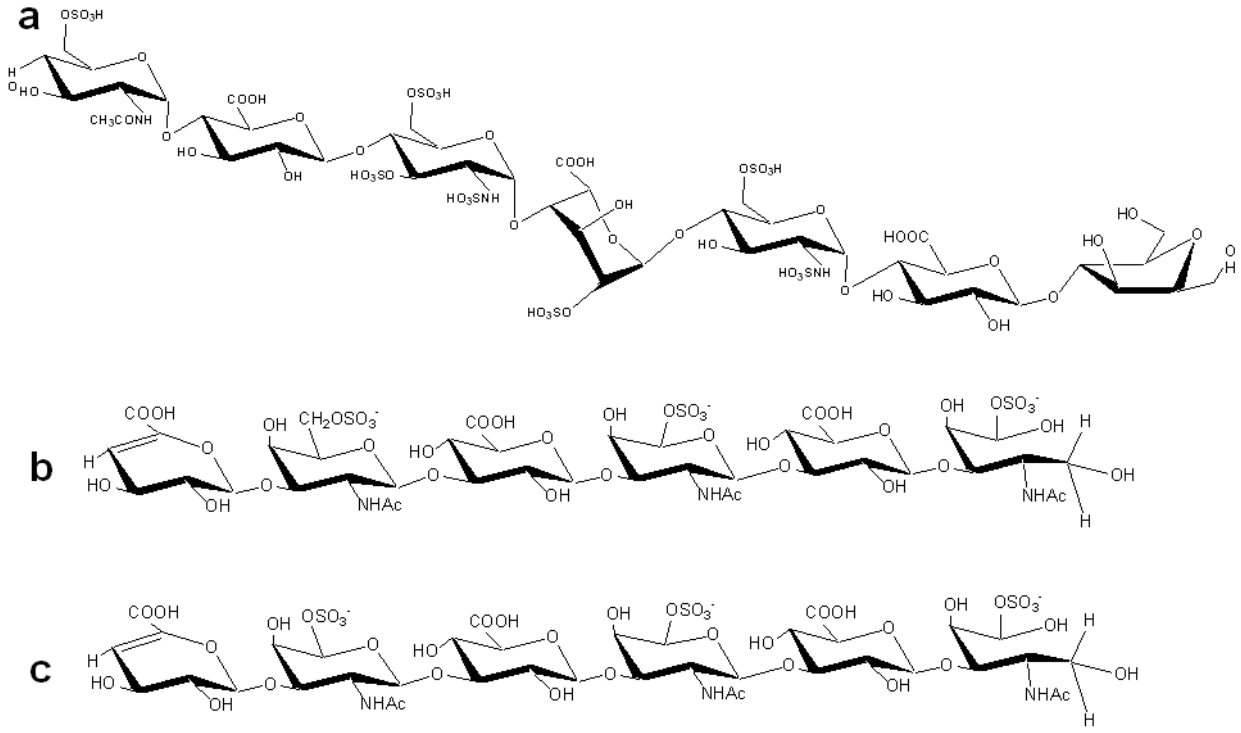


Figure 5.1. Structures of Glycosaminoglycans. Structures of ARIXTRA-like heptamer (**a**), 6,4,4-sulfated (**b**) and 4,4,4-sulfated (**c**) chondroitin sulfate.

MATERIALS AND METHODS

Materials

ARIXTRA-like heptamer was donated courtesy Professor Jian Liu at UNC-Chapel Hill. Chondroitin sulfate-A sodium salt from bovine trachea, hyaluronidase from sheep testes type V, Sephadex G-15 resin (fractionation range of dextrans < 1.5 kDa), Dowex 50WX8-100 ion-

exchange resin, acetonitrile, catalase from bovine liver, ammonium bicarbonate, and L-glutamine were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO). Bio-Gel P-10 gel resin (fractionation range of dextrans from ~1.5 to ~20 kDa) and polypropylene chromatographic columns (1.5 x 120 cm for size-exclusion chromatography, and 50 x 1.0 cm for desalting) were obtained from Bio-Rad Laboratories (Hercules, CA). 30% hydrogen peroxide and formic acid were purchased from J.T. Baker (Phillipsburg, NJ). Methionine amide (MA) was purchased from Bachem (Torrace, CA). Dithiothreitol (DTT) was purchased from Fisher Biotech (Fair Lawn, NJ). Sequencing grade modified trypsin was purchased from Promega Corporation (Madison, WI). Purified water (18 MΩ) was obtained from a Milli-Q Synthesis system (Millipore, Billerica, MA). All chemicals were used without further purification unless noted.

Expression of and E66S CCL5

E66S mutant was expressed in *E. coli* using a pET23a vector with conventional $^{15}NH_4Cl$ or $^{14}NH_4Cl$ containing M9 media for labeled and unlabeled CCL5. They were purified according to Czaplewski et al. [134]. Briefly, inclusion bodies of CCL5 were solubilized in 6 M guanidine hydrochloride and purified on a Superdex 75 column. Purified protein was then subjected to fast dilution refolding with refolding buffer consisting of 100 mM Tris, pH 8.0, 100 μM reduced glutathione and 10 μM oxidized dimeric glutathione. The refolded protein was then dialyzed against 0.1% TFA and purified further using strong anion exchange chromatography.

Preparation of 4,4,4- and 6,4,4-Sulfated Chondroitin Sulfate Hexasaccharides

CS hexasaccharides were prepared as previously described [164]. Briefly, hyaluronidase (11.5 mg) was used for glycosidic bond cleavage of chondroitin sulfate-A (153 mg) in buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.0) at 37 °C for 48 hours. The depolymerized samples were subjected to size exclusion chromatography on a Bio-Gel P-10 column using 1 M

NaCl containing 10% ethanol as the mobile phase to obtain oligosaccharide fractions with uniform length. The fraction containing hexasaccharides was concentrated, re-suspended in 2 mL of nano-pure water for desalting on a Sephadex G-15 column using nano-pure water mobile phase, and lyophilized. The desalted, dried hexasaccharide sample was reduce to corresponding galactitol forms (-ol) using an equivalent weigh of sodium borohydride ($NaBH_4$) in 2 mL water for 3 h. The reduction reaction was stopped by adding a molar equivalent of acetic acid and incubating in an ice bath for 1 h, followed by desalting on a Sephadex G-15 column and lyophilization [164]. The reduced hexasaccharide fraction was further separated by strong anion exchange (SAX) chromatography using a linear gradient of NaCl (10 mM/min) in H_2O (pH 5.0). Previous work by Huang et al. [163] and Pomin et al. [164] were consulted to determine the SAX chromatography peak that corresponds to the hexasaccharides of interest. NMR by Younghee Park determined that each chondroitin sulfate ligand was pretty pure.

Dynamic Light Scattering

A DynaPro NanoStar system (Wyatt Technology Corporation, Santa Barbara, CA) was used to conduct all dynamic light scattering (DLS) measurements. The DLS is equipped with a 658 nm laser. All samples were prepared at 100 μM and the measured sample was held at 25 °C by a temperature controlled sample holder and allowed to equilibrate for 60 seconds prior to analysis. Each size measurement was determined from 20 scans, 5 seconds per scan. All DLS data were collected and analyzed using DYNAMIC V6™ Software – Version 6.3.01. All reported particle radii are calculated from intensity based particle size distributions.

Electron Microscopy

Electron microscopy (EM) was performed and images were captured by Hong Yi and Elizabeth Wright at the Robert P. Apkarian Integrated Electron Microscopy Core at Emory

University using a JEOL JEM-1400 Transmission Electron Microscope. WT CCL5+ARIXTRA-like heptamer at pH 4.5 samples were prepared at 10 and 100 μM protein concentrations with 120% heptamer. In each case, a carbon grid was placed face down onto 5 μL of sample on a piece of laboratory film for 1 – 3 minutes, the grid is blotted with filter paper, placed on a drop of 1% phosphotungstic acid (PTA) negative staining dye for 30 seconds, and blotted again. The resulting grid is ready for EM imaging.

Hydroxyl radical footprinting MS

Samples of ^{15}N-labeled E66S mutant CCL5 (20 μM or 100 μM) in acetate buffer (pH 4.5) containing 100 mM H_2O_2 with 20 mM glutamine to limit radical half-life were flowed through the beam path of a KrF laser at 248 nm, and pulsed so that each segment of sample was irradiated with a single ~20 ns UV pulse with a small buffer region between irradiated segments to help account for sample diffusion and laminar flow effects. The oxidation was immediately quenched using methionine amide (0.5 μg/μL) and catalase (0.5 μg/μL). Prior to trypsin digestion, unoxidized and unlabeled E66S mutant (approximately equimolar) was added to the corresponding irradiated samples to serve as an internal standard for quantitation. Ammonium bicarbonate (50 mM) and DTT (5 mM) were added to the CCL5 samples (25 μL) and the samples were incubated at 55°C for 3 hours to denature and reduce the protein, respectively. Sequencing grade modified trypsin was added (0.2 μg/μL) and incubated at 37ºC for 48 hours while rotating to digest the protein samples. Tryptic peptides were analyzed in triplicate by UPLC-MS using a Waters Synapt G2 mass spectrometer to quantify the change in oxidation using the percent fractional oxidation. MS was used to quantitate the unmodified, unlabeled peptide peak to unmodified, ^{15}N labeled peptide peak for each tryptic peptide, and the fractional oxidation was calculated from the reduction of signal for the unmodified peptide in the oxidized

sample relative to that for the identical light isotopomer in the internal standard. LC-MS/MS on a LTQ-FTMS of oxidized peptides was used to manually identify sites of oxidation.

RESULTS AND DISSCUSSION

Dynamic Light Scattering

Dynamic light scattering (DLS) is a relatively fast method of characterizing biomolecules to determine size distribution of small particles in solution [165]. DLS can be used to distinguish between a homogeneous monodisperse and an aggregated sample. DLS detects the fluctuations of scattering laser light intensity due to Brownian motion of molecules in solution. DLS data primarily yield diffusion coefficient distributions of particles, which can be processed into distributions of hydrodynamic radius [165]. Hydrodynamic radius or "size" refers to the radius of a particle modeled as a sphere that diffuses in solution. Molecular weight calculations are based on a spherical model, which does not accurately represent the elongated structure of larger CCL5 oligomers.

While dynamic light scattering is capable of distinguishing between monomer or dimer, it is unable to resolve monomer from small oligomers. In general two different species must differ in hydrodynamic radius by a factor of 2 or more to be resolved into separate peaks. To our advantage in the case of CCL5, a factor of 2 in radius corresponds roughly to a factor of 8 in molecular mass. Typically monomer to octamer will be merged into one average peak, allowing detection of very small amounts of higher mass aggregates. DLS should not be used to quantitate the amount of any aggregates but rather used as a technique for relative comparisons.

DLS was used to estimate the size and distribution of aggregates in various CCL5 samples. The results in **Figure 5.2** show the hydrodynamic radii for E66S at pH 4.5, pH 7, +CS4,4,4, and +CS6,4,4, correspond to a calculated molecular weight of 19, 16, 19, and 17 kDa,

respectively; the molecular weight of CCL5 dimer is ~16 kDa. Dynamic light scattering (DLS) data suggests CCL5 in the presence of CS ligands does not induce oligomerization; therefore, changes in oxidation identified by hydroxyl radical footprinting (HRF) should be a result of ligand binding only and not ligand induced oligomerization. Addition of the ARIXTRA-like heptamer to E66S mutant or WT CCL5 results in a cloudy solution that is unable to be investigated by DLS. The DLS data for the WT samples suggests what was determined in the Chapter 4, which is that WT CCL5 forms mostly tetramer at pH 4.5 and higher order aggregates at pH 7. As mentioned previously, the molecular weight attributed to WT CCL5 are not completely accurate because molecular weight is calculated based on a spherical model and we have determine that CCL5 oligomerizes in an extended linear fashion.

When the ARIXTRA-like heptamer is added to E66S CCL5, the protein solution immediately turns cloudy. DLS requires particles to be randomly diffusing through the solution; therefore, the point where sedimentation of the particle dominates the diffusion process creates an upper size limit on DLS analysis. The presence of a few large particles or 'dust' can cause the scattering intensity to fluctuate significantly resulting in unusable measurements. In the case of 100 μM E66S CCL5 protein plus excess heptamer, very large protein aggregates are formed that should not be removed from the sample for analysis because aggregate size is of interest. The scattering intensity is proportional to the square of the molecular weight, and as the molecular weight of the aggregate increases the scattering intensity increase to a point that it cannot be accurately detected by the DLS instrumentation used here. A 10-fold dilution of the 100 μM sample still resulted in a turbid solution. The turbid protein – ligand solution scatters the laser light like 'dust' particles resulting in the software halting the analysis suggesting the aggregates in the solution are greater than the 1 μm radius, the upper size range limit for the DynaPro Star.

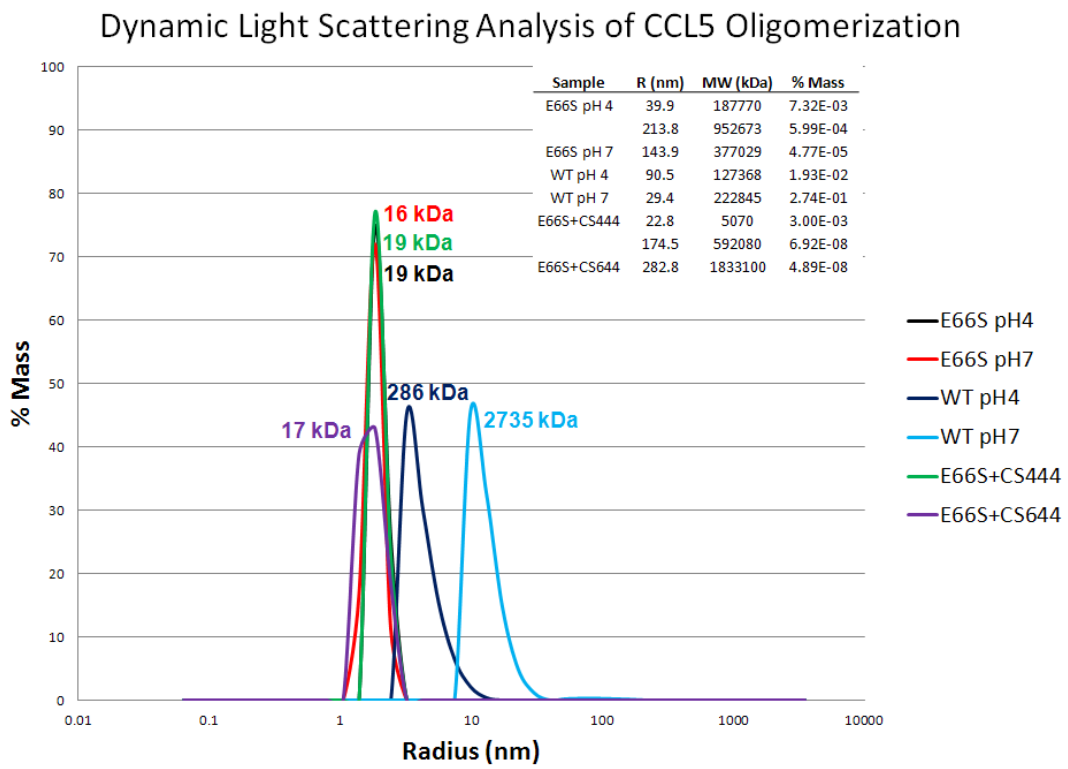


Figure 5.2. Dynamic Light Scattering Analysis of CCL5 Oligomerization. Distribution of the hydrodynamic radius of E66S at pH 4.5 (black), E66S at pH 7 (red), WT at pH 4.5 (dark blue), WT at pH 7 (light blue), E66S+CS4,4,4 (green), and E66S+CS6,4,4 (purple). Percent mass was normalized using the sum of the total mass. Tabular insert lists other detected oligomers that represent <1% of the sample by mass.

Electron Microscopy

The ARIXTRA-like heptamer caused both E66S mutant and WT CCL5 solutions to become cloudy suggesting aggregation; therefore, electron microscopy imaging was used to visualize the structure of CCL5 in presence of the heptamer. Looking at the WT + heptamer image (**Figure 5.3b)** compared to the negative control buffer pH 4.5 image (**Figure 5.3a**)**,** it is apparent that some higher order structure exists for WT CCL5 in the presence of the ARIXTRA-

like heptamer. The EM image of 100 µM WT CCL5 at pH 7 in **Figure 5.3c** looks similar to the EM image of 10 µM WT CCL5 + heptamer at pH 4.5 in **Figure 5.3b**. The 100 µM WT image at pH 7 indicates the presence of higher order oligomers; however, it is so densely packed at this concentration that the oligomers cannot be resolved. As mentioned in the Chapter 4, the WT CCL5 oligomers could be resolved at 10 µM protein concentration (**Figure 4.10**). Therefore, it is possible that the ARIXTRA-like heptamer induces oligomerization to such an extent that 10 µM protein concentration is too high to resolve the oligomers because they are so densely packed on the EM grid.

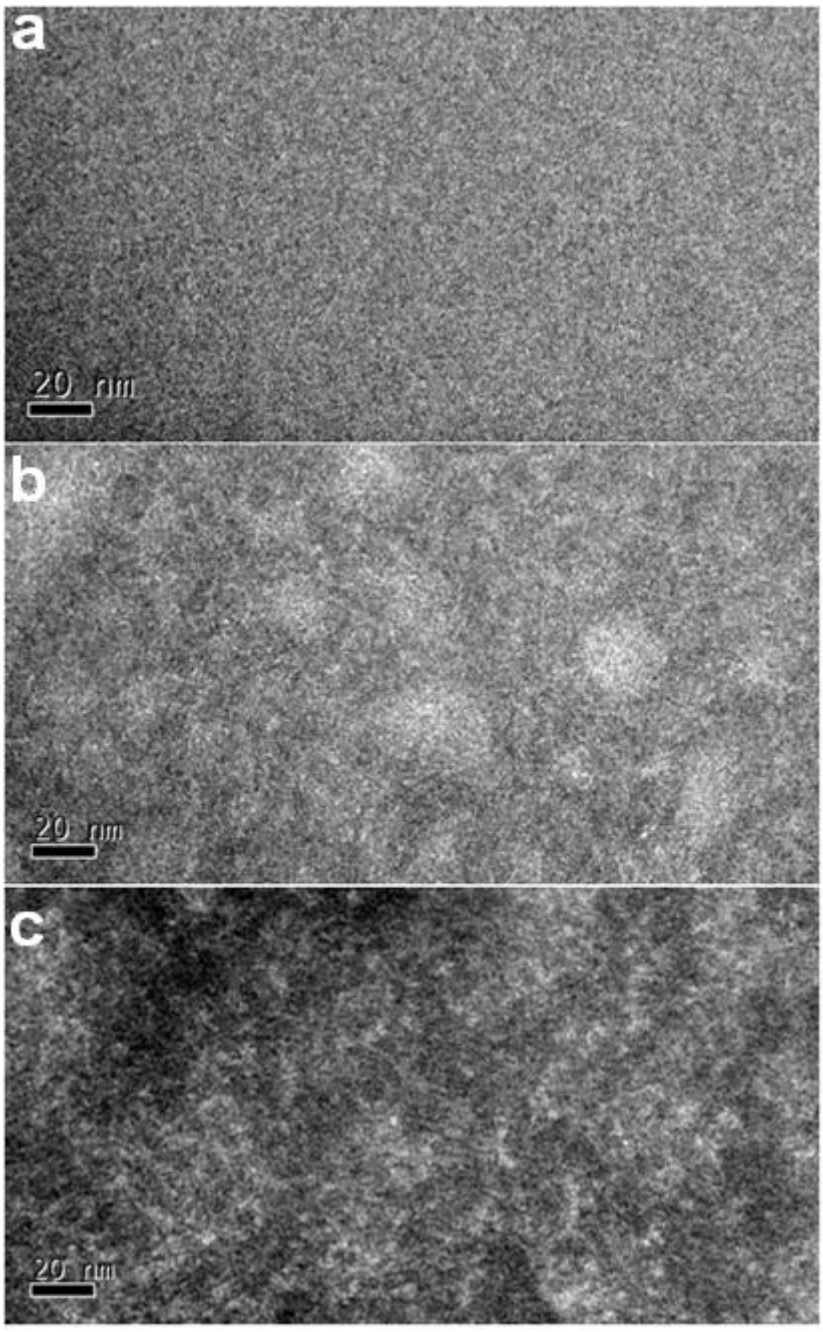


Figure 5.3. Electron Microscopy Images. (a)EM image of pH 4.5 buffer. **(b)**EM image of 10 μM WT CCL5 plus 12 μM ARIXTRA-like heptamer, and (**c**) 100 μM WT CCL5 at pH 7, bars = 20 nm.

Hydroxyl Radical Footprinting: E66S with Chondroitin Sulfate Ligands

Initially in 2009 the chondroitin sulfate hexamer binding to CCL5 was of interest to Dr. Prestegard's NMR group. A crystal structure of CCL5 dimer complexed with a heparan sulfate dimer was examined by X-ray crystallography, but NMR data with CCL5-chondroitin sulfate hexamer suggests the heparan sulfate dimer is too small to be biologically relevant. **Figure 5.4** shows the NMR HSQC data illustrating the oligomerization of E66S CCL5 in the presence of CS6,4,4 ligand (top), and the inability to obtain useful binding information on the CS6,4,4 ligand using NMR; therefore, the CCL5-CS ligand interaction was investigated using hydroxyl radical footprinting.

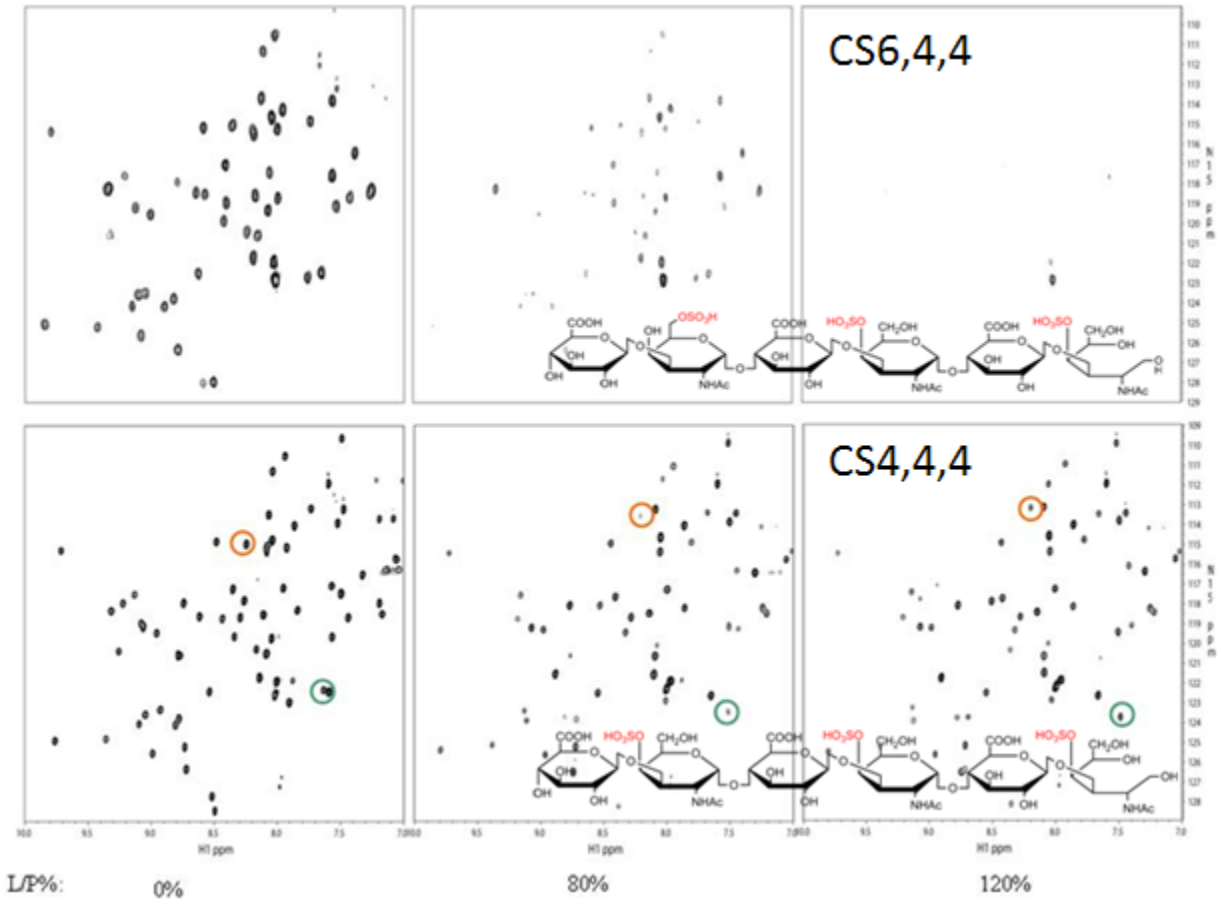


Figure 5.4. NMR HSQC Data of E66S CCL5 – CS Ligand Binding.

The NMR experiments were performed at 100 μM protein concentration; however, HRF experiments do not need protein concentrations that high. By working at between 10 – 20 μM protein, hopefully, the oligomerization induced by the CS6,4,4 binding would be reduced and allow the binding site of CS ligand to be mapped to CCL5 dimer; E66S mutant CCL5 is a non-aggregating mutant at micromolar concentrations and at pH 4.5 exists mostly as a dimer. **Figure 5.5** represents the initial HRF data obtained using the E66S CCL5 and CS ligands given to us by the NMR group. The data in **Figure 5.5** agrees well with the NMR data. CS4,4,4 shows slight protection from oxidation at peptides 1-22, 23-34, and 34-44; however, CS6,4,4 shows an overall protection from oxidation at each tryptic peptide. Peptide 34-44 experiences a significant decrease in the amount of oxidation, suggesting very tight binding. The negative number is a result of the initial ^{15}N to ^{14}N ratio in the control and the normalization. The average ^{15}N to ^{14}N ratio of each peptide in the control sample is 1.880. To get the fraction oxidized for each peptide in graph the following formula was used: 1- ((N15 peak intensity/N14 peak intensity)/1.880). The lack of oxidation at peptide 34-44 suggest very tight CS6,4,4 ligand binding in that region of the protein.

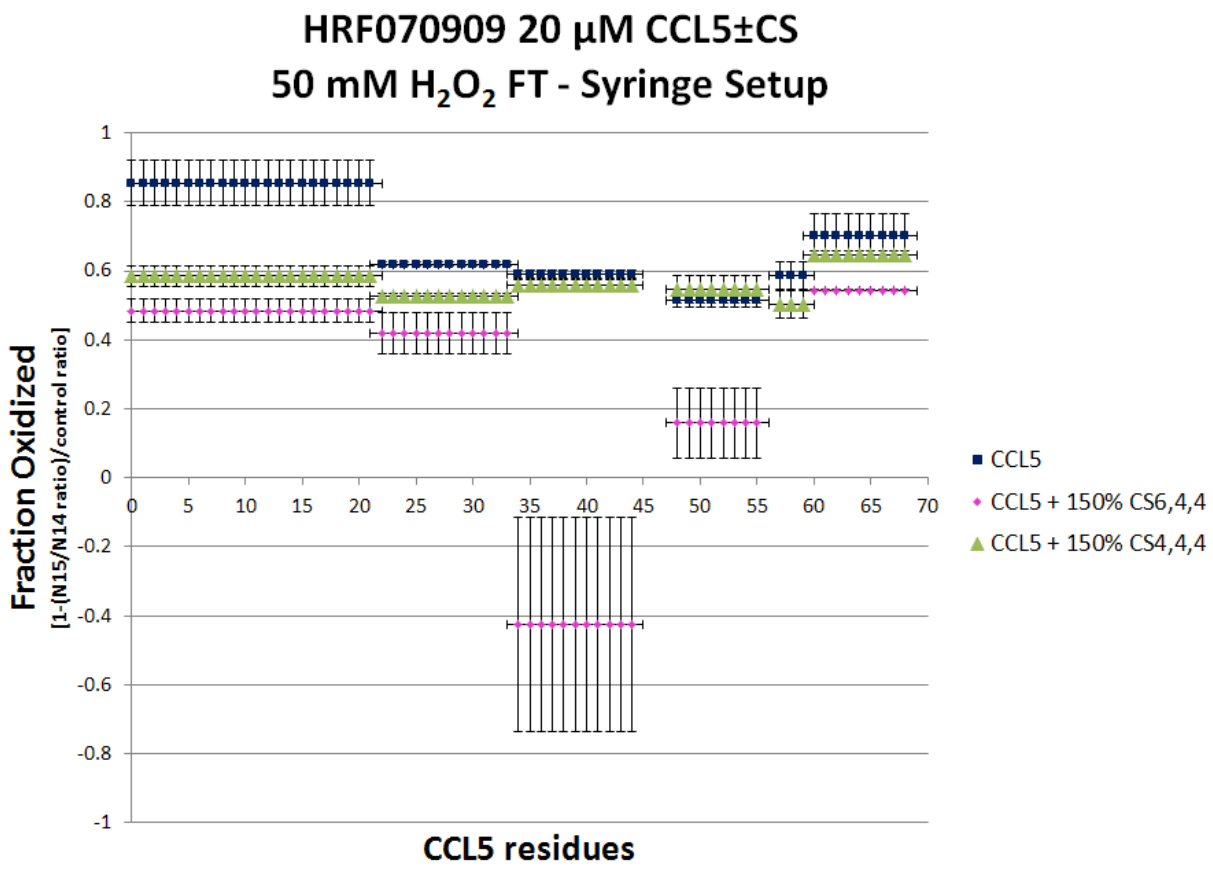


Figure 5.5. Initial CCL5-CS Ligand Binding Data.

However, shortly after the data presented in **Figures 5.4-5.5** was obtained, it was determined that the previous CS6,4,4 ligand was not a pure sample. Pure CS hexamers are hard to obtain due to the tightly spaced peaks when collecting aliquots from the strong anion exchange chromatography. Three different samples of apparently pure CS6,4,4 and CS4,4,4 hexamer samples have been used for further investigation of the CCL5-CS ligand binding. **Figure 5.6** is fairly representative of all the data obtained while studying the CCL5-CS ligand binding. Over technical triplicates there is no statistical difference in the amount of oxidation at any tryptic peptide of CCL5. Several changes have been made to the irradiation protocol, such as using one laser shot for irradiation instead of the flow system to lower the overall oxidation,

increasing the radical concentration for more protein oxidation, and increasing the protein concentration to more closely match that of NMR concentrations; however, most of the results resemble those found in **Figure 5.6**.

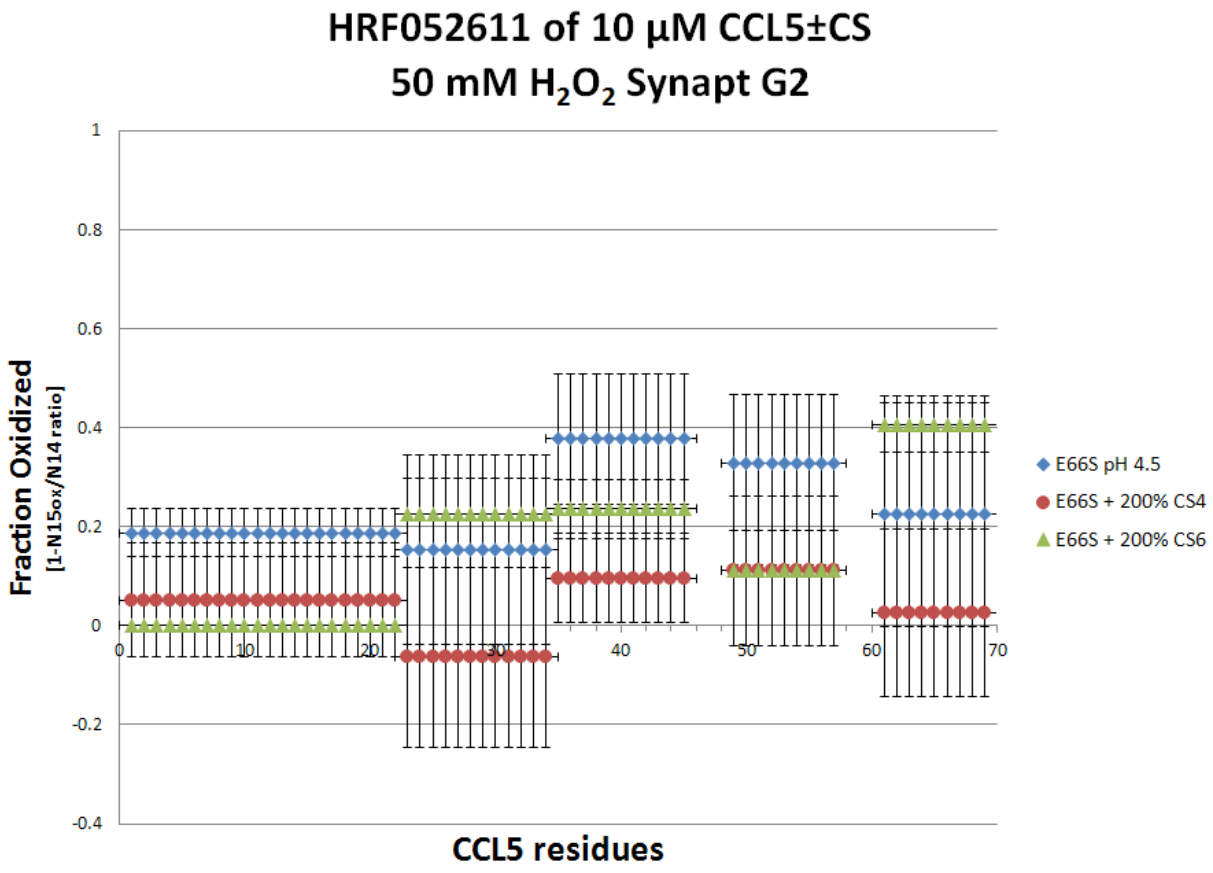


Figure 5.6. Representative HRF Data for CCL5-CS Ligand Binding.

Of mention angiotensin peptide was added to each protein solution as a dosimeter to monitor the radical dose in each protein solution. If the angiotensin under each condition is oxidized to the same extent, the results of the footprinting data are due to differences in protein solvent exposure and not radical dose. Angiotensin oxidation has been statistically the same for each oxidized replicate in all the experiments shown or mentioned (data not shown).

Of all the HRF data gathered for the CCL5-CS ligand project, there was one set of promising data from one experiment. The two of the three previously determined residues involved in heparin binding are located in the third tryptic peptide (peptide 35-45) graphed in **Figure 5.7a**. Interestingly, the third peptide is one of two peptides that experience a statistical difference in the amount of oxidation for the two CS ligands. Peptide 35-45 is more protected from oxidation in the presence of the CS4,4,4 than the CS6,4,4 hexamer. The HRF data suggests that CS4,4,4 hexamer binds more tightly to E66S CCL5. In some of HRF experiments the CS4,4,4 ligand seems to bind more tightly or result in more protection from oxidation than the CS6,4,4 ligand which does not agree with the NMR data. NMR HSQC data from CS ligand titrations with CCL5 corresponds to dissociation constants of 45 μM and 20 μM for CS6,4,4, and CS4,4,4 ligands, respectively (data not shown).

E66S CCL5 also experiences slightly less oxidation at peptide 23-34, second tryptic peptide, in the presence of CS4,4,4. Both peptide 23-34 and 35-45 experience differences in oxidation between each ligand are in the same region of the protein that is protected from oxidation by oligomerization. Since DLS data suggests that the CS ligands do not induce oligomerization of E66S CCL5, the results are must be due to changes in solvent accessibility of the peptides. Preliminary molecular modeling data (**Figure 5.7b**) done with fully flexible CS4,4,4 docked on to CCL5 dimer using AutoDock Vina [166] shows the hexamer flexibility in the groove of the dimer. Each hexamer color represents one of the top five positions determined by molecular modeling. If the hexamer position shown in green in **Figure 5.4b** is real, that particular position of the hexamer would cause protection from oxidation in peptide 23-34.

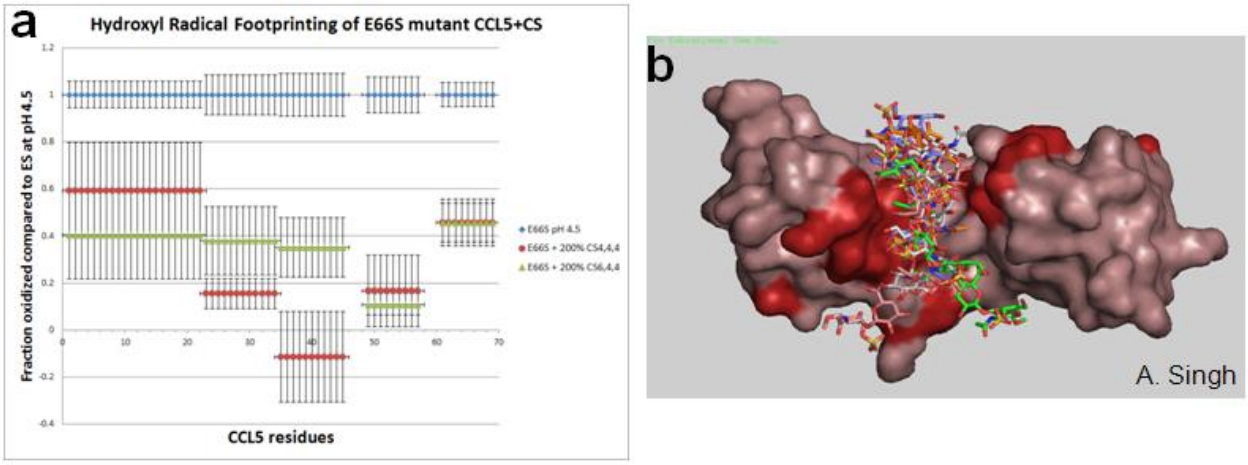


Figure 5.7. CCL5 Dimer-Chondroitin Sulfate Ligand Interaction. (**a**) HRF data showing the faction oxidized compared to E66S at pH 4.5 ± 200% CS ligands for each CCL5 tryptic peptide. (**b**) Top five poses for docking fully flexible CS4,4,4 on CCL5 dimer using AutoDock Vina [166]. Residues affected by NMR chemical shift perturbation shown in red. Each of the top five positions is designated by different colored hexamers.

The peptides most protected from oxidation are the same peptides that experienced a decrease in oxidation due to oligomerization. E66S CCL5 was used for initial experiments because it does not oligomerize; however, it is possible that the CS ligands bind differently to the CCL5 dimer than the CCL5 tetramer or higher oligomers. We have begun to investigate CS ligand binding with WT CCL5 in hopes of gathering a more comprehensive idea about how CCL5 interacts with chondroitin sulfate hexamers.

Considering the data presented and discussed in this section, we must consider the possibility that HRF is not the ideal method for studying CCL5 – chondroitin sulfate binding. CCL5 binding shows the widest discrimination between GAGs; CCL5 binds to chondroitin sulfate with the lowest affinity when compared to heparin, heparan sulfate, and dermatan sulfate

[159]. The chondroitin sulfate hexamers likely do not bind tightly enough to CCL5 to study using HRF. HRF is not the ideal method to probe arginines and lysines because they do not react very readily with hydroxyl radicals. Therefore, if the ligand binding does not involve more than the three residues mentioned, the likelihood of HRF detecting the change in oxidation is low. If it is known that two arginines and one lysine is involved in heparin binding and the same residues are expected to be involved in CS binding, other labeling techniques can be used. HDX can be used to determine the binding constant and the protein-ligand interactions [19]. There are five lysines in the CCL5 sequence that are available for acetylation to map CS ligand binding [167]. The CCL5 sequence contains five arginines that can be chemically modified to probe the ligand binding pocket [168]. Chemically modifying arginines and/or lysines may be better probes for mapping CCL5-CS ligand binding residues.

Hydroxyl Radical Footprinting: E66S with ARIXTRA-like heptamer

Mixtures of heterogeneous aggregates are not conducive to DLS or NMR analysis, but HRF technique is capable of analyzing such mixtures and was used to assess the differences in oxidation caused by the heptamer. We analyzed two concentrations of E66S CCL5 with heptamer (**Figure 5.8**). At 20 μM protein (**Figure 5.8a**), each detected tryptic peptide experiences protection in the presence of the heptamer. However, at 100 μM protein (**Figure 5.8b**) only two of the peptides experience a statistical difference in the amount of oxidation. Of note, the two peptides at 100 μM protein that show a difference in the amount of oxidation are the same two peptides that are most protected from oxidation by the heptamer at 20 μM protein. There are three possibilities to explain the changes in the amount of oxidation detected. First, since no changes in oxidation were seen with ligand binding in the previous section, the changes in oxidation could be due to the induced oligomerization and not directly attributable to ligand

binding. Second, the heptamer could be binding tighter to CCL5 than the CS ligands, and the tighter binding might have a larger effect on the footprinting results. Finally, the most likely explanation is the combination of induced oligomerization and tighter ligand binding are causing the changes in the amount of oxidation when the heptamer is present. A trend exists in the HRF data of the heptamer; however, further experiments are required to draw conclusions about CCL5-heptamer binding.

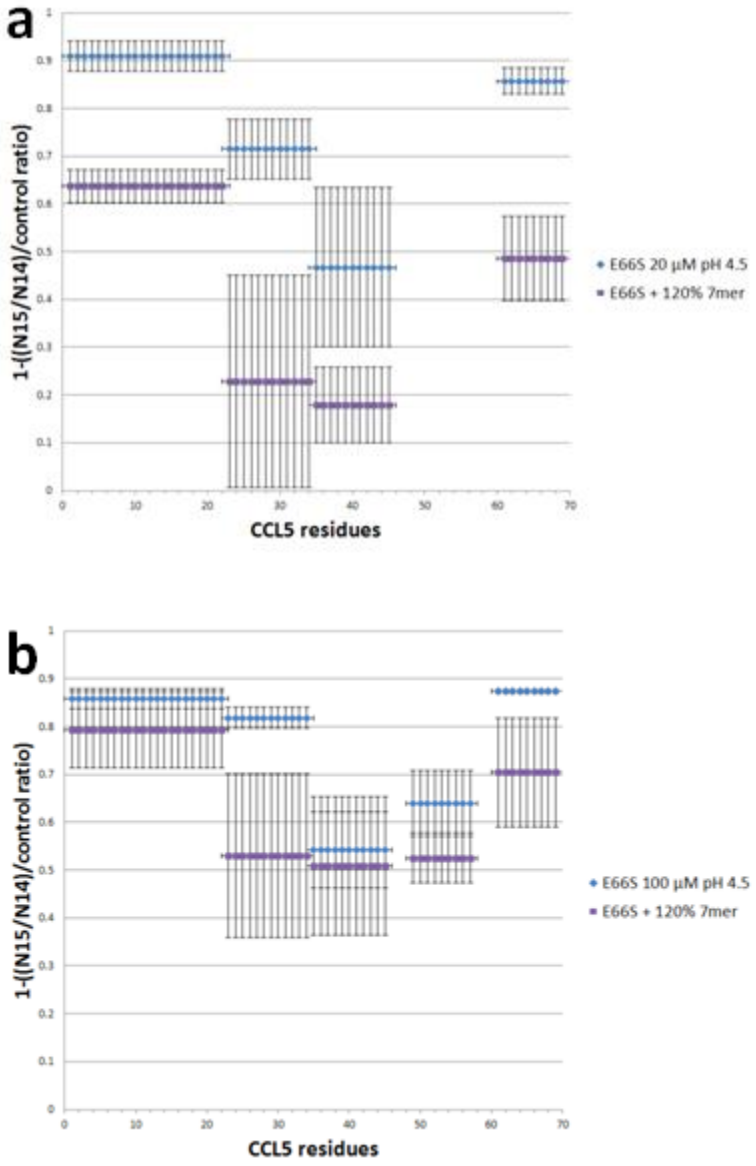


Figure 5.8. HRF of E66S CCL5 with ARIXTRA-like Heptamer. Fraction oxidized of 20 μM (**a**) and 100 μM (**b**) E66S CCL5 oxidized ± 120% heptamer is plotted.

CONCLUSIONS

After years of work on the CCL5 – ligand project, it has become apparent that hydroxyl radical footprinting is perhaps not the best protein residue probe for this particular system. Because heparin binds more tightly to CCL5, HRF could be a better probe for the CCL5 – heparin system. For future studies involving CCL5 – ligand binding, other protein residue probes should be investigated and quantified to determine the best labeling process for the system.

ACKNOWLEDGEMENTS

We thank James Prestegard for the use of his DLS and Arunima Singh for the molecular modeling. We thank Hong Yi and Elizabeth Wright at the Robert P. Apkarian Integrated Electron Microscopy Core at Emory University for collecting the EM images. This project was supported by grants from the National Center for Research Resources (5P41RR005351-23) and the National Institute of General Medical Sciences (8 P41 GM103390-23) from the National Institutes of Health. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

CHAPTER 6

CONCLUSIONS

Hydroxyl radical protein footprinting (HRF) is a robust, reproducible, and flexible method making it a promising structural analysis technique to complement the traditional high resolution techniques, X-ray crystallography and NMR. The present work demonstrates a novel method for producing hydroxyl radicals for protein footprinting, and the diversity of HRF by investigating the structure of an important biological chemokine and unveiling HRF's potential to determine conformation quality of protein pharmaceuticals.

Hydroxyl radical protein footprinting is based on the idea that the radicals must be produced and react with the protein of interest before large scale protein motion or oxidatively induced conformational changes can occur. Fast photo-oxidation of peroxide (FPOP) is one readily used technique to produce hydroxyl radicals for protein footprinting because of its fast nature; however, the presence of hydrogen peroxide can cause several conformational problems with the protein. The development of a peroxide-free technique that produces hydroxyl radicals on a submicrosecond time scale extends the applicability of HRF technique and allows peroxide-sensitive proteins, such as metal-binding proteins, to be studied while still providing extensive labeling of the native protein conformation.

With an increasing number of therapeutic proteins coming off-patent and the development of follow-on formulations of protein pharmaceuticals, methods for fast conformational analysis of biosimilars are required. Generation of an abbreviated hydroxyl radical protein footprinting technique allowed us to conformational compare a brand name

pharmaceutical, Neupogen®, to expired recombinant GCSF samples; additionally, conformational variation were determined for different samples of erythropoietin and interferon α-2A. HRF uses picomoles of sample for analysis and is an attractive technique for structural characterization during the production stage, as well as final quality control.

The ability to automate the HRF analysis prior to MS injection would greatly reduce the sampling time and increase the desirability of the HRF technique to protein pharmaceutical production labs for conformation analysis. Initially a single laser pulse method was attempted for the analysis of the protein pharmaceuticals. One well of a 384 well plate holds one replicate of one sample and a laser mirror directs the laser beam 90 degrees to hit the plate on the bench top, such that only the well of interest is irradiated. The idea lends to straightforward development of a programmable robotic platform for sample preparation and hydrogen peroxide addition immediately before irradiation by one laser pulse. The problem we found with the one laser shot technique was a decrease in amount of protein oxidation as compared to the FPOP syringe setup. We were unable to focus the beam like we do for the syringe setup due to damage threshold limitations of the laser mirror need to direct the beam at a 90 degree angle. Obtaining a laser mirror with a higher damage threshold or sufficiently increasing the laser power should allow for the development of a one laser shot FPOP technique with automation capabilities. With or without an automated setup, the abbreviated HRF method can provide a fast analysis technique with moderate resolution to labs producing the follow-on formulations of protein pharmaceuticals during each step of production. Screening with HRF during production could isolate the problematic steps and further investigation of the already oxidized protein samples would be used to acquire residue resolution.

Although CCL5 mutants and WT at low pH can be studied using NMR and X-ray crystallography, HRF and mass spectrometry were employed to investigate the oligomerization of WT CCL5 at physiological pH. Native spray MS determined the subunits of the oligomer to be concatenation of dimers. HRF was able to probe the highly oligomerized WT CCL5 at pH 7 and saw the same trend in protection from oxidation, concluding the CCL5 likely oligomerizes in a linear fashion with the residues from one dimer subunit available to orient itself with the other monomer of the next dimer subunit. The proposed oligomer will hopefully help us to understand the fascinating role CCL5 plays *in vivo* and in development of drug targets.

The difficulties presented by CCL5-ligand project forced us to explore different analytical techniques to answer questions about CCL5. Investigating CCL5 has exposed me to various analytical techniques, for example, NMR, dynamic light scattering and electron microscopy imaging. If the project is going to continue using HRF, homogeneous heparin hexamers might be a better GAG sample to use because CCL5 has a much higher binding affinity for heparin than for chondroitin sulfate. There are also other chemical labeling techniques that may better serve the examination of ligand binding; directly labeling the arginines and lysines in CCL5 will give a better picture of the ligand interaction if the previous heparin binding residues are still involved. It would be worthwhile in pursuing the EM imaging of lower concentration WT CCL5 in the presence of the ARIXTRA-like heptamer; it would be very interesting to visualize the oligomerization of CCL5 in the presence of a ligand to see if it forms similar bundles to those made in the absence of ligand, and to see any differences in oligomerization formation.

The binding site for heparin has been mapped on a CCL5 dimer. More information about how CCL5 works *in vivo* could be determined from binding site information of other GAGs in

the presence of higher order oligomers of CCL5. NMR has already shown that a CCL5 tetramer can be investigated, and ligand interactions with a CCL5 tetramer might produce different results from that of ligand binding to the dimer. The preliminary molecular modeling experiments docking CS4,4,4 onto a CCL5 dimer show a large variation in the top five positions of CS4,4,4 hexamer within the binding pocket of the dimer; the various positions suggest that the CS4,4,4 ligand is highly flexible in the presence of a dimer. The degree of flexibility could suggest that CS ligand binding to a tetramer or higher order oligomers might better constrict the GAG movement within the protein binding pocket allowing a defined binding region to be identified. The presented work gives us the knowledge that new strategies for investigating CCL5-ligand binding need to be explored.

REFERENCES

1. Sali, A., et al., *From words to literature in structural proteomics.* Nature, 2003. **422**: p. 216-225.

2. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.

3. Edwards, A.M., et al., *Protein production: feeding the crystallographers and NMR spectroscopists.* Nat. Struct. Biol, 2000. **7**: p. 970–972.

4. Yu, H., *Extending the size limit of protein nuclear magnetic resonance*, 1999, National Acad Sciences.

5. Xu, G. and M. Chance, *Hydroxyl radical-mediated modification of proteins as probes for structural proteomics.* Chem. Rev, 2007. **107**(8): p. 3514-3543.

6. Mendoza, V.L. and R.W. Vachet, *Probing protein structure by amino acid-specific covalent labeling and mass spectrometry.* Mass Spectrometry Reviews, 2009. **28**(5): p. 785-815.

7. Sinz, A., *Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes.* Journal of Mass Spectrometry, 2003. **38**(12): p. 1225-1237.

8. Back, J.W., et al., *Chemical cross-linking and mass spectrometry for protein structural modeling.* Journal of Molecular Biology, 2003. **331**(2): p. 303-313.

9. Woods Jr, V.L. and Y. Hamuro, *High resolution, high-throughput amide deuterium exchange-mass spectrometry (DXMS) determination of protein binding site structure and dynamics: Utility in pharmaceutical design.* Journal of Cellular Biochemistry, 2001. **84**(S37): p. 89-98.

10. King, D., et al., *Use of amide exchange mass spectrometry to study conformational changes within the endopolygalacturonase II-homogalacturonan-polygalacturonase inhibiting protein system.* Biochemistry, 2002. **41**(32): p. 10225-10233.

11. Englander, J.J., et al., *Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry.* Proc Natl Acad Sci U S A, 2003. **100**(12): p. 7057-62.

12. Takamoto, K. and M.R. Chance, *Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes.* Annu Rev Biophys Biomol Struct, 2006. **35**: p. 251-76.

13. Englander, S.W. and N.R. Kallenbach, *Hydrogen exchange and structural dynamics of proteins and nucleic acids.* Quarterly reviews of biophysics, 1983. **16**(04): p. 521-655.

14. Englander, S., et al., *Hydrogen exchange: The modern legacy of Linderstrøm Lang.* Protein Science, 1997. **6**(5): p. 1101-1109.

15. Bai, Y. and W. Englander, *Primary Structure Effects Hydrogen Exchange.* PROTEINS Structure, Function, and Genetics, 1993. **17**: p. 75-86.

16. Garcia, R.A., D. Pantazatos, and F.J. Villarreal, *Hydrogen/deuterium exchange mass spectrometry for investigating protein-ligand interactions.* Assay and drug development technologies, 2004. **2**(1): p. 81-91.

17. Engen, J.R. and D.L. Smith, *Investigating protein structure and dynamics by hydrogen exchange MS.* Anal Chem, 2001. **73**(9): p. 256A-265A.

18. Engen, J.R. and D.L. Smith, *Peer Reviewed: Investigating Protein Structure and Dynamics by Hydrogen Exchange MS.* Analytical Chemistry, 2001. **73**(9): p. 256-265.

19. Zhu, M.M., et al., *Quantification of protein-ligand interactions by mass spectrometry, titration, and H/D exchange: PLIMSTEX.* J Am Chem Soc, 2003. **125**(18): p. 5252-3.

20. Baichoo, N. and T. Heyduk, *Mapping conformational changes in a protein: application of a protein footprinting technique to cAMP-induced conformational changes in cAMP receptor protein.* Biochemistry, 1997. **36**(36): p. 10830-6.

21. Baichoo, N. and T. Heyduk, *Mapping cyclic nucleotide induced conformational changes in cyclicAMP receptor protein by a protein footprinting technique using different chemical proteases.* Protein Science, 1999. **8**(3): p. 518-528.

22. Heyduk, E. and T. Heyduk, *Mapping protein domains involved in macromolecular interactions: a novel protein footprinting approach.* Biochemistry, 1994. **33**(32): p. 9643-50.

23. Heyduk, T., et al., *Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits: hydroxyl-radical protein footprinting.* Proceedings of the National Academy of Sciences, 1996. **93**(19): p. 10162.

24. Garrison, W.M., *Reaction-Mechanisms in the Radiolysis of Peptides, Polypeptides, and Proteins.* Chemical Reviews, 1987. **87**(2): p. 381-398.

25. Garrison, W.M., Kland-English, M., Sokol, H. A., Jayko, M. E., *Radiolytic degradation of the peptide main chain in dilute aqueous solution containing oxygen.* J. Phys. Chem., 1970. **74**(26): p. 4506-4509.

26. Maleknia, S.D., M. Brenowitz, and M.R. Chance, *Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry.* Anal Chem, 1999. **71**(18): p. 3965-73.

27. Xu, G. and M.R. Chance, *Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting.* Anal Chem, 2005. **77**(14): p. 4549-55.

28. Aye, T.T., T.Y. Low, and S.K. Sze, *Nanosecond laser-induced photochemical oxidation method for protein surface mapping with mass spectrometry.* Anal Chem, 2005. **77**(18): p. 5814-22.

29. Hambly, D.M. and M.L. Gross, *Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale.* J Am Soc Mass Spectrom, 2005. **16**(12): p. 2057-63.

30. Hambly, D. and M. Gross, *Laser flash photochemical oxidation to locate heme binding and conformational changes in myroglobin.* International Journal of Mass Spectrometry, 2007. **259**(1-3): p. 124-129.

31. Maleknia, S.D., M.R. Chance, and K.M. Downard, *Electrospray-assisted modification of proteins: a radical probe of protein structure.* Rapid Commun Mass Spectrom, 1999. **13**(23): p. 2352-8.

32. Sharp, J.S., et al., *Photochemical surface mapping of C14S-Sml1p for constrained computational modeling of protein structure.* Anal Biochem, 2005. **340**(2): p. 201-12.

33. Sharp, J.S., J.M. Becker, and R.L. Hettich, *Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry.* Anal Chem, 2004. **76**(3): p. 672-83.

34. Sharp, J.S., J.M. Becker, and R.L. Hettich, *Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry.* Anal Biochem, 2003. **313**(2): p. 216-25.

35. Goldsmith, S.C., et al., *Synchrotron protein footprinting: a technique to investigate protein-protein interactions.* J Biomol Struct Dyn, 2001. **19**(3): p. 405-18.

36. Guan, J.Q., et al., *Mapping the G-actin binding surface of cofilin using synchrotron protein footprinting.* Biochemistry, 2002. **41**(18): p. 5765-75.

37. Liu, R., et al., *Structural reorganization of the transferrin C-lobe and transferrin receptor upon complex formation: the C-lobe binds to the receptor helical domain.* Biochemistry, 2003. **42**(43): p. 12447-54.

38. Chance, M.R., *Unfolding of apomyoglobin examined by synchrotron footprinting.* Biochem Biophys Res Commun, 2001. **287**(3): p. 614-21.

39. Maleknia, S.D., et al., *Determination of macromolecular folding and structure by synchrotron x-ray radiolysis techniques.* Anal Biochem, 2001. **289**(2): p. 103-15.

40. Kiselar, J.G., et al., *Visualizing the Ca2+-dependent activation of gelsolin by using synchrotron footprinting.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3942-7.

41. Schantz, J.T., et al., *Repair of calvarial defects with customised tissue-engineered bone grafts - II. Evaluation of cellular efficiency and efficacy in vivo.* Tissue Engineering, 2003. **9**: p. S127-S139.

42. Smedley, J.G., 3rd, et al., *Probing the pH-dependent prepore to pore transition of Bacillus anthracis protective antigen with differential oxidative protein footprinting.* Biochemistry, 2008. **47**(40): p. 10694-704.

43. Franchet-Beuzit, J., et al., *Radiolytic footprinting. Beta rays, gamma photons, and fast neutrons probe DNA-protein interactions.* Biochemistry, 1993. **32**(8): p. 2104-10.

44. Armstrong, R.C., Swallow, A. J., *Pulse- and gamma-radiolysis of aqueous solutions of tryptophan.* J Radiat Res, 1969. **40**(3): p. 563-579.

45. Kopoldova, J., Hrncir, S., *Gamma-radiolysis of aqueous solution of histidine.* Z. Naturforsch. C.: Biosci., 1977. **32C**(7-8): p. 482-487.

46. Winchester, R.V., Lynn, K. R., *X- and gamma-radiolysis of some tryptophan dipeptides.* Int J Radiat Biol, 1970. **17**(6): p. 541-548.

47. Kiselar, J.G., et al., *Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry.* Int J Radiat Biol, 2002. **78**(2): p. 101-14.

48. Rashidzadeh, H., et al., *Solution structure and interdomain interactions of the Saccharomyces cerevisiae "TATA binding protein" (TBP) probed by radiolytic protein footprinting.* Biochemistry, 2003. **42**(13): p. 3655-65.

49. Watson, C., et al., *Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting.* Anal. Chem, 2009. **81**(7): p. 2496-2505.

50. Buxton, G.V., F.C.R. Cattell, and F.S. Dainton, *Application of Time-Dependent Rate Constant Theory to Reactions of Solvated Electrons - Reaction Distances, Rate Constants and Diffusion-Coefficients in Concentrated Aqueous-Solutions.* Journal of the Chemical Society-Faraday Transactions I, 1975. **71**(1): p. 115-122.

51. Sharp, J.S. and K.B. Tomer, *Analysis of the oxidative damage-induced conformational changes of apo- and holocalmodulin by dose-dependent protein oxidative surface mapping.* Biophys J, 2007. **92**(5): p. 1682-92.

52. Sharp, J.S., et al., *Measurement of multisite oxidation kinetics reveals an active site conformational change in Spo0F as a result of protein oxidation.* Biochemistry, 2006. **45**(20): p. 6260-6.

53. Venkatesh, S., K.B. Tomer, and J.S. Sharp, *Rapid identification of oxidation-induced conformational changes by kinetic analysis.* Rapid Commun Mass Spectrom, 2007. **21**(23): p. 3927-36.

54. Hambly, D.M. and M.L. Gross, *Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale.* Journal of the American Society for Mass Spectrometry, 2005. **16**(12): p. 2057-2063.

55. Gau, B.C., et al., *Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding.* Analytical Chemistry, 2009. **81**(16): p. 6563-6571.

56. Charvátová, O., et al., *Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1*. J. Am. Soc. Mass Spectr., 2008. **19**: p. 1692-1705.

57. Garrison, W.M., *Reaction mechanisms in the radiolysis of peptides, polypeptides, and proteins.* Chemical Reviews, 1987. **87**(2): p. 381-398.

58. Xu, G., K. Takamoto, and M.R. Chance, *Radiolytic modification of basic amino acid residues in peptides: probes for examining protein-protein interactions.* Anal Chem, 2003. **75**(24): p. 6995-7007.

59. Xu, G. and M.R. Chance, *Radiolytic modification of acidic amino acid residues in peptides: probes for examining protein-protein interactions.* Anal Chem, 2004. **76**(5): p. 1213-21.

60. Xu, G., et al., *Secondary reactions and strategies to improve quantitative protein footprinting.* Anal Chem, 2005. **77**(10): p. 3029-37.

61. Xu, G. and M.R. Chance, *Radiolytic modification of sulfur-containing amino acid residues in model peptides: fundamental studies for protein footprinting.* Anal Chem, 2005. **77**(8): p. 2437-49.

62. Nukuna, B.N., M.B. Goshe, and V.E. Anderson, *Sites of hydroxyl radical reaction with amino acids identified by (2)H NMR detection of induced (1)H/(2)H exchange.* J. Am. Chem. Soc., 2001. **123**: p. 1208–14.

63. Goshe, M.B., Y.H. Chen, and V.E. Anderson, *Identification of the sites of hydroxyl radical reaction with peptides by hydrogen/deuterium exchange: prevalence of reactions with the side chains.* Biochemistry, 2000. **39**: p. 1761–1770.

64. Buxton, G.V., et al., *Critical-Review of Rate Constants for Reactions of Hydrated Electrons, Hydrogen-Atoms and Hydroxyl Radicals (.Oh/.O-) in Aqueous-Solution.* Journal of Physical and Chemical Reference Data, 1988. **17**(2): p. 513-886.

65. Appay, V., et al., *Aggregation of RANTES is responsible for its inflammatory properties - Characterization of nonaggregating, noninflammatory RANTES mutants.* Journal of Biological Chemistry, 1999. **274**(39): p. 27505-27512.

66. Campanella, G.S., et al., *Oligomerization of CXCL10 is necessary for endothelial cell presentation and in vivo activity.* J Immunol, 2006. **177**(10): p. 6991-8.

67. Proudfoot, A.E.I., et al., *Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(4): p. 1885-1890.

68. Johnson, Z., et al., *Interference with heparin binding and oligomerization creates a novel anti-inflammatory strategy targeting the chemokine system.* Journal of Immunology, 2004. **173**(9): p. 5776-5785.

69. Murooka, T.T., et al., *CCL5-CCR5-mediated apoptosis in T cells - Requirement for glycosaminoglycan binding and CCL5 aggregation.* Journal of Biological Chemistry, 2006. **281**(35): p. 25184-25194.

70. Charvátová, O., et al., *Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1.* Journal of the American Society for Mass Spectrometry, 2008.

71. Heyduk, T., N. Baichoo, and E. Heyduk, *Hydroxyl radical footprinting of proteins using metal ion complexes.* Met Ions Biol Syst, 2001. **38**: p. 255-87.

72. Bridgewater, J.D. and R.W. Vachet, *Metal-catalyzed oxidation reactions and mass spectrometry: The roles of ascorbate and different oxidizing agents in determining Cu–protein-binding sites.* Analytical Biochemistry, 2005. **341**(1): p. 122-130.

73. Bridgewater, J.D., J. Lim, and R.W. Vachet, *Using Metal-Catalyzed Oxidation Reactions and Mass Spectrometry to Identify Amino Acid Residues Within 10 L of the Metal in Cu-Binding Proteins.* Journal of the American Society for Mass Spectrometry, 2006. **17**(11): p. 1552-1559.

74. Qin, B.Y., et al., *Structural basis of the tanford transition of bovine beta-lactoglobulin.* Biochemistry, 1998. **37**(40): p. 14014-14023.

75. Zheng, X., P. Wintrode, and M. Chance, *Complementary Structural Mass Spectrometry Techniques Reveal Local Dynamics in Functionally Important Regions of a Metastable Serpin.* Structure, 2008. **16**(1): p. 38-51.

76. Zhuang, T., H. Leffler, and J.H. Prestegard, *Enhancement of bound-state residual dipolar couplings: Conformational analysis of lactose bound to Galectin-3.* Protein Science, 2006. **15**(7): p. 1780.

77. Delaglio, F., et al., *NMRPipe: A multidimensional spectral processing system based on UNIX pipes.* Journal of Biomolecular NMR, 1995. **6**(3): p. 277-293.

78. Klassen, N.V., et al., *Fricke dosimetry: the difference between G (Fe3+) for 60Co gamma-rays and high-energy x-rays.* Physics in Medicine and Biology, 1999. **44**: p. 1609-1624.

79. Janik, I., D.M. Bartels, and C.D. Jonah, *Hydroxyl Radical Self-Recombination Reaction and Absorption Spectrum in Water Up to 350 C.* J. Phys. Chem. A, 2007. **111**(10): p. 1835-1843.

80. Nielsen, S.O., B.D. Michael, and E.J. Hart, *Ultraviolet Absorption Spectra of eaq-, H, OH, D, and OD from Pulse Radiolysis of Aqueous Solutions.* Journal of Physical Chemistry, 1976. **80**(22): p. 2482-2488.

81. Weatherly, D.B., et al., *A Heuristic Method for Assigning a False-discovery Rate for Protein Identifications from Mascot Database Search Results*.* Molecular & Cellular Proteomics, 2005. **4**(6): p. 762-772.

82. Bern, M., Y. Cai, and D. Goldberg, *Lookup peaks: a hybrid of de novo sequencing and database search for protein identification by tandem mass spectrometry.* Anal. Chem, 2007. **79**(4): p. 1393-1400.

83. Berman, H.M., et al., *The Protein Data Bank.* Acta Crystallogr. D, 2002. **58**(Pt 6 No 1): p. 899-907.

84. Case, D.A., et al., *AMBER 8*, 2004, University of California: San Francisco.

85. Wang, J., P. Cieplak, and P.A. Kollman, *How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules?* J. Comput. Chem., 2000. **21**(12): p. 1049-1074.

86. Jorgensen, W.L., et al., *Comparison of Simple Potential Functions for Simulating Liquid Water.* J. Phys. Chem., 1983. **79**: p. 926-935.

87. Group, W. *GLYCAM Web*. 18 July 2008 [cited (http://www.glycam.com).

88. Ryckaert, J.-P., G. Ciccotti, and H.J. Berendsen, *Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes.* J. Comput. Phys., 1977. **23**: p. 327-341.

89. Hubbard, S.J. and J.M. Thornton, *NACCESS*, 1993, University College London: London.

90. Cabelli, D.E., *The Reactions of HO2/O2- Radicals in Aqueous Solution*, in *Peroxyl Radicals*, Z.B. Alfassi, Editor 1997, John Wiley and Sons Ltd: New York. p. 407-437.

91. Bielski, B.H.J. and G.G. Shiue, *Reacion rate of superoxide radicals with the essential aminoacids*, in *Oxygen Free Radicals and Tissue Damage*1979, Excerpta Medica: Amsterdam. p. 43-56.

92. Bielski, B.H.J. and A.O. Allen, *Mechanism of Disproportionation of Superoxide Radicals.* Journal of Physical Chemistry, 1977. **81**(11): p. 1048-1050.

93. Bielski, B.H.J., et al., *Reactivity Of Ho2/O-2 Radicals In Aqueous-Solution.* Journal Of Physical And Chemical Reference Data, 1985. **14**(4): p. 1041-1100.

94. von Sonntag, C., *Amino acids, oligopeptides and proteins*, in *The Chemical Basis of Radiation Biology*1987, Tylor & Francis: London. p. 394-457.

95. Bobrowski, K., C. Houee-Levin, and B. Marciniak, *Stabilization and reactions of sulfur radical cations: Relevance to one-electron oxidation of methionine in peptides and proteins.* Chimia, 2008. **62**(9): p. 728-734.

96. von Sonntag, C. and H.P. Schuchmann, *Peroxyl Radicals in Aqueous Solutions*, in *Peroxyl Radicals*, Z.B. Alfassi, Editor 1997, John Willey & Sons: New York. p. 173-234.

97. Takamoto, K. and M.R. Chance, *Radiolytic protein footprinting with mass Spectrometry to probe the structure of macromolecular complexes.* Annual Review of Biophysics and Biomolecular Structure, 2006. **35**: p. 251-276.

98. Janik, I., et al., *Hydroxyl-radical-induced reactions of the poly(vinyl methyl ether) model 2,4-dimethoxypentane in the absence and presence of dioxygen: a pulse radiolysis and product study.* Journal of the Chemical Society-Perkin Transactions 2, 2000(10): p. 2034-2040.

99. Tong, X., J.C. Wren, and L. Konermann, *Effects of Protein Concentration on the Extent of-Ray-Mediated Oxidative Labeling Studied by Electrospray Mass Spectrometry.* Anal. Chem, 2007. **79**(16): p. 6376-6382.

100. Harvey, A.H., J.S. Gallagher, and J. Sengers, *Revised formulation for the refractive index of water and steam as a function of wavelength, temperature and density.* Journal of Physical and Chemical Reference Data, 1998. **27**(4): p. 761-774.

101. Locatelli, F., L. Del Vecchio, and P. Pozzoni, *PURE RED-CELL APLASIA "EPIDEMIC"—MYSTERY COMPLETELY REVEALED?* Peritoneal Dialysis International, 2007. **27**(Supplement 2): p. S303-S307.

102. McKoy, J.M., et al., *Epoetin-associated pure red cell aplasia: past, present, and future considerations.* Transfusion, 2008. **48**(8): p. 1754-1762.

103. Kessler, M., D. Goldsmith, and H. Schellekens, *Immunogenicity of biopharmaceuticals.* Nephrology Dialysis Transplantation, 2006. **21**(suppl 5): p. v9.

104. Srebalus Barnes, C.A. and A. Lim, *Applications of mass spectrometry for the structural characterization of recombinant protein pharmaceuticals.* Mass Spectrom Rev, 2007. **26**(3): p. 370-88.

105. Bobst, C.E., et al., *Detection and characterization of altered conformations of protein pharmaceuticals using complementary mass spectrometry-based approaches.* Anal Chem, 2008. **80**(19): p. 7473-81.

106. Kaltashov, I.A., et al., *Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics.* Biotechnology Advances, 2011.

107. Konermann, L., J. Pan, and Y.H. Liu, *Hydrogen exchange mass spectrometry for studying protein structure and dynamics.* Chem. Soc. Rev., 2010.

108. Kaltashov, I.A., et al., *Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics.* Biotechnol Adv, 2011.

109. Konermann, L., et al., *Mass spectrometry combined with oxidative labeling for exploring protein structure and folding.* Mass Spectrometry Reviews, 2010. **29**(4): p. 651-667.

110. Desjardins, P., J.B. Hansen, and M. Allen, *Microvolume protein concentration determination using the NanoDrop 2000c spectrophotometer.* Journal of visualized experiments: JoVE, 2009(33).

111. Zink, T., et al., *Structure and dynamics of the human granulocyte colony-stimulating factor determined by NMR spectroscopy. Loop mobility in a four-helix-bundle protein.* Biochemistry, 1994. **33**(28): p. 8453-8463.

112. Ross, C.A. and M.A. Poirier, *Protein aggregation and neurodegenerative disease.* Nat Med, 2004. **10 Suppl**: p. S10-7.

113. Klaus, W., et al., *The three-dimensional high resolution structure of human interferon alpha-2a determined by heteronuclear NMR spectroscopy in solution.* J Mol Biol, 1997. **274**(4): p. 661-75.

114. Cheetham, J.C., et al., *NMR structure of human erythropoietin and a comparison with its receptor bound conformation.* Nat Struct Biol, 1998. **5**(10): p. 861-6.

115. Wang, X., et al., *Oligomeric structure of the chemokine CCL5/RANTES from NMR, MS, and SAXS data.* Structure, 2011. **19**(8): p. 1138-48.

116. Czaplewski, L.G., et al., *Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1 alpha, MIP-1 beta, and RANTES - Characterization of active disaggregated chemokine variants.* Journal of Biological Chemistry, 1999. **274**(23): p. 16077-16084.

117. Laurence, J.S., et al., *CC chemokine MIP-1 beta can function as a monomer and depends on Phe13 for receptor binding.* Biochemistry, 2000. **39**(12): p. 3401-9.

118. Paavola, C.D., et al., *Monomeric monocyte chemoattractant protein-1 (MCP-1) binds and activates the MCP-1 receptor CCR2B.* J Biol Chem, 1998. **273**(50): p. 33157-65.

119. Rajarathnam, K., et al., *Neutrophil activation by monomeric interleukin-8.* Science, 1994. **264**(5155): p. 90-2.

120. Braunersreuther, V., et al., *Chemokine CCL5/RANTES inhibition reduces myocardial reperfusion injury in atherosclerotic mice.* J Mol Cell Cardiol, 2010. **48**(4): p. 789-98.

121. Hoover, D.M., et al., *The crystal structure of the chemokine domain of fractalkine shows a novel quaternary arrangement.* J Biol Chem, 2000. **275**(30): p. 23187-93.

122. Jabeen, T., et al., *Structure of mouse IP-10, a chemokine.* Acta Crystallographica Section D-Biological Crystallography, 2008. **64**: p. 611-619.

123. Jin, H., et al., *Structural and functional studies of the potent anti-HIV chemokine variant P2-RANTES.* Proteins, 2010. **78**(2): p. 295-308.

124. Lau, E.K., et al., *Identification of the glycosaminoglycan binding site of the CC chemokine, MCP-1 - Implications for structure and function in vivo.* Journal of Biological Chemistry, 2004. **279**(21): p. 22294-22305.

125. Murphy, J.W., et al., *Heterologous quaternary structure of CXCL12 and its relationship to the CC chemokine family.* Proteins, 2010. **78**(5): p. 1331-7.

126. Ren, M., et al., *Polymerization of MIP-1 chemokine (CCL3 and CCL4) and clearance of MIP-1 by insulin-degrading enzyme.* EMBO J, 2010.

127. Swaminathan, G.J., et al., *Crystal structures of oligomeric forms of the IP-10/CXCL10 chemokine.* Structure, 2003. **11**(5): p. 521-532.

128. Keov, P., P.M. Sexton, and A. Christopoulos, *Allosteric modulation of G protein-coupled receptors: a pharmacological perspective.* Neuropharmacology, 2011. **60**(1): p. 24-35.

129. Roscic-Mrkic, B., et al., *RANTES (CCL5) uses the proteoglycan CD44 as an auxiliary receptor to mediate cellular activation signals and HIV-1 enhancement.* Blood, 2003. **102**(4): p. 1169-77.

130. Chung, C.W., et al., *The 3-Dimensional Solution Structure of Rantes.* Biochemistry, 1995. **34**(29): p. 9307-9314.

131. Skelton, N.J., F. Aspiras, and T.J. Schall, *Solution Structure of Rantes, a C-C Chemokine.* Journal of Cellular Biochemistry, 1995: p. 37-37.

132. Shaw, J.P., et al., *The X-ray structure of RANTES: Heparin-derived disaccharides allows the rational design of chemokine inhibitors.* Structure, 2004. **12**(11): p. 2081-2093.

133. Wilken, J., et al., *Total chemical synthesis and high-resolution crystal structure of the potent anti-HIV protein AOP-RANTES.* Chemistry & Biology, 1999. **6**(1): p. 43-51.

134. Czaplewski, L.G., et al., *Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1 , MIP-1 , and RANTES.* Journal of Biological Chemistry, 1999. **274**(23): p. 16077.

135. Cierpicki, T. and J.H. Bushweller, *Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins.* Journal of the American Chemical Society, 2004. **126**(49): p. 16259-16266.

136. Liu, Y.Z. and J.H. Prestegard, *A device for the measurement of residual chemical shift anisotropy and residual dipolar coupling in soluble and membrane-associated proteins.* Journal of Biomolecular Nmr, 2010. **47**(4): p. 249-258.

137. Takahashi, H., et al., *A novel NMR method for determining the interfaces of large protein-protein complexes.* Nature Structural Biology, 2000. **7**(3): p. 220-223.

138. Wang, X., et al., *Structural NMR of protein oligomers using hybrid methods.* J Struct Biol, 2010.

139. Svergun, D.I., *DETERMINATION OF THE REGULARIZATION PARAMETER IN INDIRECT-TRANSFORM METHODS USING PERCEPTUAL CRITERIA.* Journal of Applied Crystallography, 1992. **25**: p. 495-503.

140. Wang, X., et al., *RDC assisted modeling of symmetric protein homo oligomers.* Protein Science, 2008. **17**(5): p. 899-907.

141. Svergun, D., C. Barberato, and M.H.J. Koch, *CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates.* Journal of Applied Crystallography, 1995. **28**: p. 768-773.

142. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**: p. 1277-1282.

143. Moont, G., H.A. Gabb, and M.J. Sternberg, *Use of pair potentials across protein interfaces in screening predicted docked complexes.* Proteins, 1999. **35**(3): p. 364-73.

144. Al-Hashimi, H.M., P.J. Bolon, and J.H. Prestegard, *Molecular symmetry as an aid to geometry determination in ligand protein complexes.* J Magn Reson, 2000. **142**(1): p. 153-8.

145. Prestegard, J.H., C.M. Bougault, and A.I. Kishore, *Residual dipolar couplings in structure determination of biomolecules.* Chemical Reviews, 2004. **104**(8): p. 3519-3540.

146. Clore, G.M. and A.M. Gronenborn, *Determining the structures of large proteins and protein complexes by NMR.* Trends in Biotechnology, 1998. **16**(1): p. 22-34.

147. Tjandra, N. and A. Bax, *Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium.* Science, 1997. **278**(5340): p. 1111-4.

148. Tolman, J.R., et al., *Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution.* Proc Natl Acad Sci U S A, 1995. **92**(20): p. 9279-83.

149. Duma, L., et al., *Recognition of RANTES by extracellular parts of the CCR5 receptor.* Journal of Molecular Biology, 2007. **365**(4): p. 1063-1075.

150. Hoogewerf, A.J., et al., *Glycosaminoglycans mediate cell surface oligomerization of chemokines.* Biochemistry, 1997. **36**(44): p. 13570-13578.

151. VivÃ¨s, R.R., et al., *A Kinetics and Modeling Study of RANTES(9âˆ’68) Binding to Heparin Reveals a Mechanism of Cooperative Oligomerizationâ€* Biochemistry, 2002. **41**(50): p. 14779-14789.

152. Yu, Y., et al., *Chemokine-glycosaminoglycan binding: specificity for CCR2 ligand binding to highly sulfated oligosaccharides using FTICR mass spectrometry.* J Biol Chem, 2005. **280**(37): p. 32200-8.

153. Gorman, J. and E.C. Greene, *Visualizing one-dimensional diffusion of proteins along DNA.* Nature Structural & Molecular Biology, 2008. **15**(8): p. 768-774.

154. Vale, R.D., D.R. Soll, and I.R. Gibbons, *ONE-DIMENSIONAL DIFFUSION OF MICROTUBULES BOUND TO FLAGELLAR DYNEIN.* Cell, 1989. **59**(5): p. 915-925.

155. Proudfoot, A.E.I., *Chemokine receptors: multifaceted therapeutic targets.* Nature Reviews Immunology, 2002. **2**(2): p. 106-115.

156. Proudfoot, A.E.I., et al. *Strategies for chemokine antagonists as therapeutics*. 2003. Elsevier.

157. Proudfoot, A.E.I., et al., *The BBXB motif of RANTES is the principal site for heparin binding and controls receptor selectivity.* Journal of Biological Chemistry, 2001. **276**(14): p. 10620.

158. Witt, D.P. and A.D. Lander, *Differential binding of chemokines to glycosaminoglycan subpopulations.* Current Biology, 1994. **4**(5): p. 394-400.

159. Kuschert, G.S.V., et al., *Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses.* Biochemistry, 1999. **38**(39): p. 12959-12968.

160. Middleton, J., et al., *Transcytosis and surface presentation of IL-8 by venular endothelial cells.* Cell, 1997. **91**(3): p. 385-395.

161. Handel, T., et al., *Regulation of protein function by glycosaminoglycans-as exemplified by chemokines.* Annu. Rev. Biochem., 2005. **74**: p. 385-410.

162. Wang, X., et al., *Oligomeric Structure of the Chemokine CCL5/RANTES from NMR, MS, and SAXS Data.* Structure, 2011. **19**(8): p. 1138-1148.

163. Huang, R., V.H. Pomin, and J.S. Sharp, *LC-MS n Analysis of Isomeric Chondroitin Sulfate Oligosaccharides Using a Chemical Derivatization Strategy.* Journal of the American Society for Mass Spectrometry, 2011: p. 1-11.

164. Pomin, V.H., et al., *Characterization of Glycosaminoglycans by 15N NMR Spectroscopy and in Vivo Isotopic Labeling.* Analytical Chemistry, 2010. **82**(10): p. 4078-4088.

165. Berne, B.J. and R. Pecora, *Dynamic light scattering: with applications to chemistry, biology, and physics*2000: Dover Pubns.

166. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* Journal of computational chemistry, 2010. **31**(2): p. 455-461.

167. Wang, X., et al., *Probing rhodopsin-transducin interactions by surface modification and mass spectrometry.* Biochemistry, 2004. **43**(35): p. 11153-11162.

168. Hager-Braun, C. and K.B. Tomer, *Characterization of the tertiary structure of soluble CD4 bound to glycosylated full-length HIVgp120 by chemical modification of arginine residues and mass spectrometric analysis.* Biochemistry, 2002. **41**(6): p. 1759-66.