Plasma-Generated Hydroxyl Radicals for Epitope Mapping

By

Daniel Benjamin

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Electrical Engineering)

at the

University of Wisconsin – Madison

2021

Date of final oral examination: March 30, 2021

This dissertation is approved by the following members of the Final Oral Committee:

J. Leon Shohet, Professor Emeritus, Electrical and Computer Engineering

David Anderson, Professor, Electrical and Computer Engineering

Mark Richards, Professor, Animal Science

Joseph Andrews, Assistant Professor, Mechanical Engineering and Electrical and Computer Engineering

Plasma-Generated Hydroxyl Radicals for Epitope Mapping

Daniel Benjamin

Under the supervision of Professor J. L. Shohet

at the University of Wisconsin-Madison

Abstract

A technology, Plasma Induced Modification of Biomolecules (PLIMB), has been developed for measuring three-dimensional (3-D) structural characteristics of proteins. It is used to perform hydroxyl-radical protein footprinting (HRPF) wherein a protein's solvent-accessible side chains are oxidized in solution using a hydroxyl-radical source and then analyzed with mass spectrometry.

Although a technology like PLIMB can be used for many different applications, the unique characteristics of the technology and data it can produce may prove it to be uniquely qualified for the role of mapping the epitopes, or binding locations, of protein therapeutics to their target molecules. This is a crucial aspect of drug development and is important for drug optimization and intellectual-property filings.

Although several proof-of-concept experiments demonstrating PLIMB's utility have been performed, additional research and development must be completed before PLIMB can be considered a viable option for epitope mapping. We hypothesize that by implementing technological improvements, processes, and methods, the epitopes for protein therapeutics can be mapped with PLIMB.

This work first demonstrates the use of PLIMB for mapping the epitope of an antibody/antigen pair. Experimentation and data analysis using this procedure was effective, as several regions of the antigen were identified as the epitope region, which agree with previously published results.

To address the recurring issue of variability in PLIMB data, a feedback-control system was developed, which uses UV-spectroscopy to measure hydroxyl radical generation in real time. Samples of the amino acid methionine were exposed to plasma and the amount of oxidation was measured with mass spectrometry. The results showed a reduction in the coefficient of variation between replicate samples from 15% to 6% when using the feedback control system, thus demonstrating its utility.

To extract solvent accessibility measurements from PLIMB data, a procedure developed by other HRPF researchers was used. Samples of the protein myoglobin were exposed to various plasma-exposure times. Then a reactivity-normalized oxidation rate was calculated for several myoglobin amino acids. Results showed a linear correlation between reactivity-normalized oxidation rates and solvent-accessibility measurements generated from a published 3-D structure of myoglobin, and a root-mean-squared error in calculated solvent accessible surface areas of 24.98 $Å^2$.

Table of Contents

Chapter 1: Introduction

Protein therapeutics is one of the fastest-growing segments of the pharmaceutical industry and has proven to be a source of successful treatments for a range of indications including cancer and autoimmune disorders. Proteins have notable advantages over traditional small-molecule therapeutics, including their wide applicability and high specificity, *i.e.* their ability to bind to a specific protein molecule in the body known as the *target*, to treat a disease while minimizing side effects. [1]

1.1 Epitope mapping

One important process that is conducted as part of the development of a protein therapeutic is mapping the specific location where a protein therapeutic binds to its target protein. [2] This process is known as epitope mapping and is often performed during the lead-validation stage of the pharmaceutical pipeline (see Figure 1-1). Data on the epitope, or binding location, is used for drug optimization, whereby the protein therapeutic is modified in order to enhance its binding properties at that location, [3-5] and for intellectual property fillings. [6]

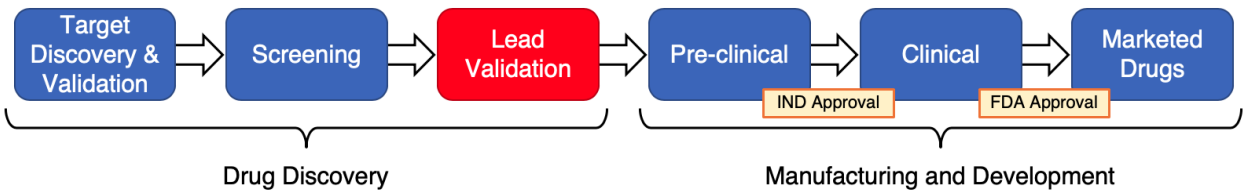


Figure 1-1: Lead validation stage of pharmaceutical pipeline

There are currently several techniques which can be utilized to fill the epitope-mapping niche, but each has their own set of drawbacks such as poor data reproducibility and low resolution.

1.1 Protein footprinting

Protein footprinting is a technique for characterizing protein structure and has shown significant promise for development of protein therapeutics. [2, 7-9] Protein footprinting works by labelling a protein molecule in solution and analyzing the labelled protein with mass spectrometry. [10-13] One form of protein footprinting, called hydrogen-deuterium exchange (HDX), uses a heavy-water solution to label proteins with deuterium atoms along their amide[1] backbone. [13] HDX has gained significant traction in the pharmaceutical industry for protein-therapeutics development [9, 14, 15] including its use for mapping epitopes. [16-18] Despite the successful adoption of HDX, the technology has some inherent disadvantages, including its potentially poor data reproducibility due to a phenomenon called hydrogen back-exchange, in which the deuterium atoms that have bound to the antigen exchange back with hydrogen atoms. [19, 20] Currently, there are also limitations to the post-labelling sample-handling procedures that can be performed. [21, 22]

Another form of protein footprinting is known as hydroxyl-radical protein footprinting (HRPF). HRPF techniques work by labelling proteins in solution using hydroxyl (·OH) radicals, which are highly reactive chemical species that readily oxidize regions of proteins that are solvent accessible. [10, 11, 23] As with HDX, the labelled regions can then be analyzed with mass spectrometry to characterize various aspects of a protein's three-dimensional structure. HRPF has some distinct advantages over HDX including its wide range of post-labelling sample-handling procedures that can be performed because the labels are an irreversible, covalent modification. [11] This is unlike HDX, which relies on a modification that is subject to back-exchange after labelling.

[1] Amide refers to a group of compounds characterized by the N-C=O moiety. A chain of amides makes up a protein's "backbone."

HRPF techniques have been developed and utilized in an academic setting over the past 30 years, [2, 11] but have failed to gain significant traction in the pharmaceutical industry, especially when compared to HDX. There are likely a number of reasons for the lack of industry adoption, but some include drawbacks of some of the currently developed methods for generating hydroxyl radicals for labelling proteins. These include the inability to produce hydroxyl radicals without additional chemicals [11] and the inability to oxidize proteins on a sub-microsecond timescale, [24] both of which are likely to be important for capturing the native structure of a protein in solution [25] (see Section 2.3.2 for details), and the need for expensive and/or potentially hazardous equipment.

1.2 PLIMB

Plasma Induced Modification of Biomolecules (PLIMB) is a newly developed technology for performing HRPF. PLIMB uses an atmospheric-pressure plasma to generate hydroxyl radicals in sub-microsecond bursts. [26, 27] The system that has been developed has some unique features including the ability to control, among other things, the radical dose and generation rate by modifying the plasma parameters. Furthermore, a temperature-control unit has been integrated into the system for maintaining a constant sample temperature from 2-95° C with ± 2° C variation. The complete system is built as a fully automated benchtop instrument with software that allows the system to sequentially expose an entire 96-well plate of samples. PLIMB technology addresses many of the limitations of existing HRPF technologies by generating hydroxyls from water without the need for additional chemicals and does so on a sub-microsecond timescale, [26] both of which are important for minimizing structural changes due to labelling. [12, 25, 28] The combination of these features is unique to PLIMB.

Although PLIMB can be used for many different applications, the unique characteristics of the technology and the data it can produce may prove it to be especially qualified for the role of epitope mapping of protein therapeutics. These include the fact that PLIMB can generate single amino-acid level resolution, and an entire experiment can be completed in about two weeks, both of which are important attributes for epitope mapping of protein therapeutics during the lead validation stage of the pharmaceutical pipeline. [3, 29, 30]

1.3 Purpose of this work

We hypothesize that PLIMB technology can be used to map the epitopes of potential protein therapeutics during development. Although several proof-of-concept experiments have been performed which highlight PLIMB's utility for protein footprinting, [26] additional research and development must be completed before PLIMB can be considered a viable option for epitope mapping. In order to show viability for this application, the following tasks will need to be completed.

1. **A process will be established that can be followed for experimentally mapping epitopes using PLIMB labelling.** This process will include protein digestion and other sample-handling procedures in order to analyze the PLIMB-exposed samples with mass spectrometry. Then, a data-analysis procedure will be established for taking raw mass-spectrometry data and using it to find the epitope locations. Consideration will also be taken in order to separate indications of epitopes from other conformational changes that can happen to a protein upon binding. We will begin by utilizing processes that have been developed to map epitopes using other HRPF techniques, [23, 31, 32] which can be followed as a guide for experimentation using PLIMB. We propose that the resulting data

can be used to identify epitope regions, and these regions can be distinguished from conformational changes upon binding.

2. **Reduce data variability between replicate samples.** [26] An additional system will be added to the PLIMB system, which accomplishes the goal of reducing variability between sample replicates. By constructing a system that measures hydroxyl-radical generation in real time, we can implement a feedback-control system that reduces the variability in hydroxyl-radical generation, thus reducing the overall variability of the protein-footprinting data. Additionally, a method for quantifying the rate of self-oxidation of samples after PLIMB exposure will be investigated.
3. **Develop a data analysis protocol for extracting solvent accessibility measurements from PLIMB data**. The data that results from protein-footprinting experiments is typically a measure of percent oxidation over various regions of the protein. These levels of oxidation can be used to compare relative levels of solvent accessibility of a particular region of protein between samples under varying conditions, such as the presence or absence of a bound antibody. However, measurements of percent oxidation are a result of a *single* experiment, and will vary depending on many factors, including the relative reactivity of the region of protein being analyzed. [11, 33, 34] Our goal in this task is to develop a method that can be used to extract solvent accessibility from PLIMB data, which can be directly compared with 1) other regions of the same protein in a PLIMB experiment, 2) with other proteins that were analyzed in a different PLIMB experiment, or 3) with data that was generated using other protein-structural techniques. We hypothesize that, by exposing samples of a protein to a range of hydroxyl doses and calculating a rate of oxidation over different regions of the protein, and then normalizing the rate of oxidation

by the relative reactivity of the region, a measurement of solvent accessibility can be obtained.

1.4 Organization of thesis

This thesis is organized as follows. Chapter 2 provides the background information necessary to understand the purpose and scope of this thesis. Protein therapeutics, as well as a particular class of protein therapeutics, namely antibody therapeutics, are outlined and described. Techniques to characterize the structure of proteins are then described, including high-resolution technologies such as x-ray crystallography and cryo-EM, as well as protein-footprinting techniques including HDX and various other methods for performing HRPF including PLIMB. Next, the pharmaceutical pipeline will be introduced, followed by a discussion of which stages of the pipeline PLIMB can be most valuable for and the reasons why. Finally, epitope mapping will be described in greater detail, as well as the reasons for which PLIMB is an excellent choice of technology for this particular application.

Chapter 3 describes the PLIMB setup in greater detail, as well as the various diagnostics that are used for characterizing the electrical and chemical properties of the plasma. First, the various methods for measuring hydroxyl-radical generation will be discussed, including each of their strengths and weaknesses. Next, the various iterations of the PLIMB setup since its inception will be described, including the problems and challenges that were faced during development, and how each of these were addressed before the current system was constructed. Lastly, the methods and results of characterizing the electrical properties of the plasma, including the plasma current, will be discussed, followed by the methods and results of photoemission-spectroscopy measurements.

Chapter 4 demonstrates the use of PLIMB for epitope mapping of an antibody to its target antigen. A procedure similar to that by used by other HRPF researchers for mapping epitopes was followed. Several regions of the target protein were found to undergo protection, or a decrease in solvent accessibility, which is an indication of epitope interactions. Upon analyzing data generated using two plasma-exposure times, epitope regions were able to be identified from the list of regions that showed protection, and the results agree with previously published results.

Chapter 5 describes the experimental setup for a feedback-control system that was developed to reduce data variability for PLIMB experiments. The setup works by measuring the intensity of light over a UV-band which covers an emission line of OH radicals. By integrating the light intensity over the entire plasma exposure, a measurement of the hydroxyl-radical dose can be obtained. The setup was first tested by exposing samples of coumarin, a fluorescent chemical reporter, and was shown to produce a measurement of radical dose that correlated with percent coumarin oxidation. Samples of the amino-acid methionine were then exposed to either a set plasma-exposure time or a fixed radical dose set by the feedback control system. The results show a significant decrease in variability of percentage oxidation between replicate samples for those exposed to the fixed radical dose rather than a fixed plasma-exposure time. A method for measuring the rate of self-oxidation of PLIMB-exposed samples was investigated as an additional cause of data variability. Samples of coumarin were exposed to PLIMB and then stored at various temperatures for several days. The fluorescence, which is a measure of the amount of oxidation in the coumarin samples, was measured after storage. The results suggest several strategies for minimizing the amount of self-oxidation after PLIMB exposure, thus minimizing the potential for data variability due to self-oxidation.

Chapter 6 describes a method for extracting solvent-accessibility measurements from PLIMB data. Samples of the protein myoglobin were prepared and exposed to PLIMB for a range of timepoints. Rates of percent oxidation per unit of exposure time were calculated on an amino-acid level, and then normalized by the reactivity rate to hydroxyl radicals for each amino acid.

The resulting measurements for each identified amino acid of myoglobin were compared to a solvent-accessibility measurement generated from a three-dimensional x-ray crystallography structure of myoglobin. [35] The results show a linear correlation between the solvent accessibility measurements calculated from PLIMB data and from the crystal structure. Solvent-accessible surface areas (SASAs) were then calculated from PLIMB-generated data and compared to those generated from the crystal structure. The results showed a 24.98 $Å^2$ root-mean-squared (RMS) error, which is comparable to the errors reported for SA models reported by other HRPF researchers. [36]

Lastly, Chapter 7 summarizes the conclusions of this work and describes a direction for future work. The Appendices also provide supplementary information including circuit diagrams and technical drawings for custom-made parts.

1.5 References

[1] B. Leader, Q. Baca, and D. Golan, "Protein therapeutics: a summary and pharmacological classification," *Nature Reviews Drug Discovery,* vol. 7, no. 1, pp. 21-39, 2008, doi: 10.1038/nrd2399.

[2] L. Wang and M. Chance, "Protein Footprinting Comes of Age: Mass Spectrometry for Biophysical Structure Assessment," in *Mol. Cell. Proteomics* vol. 16, ed, 2017, pp. 706-716.

[3] G. Raghunathan, E. Hsieh, and G. Ermakov, "Epitope Characterization and Isotype Selection," in *Development of Antibody-Based Therapeutics*: Springer, 2018, pp. 41-69.

[4] G. Nimrod *et al.*, "Computational Design of Epitope-Specific Functional Antibodies," *Cell Reports,* vol. 25, no. 8, pp. 2121-2131.e5, 2018/11/20/ 2018, doi: https://doi.org/10.1016/j.celrep.2018.10.081.

[5] U. Reineke *et al.*, "Identification of distinct antibody epitopes and mimotopes from a peptide array of 5520 randomly generated sequences," *Journal of Immunological Methods,* vol. 267, no. 1, pp. 37-51, 2002/09/01/ 2002, doi: https://doi.org/10.1016/S0022-1759(02)00139-4.

[6] X. Deng, U. Storz, and B. J. Doranz, "Enhancing antibody patent protection using epitope mapping information," *mAbs,* vol. 10, no. 2, pp. 204-209, 2018, doi: 10.1080/19420862.2017.1402998.

[7] G. Deperalta, M. Alvarez, C. Bechtel, K. Dong, R. McDonald, and V. Ling, "Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting," *mAbs,* vol. 5, no. 1, pp. 86-101, 2013/01/01 2013, doi: 10.4161/mabs.22964.

[8] P. Kaur *et al.*, "Characterizing monoclonal antibody structure by carboxyl group footprinting," *mAbs,* vol. 7, no. 3, pp. 540-552, 2015/05/04 2015, doi: 10.1080/19420862.2015.1023683.

[9] G. R. Masson, M. L. Jenkins, and J. E. Burke, "An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery," *Expert Opinion on Drug Discovery,* vol. 12, no. 10, pp. 981-994, 2017/10/03 2017, doi: 10.1080/17460441.2017.1363734.

[10] K. Takamoto and M. R. Chance, "Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes," *Annu. Rev. Biophys. Biomol. Struct.,* vol. 35, pp. 251-276, 2006.

[11] X. R. Liu, M. M. Zhang, and M. L. Gross, "Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications," *Chemical Reviews,* vol. 120, no. 10, pp. 4355-4454, 2020/05/27 2020, doi: 10.1021/acs.chemrev.9b00815.

[12] C. edited by Mark, *Mass spectrometry analysis for protein-protein interactions and dynamics*. Hoboken, N.J. : Wiley, ©2008., 2008.

[13] J. R. Engen, "Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS," ed: ACS Publications, 2009.

[14] Z. E. Nazari, M. van de Weert, G. Bou-Assaf, D. Houde, A. Weiskopf, and K. D. Rand, "Rapid conformational analysis of protein drugs in formulation by hydrogen/deuterium exchange mass spectrometry," *Journal of pharmaceutical sciences,* vol. 105, no. 11, pp. 3269-3277, 2016.

[15] J. Pan, S. Zhang, and C. H. Borchers, "Comparative higher-order structure analysis of antibody biosimilars using combined bottom-up and top-down hydrogen-deuterium exchange mass spectrometry," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* vol. 1864, no. 12, pp. 1801-1808, 2016/12/01/ 2016, doi: https://doi.org/10.1016/j.bbapap.2016.08.013.

[16] K. F. M. Opuni, M. Al-Majdoub, Y. Yefremova, R. F. El-Kased, C. Koy, and M. O. Glocker, "Mass spectrometric epitope mapping," vol. 37, ed, 2018, pp. 229-241.

[17] C. Puchades *et al.*, "Epitope mapping of diverse influenza Hemagglutinin drug candidates using HDX-MS," *Scientific reports,* vol. 9, no. 1, pp. 1-10, 2019.

[18] G. Hao, J. S. Wesolowski, X. Jiang, S. Lauder, and V. D. Sood, "Epitope characterization of an anti-PD-L1 antibody using orthogonal approaches," *Journal of Molecular Recognition,* vol. 28, no. 4, pp. 269-276, 2015.

[19] B. T. Walters, A. Ricciuti, L. Mayne, and S. W. Englander, "Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment," *Journal of the American Society for Mass Spectrometry,* vol. 23, no. 12, pp. 2132-2139, 2012/12/01 2012, doi: 10.1021/jasms.8b04189.

[20] J. G. Sheff, M. Rey, and D. C. Schriemer, "Peptide–Column Interactions and Their Influence on Back Exchange Rates in Hydrogen/Deuterium Exchange-MS," *Journal of the American Society for Mass Spectrometry,* vol. 24, no. 7, pp. 1006-1015, 2013/07/01 2013, doi: 10.1021/jasms.8b04561.

[21] G. R. Masson *et al.*, "Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments," *Nature Methods,* vol. 16, no. 7, pp. 595-602, 2019/07/01 2019, doi: 10.1038/s41592-019-0459-y.

[22] E. Trabjerg, Z. E. Nazari, and K. D. Rand, "Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): challenges and emerging solutions," *TrAC Trends in Analytical Chemistry,* vol. 106, pp. 125-138, 2018.

[23] J. Kiselar and M. R. Chance, "High-Resolution Hydroxyl Radical Protein Footprinting: Biophysics Tool for Drug Discovery.," *Annual Review of Biophysics,* vol. 47, no. 1, p. 315, 2018, doi: 10.1146/annurev-biophys-070317-033123.

[24] S. Vahidi and L. Konermann, "Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds," *Journal of the American Society for Mass Spectrometry,* vol. 27, no. 7, pp. 1156-1164, 2016/07/01 2016, doi: 10.1021/jasms.8b05311.

[25] D. M. Hambly 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,* vol. 16, no. 12, pp. 2057-2063, 2005/12/01 2005, doi: 10.1016/j.jasms.2005.09.008.

[26] B. Minkoff, J. Blatz, F. Choudhury, D. Benjamin, J. Shohet, and M. Sussman, "Plasma-Generated OH Radical Production for Analyzing Three- Dimensional Structure in Protein Therapeutics," *Sci Rep,* vol. 7, no. 1, pp. 12946-12946, 2017, doi: 10.1038/s41598-017-13371-7.

[27] J. M. Blatz, D. Benjamin, F. A. Choudhury, B. B. Minkoff, M. R. Sussman, and J. L. Shohet, "Effect of frequency and applied voltage of an atmospheric-pressure dielectric-barrier discharge on breakdown and hydroxyl-radical generation with a liquid electrode," *Journal of Vacuum Science & Technology A,* vol. 38, no. 4, p. 043001, 2020, doi: 10.1116/6.0000125.

[28] J. G. Kiselar and M. R. Chance, "Future directions of structural mass spectrometry using hydroxyl radical footprinting," *Journal of Mass Spectrometry,* vol. 45, no. 12, pp. 1373-1382, 2010, doi: https://doi.org/10.1002/jms.1808.

[29] X. Deng, U. Storz, and B. J. Doranz, "Enhancing antibody patent protection using epitope mapping information," in *MAbs*, 2018, vol. 10, no. 2: Taylor & Francis, pp. 204-209.

[30] M. V. Chaubal, "Application of formulation technologies in lead candidate selection and optimization," *Drug Discovery Today,* vol. 9, no. 14, pp. 603-609, 2004/07/01/ 2004, doi: https://doi.org/10.1016/S1359-6446(04)03171-X.

[31] L. M. Jones, J. B Sperry, J. A Carroll, and M. L. Gross, "Fast photochemical oxidation of proteins for epitope mapping," *Analytical chemistry,* vol. 83, no. 20, p. 7657, 2011, doi: 10.1021/ac2007366.

[32] Y. Yan, "Fast photochemical oxidation of proteins (FPOP) maps the epitope of EGFR binding to adnectin," *Journal of the American Society for Mass Spectrometry.,* vol. 25, no. 12, p. 2084, 2014, doi: 10.1007/s13361-014-0993-x.

[33] G. Xu and M. R. Chance, "Hydroxyl radical-mediated modification of proteins as probes for structural proteomics," *Chemical reviews,* vol. 107, no. 8, pp. 3514-3543, 2007.

[34] K. S. Li, L. Shi, and M. Gross, "Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization," in *Accounts Chem. Res.* vol. 51, ed, 2018, pp. 736-744.

[35] R. Maurus *et al.*, "A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket," *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology,* vol. 1341, no. 1, pp. 1-13, 1997/08/15/ 1997, doi: https://doi.org/10.1016/S0167-4838(97)00064-2.

[36] B. Xie, A. Sood, R. Woods, and J. Sharp, "Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection," *Sci Rep,* vol. 7, no. 1, pp. 4552-4552, 2017, doi: 10.1038/s41598-017-04689-3.

Chapter 2: Background

Protein therapeutics are an important class of medicines that now represent a rapidly growing multibillion-dollar industry that is currently projected to exceed $300 billion by the year 2025. [1, 2] It is currently estimated to be around 10% of the global pharmaceutical industry and expected to make up an even larger portion of the market in the future. [3] The success of these treatments is due in part to the various technologies which have been developed to help scientists understand proteins on a biochemical level. [4]

This chapter will explore the protein therapeutics industry, and what technologies are specifically being used to study protein structure. A new technology for protein-structural analysis, known as Plasma Induced Modification of Biomolecules (PLIMB), will then be introduced and compared to technologies that are currently available. Next, different applications of PLIMB will be explored, and finally, the application that is central to this thesis, epitope mapping, will be explained in greater detail, and why PLIMB is well situated to fill the gap in technologies available for epitope mapping for pharmaceutical development.

2.1 Protein therapeutics

Proteins are large biological molecules, which consist of long chains of amino-acid residues.[2] These chains fold together to create a complicated three-dimensional structure, or conformation, which, along with its chemical composition, dictates a protein's biological function. Proteins are responsible for a wide array of functions including catalyzing metabolic and enzymatic reactions,

[2] A residue here refers to a single amino acid of the chain of many amino acids that make up a protein.

molecular transportation, DNA replication and transcription, and for providing structure to cells. [5]

Protein therapeutics are proteins specifically designed to treat a wide range of indications including cancer, infections, and autoimmune disorders. [2, 6] Protein therapeutics generally work by binding to a specific biomolecule in the human body, usually another protein, which is referred to as a therapeutic *target.* The target protein is selected based on its involvement in the biological processes associated with a disease. The binding of a therapeutic to a target is generally used to inhibit the biological function of the target protein, in order to correct for the misbehavior which is causing a disease. [2, 6]

Because of their unique properties, proteins have many advantages when compared to traditional small-molecule therapeutics. Protein therapeutics can be engineered to be highly specific-*i.e*., the protein binds to a single region of a specific protein in the body without causing many unwanted interactions or interferences with other biological functions. [6] Also, because proteins are naturally produced in the body, protein therapeutics are well tolerated and are less likely to produce unwanted effects such as unintended immune responses. [6]

Because of these advantages, protein therapeutics have continued to rise in popularity since the first recombinant[3] protein therapeutic, human insulin, was approved by the FDA in 1982. [6, 7] Currently, monoclonal antibodies are among the most widely utilized form of protein therapeutics. [8] Antibodies are large protein molecules that are produced by the plasma cells in the human body, and are used by the immune system to fight off infectious pathogens. [9, 10] Antibodies work by targeting and binding to a particular antigen, *i.e.*, a molecule to which an antibody binds, in order to neutralize its effect. [11] Pharmaceutical researchers take advantage of

[3] Recombinant proteins are those that are produced through expression of a single DNA sequence, which results in proteins of an exact and homogeneous amino-acid sequence.

the unique properties of antibodies to engineer therapeutics that recognize a specific target antigen in the body associated with a particular disease and neutralize its effects.

2.2 Structural characterization

Despite their advantages, there are significant challenges associated with developing protein therapeutics. One of the major challenges involves characterizing the three-dimensional structure of proteins. [12, 13] Because structure largely determines a protein's biological function, characterizing structural aspects of proteins is essential for protein therapeutics development and is often quite challenging. Most protein-structural characterization techniques either provide detailed and high-resolution data, yet are quite expensive and time consuming, or are inexpensive and fast, but provide low-resolution data.[4]

2.2.1 X-ray crystallography

X-ray crystallography is a structural technique that gives atomic-level resolution. [14] The process involves taking an aqueous solution of protein and drying it down into a solid made of protein molecules aligned in a crystal lattice. The solid is then analyzed using x-ray diffraction, often generated with a synchrotron-radiation source, in order to deduce the three-dimensional structure of the protein molecule with atomic-level resolution. [13]

Although x-ray crystallography is used heavily in the pharmaceutical industry and is often considered the gold standard in protein structural characterization, it does exhibit some drawbacks and limitations. Because the technique analyzes the protein in a crystalline structure, there is often a concern that the technique may yield a structure which does not resemble that of the protein in

[4] Resolution here refers to the smallest unit on which a particular technology can generate data. In the context of proteins, low resolution usually refers to techniques which generate data on an entire protein or large regions of a protein. High resolution usually refers to techniques that can generate data on individual amino acids (amino-acid level resolution) or on the positions of each atom (atomic-level resolution) of a protein.

an aqueous solution, which is most characteristic of its natural environment in the human body. [15, 16] Furthermore, for some proteins, it is often difficult or even impossible to create a protein crystal suitable for x-ray analysis. Membrane-bound proteins, for example, which are a class of proteins that are studied as potential drug-target candidates, are notoriously difficult to crystalize. [17-19] Lastly, crystallography is a time-consuming and expensive process. It can often take several months and upwards of $500,000 to obtain the crystal structure of a single protein. [15, 20] Because of this, pharmaceutical companies limit the number of protein-therapeutic candidates that are analyzed with x-ray crystallography to only those that are far enough along the drug-discovery pipeline and show promise as a pharmaceutical candidate.

2.2.2 Cryo-electron microscopy

Cryo-electron microscopy (cryo-EM) is another structural technique that can give atomic-level protein structural resolution, like that of x-ray crystallography. [21-23] Cryo-EM works by flash freezing a liquid sample containing proteins and analyzing their structure with electron microscopy. Though it is a newer technique, it has quickly gained adoption in the pharmaceutical industry. Unlike x-ray crystallography, cryo-EM can give structural information about a protein in an aqueous solution, yielding a structure similar to that of the protein's natural conformation. [24] Furthermore, it can be used to analyze virtually all types of proteins given they are of sufficient size to yield atomic-level resolution. [21, 22] However, like crystallography, it is quite expensive and time consuming, and is reserved for use on only select proteins of interest in the pharmaceutical pipeline.

2.3 Protein footprinting

As the need for fast structural characterization has continued to grow, pharmaceutical companies have explored more mass-spectrometry-based techniques to fill this need. [25] When compared to x-ray crystallography, mass-spectrometry analysis is often much faster and can additionally provide structural information about a protein in solution rather than in a crystalline form. [26] In-solution structure is important for pharmaceutical research because it mimics the conditions of a protein in the human body. Specifically, a mass-spectrometry technique known as protein footprinting has played an important role in structural characterization. [27, 28] Protein footprinting works by labelling solvent-accessible regions of a protein in solution with a reactive molecule. These labels can then be detected with a mass spectrometer, allowing for a measurement of solvent accessibility. [29, 30] Solvent accessibility is a measurement of how exposed a given region of protein is to its surrounding solvent (see Figure 2-1) and is generally considered to be a characteristic of a protein's higher-order structure. Higher-order structure generally refers to the *tertiary* and *quaternary* structure (see Figure 2-2). [31] Tertiary structure here refers to the three-dimensional folding of a protein, and quaternary refers to how a protein binds or interacts with other proteins.

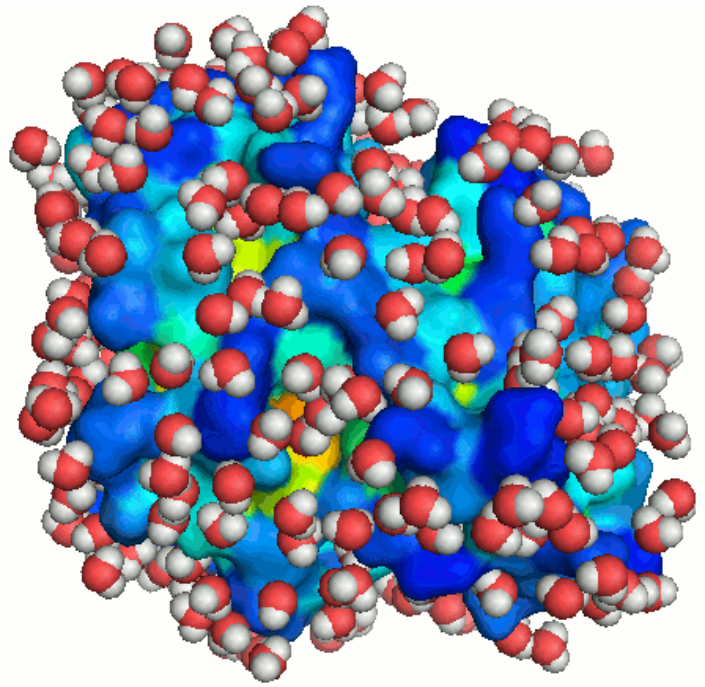


Figure 2-1: Protein solvent accessibility

This image depicts a protein with water molecules (red and white) surrounding its outer surfaces. Regions of the protein that are in contact with solvent are referred to as “solvent accessible.” The degree of solvent accessibility is shown in this image on a color scale from blue (most solvent accessible) to orange (least solvent accessible). [32]

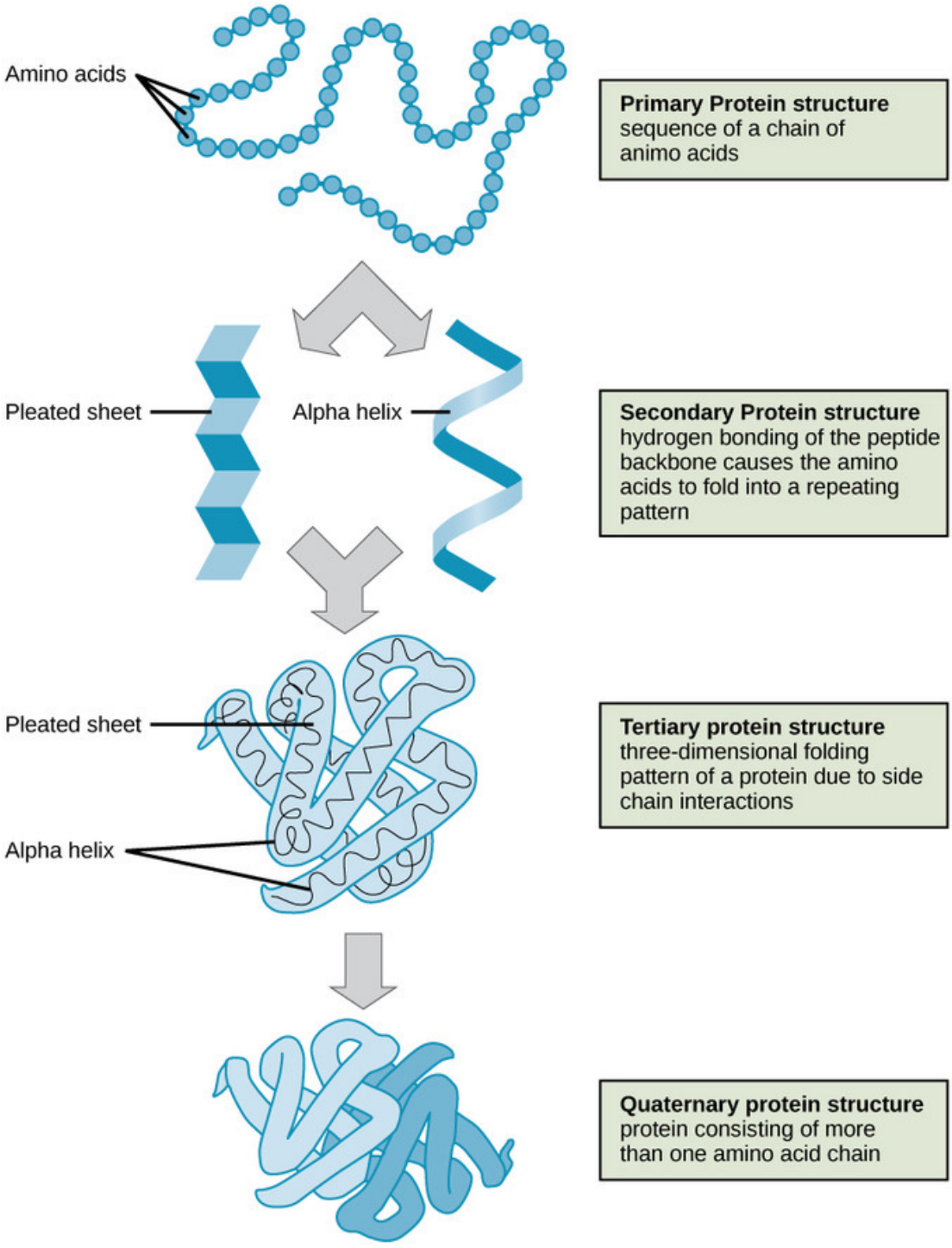


Figure 2-2: Levels of protein structure [31]

2.3.1 Hydrogen-deuterium exchange

One of the most widely used protein-footprinting techniques is called hydrogen-deuterium exchange (HDX). [33] A workflow for HDX is shown in Figure 2-3. [30] The process for performing HDX is as follows: a sample of protein is submerged in a deuterated-water solution, and over time, some of the protein's solvent-accessible hydrogen atoms will exchange with deuterium atoms. The rate of exchange is dependent on several factors including the pH, temperature, kinetics of the protein in solution, and the solvent accessibility of a given exchange site. [34] The deuterium atoms that exchange with the amide[5] hydrogens along the backbone of the protein are used as labels for measuring solvent accessibility, as their exchange rate is on the order of seconds to hours. The exchange rate of deuterium with hydrogen on the amino-acid side chains is too fast (half-life of μs) to be probed by HDX, while the exchange rate of hydrogen attached to carbon is too slow. [26, 34]

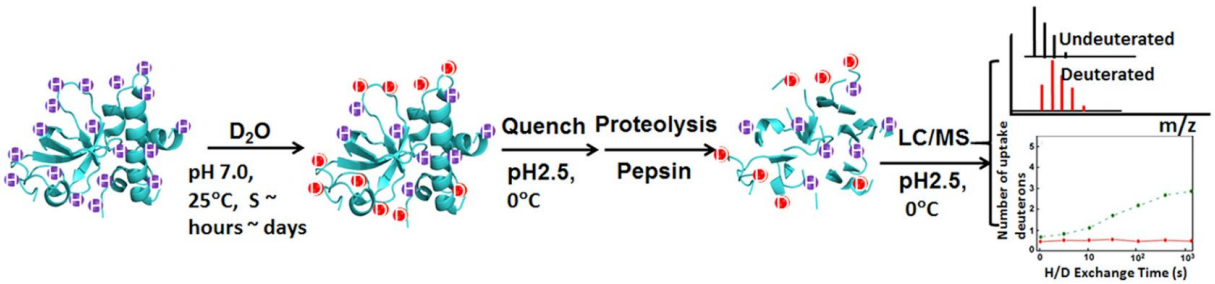


Figure 2-3: Hydrogen-deuterium exchange (HDX) workflow [30]

After labelling the protein with deuterium, the solution is then quenched in a low-pH (~2.5), low-temperature (0° C) buffer, and then the protein is rapidly digested, or broken into smaller chains of amino acids called peptides using the protease[6] pepsin, and analyzed with a mass

[5] Amide refers to a group of compounds characterized by the N-C=O moiety. A chain of amides makes up a protein's "backbone."

[6] A protease is an enzyme that catalyzes proteolysis, the breakdown of proteins into smaller polypeptides or single amino acids. It does this by cleaving the peptide bonds within proteins by hydrolysis, a reaction where water breaks bonds.

spectrometer. [33, 34] Low-PH and low-temperature conditions must be maintained throughout these steps in order to minimize the back-exchange rate, *i.e*., the exchange of deuterium labels in the protein back to hydrogen. Back exchange reduces the labelling "signal" and therefore should be minimized in order to maximize dynamic range in the analysis. In order to fully minimize the overall back exchange, the time between labelling and analysis should also be minimized.

While HDX is a powerful technique because it can analyze structure in a protein's native state, it has some significant drawbacks. Because proteolysis and other sample preparation steps that are performed between labelling and injection of the sample into the mass spectrometer must be done at low temperature and low pH conditions and in a short amount of time to reduce back exchange. Thus, there is limited flexibility in the choice of proteases, enzymes, or other reagents that can be used effectively. Pepsin is one of the few proteases that is effective under these conditions but it can often yield mass-spectrometry sequence coverages that are less than ideal. [35] Furthermore, it may be difficult or technically impracticable to remove chemical additives that are necessary for capturing accurate structural data yet are incompatible with liquid-chromatography mass spectrometry (LC-MS). For example, membrane proteins usually need to be bound in nano-discs (see Figure 2-4) or in a lipid membrane with detergents added to the solution to increase solubilization, in order to retain a conformation like that of the protein *in-vitro* or *in-vivo*. Therefore, an accurate analysis of their structure with protein footprinting would require labelling the sample protein with one or more of these reagents, and then purifying the protein sample before mass-spectrometry injection. With an HDX experimental setup, this is technically challenging. [35, 36]

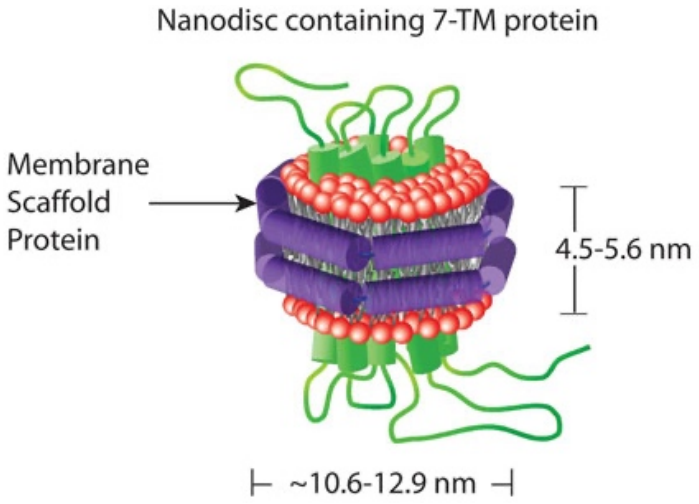


Figure 2-4: Nanodisc

Schematic illustration of a membrane scaffold protein (MSP), also known as a nanodisc, with a 7-transmembrane protein, a protein which spans the entirety of a cell membrane *in vivo*, embedded. Diameter is about 10 nm. Picture is from Sligar Lab. [37]

Another drawback of HDX is the potential for poor data reproducibility. Because of the deuterium back exchange that is inevitable for any HDX experiment, great care must be taken to ensure that the sample conditions (pH, temperature, and proteolysis) that occur between labelling and injection must be quick (seconds to minutes) and must be repeatable between samples in order to ensure accurate and reproducible results. One solution to for this problem is to utilize an automated liquid-handling system for performing HDX experiments, rather than performing them by hand. LEAP Technologies manufactures an instrument to automate HDX experiments in this way. [38] While this instrument is utilized heavily in the pharmaceutical industry, the cost of such an instrument is significant, and may not be a realistic option for many labs. Furthermore, data reproducibility can present issues even when utilizing a liquid-handling system, and thus may not be an adequate tool for gathering reliable structural data for certain applications. Despite the drawbacks to HDX, it is still heavily used in biotechnology and pharmaceutical research. [29]

2.3.2 Hydroxyl-radical protein footprinting

Another protein-footprinting technique, hydroxyl-radical protein footprinting (HRPF), uses hydroxyl (·OH) radicals as a labelling agent rather than deuterium. [30, 39-41] Hydroxyl radicals are highly reactive chemical species that readily oxidize solvent-accessible side chains of a protein in solution, and cause a permanent, covalent oxidative label, which allows for more flexibility in the post-labelling, pre-mass-spectrometry sample-preparation and cleanup steps of the experiment. Enzymes for removing glycans[7] may be used to allow better detection of naturally glycosylated peptides, [42, 43] and detergents or nano-discs may be used in solution during labelling and then removed before mass-spectrometry analysis. This allows for structural analysis of a protein in its native conformation, and then purification for compatibility with LC-MS. For protein digestion, any protease or combination of proteases may be used to maximize mass-spectrometry sequence coverage. Furthermore, the amount of time needed to complete these steps will not negatively affect the reproducibility or sensitivity of structural data because the hydroxyl-radical labels are permanent modifications. [29] Ultimately, the flexibility in these steps allows for accurate structural analysis in a protein's native conformation, improved reproducibility, and potentially higher resolution when compared with HDX.

Figure 2-5 shows a general workflow diagram for HRF. [30] The process is performed as follows: 1) several samples of a protein are prepared in an aqueous solution, 2) each protein sample is exposed to hydroxyl radicals over a range of exposure times, 3) proteolysis is performed in which the protein is digested into small peptides with a protease, and 4) the peptides are analyzed with mass spectrometry in order to identify unmodified and modified quantities for each exposure

[7] Glycans are carbohydrates linked to a protein's amino acid residues. [42] Because of the heterogenicity of glycan masses, mass spectrometry analysis becomes quite challenging on glycosylated proteins. For this reason glycans are generally cleaved, or removed, enzymatically prior to mass-spectrometry analysis. [42, 43]

time.[8] An oxidation rate for each peptide can then be determined, and differences in a protein's solvent accessibility in varying conditions can be inferred through differential analysis. Decreases in the rate of oxidation correspond to decreases in solvent accessibility, while increases correspond to increases in solvent accessibility.

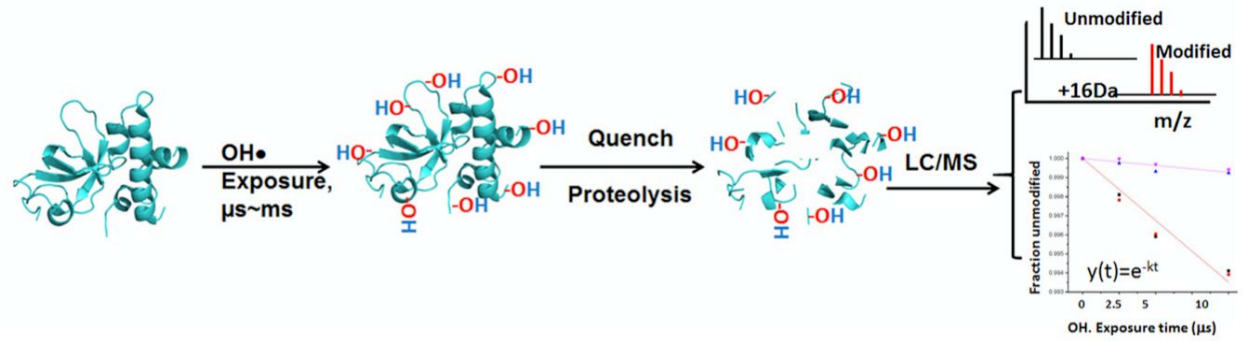


Figure 2-5: Hydroxyl-radical protein footprinting (HRPF) workflow [30]

Hydroxyl-radical protein footprinting relies on the generation or injection of hydroxyl radicals in solution. One potential concern with oxidative labelling (labelling with hydroxyl radicals) is the possibility that the process of oxidizing a protein may perturb its structure. Protein motion, or *breathing*, generally occurs on timescales between several microseconds and seconds. [44] It is thought that if a protein molecule is left in an oxidative environment for timescales longer than the breathing period, significant oxidation may occur in crevices that become solvent exposed as the protein opens up or unfolds, thus altering the protein structure. [45] The most common strategy for minimizing structural changes due to oxidation is to generate hydroxyl radicals on a timescale that is less than the fastest likely breathing rate of most proteins. [26, 46, 47] Since fast

[8] Exposure time here refers to the amount of time that a protein sample was exposed to hydroxyl radicals. It is used as a measure of the hydroxyl-radical dose that was used to treat a sample. Note that for some labelling techniques such as FPOP, the radical dose is controlled by the concentration of radical scavengers added to solution, rather than exposure time.

motion of protein is reported to be on the timescale of several microseconds or faster, systems which generate hydroxyl radicals on timescales less than one microsecond are ideal. [44, 48]

One approach to sub-microsecond radical generation involves radiolysis of water by a synchrotron x-ray source. [30, 48, 49] Aqueous-buffered samples are exposed to the synchrotron radiation for microseconds to milliseconds to generate hydroxyl radicals. [48, 50] The advantages of this experimental setup are that it does not require additional chemicals to generate radicals and the hydroxyl dose, *i.e.* the number of hydroxyl radicals generated in a given exposure, can be set and adjusted by varying the exposure time. [50] However, this method requires access to a synchrotron, specialized equipment, and technical expertise for exposing samples.

Another approach for generating hydroxyl radicals is done by photolysis of hydrogen peroxide. The technology is known as Fast Photochemical Oxidation of Proteins (FPOP) and is a widely used technology for hydroxyl-radical footprinting. [30, 51] It is performed using a flow system designed to mix an aqueous-buffered protein sample with hydrogen peroxide into a transparent capillary. The solution then is then irradiated with an ultraviolet (UV) excimer laser[9] to generate hydroxyl radicals. [51] The timescales on which hydroxyl radicals are generated is sub-microsecond, although the radical lifetime, which is limited to the recombination time, is reported to be on the order of 100 microseconds. [51] For this reason, a radical scavenger, such as 20 millimolar glutamine, is added to the protein solution in order to eliminate excess hydroxyl radicals in solution, and reduce the effective labelling timescale to around 1 microsecond, [51] though longer-lived reactive species have shown be present for times on the order of milliseconds. [53] These reactive species also have the potential to induce oxidation, thus, the effective labelling time of FPOP is likely on the order of milliseconds to seconds.

[9] Although they can vary, typical operating parameters for the excimer laser are as follows: wavelength is 248nm, pulse repetition frequency is ~10 Hz, and the energy is about 44 mJ/pulse. [52]

While used at several academic research institutions, FPOP has not gained much traction for use in the pharmaceutical industry because of several drawbacks associated with the system. These include the necessity to add hydrogen peroxide in solution [29] and the inability to label proteins on a sub-microsecond timescale, both of which are important for capturing the native structure of a protein in solution. [54] Furthermore, the apparatus is technically complicated to set up and use, expensive, and because it uses a high-powered laser, there are concerns about safety.

2.4 PLIMB

Plasma-Induced Modification of Biomolecules (PLIMB) technology was developed to generate hydroxyl radicals for performing hydroxyl-radical footprinting. [37] PLIMB uses a dielectric-barrier discharge plasma to generate hydroxyl and other radicals to modify a protein sample in an aqueous-buffered protein solution. [37] Radicals are generated from water and PLIMB does not require additional chemicals. [37] The plasma is produced by applying a time-varying high-voltage signal across two conducting electrodes. They are 1) a needle[10] and 2) a plane electrode, separated by an air gap. In addition, a sample tube that acts as a dielectric layer is needed. The system used here can be operated over a range of voltages from 1–30 kV, and a range of frequencies from 0 to 15 kHz. [37] Figure 2-6 shows a diagram of the PLIMB instrument, and Figure 2-7 shows a photograph of the setup and discharge.

[10] The PLIMB system uses a 2.5 inch-long, 0.06-inch-thick nickel-plated steel needle electrode.

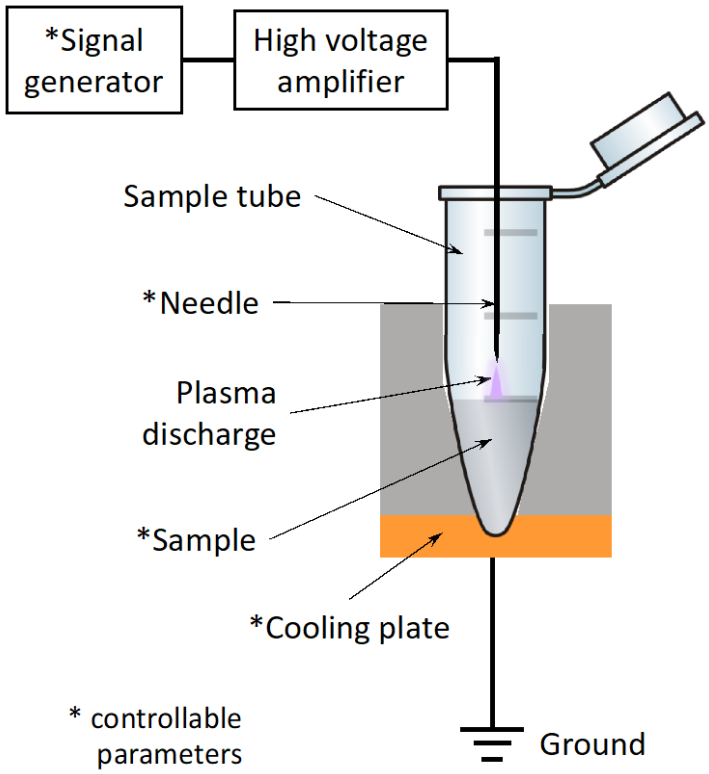


Figure 2-6: Diagram of the PLIMB instrument

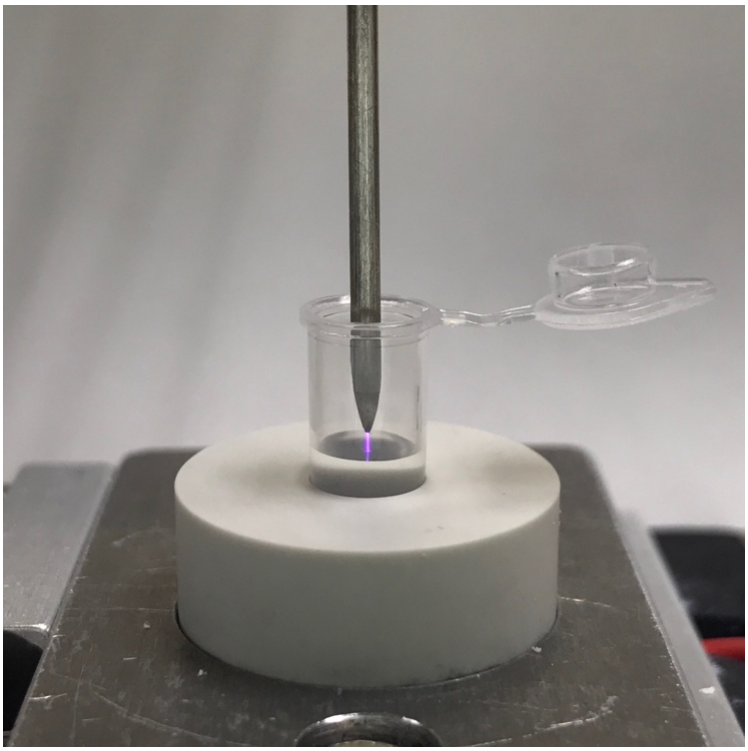


Figure 2-7: Photograph of PLIMB

PLIMB has several inherent advantages over other HRF techniques including but not limited to the following:

1. **Multiple tunable parameters**: Many of PLIMB's parameters such as applied voltage, frequency, and needle spacing can be independently controlled or varied in order to optimize the plasma for a given experiment or application.
2. **No added chemicals:** Unlike FPOP, PLIMB does not require addition of hydrogen peroxide or radical scavengers that can react with the protein or induce conformational changes.
3. **Microsecond radical generation:** Radicals can be generated in bursts on a microsecond timescale.[11] [37] Although oxidation via hydroxyl radicals can induce unfolding of proteins on a timescale of milliseconds or longer, PLIMB produces sub-microsecond bursts of hydroxyl radicals (see sections 3.7 and 3.8) in order to label the protein in its native conformation and preserve its structural integrity during exposure.
4. **Benchtop instrument:** PLIMB is a compact, benchtop instrument and is simple to operate.
5. **Temperature control:** PLIMB has a built-in Peltier cooler which allows the temperature of the sample to be varied from 2-95^{O} C with a variation of +/- 2 ^{O} C over a 60-second exposure.[12] This not only allows the sample temperature to be kept low enough to prevent protein denaturation, but it can also be used to study the effects of temperature on protein structure and conformation.
6. **High-throughput, automated, multiplexed system:** The instrument is fully automated, and the instrument is fully capable of sequentially processing a full 96-well plate.

[11] This proposition is based on the plasma current measurements described in Chapter 3.
[12] This proposition assumes that the temperature does not change dramatically after each plasma pulse.

2.5 Pharmaceutical pipeline

Although PLIMB may be used for many different applications, the goal of this work is to find the most compelling application of this technology for use in an industrial setting. As protein therapeutics candidates move through the pharmaceutical pipeline, there are many different tests that are performed on potential protein therapeutics to determine their suitability for treating a given disease.

Figure 2-8 shows a high-level diagram of the pharmaceutical pipeline. In the Target Discovery and Validation stage, researchers study the mechanisms underlying disease on a biochemical level. [55] In many cases, they look for a specific protein that is involved in a key aspect of a disease, which, if targeted with a therapeutic, will disrupt biological activity to halt the disease progression. [55, 56] This protein "target" is then used for the next stage: Screening. In this stage of the pipeline, thousands of molecules are tested for, in most cases, their ability to bind to the target. The molecules that bind are called "binders" or "hits." [56] After generating a collection of hits, further screening will be done to narrow this down to a smaller selection of "lead" candidates. These are molecules which look most promising based on many different tests that can be performed with relatively high throughput, usually multiplexed on 96 or 256-well plates. The properties of a lead molecule that are tested for when narrowing a selection of hits to a smaller selection of lead candidates includes dose response[13], binding affinity[14], solubility[15],

[13] The dose–response relationship, or exposure–response relationship, describes the magnitude of the response of an organism, as a function of exposure (or doses) to a stimulus or stressor (usually a chemical) after a certain exposure time.
[14] Binding affinity is the strength of the binding interaction between a single biomolecule (e.g., protein or DNA) to its ligand/binding partner (e.g., drug or inhibitor).
[15] Solubility here refers to the amount (mass or molar quantity) of a drug that can be dissolved in a liquid solution per unit volume.

permeability[16], biological function[17], toxicity[18], and stability[19]. [56] Further steps are then performed on lead candidates before they move to pre-clinical studies. These include modifying the lead candidates, often called molecular engineering [57], and in the case of antibody therapeutics, may include humanization[20], Fc[21] engineering, and affinity maturation[22]. [58] After lead validation, lead candidates will go on to pre-clinical or animal studies to test their effect *in-vivo*. Once successful data has been shown in pre-clinical studies, candidates must obtain an approved Investigational New Drug (IND) application from the FDA. [59] IND approval shows that the drug candidate has been shown to be safe for human clinical trials. The drug candidate will then need to go through several phases of clinical trials to test for its effectiveness and safety in humans. After clinical trials have been completed and approved by FDA, the drug can be manufactured and marketed.

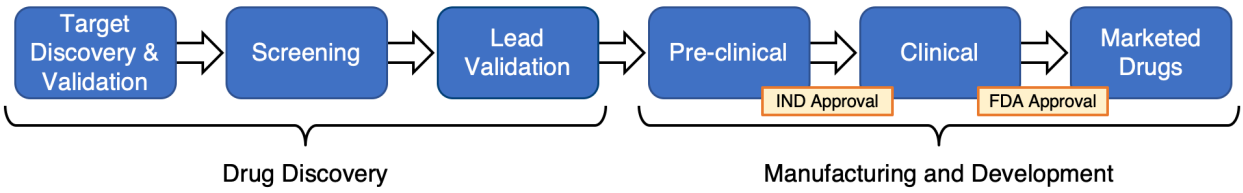


Figure 2-8: Pharmaceutical Pipeline

One key aspect of the pharmaceutical pipeline that should be pointed out is that the number of potential drug candidates being considered by a pharmaceutical company for treating a disease

[16] Permeability here refers to the ability of a drug to pass through a biological membrane.
[17] Function here refers to the specific biological activity that a drug induces.
[18] Toxicity is the degree to which a chemical substance or a particular mixture of substances can damage an organism.
[19] Drug stability refers to the ability of a drug to maintain the physical, chemical, and therapeutic properties during the time of storage and usage by the patient.
[20] Humanization is the process of replacing all regions of a non-human antibody sequence, the binding regions, onto a human antibody scaffold.
[21] The fragment crystallizable region, or Fc region, of an antibody is the tail (c-terminus) region of an antibody that reacts with the cell surface receptors of an organism.
[22] Affinity maturation is the process of optimizing an antibody's binding region for increased affinity, or strength of binding.

sharply decreases during the screening phase of drug discovery and continues to decrease as a pharmaceutical program moves further along in the pipeline. The number of possible drug molecules is estimated to be about 10^{60} for small molecule therapeutics, and 10^{390} for protein therapeutics, [60] yet the number of possible molecules that can serve as a viable drug for a given disease is just a minute fraction of these numbers. Therefore, finding a viable drug is akin to finding a needle in a haystack. The general strategy of most pharmaceutical programs is to start with a large selection of possible drug molecules, and systematically narrow this down by performing various tests as the candidates move through the pharmaceutical pipeline. For this reason, the order in which each test is performed is based largely on throughput. For example, binding assays such as an enzyme-linked immunosorbent assay (ELISA) [61] can be performed on hundreds or even thousands of different molecules in parallel, [62] and the readout for these tests is generally very accurate and requires very little human data validation. In addition, these tests give information about a very important aspect of a therapeutic's potential success, namely, whether it binds to the intended target. Therefore, they are excellent for narrowing down a large selection of potential drug candidates. [63] For this reason, binding assays are generally one of the first tests that are performed. On the other hand, *in-vivo* testing is also very important, as it tests the effectiveness and toxicity of a therapeutic in the conditions of a whole organism, [64] yet it is a much more expensive and time-consuming process. It is for this reason that pre-clinical studies are generally the last tests that are performed *before* clinical studies. During pre-clinical testing, the number of candidates has been narrowed down significantly since screening, and therefore, pre-clinical studies are much more manageable.

2.6 Applications of PLIMB

Here we look at the different features and characteristics of PLIMB technology and the data it produces and compare them to technologies currently in use today. The nature of the data in the case of PLIMB is mass-spectrometry (MS) based. The mass-spectrometry process performed after PLIMB labelling is done in a bottom-up proteomics fashion [65], whereby the protein is digested into peptides prior to analysis. This produces data on segments of the protein's amino-acid (AA) sequence and side-chain modifications on a sample protein. The addition of oxidative labeling using plasma treatment gives the ability to capture structural aspects of a protein at the time that it was being exposed to plasma in the PLIMB reactor. These structural aspects are in the form of solvent accessibility across different regions of the protein. Measurements of solvent accessibility are often used for the following applications:

1. **Conformational changes and allostery:** Protein conformation, or three-dimensional structure, is dependent on a protein's environment. As proteins are heated close to their melting temperature, they start to unfold and can even start to aggregate. [66, 67] Proteins may also exhibit different conformations depending on buffer conditions such as pH or salt content. [68-71] Furthermore, proteins exhibit allosteric effects, or binding-induced conformational changes, when interacting with other biomolecules. [5] Allosteric effects are used in biology for cell signaling, feedback, and enzyme-activity control. [5] In the context of pharma, choosing a protein to target often involves understanding the different allosteric effects associated with a disease. [72, 73] Furthermore, when a protein-drug target is bound to a pharmaceutical agent, it often will change conformation. This conformational change may be significant in terms of the protein's function in a biological system, and therefore, it is often necessary to be able to measure and characterize

conformational changes during the drug discovery and development processes. [72, 74, 75] Conformational changes can be detected by protein footprinting simply by assessing differences when capturing solvent accessibility information on a protein in different environments, such as temperature, pH, buffer composition and concentration, or presence or absence of a ligand or antibody binder.

2. **Protein structural dynamics:** Studying protein structural dynamics, or protein motion, can be critically important for understanding the various interactions between proteins and other biomolecules, especially for pharmaceutical researchers studying the various biomolecular mechanisms involved in a disease. [76, 77] Structural flexibility is a protein characteristic that can be assessed to help predict the binding interactions between a drug target and its target. [78] HDX can be an especially effective protein-footprinting method for characterizing dynamics by looking at the amount of labelling that occurs on a region of a protein after many different labelling times. This information can be used to determine the fraction of time that a protein spends in an exposed conformation versus a folded conformation, and the timescale over which the protein shifts between these conformations. [79, 80] Permanent conformational changes, due to binding or other causes, may also be characterized with HDX as well. [80-82]
3. **Comparability studies**: A comparability study refers to determining whether different samples of the same protein or protein therapeutic possess the same or similar chemical and structural characteristics. [83] Comparability is used as a quality-control measure to ensure biological activity of protein therapeutics throughout manufacturing, storage, and administration. [83] The FDA has extensive recommendations regarding comparability protocols for the pharmaceutical industry. [84] Comparability studies are also quite

important for biosimilar development. A biosimilar, as defined by the U.S. FDA, is a biological product (biologic) that is highly similar to and has no clinically meaningful differences from an existing approved reference product. [85] In order to show that a new biologic such as an antibody therapeutic is highly similar to that of the approved antibody, comparability of the protein's higher-order structures must be shown. HDX is used quite often for this application, as it can detect slight differences in solvent accessibility and structural flexibility between samples of the biosimilar and the innovator therapeutic. [86-88]

4. **Protein interactions**: One of the most important aspects of protein function is their ability bind or interacting with other biomolecules such as other molecules of the same protein (*e.g.,* dimers), other proteins, DNA/RNA, cell membranes, etc. Mapping how a protein interacts with other biomolecules is often critical to understanding the function of an interaction and how to develop a therapeutic to enhance or inhibit a given interaction. [89]

Given these potential applications for solvent-accessibility techniques, it is worth pointing out the unique features and characteristics of PLIMB, in order to provide a basis of comparison over other techniques which are currently used in industry. The first notable characteristic of PLIMB is that it, like other HRPF techniques, captures solvent accessibility information on the amino-acid side chains of a protein. This is unlike HDX, which captures solvent accessibility on the amide backbone of the protein. [33] Secondly, a PLIMB experiment and analysis can completed in about 48 hours, provided a mass spectrometer is available for analysis. When compared to techniques such as gel electrophoresis or enzyme-linked immunosorbent assay (ELISA), which can take as little as several hours, this is a relatively slow process. However, when

compared to x-ray crystallography or cryo-EM, which usually take several months to complete, [15] the PLIMB process is quite rapid. Another notable characteristic is that the minimum protein-sample concentration requirements are generally about two micrograms per sample.[23] Again, this can be a small or a large quantity depending on with what it is being compared, but it is generally in the acceptable range of protein quantities available for drug-discovery programs. PLIMB also works best with purified (>95%), recombinant sources of protein. This is not an absolute requirement, but allows for much quicker and easier data analysis, and will yield mass-spectrometer data with a much stronger signal intensity giving a greater potential for detecting statistically significant differences in solvent accessibility.[24]

As noted previously, high-resolution structural measurement techniques such as x-ray crystallography and cryo-EM give extremely high (atomic-level) resolution. However, the major constraint associated with these techniques is their turnaround time, which is on the order of several weeks to months. Therefore, because they are not feasible to be performed on a large number of samples, these techniques are used sparingly in the pharmaceutical pipeline only when high-resolution data is needed. Crystallography and cryo-EM are typically used in target identification, as it is often beneficial to have a complete three-dimensional structure of the specific protein which is to be targeted with a therapeutic. [90, 91] They are also often used in the lead-validation stages for assessing the structure of the antibody-antigen complex. [89] This is often referred to in the

[23] The sample requirement dependent on many variables including but not limited to 1) the type of mass spectrometer that is used for analysis, 2) the type of liquid chromatography mass spectrometer (LC-MS) analysis that is used, 3) the efficiency of a digestion protocol, and 4) what mass-spectrometry signal intensities are needed for a given experimental analysis.

[24] This is because the mass spectrometry data will only consist of a set of peptides, and their mass-shifted variants due to plasma modification, which are predetermined by the digestion protocol and the primary sequence of the protein being analyzed. A heterogeneous population of proteins will yield to a much larger set of peptides, which makes data analysis much more difficult. Furthermore, if the same amount of total protein is injected as in a homogeneous-protein sample, the amounts of each native peptide and their mass-shifted variants will be much less. This will lead to decreased signal intensity, as well as decreased sensitivity for detecting statistically significant differences between samples.

case of crystallography as co-crystallization and is used to understand the structural interactions between a drug and its target.

Because there are relatively few drug candidates left at the lead-validation stage of the pharmaceutical pipeline compared to previous stages, x-ray crystallography and cryo-EM are feasible at these stages. However, there are several reasons why these techniques are still not ideal for this stage of pharmaceutical development. The first is that, as stated previously, these techniques can have quite a significant turnaround time and cost associated with them. Although it may be feasible to perform an analysis on the antibody-antigen complex before moving on to later stages of development, it may not be feasible to use these techniques in an iterative process whereby the lead candidates are engineered or modified to improve their binding properties several times using the generated structural data as feedback. The second reason is that crystallography is not always possible with all drug targets, or with all therapeutic/target complexes. Furthermore, if crystallography is possible, the data that the technique generates comes from the crystalline structure of the complex, which may not exactly mimic the protein structure or dynamics *in vivo*.

PLIMB, on the other hand, is a much more rapid and cost-effective technique and generates data on proteins in solution. For these reasons, PLIMB is specifically suited to address the needs of generating structural data on protein therapeutics in the lead-validation stage of the pharmaceutical pipeline.

2.7 PLIMB for epitope mapping

One of the most appropriate and pressing applications for fast protein-structural technologies, such as PLIMB, is epitope mapping. [92, 93] Epitope mapping is the process of mapping the location at which a protein therapeutic binds with its therapeutic target. Figure 2-10 shows a diagram of a therapeutic antibody bound to a target, or antigen, and the epitope associated with this system.

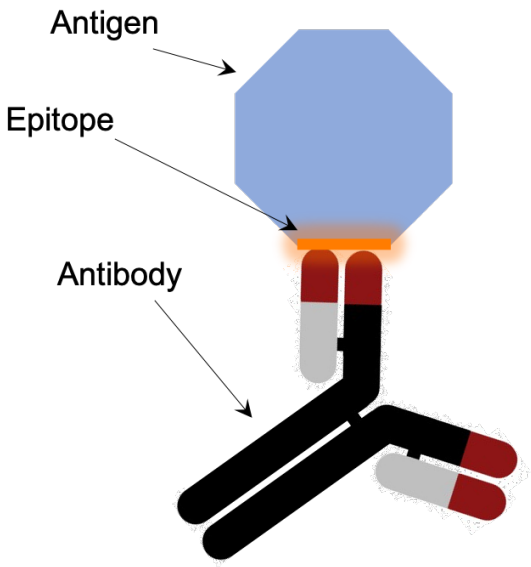


Figure 2-10: Epitope Diagram

The **blue** shape represents an antigen protein molecule, while the **orange** represents the epitope, or region where the antibody binds. The **red** region of the antibody represents the variable regions, or the regions that are unique to a particular antibody, which bind specifically to a particular epitope. The **black** portion of the antibody represents the heavy chain of the constant domain, or the domain of an antibody that is constant between all antibodies produced by a particular species, and the **grey** represents the light chain of the constant domain.

Epitope mapping is typically done in the lead-validation stage of the pharmaceutical pipeline. As mentioned previously, lead validation is the stage of the drug-discovery pipeline in which several leads, or compounds which have been found to bind to a therapeutic target, undergo further analysis before they continue towards pre-clinical and clinical trials.

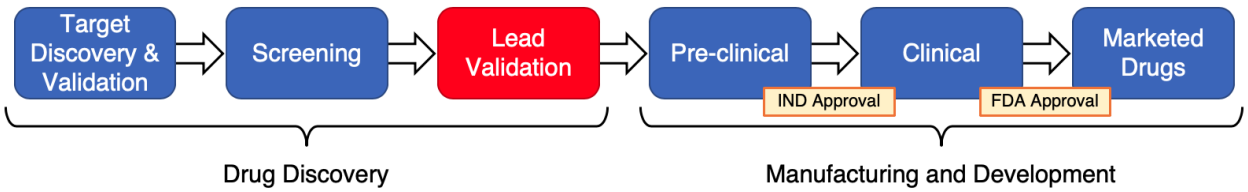


Figure 2-9: Lead Validation

One of the key uses of epitope data is for intellectual-property (IP) purposes. A patent for a protein therapeutic which includes detailed epitope data can be much easier to defend against infringement. Furthermore, pharmaceutical companies often file an “epitope patent” which will

protect against *any* molecule which binds to a specific protein at a specific epitope. [94] Over the past ten years, 10 to 24% of all antibody patents filed included epitope claims. [94, 95]

Epitope data can also be used to understand the mechanism of binding. Analyzing both the epitope (the binding location on the antigen), and the paratope (the binding location on the antibody) can give insight into the binding interactions. This information can be used to modify an antibody's chemical composition to create a new antibody with enhanced binding characteristics, such as a higher binding affinity[25]. [96-98]

Additionally, epitope mapping can be used as a screening assay. For example, it can be used to narrow down a large selection of potential antibodies by selecting only those that bind to a specific functional epitope, *i.e.*, a specific region that causes a disruption in disease activity. This can increase the probability of success in Phase I and II trials, reducing overall costs and increasing R&D productivity. [92]

Technologies for epitope mapping can be categorized by their level of resolution and throughput/turnaround time. High-resolution techniques such as cryo-EM and x-ray crystallography can be used for epitope mapping, yet they have the downsides previously discussed, including cost and turnaround time being significant for epitope mapping. Low-resolution techniques on the other hand, provide less detailed epitope data, but can be performed with much higher throughput. These techniques include epitope binning, whereby a competition assay is used to determine whether two or more antibodies bind to the same epitope, and therefore block each other from binding. With an increased number of antibodies with unique epitopes that are used for epitope binning, the higher the resolution that can be achieved. However, the highest possible resolution is still theoretically limited to the total solvent-accessible surface area of a

[25] The affinity of a binding interaction is a measure of the rate of binding, and generally implies the strength of the interaction.

protein, divided by the number of antibody binders with non-overlapping epitopes. Another method for low-resolution epitope mapping is called peptide scanning. This method tests whether the antibody binds to small pieces, or peptides, of the antigen protein. Although this technique can in theory give amino-acid level resolution, the practical resolution is about 5-10 amino acids in length. Furthermore, peptide scanning does not give information about conformationally dependent or discontinuous epitopes. For these reasons, peptide scanning and epitope binning are generally not suitable for epitope mapping in the lead-validation stages of the pharmaceutical pipeline and are typically used only for screening (see Section 2.5).

HDX has also become a popular technology for epitope mapping, [28, 99, 100] and HDX, like HRPF, has the benefit of providing peptide or amino-acid level resolution, [29] yet can be performed with much faster turnaround time than other high-resolution techniques. Nevertheless, HDX does still face some of the drawbacks discussed previously (Section 2.3.1). Furthermore, epitopes are highly influenced by side-chain interactions. [101, 102] Because HDX probes the protein backbone, [29] the technology generally is better suited to analysis of protein dynamics, whereas HRPF technologies, which probe the protein's side chains, [29] better lend themselves to protein/protein interactions over HDX.

Given the drawbacks that are associated with currently available techniques for epitope mapping, pharmaceutical development could benefit from the development of new technologies which overcome these drawbacks. Because PLIMB can generate data in a high-throughput fashion with turnaround times as little as 48-hours, it is considerably faster than high-resolution techniques, including x-ray crystallography and cryo-EM. It is also considerably less expensive and can be potentially applied to a wider range of target proteins than crystallography. Furthermore, the maximum resolution of PLIMB is amino-acid level, which is significantly higher than the low-

resolution techniques that are currently available. The combination of these features makes it well-suited for this role.

2.8 Summary

Protein therapeutics is a large and important segment of the pharmaceutical industry. [4] The use of proteins as therapeutics allows for much higher specificity than traditional small-molecule pharmaceuticals and can achieve a wide range of biological functions. [6] Because the structure of proteins largely determines their function, structural characterization of proteins is an essential process in the drug-development pipeline. Although techniques such as x-ray crystallography and cryo-EM can provide high-resolution structural data, they each exhibit disadvantages, most notably high cost and time requirements, which often make them unsuitable techniques for some analyses which require structural data.

Protein footprinting is a structural technique which can be much more time and cost effective than crystallography and cryo-EM. Protein footprinting works by chemically labelling the solvent-accessible regions of a protein sample in an aqueous solution, analyzing the protein samples with mass spectrometry, and quantifying the degree of labelling over various regions of the protein molecule to obtain a measurement of solvent accessibility. This information can be used to study protein/protein interactions, conformational changes, and higher-order structure comparability.

The two most common types of protein footprinting are hydrogen-deuterium exchange (HDX), and hydroxyl-radical protein footprinting (HRPF). HDX uses deuterium as a chemical label: proteins are submerged in a deuterated water solution, and deuterium exchanges with hydrogen atoms along the amide backbone of the protein, and the samples are analyzed with mass spectrometry. The rate of exchange is proportional to the solvent accessibility of the backbone

region of the protein. Hydroxyl-radical protein footprinting on the other hand, measures solvent accessibility of a protein's side chains. It is done by exposing a protein sample in solution to hydroxyl radicals, which oxidize the solvent-accessible side chains. One method of generating hydroxyl radicals for HRPF called fast photochemical oxidation of proteins (FPOP) works by irradiating an aqueous solution of hydrogen peroxide with a high-powered excimer laser. Although FPOP can be a very effective method for acquiring structural data, there are several downsides of the technique as discussed in Section 2.3.2 of this chapter, which may be hindering its adoption for widespread use in the pharmaceutical industry.

Plasma Induced Modification of Biomolecules (PLIMB) is a method for generating hydroxyl radicals in solution, which overcomes many of the downsides of FPOP. It uses a dielectric-barrier discharge plasma to generate sub-microsecond bursts of hydroxyl radicals in solution, without the need for additional chemicals such as hydrogen peroxide or a high-powered laser.

One of the most appropriate applications for PLIMB technologies is for a process called epitope mapping. This is because PLIMB can generate structural data for epitope mapping on an amino-acid level resolution in as little as 48 hours, whereas high-resolution techniques such as cryo-EM are significantly more expensive and time consuming.

2.9 References

[1] A. R. LLP, "Protein Therapeutics Market Analysis and Trends - Therapeutic Proteins, Application, Function - Forecast to 2025," @researchmarkets, Research and Markets, 2016. [Online]. Available: https://www.researchandmarkets.com/reports/3951937/protein-therapeutics-market-analysis-and-trends

[2] D. S. Dimitrov, "Therapeutic proteins," (in eng), *Methods Mol Biol,* vol. 899, pp. 1-26, 2012, doi: 10.1007/978-1-61779-921-1_1.

[3] B. J. Bruno, G. D. Miller, and C. S. Lim, "Basics and recent advances in peptide and protein drug delivery," (in eng), *Ther Deliv,* vol. 4, no. 11, pp. 1443-1467, 2013, doi: 10.4155/tde.13.104.

[4] H. A. D. Lagassé *et al.*, "Recent advances in (therapeutic protein) drug development," (in eng), *F1000Res,* vol. 6, pp. 113-113, 2017, doi: 10.12688/f1000research.9970.1.

[5] J. H. Wilson and T. Hunt, *Molecular biology of the cell, 4th edition : a problems approach*. New York; London: Garland Science (in English), 2002.

[6] B. Leader, Q. Baca, and D. Golan, "Protein therapeutics: a summary and pharmacological classification," *Nature Reviews Drug Discovery,* vol. 7, no. 1, pp. 21-39, 2008, doi: 10.1038/nrd2399.

[7] C. C. Quianzon and I. Cheikh, "History of insulin," (in eng), *J Community Hosp Intern Med Perspect,* vol. 2, no. 2, p. 10.3402/jchimp.v2i2.18701, 2012, doi: 10.3402/jchimp.v2i2.18701.

[8] N. A. P. S. Buss, S. J. Henderson, M. McFarlane, J. M. Shenton, and L. de Haan, "Monoclonal antibody therapeutics: history and future," *Current Opinion in Pharmacology,* vol. 12, no. 5, pp. 615-622, 2012/10/01/ 2012, doi: https://doi.org/10.1016/j.coph.2012.08.001.

[9] R. Rhoades and R. Pflanzer, *Human Physiology*. Brooks/Cole, 2002.

[10] C. Janeway, *Immunobiology Five*. Garland Pub., 2001.

[11] R.-M. Lu *et al.*, "Development of therapeutic antibodies for the treatment of diseases," *Journal of Biomedical Science,* vol. 27, no. 1, p. 1, 2020/01/02 2020, doi: 10.1186/s12929-019-0592-z.

[12] R. Y.-C. Huang *et al.*, "Hydrogen/deuterium exchange mass spectrometry and computational modeling reveal a discontinuous epitope of an antibody/TL1A Interaction," in *MAbs*, 2018, vol. 10, no. 1: Taylor & Francis, pp. 95-103.

[13] S. Aitipamula and V. R. Vangala, "X-ray crystallography and its role in understanding the physicochemical properties of pharmaceutical cocrystals," *Journal of the Indian Institute of Science,* vol. 97, no. 2, pp. 227-243, 2017.

[14] M. Smyth and J. Martin, "x Ray crystallography," *Molecular Pathology,* vol. 53, no. 1, p. 8, 2000.

[15] S. J. Coles and P. A. Gale, "Changing and challenging times for service crystallography," *Chemical Science,* vol. 3, no. 3, pp. 683-689, 2012.

[16] S. O. Garbuzynskiy, B. S. Melnik, M. Y. Lobanov, A. V. Finkelstein, and O. V. Galzitskaya, "Comparison of X-ray and NMR structures: is there a systematic difference in residue contacts between X-ray-and NMR-resolved protein structures?," *Proteins: Structure, Function, and Bioinformatics,* vol. 60, no. 1, pp. 139-147, 2005.

[17] D. E. Budil, P. Gast, C. Chang, M. Schiffer, and J. R. Norris, "Three-dimensional X-ray crystallography of membrane proteins: insights into electron transfer," *Annual Review of Physical Chemistry,* vol. 38, no. 1, pp. 561-583, 1987.

[18] H. Michel, "Crystallization of membrane proteins," in *Techniques and New Developments in Photosynthesis Research*: Springer, 1989, pp. 11-15.

[19] W. Kühlbrandt, "Three-dimensional crystallization of membrane proteins," *Quarterly Reviews of Biophysics,* vol. 21, no. 4, pp. 429-477, 1988.

[20] R. C. Stevens, "The cost and value of three-dimensional protein structure," *Drug Discovery World,* vol. 4, no. 3, pp. 35-48, 2003.

[21] R. Fernandez-Leiro and S. H. Scheres, "Unravelling biological macromolecules with cryo-electron microscopy," *Nature,* vol. 537, no. 7620, pp. 339-346, 2016.

[22] K. Murata and M. Wolf, "Cryo-electron microscopy for structural analysis of dynamic biological macromolecules," *Biochimica et Biophysica Acta (BBA)-General Subjects,* vol. 1862, no. 2, pp. 324-334, 2018.

[23] P. L. Stewart, "Cryo-electron microscopy and cryo-electron tomography of nanoparticles," *WIREs Nanomedicine and Nanobiotechnology,* vol. 9, no. 2, p. e1417, 2017, doi: 10.1002/wnan.1417.

[24] Y. Cheng, "Single-Particle Cryo-EM at Crystallographic Resolution," (in eng), *Cell,* vol. 161, no. 3, pp. 450-457, 2015, doi: 10.1016/j.cell.2015.03.049.

[25] S. Rogstad, A. Faustino, A. Ruth, D. Keire, M. Boyne, and J. Park, "A Retrospective Evaluation of the Use of Mass Spectrometry in FDA Biologics License Applications," *Journal of The American Society for Mass Spectrometry,* vol. 28, no. 5, pp. 786-794, 2017/05/01 2017, doi: 10.1007/s13361-016-1531-9.

[26] J. G. Kiselar and M. R. Chance, "Future directions of structural mass spectrometry using hydroxyl radical footprinting," *Journal of Mass Spectrometry,* vol. 45, no. 12, pp. 1373-1382, 2010, doi: https://doi.org/10.1002/jms.1808.

[27] P. Kaur *et al.*, "Characterizing monoclonal antibody structure by carboxyl group footprinting," *mAbs,* vol. 7, no. 3, pp. 540-552, 2015/05/04 2015, doi: 10.1080/19420862.2015.1023683.

[28] G. R. Masson, M. L. Jenkins, and J. E. Burke, "An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery," *Expert Opinion on Drug Discovery,* vol. 12, no. 10, pp. 981-994, 2017/10/03 2017, doi: 10.1080/17460441.2017.1363734.

[29] X. R. Liu, M. M. Zhang, and M. L. Gross, "Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications," *Chemical Reviews,* vol. 120, no. 10, pp. 4355-4454, 2020/05/27 2020, doi: 10.1021/acs.chemrev.9b00815.

[30] L. Wang and M. Chance, "Protein Footprinting Comes of Age: Mass Spectrometry for Biophysical Structure Assessment," in *Mol. Cell. Proteomics* vol. 16, ed, 2017, pp. 706-716.

[31] J. Zedalis and J. Eggebrecht, *Biology for AP® Courses*, Houston, Texas: OpenStax, 2018. [Online]. Available: https://openstax.org/books/biology-ap-courses/pages/3-4-proteins.

[32] I. C. Computational and Structural Biology Lab at Pune. "DEPTH Algorithm." http://cospi.iiserpune.ac.in/depth/htdocs/intro.html#depth_algorithm (accessed 2019).

[33] R. Huang and G. Chen, "Higher order structure characterization of protein therapeutics by hydrogen/deuterium exchange mass spectrometry," *Analytical and Bioanalytical Chemistry,* vol. 406, no. 26, pp. 6541-6558, 2014, doi: 10.1007/s00216-014-7924-3.

[34] J. R. Engen, "Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS," ed: ACS Publications, 2009.

[35] I. Oganesyan, C. Lento, and D. J. Wilson, "Contemporary hydrogen deuterium exchange mass spectrometry," *Methods,* vol. 144, pp. 27-42, 2018/07/15/ 2018, doi: https://doi.org/10.1016/j.ymeth.2018.04.023.

[36] E. Trabjerg, Z. E. Nazari, and K. D. Rand, "Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): challenges and emerging solutions," *TrAC Trends in Analytical Chemistry,* vol. 106, pp. 125-138, 2018.

[37] B. Minkoff, J. Blatz, F. Choudhury, D. Benjamin, J. Shohet, and M. Sussman, "Plasma-Generated OH Radical Production for Analyzing Three- Dimensional Structure in Protein Therapeutics," *Sci Rep,* vol. 7, no. 1, pp. 12946-12946, 2017, doi: 10.1038/s41598-017-13371-7.

[38] T. S. a. Medical, "LEAP HDX automation," 2020. [Online]. Available: https://www.trajanscimed.com/products/leap-hdx-automation.

[39] K. Takamoto and M. R. Chance, "Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes," *Annu. Rev. Biophys. Biomol. Struct.,* vol. 35, pp. 251-276, 2006.

[40] G. Xu and M. R. Chance, "Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting," *Analytical chemistry,* vol. 77, no. 14, pp. 4549-4555, 2005.

[41] G. Xu and M. R. Chance, "Hydroxyl radical-mediated modification of proteins as probes for structural proteomics," *Chemical reviews,* vol. 107, no. 8, pp. 3514-3543, 2007.

[42] R. A. Dwek, "Glycobiology: Toward Understanding the Function of Sugars," *Chemical Reviews,* vol. 96, no. 2, pp. 683-720, 1996/01/01 1996, doi: 10.1021/cr940283b.

[43] L. Han and C. E. Costello, "Mass spectrometry of glycans," *Biochemistry (Moscow),* vol. 78, no. 7, pp. 710-720, 2013/07/01 2013, doi: 10.1134/S0006297913070031.

[44] Z. Bu and D. J. E. Callaway, "Chapter 5 - Proteins MOVE! Protein dynamics and long-range allostery in cell signaling," in *Advances in Protein Chemistry and Structural Biology*, vol. 83, R. Donev Ed.: Academic Press, 2011, pp. 163-221.

[45] Y. Nabuchi, Y. Asoh, and M. Takayama, "Folding analysis of hormonal polypeptide calcitonins and the oxidized calcitonins using electrospray ionization mass spectrometry combined with H/D exchange," *Journal of the American Society for Mass Spectrometry,* vol. 15, no. 11, pp. 1556-1564, 2004/11/01 2004, doi: 10.1016/j.jasms.2004.07.007.

[46] M. R. Chance, B. Sclavi, S. A. Woodson, and M. Brenowitz, "Examining the conformational dynamics of macromolecules with time-resolved synchrotron X-ray 'footprinting'," *Structure,* vol. 5, no. 7, pp. 865-869, 1997.

[47] B. B. Stocks and L. Konermann, "Time-Dependent Changes in Side-Chain Solvent Accessibility during Cytochrome c Folding Probed by Pulsed Oxidative Labeling and Mass Spectrometry," *Journal of Molecular Biology,* vol. 398, no. 2, pp. 362-373, 2010/04/30/ 2010, doi: https://doi.org/10.1016/j.jmb.2010.03.015.

[48] J. Kiselar and M. Chance, "Future Directions of Structural Mass Spectrometry using Hydroxyl Radical Footprinting," *Journal of Mass Spectrometry,* vol. 45, 2011.

[49] S. Gupta, M. Sullivan, J. Toomey, J. Kiselar, and M. R. Chance, "The Beamline X28C of the Center for Synchrotron Biosciences: a National Resource for Biomolecular Structure and Dynamics Experiments Using Synchrotron Footprinting," *Journal of Synchrotron Radiation,* vol. 14, no. 3, pp. 233-243, 2007, doi: doi:10.1107/S0909049507013118.

[50] A. L. Klinger, J. Kiselar, S. Ilchenko, H. Komatsu, M. R. Chance, and P. H. Axelsen, "A Synchrotron-Based Hydroxyl Radical Footprinting Analysis of Amyloid Fibrils and Prefibrillar Intermediates with Residue-Specific Resolution," *Biochemistry,* vol. 53, no. 49, pp. 7724-7734, 2014/12/16 2014, doi: 10.1021/bi5010409.

[51] K. S. Li, L. Shi, and M. Gross, "Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization," in *Accounts Chem. Res.* vol. 51, ed, 2018, pp. 736-744.

[52] Y. Yan, "Fast photochemical oxidation of proteins (FPOP) maps the epitope of EGFR binding to adnectin," *Journal of the American Society for Mass Spectrometry.,* vol. 25, no. 12, p. 2084, 2014, doi: 10.1007/s13361-014-0993-x.

[53] S. Vahidi and L. Konermann, "Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds," *Journal of the American Society for Mass Spectrometry,* vol. 27, no. 7, pp. 1156-1164, 2016/07/01 2016, doi: 10.1021/jasms.8b05311.

[54] D. M. Hambly 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,* vol. 16, no. 12, pp. 2057-2063, 2005/12/01 2005, doi: 10.1016/j.jasms.2005.09.008.

[55] J. B. Gibbs, "Mechanism-based target identification and drug discovery in cancer research," *Science,* vol. 287, no. 5460, pp. 1969-1973, 2000.

[56] J. Hughes, S. Rees, S. Kalindjian, and K. Philpott, "Principles of early drug discovery," *Br J Pharmacol,* vol. 162, no. 6, pp. 1239-1249, 2011, doi: https://doi.org/10.1111/j.1476-5381.2010.01127.x.

[57] J. Zyss, "Molecular engineering implications of rotational invariance in quadratic nonlinear optics: From dipolar to octupolar molecules and materials," *The Journal of Chemical Physics,* vol. 98, no. 9, pp. 6583-6599, 1993, doi: 10.1063/1.464802.

[58] H. H. Shih, "Discovery process for antibody-based therapeutics," in *Development of Antibody-Based Therapeutics*: Springer, 2012, pp. 9-32.

[59] U. Food and D. Administration, "CFR-code of federal regulations title 21," ed, 2017.

[60] C. M. Dobson, "Chemical space and biology," *Nature,* vol. 432, no. 7019, pp. 824-828, 2004/12/01 2004, doi: 10.1038/nature03192.

[61] S. D. Gan and K. R. Patel, "Enzyme immunoassay and enzyme-linked immunosorbent assay," *J Invest Dermatol,* vol. 133, no. 9, p. e12, 2013.

[62] R. P. Hertzberg and A. J. Pope, "High-throughput screening: new technology for the 21st century," *Current Opinion in Chemical Biology,* vol. 4, no. 4, pp. 445-451, 2000/08/01/ 2000, doi: https://doi.org/10.1016/S1367-5931(00)00110-1.

[63] R. Macarron *et al.*, "Impact of high-throughput screening in biomedical research," *Nature Reviews Drug Discovery,* vol. 10, no. 3, pp. 188-195, 2011/03/01 2011, doi: 10.1038/nrd3368.

[64] R. Mangipudy, J. Burkhardt, and V. J. Kadambi, "Use of animals for toxicology testing is necessary to ensure patient safety in pharmaceutical development," *Regulatory Toxicology and Pharmacology,* vol. 70, no. 2, pp. 439-441, 2014/11/01/ 2014, doi: https://doi.org/10.1016/j.yrtph.2014.07.014.

[65] R. L. Gundry *et al.*, "Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow," *Current Protocols in Molecular Biology,* vol. 90, no. 1, pp. 10.25.1-10.25.23, 2010, doi: https://doi.org/10.1002/0471142727.mb1025s88.

[66] H.-C. Mahler, W. Friess, U. Grauschopf, and S. Kiese, "Protein aggregation: Pathways, induction factors and analysis," *Journal of Pharmaceutical Sciences,* vol. 98, no. 9, pp. 2909-2934, 2009, doi: https://doi.org/10.1002/jps.21566.

[67] B. Rosenberg, G. Kemeny, R. C. Switzer, and T. C. Hamilton, "Quantitative Evidence for Protein Denaturation as the Cause of Thermal Death," *Nature,* vol. 232, no. 5311, pp. 471-473, 1971/08/01 1971, doi: 10.1038/232471a0.

[68] N. V. Di Russo, D. A. Estrin, M. A. Martí, and A. E. Roitberg, "pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4," (in eng), *PLoS Comput Biol,* vol. 8, no. 11, pp. e1002761-e1002761, 2012, doi: 10.1371/journal.pcbi.1002761.

[69] G. Thakur, K. Jiang, D. Lee, K. Prashanthi, S. Kim, and T. Thundat, "Investigation of pH-Induced Protein Conformation Changes by Nanomechanical Deflection," *Langmuir,* vol. 30, no. 8, pp. 2109-2116, 2014/03/04 2014, doi: 10.1021/la403981t.

[70] C. P. Schwartz *et al.*, "Investigation of protein conformation and interactions with salts via X-ray absorption spectroscopy," *Proceedings of the National Academy of Sciences,* vol. 107, no. 32, pp. 14008-14013, 2010, doi: 10.1073/pnas.1006435107.

[71] M. V. Nguyen, K. A. Thorarinsdottir, A. Gudmundsdottir, G. Thorkelsson, and S. Arason, "The effects of salt concentration on conformational changes in cod (Gadus morhua) proteins during brine salting," *Food Chemistry,* vol. 125, no. 3, pp. 1013-1019, 2011/04/01/ 2011, doi: https://doi.org/10.1016/j.foodchem.2010.09.109.

[72] R. Nussinov and C.-J. Tsai, "Allostery in Disease and in Drug Discovery," *Cell,* vol. 153, no. 2, pp. 293-305, 2013, doi: 10.1016/j.cell.2013.03.034.

[73] A. F. Abdel-Magid, "Allosteric modulators: an emerging concept in drug discovery," (in eng), *ACS Med Chem Lett,* vol. 6, no. 2, pp. 104-107, 2015, doi: 10.1021/ml5005365.

[74] F. Feixas, S. Lindert, W. Sinko, and J. A. McCammon, "Exploring the role of receptor flexibility in structure-based drug discovery," (in eng), *Biophys Chem,* vol. 186, pp. 31-45, 2014, doi: 10.1016/j.bpc.2013.10.007.

[75] Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw, "How Does a Drug Molecule Find Its Target Binding Site?," *Journal of the American Chemical Society,* vol. 133, no. 24, pp. 9181-9183, 2011/06/22 2011, doi: 10.1021/ja202726y.

[76] Y. Lee, R. Lazim, S. J. Y. Macalino, and S. Choi, "Importance of protein dynamics in the structure-based drug discovery of class A G protein-coupled receptors (GPCRs)," *Current Opinion in Structural Biology,* vol. 55, pp. 147-153, 2019/04/01/ 2019, doi: https://doi.org/10.1016/j.sbi.2019.03.015.

[77] F. R. Salsbury, "Molecular dynamics simulations of protein dynamics and their relevance to drug discovery," *Current Opinion in Pharmacology,* vol. 10, no. 6, pp. 738-744, 2010/12/01/ 2010, doi: https://doi.org/10.1016/j.coph.2010.09.016.

[78] M. De Vivo, M. Masetti, G. Bottegoni, and A. Cavalli, "Role of Molecular Dynamics and Related Methods in Drug Discovery," *Journal of Medicinal Chemistry,* vol. 59, no. 9, pp. 4035-4061, 2016/05/12 2016, doi: 10.1021/acs.jmedchem.5b01684.

[79] L. Konermann, J. Pan, and Y.-H. Liu, "Hydrogen exchange mass spectrometry for studying protein structure and dynamics," *Chemical Society Reviews,* 10.1039/C0CS00113A vol. 40, no. 3, pp. 1224-1234, 2011, doi: 10.1039/C0CS00113A.

[80] J. Zheng, T. Strutzenberg, B. D. Pascal, and P. R. Griffin, "Protein dynamics and conformational changes explored by hydrogen/deuterium exchange mass spectrometry," *Current Opinion in Structural Biology,* vol. 58, pp. 305-313, 2019/10/01/ 2019, doi: https://doi.org/10.1016/j.sbi.2019.06.007.

[81] X. Huang *et al.*, "Probing the conformational changes of ovalbumin after glycation using HDX-MS," *Food Chemistry,* vol. 166, pp. 62-67, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.foodchem.2014.05.155.

[82] C. S. Kielkopf, M. Ghosh, G. S. Anand, and S. H. Brown, "HDX-MS reveals orthosteric and allosteric changes in apolipoprotein-D structural dynamics upon binding of progesterone," *Protein Science,* vol. 28, no. 2, pp. 365-374, 2019.

[83] M. A. Alsenaidy, N. K. Jain, J. H. Kim, C. R. Middaugh, and D. B. Volkin, "Protein comparability assessments and potential applicability of high throughput biophysical methods and data visualization tools to compare physical stability profiles," (in eng), *Front Pharmacol,* vol. 5, pp. 39-39, 2014, doi: 10.3389/fphar.2014.00039.

[84] (2016). *Comparability Protocols for Human Drugs and Biologics: Chemistry, Manufacturing, and Controls Information Guidance for Industry*. [Online] Available: **https://www.fda.gov/files/drugs/published/Comparability-Protocols-for-Human-Drugs-and-Biologics--Chemistry--Manufacturing--and-Controls-Information-Guidance-for-Industry.pdf**

[85] U. S. Food & Administration Drug, "Biosimilar and Interchangeable Products | FDA," 2020. [Online]. Available: https://www.fda.gov/drugs/biosimilars/biosimilar-and-interchangeable-products.

[86] A. Beck *et al.*, "Analytical characterization of biosimilar antibodies and Fc-fusion proteins," *TrAC Trends in Analytical Chemistry,* vol. 48, pp. 81-95, 2013/07/01/ 2013, doi: https://doi.org/10.1016/j.trac.2013.02.014.

[87] S. A. Berkowitz, J. R. Engen, J. R. Mazzeo, and G. B. Jones, "Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars," *Nature Reviews Drug Discovery,* vol. 11, no. 7, pp. 527-540, 2012/07/01 2012, doi: 10.1038/nrd3746.

[88] J. Pan, S. Zhang, and C. H. Borchers, "Comparative higher-order structure analysis of antibody biosimilars using combined bottom-up and top-down hydrogen-deuterium exchange mass spectrometry," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* vol. 1864, no. 12, pp. 1801-1808, 2016/12/01/ 2016, doi: https://doi.org/10.1016/j.bbapap.2016.08.013.

[89] G. Raghunathan, E. Hsieh, and G. Ermakov, "Epitope Characterization and Isotype Selection," in *Development of Antibody-Based Therapeutics*: Springer, 2018, pp. 41-69.

[90] T. L. Blundell, H. Jhoti, and C. Abell, "High-throughput crystallography for lead discovery in drug design," *Nature Reviews Drug Discovery,* vol. 1, no. 1, pp. 45-54, 2002/01/01 2002, doi: 10.1038/nrd706.

[91] J.-P. Renaud *et al.*, "Cryo-EM in drug discovery: achievements, limitations and prospects," *Nature Reviews Drug Discovery,* vol. 17, no. 7, pp. 471-492, 2018/07/01 2018, doi: 10.1038/nrd.2018.77.

[92] N. T. Ditto and B. D. Brooks, "The emerging role of biosensor- based epitope binning and mapping in antibody- based drug discovery," vol. 11, ed: Taylor & Francis, 2016, pp. 925-937.

[93] K. F. M. Opuni, M. Al-Majdoub, Y. Yefremova, R. F. El-Kased, C. Koy, and M. O. Glocker, "Mass spectrometric epitope mapping," vol. 37, ed, 2018, pp. 229-241.

[94] X. Deng, U. Storz, and B. J. Doranz, "Enhancing antibody patent protection using epitope mapping information," *mAbs,* vol. 10, no. 2, pp. 204-209, 2018, doi: 10.1080/19420862.2017.1402998.

[95] WIPO PATENTSCOPE, "Search with “antibody or antibodies” in Title and/or “epitope or epitopes” in Claims," ed, 2021.

[96] G. Nimrod *et al.*, "Computational Design of Epitope-Specific Functional Antibodies," *Cell Reports,* vol. 25, no. 8, pp. 2121-2131.e5, 2018/11/20/ 2018, doi: https://doi.org/10.1016/j.celrep.2018.10.081.

[97] U. Reineke *et al.*, "Identification of distinct antibody epitopes and mimotopes from a peptide array of 5520 randomly generated sequences," *Journal of Immunological Methods,* vol. 267, no. 1, pp. 37-51, 2002/09/01/ 2002, doi: https://doi.org/10.1016/S0022-1759(02)00139-4.

[98] C. Pinilla, J. R. Appel, and R. A. Houghten, "Functional importance of amino acid residues making up peptide antigenic determinants," *Molecular Immunology,* vol. 30, no. 6, pp. 577-585, 1993/04/01/ 1993, doi: https://doi.org/10.1016/0161-5890(93)90032-7.

[99] C. Puchades *et al.*, "Epitope mapping of diverse influenza Hemagglutinin drug candidates using HDX-MS," *Scientific reports,* vol. 9, no. 1, pp. 1-10, 2019.

[100] Q. Zhang *et al.*, "Epitope mapping of a 95 kDa antigen in complex with antibody by solution-phase amide backbone hydrogen/deuterium exchange monitored by Fourier transform ion cyclotron resonance mass spectrometry," *Analytical chemistry,* vol. 83, no. 18, pp. 7129-7136, 2011.

[101] B. C. Cunningham and J. A. Wells, "Comparison of a structural and a functional epitope," *Journal of molecular biology,* vol. 234, no. 3, pp. 554-563, 1993.

[102] D. C. Benjamin and S. S. Perdue, "Site-directed mutagenesis in epitope mapping," *Methods,* vol. 9, no. 3, pp. 508-515, 1996.

Chapter 3: Experimental Setup and Diagnostics

The idea for creating PLIMB originally started as a collaboration between researchers from the Biochemistry department and the Electrical and Computer Engineering department at UW – Madison. After needing a technology to analyze structural characteristics of proteins in solution, Michael Sussman, Professor of biochemistry at UW, investigated protein footprinting. The original goal was to set up an FPOP system [1-4] (see Section 2.3.2 for details) to perform hydroxyl-radical footprinting. After discussion with Leon Shohet, Professor of Electrical and Computer engineering at UW, to help with setting up the laser required to perform FPOP, it was decided instead to try to use plasma to generate hydroxyl radicals, as plasmas are known for generating free radicals. [5-10]

Research then began by Mike Sussman, Leon Shohet, Benjamin Minkoff, Faraz Choudhury, Joshua Blatz, and me to develop a system for generating hydroxyl radicals using plasma. The idea was first explored by researching plasma systems that interface with an aqueous medium. It was found that many plasmas like this are used for the application of water treatment. [11-13] This research provided ideas for initial experimental setups. After trying out several iterations of experimental setups, a working prototype was eventually settled upon, and the technology was coined Plasma Induced Modification of Biomolecules (PLIMB). The setup was used for many experiments to test its ability to generate radicals in solution and modify proteins in sufficient quantities to be detected and quantified with mass spectrometry, in order to prove the technology's utility for protein footprinting. A manuscript serving as a proof-of-concept, was then published by all six of the above-listed inventors. [14]

After realized that this system was novel and had several benefits over existing hydroxyl-radical footprinting systems, it was decided that the technology should be patented. The Wisconsin Alumni Research Foundation (WARF), which is the technology-transfer organization for the University of Wisconsin – Madison, saw the uniqueness and utility of the system, and made the decision to move forward with filing a patent on the technology. [15]

As this technology was being developed and tested for its use for protein characterization, many diagnostic tests were performed to characterize the nature of the plasma, the radicals being produced, and how these are affected by the various plasma parameters. This chapter outlines the various iterations of the experimental setup and the setup that is currently being used, as well as the various diagnostics that were performed on the system.

3.1 Hydroxyl-radical detection

One of the first diagnostic methods that was needed for development of the PLIMB system was a method to measure hydroxyl-radical production and/or concentration in a liquid sample. Our first approach was to use a chemical reporter that changed its optical properties as it was exposed to and oxidized by hydroxyl radicals. The chemicals that were used were Alexa Fluor 488, [5] terephthalic acid, [16] and coumarin. [17] After evaluating each of these chemicals for their effectiveness as a hydroxyl reporter, coumarin was determined to be most effective for the following reasons: 1) it showed a dose-dependent increase in fluorescence[26] over a wide range of plasma exposure times (up to about 30 seconds), 2) it is relatively inexpensive and simple to prepare in an aqueous-buffered solution, and 3) its fluorescence could be easily measured with a

[26] Fluorescence measurements used an excitation wavelength of 360 nm and a detection wavelength of 465 nm.

Tecan SPECTRAFluor Plus fluorometer, an instrument which was available for use for this project.

A second approach that was used to measure hydroxyl-radical generation by PLIMB was to expose samples of aqueous-buffered free methionine in solution to the plasma and measure the degree of oxidation with mass spectrometry. Methionine is an amino acid that converts to methionine sulfoxide in the presence of hydroxyl radicals with a high reactivity rate compared to other amino acids. [18, 19] Samples of plasma-exposed methionine can be analyzed with a mass spectrometer to detect the relative intensity of modified methionine (+16 Da) with respect to unmodified methionine. While this method requires a more labor-intensive analysis protocol than that of coumarin, it has some advantages. First, it shows a linear oxidation versus dose relationship over a longer range of plasma-exposure times compared to coumarin, as can be seen by comparing Figure 3-1 and Figure 3-2. Second, the method of detection can show less variability between replicate samples, particularly in exposure times less than 15 seconds.

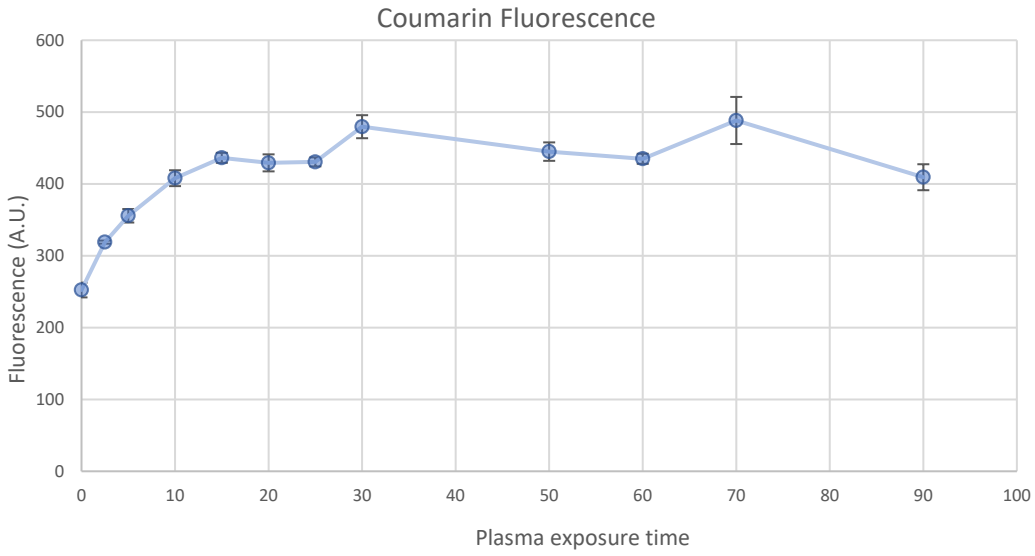


Figure 3-1: Coumarin fluorescence over a range of PLIMB exposure times

This graph shows the fluorescence intensity and variability between replicate coumarin samples exposed to a range of plasma exposure times. Coumarin was prepared in triplicates with a 10mM concentration in water and was exposed to a 10 kHz, 10 kV plasma at room temperature. Fluorescence measurement was done using a Tecan SPECTRAFluor Plus fluorometer with excitation wavelength of 360 nm and a detection wavelength of 465 nm.

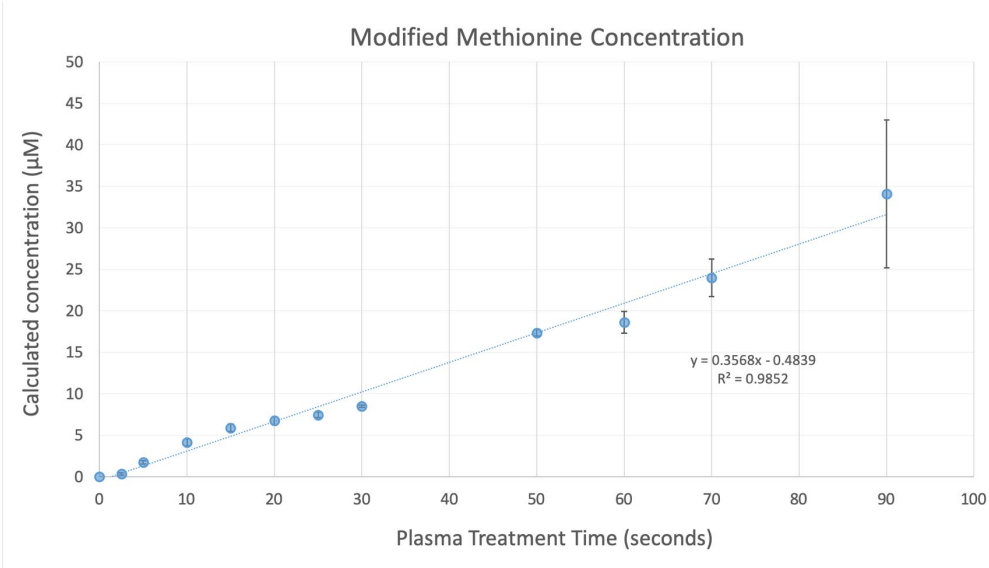


Figure 3-2: Production of oxidized methionine over a range of PLIMB exposure times

This graph shows the calculated concentration of methionine sulfoxide and variability between replicate samples of free methionine exposed to a range of plasma exposure times. Samples of 10µM methionine were prepared in triplicates in 50 mM ammonium acetate, and were exposed to a 10 kHz, 10 kV plasma at 4° C. Oxidized methionine was measured with an ESI-TOF mass spectrometer. The experimentation and data shown here was performed and prepared by Benjamin Minkoff. [14]

A third approach to measuring hydroxyl-radical production is by exposing samples of protein in solution, and after digestion and analysis by mass spectrometry, analyzing the proportion of oxidized peptides versus unmodified peptides. This approach represents an accurate metric for testing the ability of a given plasma system to generate hydroxyl radicals to modify a protein in solution for the purpose of protein-footprinting analysis. The model proteins that were used for testing PLIMB throughout its various iterations included cytochrome c (cyt-c), bovine-serum albumin (BSA), and lysozyme. These were used because they are readily available, inexpensive, well characterized, and familiar to most researchers in the field of protein biochemistry.

3.2 PLIMB evolution

Since the use of plasma as a method of in-solution hydroxyl radical generation for protein footprinting had not been previously investigated, a considerable amount of development was needed to build a system that was suitable for this application. The following sections describe the various stages of development of the PLIMB system that were built and evaluated before arriving at a system that was safe, easy to use, and could reliably oxidize protein samples to levels that are sufficient for protein footprinting.

3.2.1 Capacitor discharge

The first concept that was tested for hydroxyl radical generation was a capacitor-discharge plasma. A diagram for the setup that was built and tested by Josh Blatz, Benjamin Minkoff, Faraz Choudhury, Leon Shohet, and me, is shown in Figure 3-3. It includes a needle[27] placed about 1 mm above the surface of a liquid sample contained in sample holder consisting of a glass tube bonded to a sheet of metal which acted as a ground plate. A large capacitor (0.1-5μF, ≥15kV rated)

[27] The needle used here and throughout the rest of the experimental iterations was a nickel-plated steel needle, with a diameter of 0.06 inches and length of 2.5 inches.

was connected through a simple switch to the needle, and a high voltage DC power supply was connected to the capacitor in parallel. Note that this configuration allows current to flow directly from high potential to ground through the liquid sample. This allows for the high voltage potential (10 to 30 kV) to be DC and is stored in a capacitor before discharge.

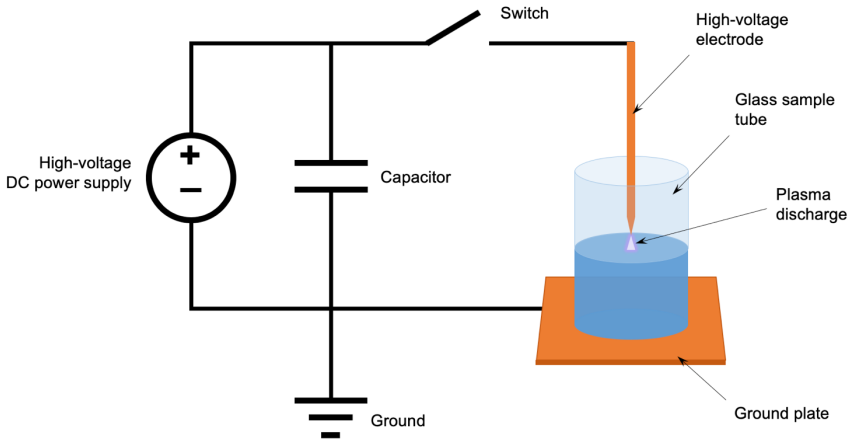


Figure 3-3: Capacitor discharge setup

The initial design produced a plasma discharge that was quite violent and uncontrollable. The glass container that held the protein sample would often shatter if the capacitor voltage was too high, and the amount of oxidation seen in a sample would be low if the capacitor voltage was set too low. To address this problem, a set of resistors was added to the electrical setup in series with the high-voltage line as shown in Figure 3-4. This would allow the capacitor to be charged to a high voltage but limited the amount of current that flowed through the plasma discharge. The resulting setup was much more controllable, although it still resulted in considerable sample loss (10-75% by volume) after exposure.

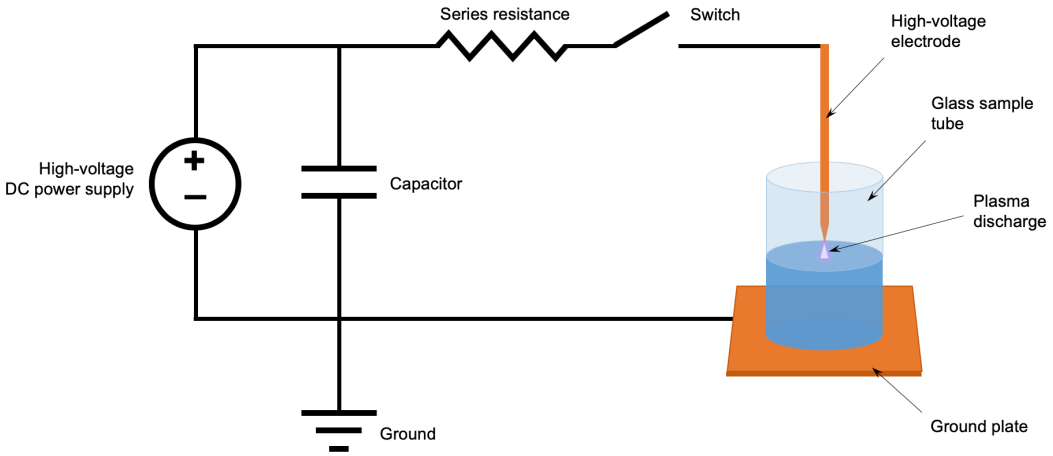


Figure 3-4: Capacitor discharge setup with series resistance

A combination of 1 kΩ, 1W-rated resistors would be connected both in parallel, allowing for a larger dissipation of power in the overall resistor network, and series, in order to increase the overall resistance. Total resistance values ranged from about 5 to 50 kΩ, limiting the discharge current produced from the capacitor, charged to 10 to 30 kV, to ≲ 1mA.

The radical dose was set with this system by exposing each sample to a specific number (1-20) of plasma discharges, with about 1-10 seconds between exposures. A solenoid switch was developed to complete the high-voltage circuit and was controlled with a computer-controlled function generator as seen in Figure 3-5. The circuit diagram for the solenoid controller can be found in Appendix A. This setup allowed the time between pulses to be precisely set and did not require a human operator to physically operate a switch while the high voltage power supply was running.

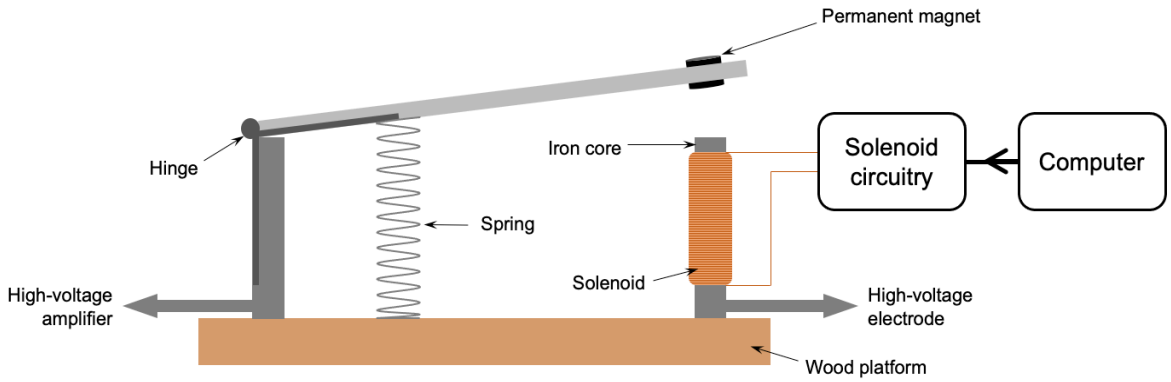


Figure 3-5: Solenoid switch

This setup was tested by Josh Blatz, Benjamin Minkoff, Faraz Choudhury, and me, by exposing samples of the protein cytochrome-c (see Section 3.1) to various numbers of plasma discharges. Samples of cytochrome-c were prepared in 200-300 µL volumes per sample at 50 µM concentration in 25 mM PBS. After exposure, the protein samples were prepared for mass spectrometry analysis. The process first involved digesting the protein with a protease (see Section 4.1.3 for a detailed description of the typical digestion protocol that was used), and the resulting peptides were analyzed with an Orbitrap LTQ-XL mass spectrometer. The results showed a positive correlation between the percentage of oxidized product from a specific peptide in the protein, 'MIFAGIK'[28], and the number of plasma discharges the sample was exposed to (see Figure 3-6). Despite the optimistic results shown here, this plasma setup proved to be less than ideal for several reasons, including the following: 1) it caused a significant amount of sample loss (10-75% by volume) after exposure, 2) the sample had to be in contact with the metal grounding electrode, which became significantly oxidized after several experiments, and there were concerns

[28] The notation here denotes the chain of amino acids, listed by their single letter abbreviation, that make up the peptide.

about sample contamination, and 3) the results were not sufficiently reproducible between replicates.

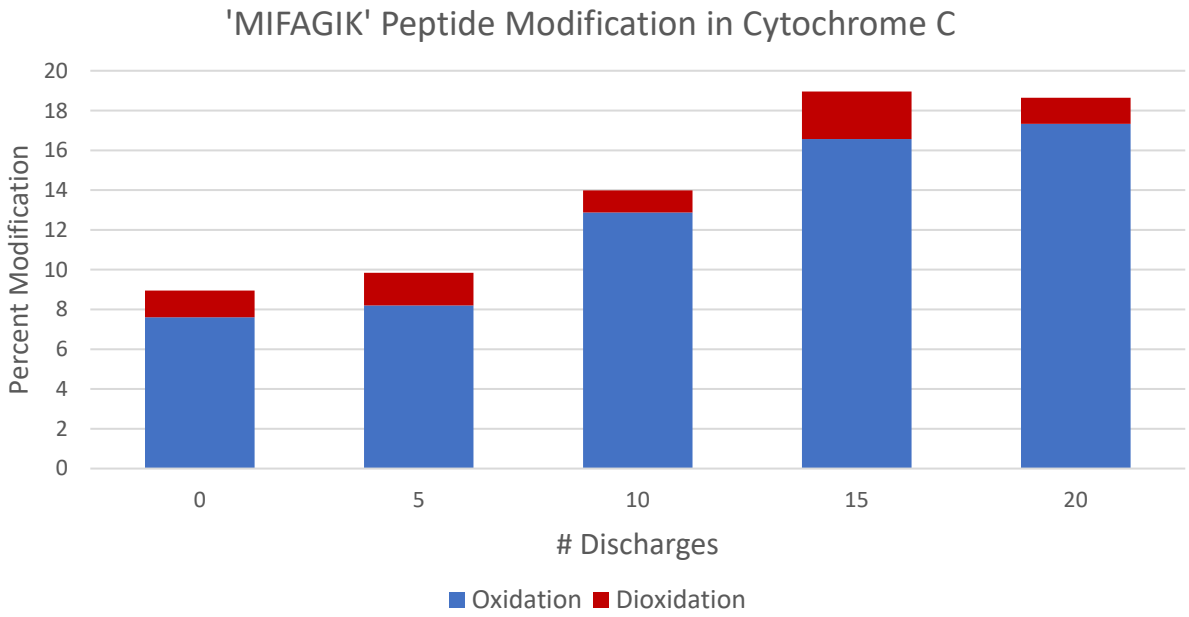


Figure 3-6: Capacitor discharge protein modification

3.2.2 Dielectric-barrier discharge

The next iteration of the PLIMB system involved a dielectric-barrier discharge (DBD). [20-22] This setup was different than the capacitor-discharge setup, because the sample was not in direct contact with the grounding electrode. Because of this, it does not allow electric current to flow directly from the high potential at the needle to ground potential. Rather, current flows from the needle into the sample where charge can build up. Because the applied voltage is AC, the changing voltage at the needle causes plasma breakdowns to occur repeatedly as the potential between the needle and the surface of the liquid solution is high enough to cause breakdown. The resulting plasma appeared visually as a continuous plasma discharge between the needle and the solution. However, upon further electrical diagnostics, the nature of the plasma was determined to exhibit

several rapid plasma discharges over one AC-voltage cycle. More discussion of the plasma diagnostics and timescales can be seen in Sections 3.7 and 3.8.

The first iteration of the DBD system used a 60 Hz neon-sign transformer as the high voltage source (see Figure 3-7). The applied voltage necessary to consistently produce a plasma with a needle spacing of about 1 mm was about 15kV. After testing with coumarin, it was found that this system generated a dose of hydroxyl radicals that was proportional to the exposure time.

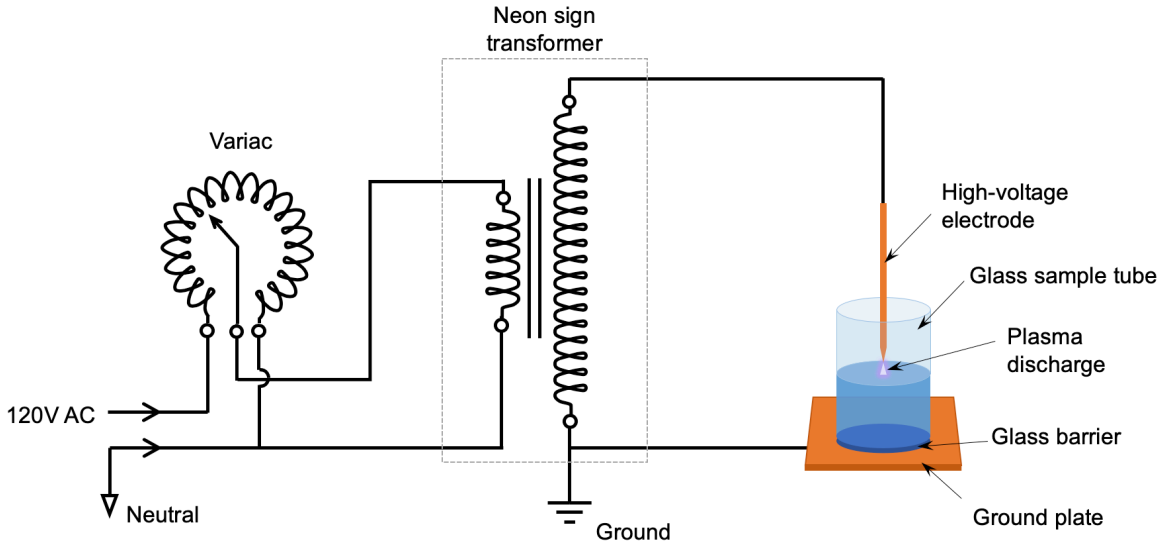


Figure 3-7: Neon sign transformer dielectric-barrier discharge

This system had several benefits over the original capacitive-discharge system. The most important was that it produced higher radical doses in samples in a shorter amount of time, and the radical dose itself could be more finely adjusted by setting an exact exposure time rather than by setting an integer number of individual discharges as in the capacitor discharge system. Additionally, significant sample loss during exposure was no longer observed.

3.2.3 Audio-frequency amplifier system

After showing that a DBD could produce hydroxyl radicals that oxidized proteins using a simple and inexpensive DBD that used a neon-sign transformer as the high voltage source, the system was modified to have the capability to produce variable applied-high-voltages and frequencies. As seen in Figure 3-8, this involved using a Trek Model P0621P high voltage amplifier with a signal generator used at the input to set the amplitude, frequency, and wave shape of the output voltage. This particular amplifier is a single sided (positive) voltage supply with a gain of 30,000 V/V, a maximum voltage of 30 kV, a large signal bandwidth (<1% distortion) of 3.5 kHz, and a small signal bandwidth (-3 dB) of 25 kHz.

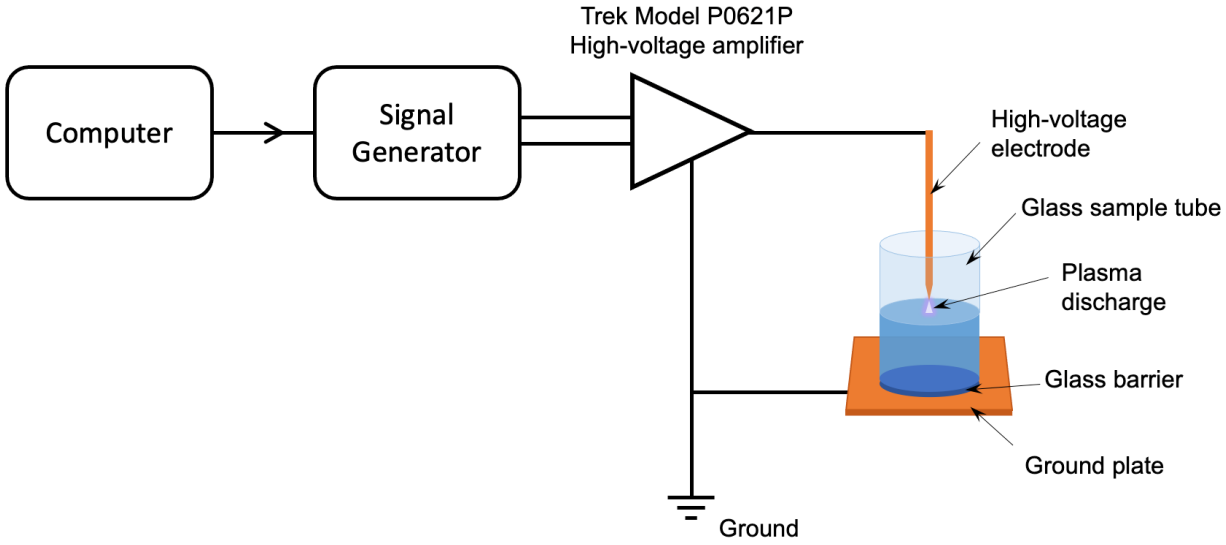


Figure 3-8: High-voltage amplifier setup

3.3 Plasma parameters

It was found that there were certain combinations of needle spacings, applied voltages and frequencies that produced a stable plasma that was less likely to result in unwanted arcing and generated high doses of radicals over a given exposure time. One of these combinations was the following:

Needle spacing: 1 mm

Frequency: 10 kHz

Wave shape: Sinusoidal

Signal generator amplitude: 10 V peak-to-peak

D.C. Offset[29]: 0 V

This produced an output high voltage of 10 kV peak-to-peak[30] with a D.C. offset of about 3 kV with respect to ground potential and a triangular waveshape (see Figure 3-16). The plasma produced at these settings is visually and audibly very smooth and consistent and behaves like a corona discharge rather than an arc discharge. Another minor benefit to this combination of settings, especially with respect to the neon-sign transformer, is that the plasma is much less audibly noisy.

For the sake of consistency between experiments, the above-listed settings and system were used throughout almost all future experiments. Throughout this thesis, the plasma parameters listed above should be assumed to be used unless otherwise stated.

3.4 Temperature control

It is known that solution temperature affects protein conformation. [23-26] Since the purpose of exposing samples to plasma in the PLIMB system is to capture structural information, it is important to be able to precisely control a protein sample's temperature during exposure so as to not induce temperature-driven protein structural changes during exposure. Since it is often ideal to keep protein samples at temperatures as low as ~2° C during experiments in order to prevent

[29] D.C. Offset refers to the voltage offset from the sinusoid above or below ground potential.
[30] Because the frequency is outside the large-signal bandwidth of the amplifier, the gain of the amplifier at this frequency is about 10,000 V/V whereas its gain at frequencies below its bandwidth is about 30,000 V/V.

the protein's tightly bundled conformation from unfolding, it was determined that a cooling system would need to be added to the PLIMB system to keep a protein sample at low temperatures during exposure. Furthermore, any plasma-generated heat that could cause an increase in sample temperature would also need to be removed.

The amount of temperature increase in samples of various volumes was tested over a range of exposure times in order to quantify the effect of plasma-induced heating. This was done by placing a thermistor-based temperature probe into the liquid sample and recording the temperature both before and immediately after plasma exposure. The results for various sample volumes can be seen in Figure 3-9. Since it was found that plasma exposure would transfer a non-negligible amount of heat to a protein sample, it was determined that the PLIMB system would require some form of active sample cooling during plasma exposure.

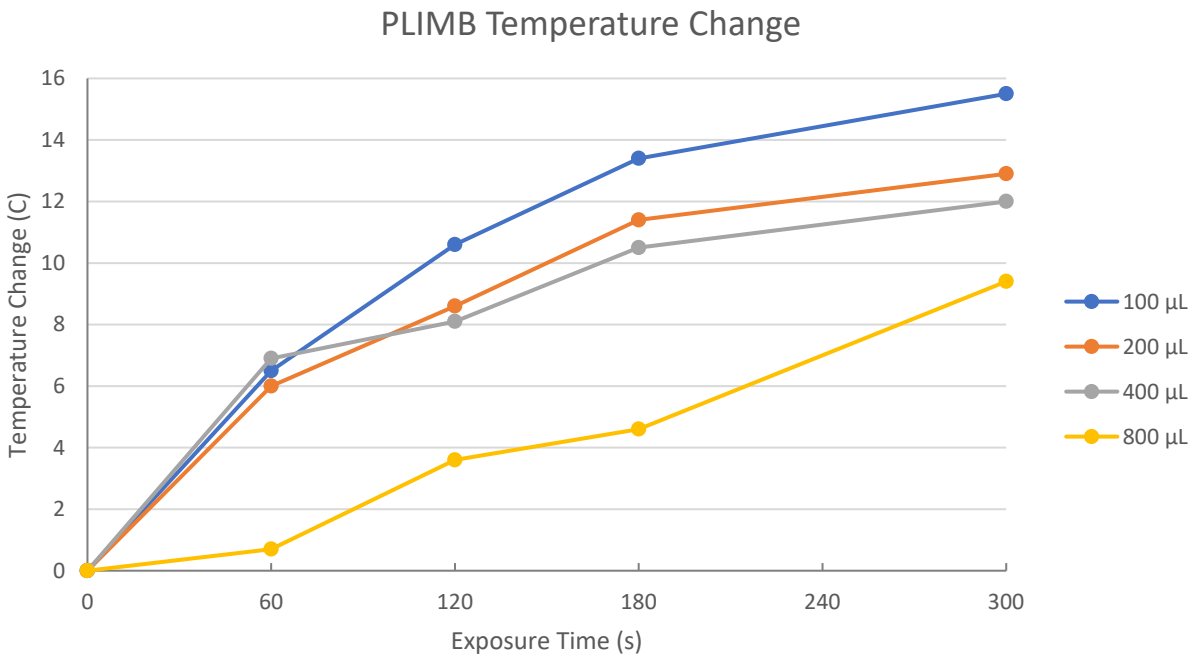


Figure 3-9: Plasma-induced temperature change

The cooling system that was chosen was a Peltier cooler, which is a thermoelectric device that directs heat away from a targeted region. [27, 28] A Peltier cooler was chosen because the

amount of cooling can be precisely controlled using a pulse-width modulated (PWM) power supply, and it allows for both cooling and heating of a sample. A schematic of the setup that was implemented is shown in Figure 3-10. It included a thermistor-based temperature sensor connected to the sample holder, which allows for feedback into a temperature controller, allowing for precise temperature control. The temperature controller, a TE-Tech model TC-720, works by switching the DC voltage applied to the Peltier cooler in a PWM fashion, and adjusting the duty cycle to precisely control the amount of cooling. The controller connects to an external computer, which allows for useful features such as temperature monitoring and ramping functions.

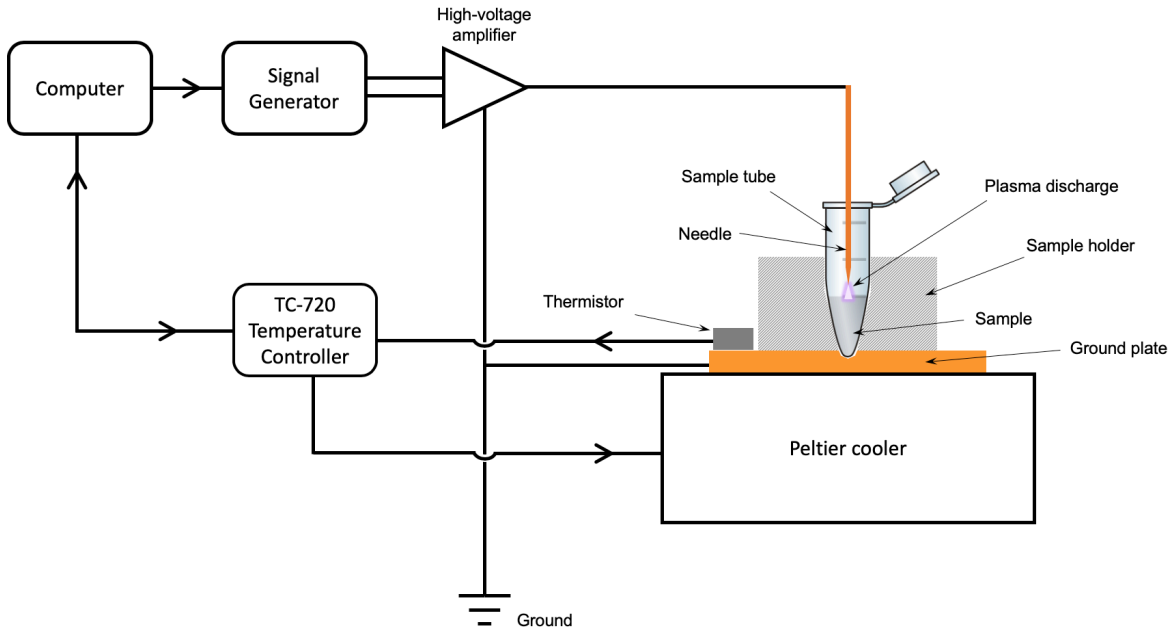


Figure 3-10: Cooling diagram

One of the challenges associated with designing a cooling system was in developing a sample holder that was able to conduct heat away from the protein sample while preventing electrical arcing. Early cooling prototypes that were built used a polymerase chain reaction (PCR) temperature cycler, which had a sample holder made of aluminum with an oxidized (alumina) coating. Although alumina has high heat conducting capabilities and provides some electrical

insulation from the bulk aluminum, the alumina-coated sample holder would often cause electrical arcing when used in conjunction with PLIMB. The alumina coating in this case did not provide adequate electrical insulation to hold off the > 10 kV of electrical potential. The sample holder that was eventually used in this system is made of a ceramic material called Shapal Hi-M Soft. This ceramic has the unique characteristics of being both highly electrically insulating and highly thermally conductive. Furthermore, the material is considered a soft ceramic, making it much easier to machine into any desired shape. The sample holder, as seen in Figure 3-11, was made of a single piece of machined ceramic. A technical drawing of the sample holder can be seen in Appendix B. It was designed to hold a 250 µL thin-walled plastic tube, which is often used for PCR applications. These tubes provide a high amount of surface area in contact with the sample holder for good thermal conductivity and hold a liquid volume that is reasonable for most protein-footprinting experiments. In order to allow for the sample tube to be touching the ground electrode while maintaining sufficient insulation between the high voltage needle and ground, a grounding plate, made of either copper (first used) or stainless steel (later used), was added to the setup, and was positioned below the ceramic sample holder. A technical drawing of the grounding plate can be seen in Appendix C. The grounding plate was attached to the Peltier cooling plate.

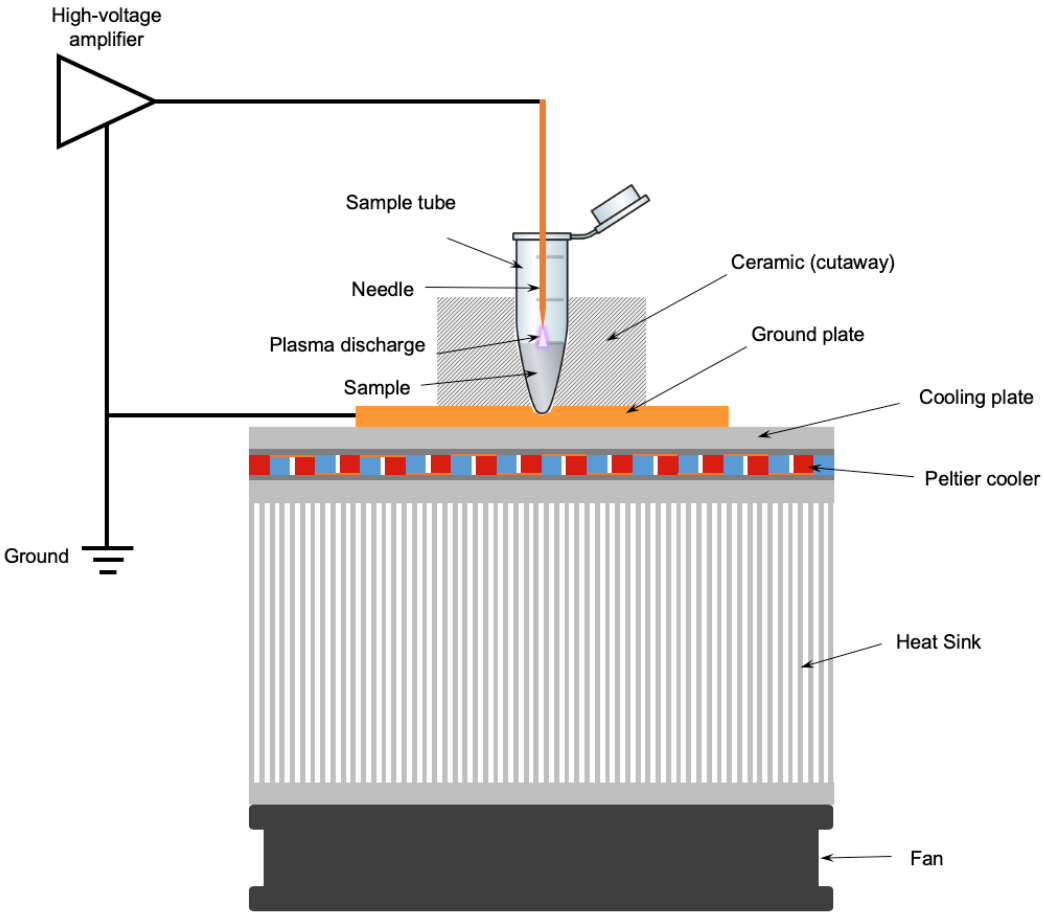


Figure 3-11: Peltier cooling setup

3.5 Single-well setup

After implementation of the temperature control system, a "single-well" PLIMB setup, which allows for exposure of a single protein sample at a time, was constructed and is shown in Figure 3-12 and Figure 3-13. This setup includes the following components: 1) a mechanical breadboard for holding and positioning the cooling system, 2) a Peltier cooler, 3) a TC-720 temperature controller, 4) the high thermal conductivity and high electrical insulation ceramic sample holder,

5) an acrylic needle holder,[31] 6) a large translation stage[32] for quickly moving the high-voltage needle vertically several inches in order to place a new sample into the setup, 7) a small translation stage[33] for adjusting the length of the gap between the needle and the sample surface, 8) an acrylic enclosure[34] to prevent a user from coming in contact with the high voltage while it is enabled, 9) an interlock switch which disables the high-voltage amplifier when the acrylic enclosure is open, 10) a Trek Model P0621P high voltage amplifier, 11) an Agilent 8116A signal generator whose output is fed into the input of the high-voltage amplifier, 12) and a computer which connects to and controls both the temperature controller and the signal generator. The computer runs a LabView program that is specifically designed to set the desired plasma parameters and turn on the plasma for a given amount of time by controlling the output of the signal generator, waiting for a specified amount of time, and then disabling the output of the signal generator.

[31] The needle holder is made of a solid piece of cast acrylic, and has dimensions of 6 inches long, approximately 1 inch wide, and 1 inch thick. A technical drawing can be seen in Appendix D.
[32] ThorLabs model VAP10, 10" Travel Vertical Translation Stage.
[33] ThorLabs model PT1/M 25.0 mm Translation Stage with Standard Micrometer
[34] The enclosure is made of cast acrylic that is approximately 3/16 inches thick.

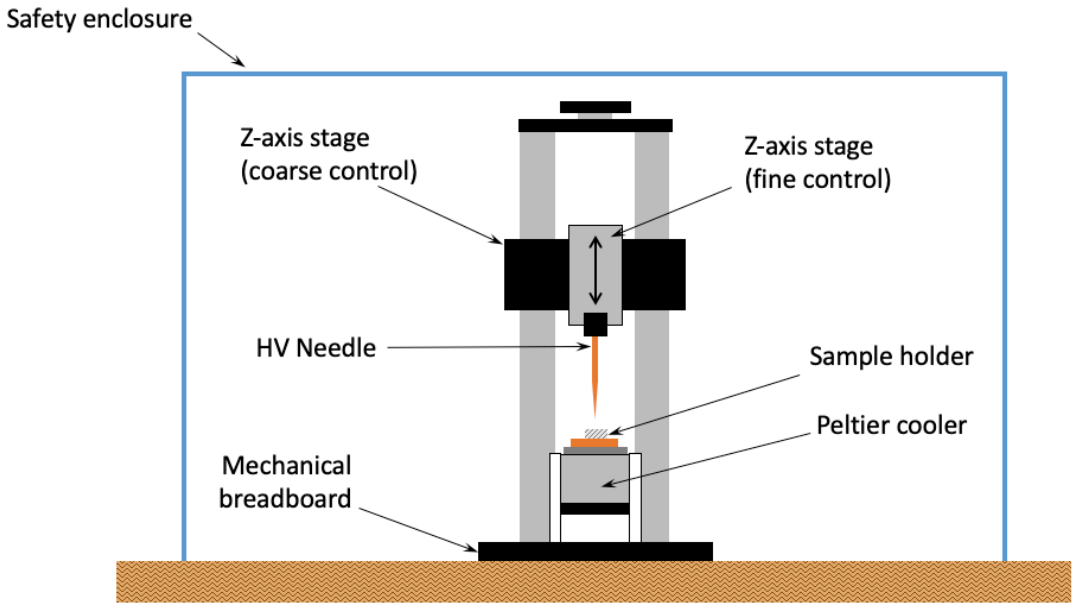


Figure 3-12: Single-well diagram

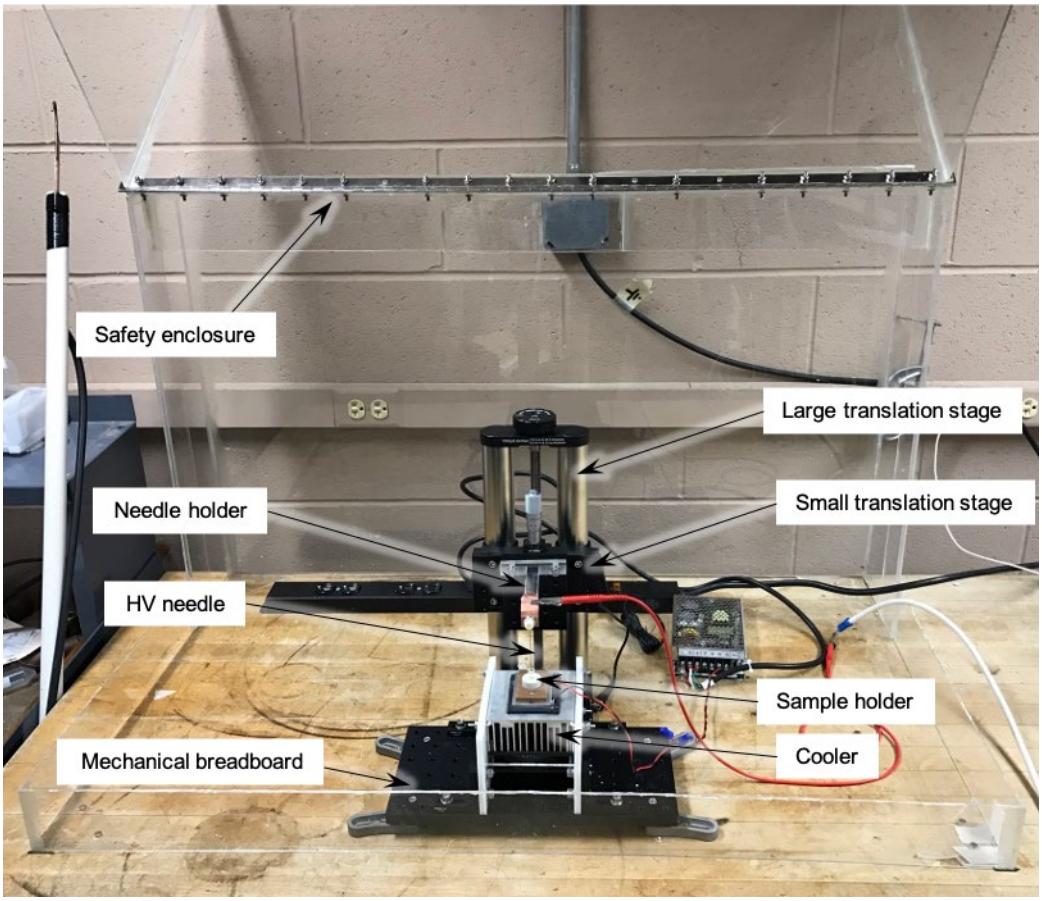


Figure 3-13: Single-well picture

3.6 Multiplexing system

The purpose of the next iteration of the PLIMB instrument is to increase the throughput of the system. Using the single-well setup, a user must manually replace each sample after exposure. Although this may be practicable for experiments which only require up to about 12 samples, it would be quite laborious to do this for large numbers of samples. An example of a case where large numbers of samples would need to be analyzed is for the screening phase of the drug-discovery process. A potential application of PLIMB for drug discovery would be to screen antibody therapeutic hits by their epitope interactions (see Section 2.7 for details). Given that a pharmaceutical program could in theory have on the order of 100 hits or more to screen, a PLIMB instrument with higher throughput than a single-well setup would be more suitable for this type of application.

The approach for increasing the throughput of the PLIMB instrument was to construct a system which would automatically expose a 96-well plate of samples in a sequential fashion. Ninety-six-well plates are ubiquitous in many biochemistry and medical applications, especially for PCR, which is why this format was chosen. A new ceramic sample holder was constructed to hold a standard 96-well plate and was fitted to a larger Peltier cooler. A technical drawing of the 96-well plate can be seen in Appendix E. Three motorized translation stages were then used to provide three dimensions of motion to allow the needle to be positioned directly above any of the 96 wells in the plate. Figure 3-14 shows a diagram and Figure 3-15 shows a picture of the multiplexing setup. It includes the following components: 1) a 100-mm-travel motorized stage[35] which moves the entire cooling system and 96-well plate in the y-direction (forward and

[35] ThorLabs model NRT100/M 100mm Motorized Linear Translation Stage

backwards), 2) a Thorlabs 150-mm travel motorized stage[36] which moves the needle along the x-direction (left and right), 3) a 50-mm travel motorized stage[37] which moves the needle in the z-direction (up and down), 4) a TE-Tech model TC-720 temperature controller, 5) a TE-Tech model CP-061HT Peltier-cooling plate, 6) a 96-well ceramic sample holder plate, 7) an aluminum breadboard to fix all of the components in place and in alignment, 8) an enclosure[38] with acrylic panels[39] which ensures user safety from the high voltage needle, and 9) an interlock switch which disables the high-voltage amplifier when the enclosure is open.

[36] ThorLabs model NRT150/M 150mm Motorized Linear Translation Stage
[37] Thorlabs model PT1B/M 25 mm Translation Stage with a model ZFS25B 25 mm Travel, Compact Stepper Motorized Actuator
[38] The enclosure was constructed using ThorLabs 25mm Construction Rails.
[39] The panels were made from extruded acrylic with thickness of approximately 3/16 inches.

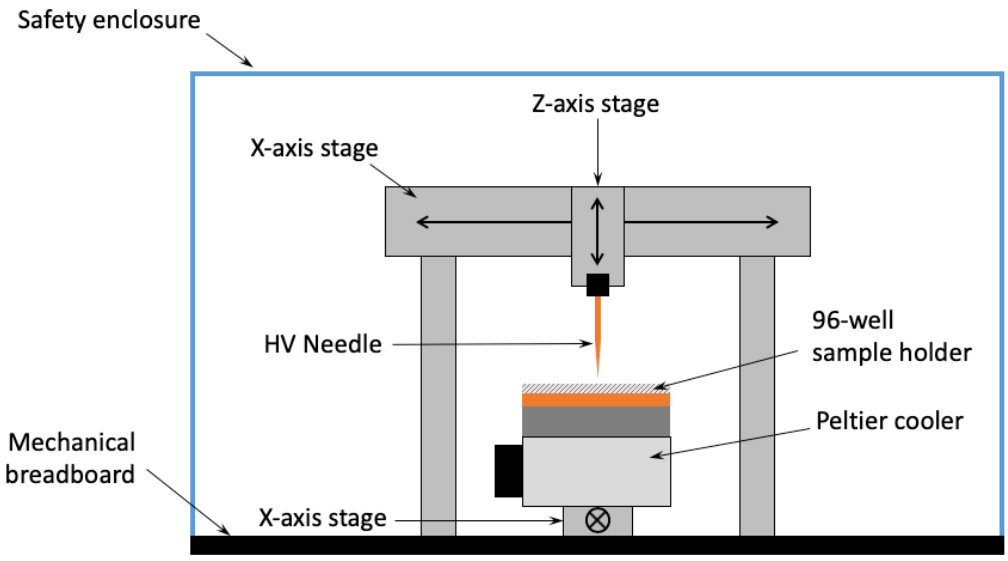


Figure 3-14: Multiplexing system diagram

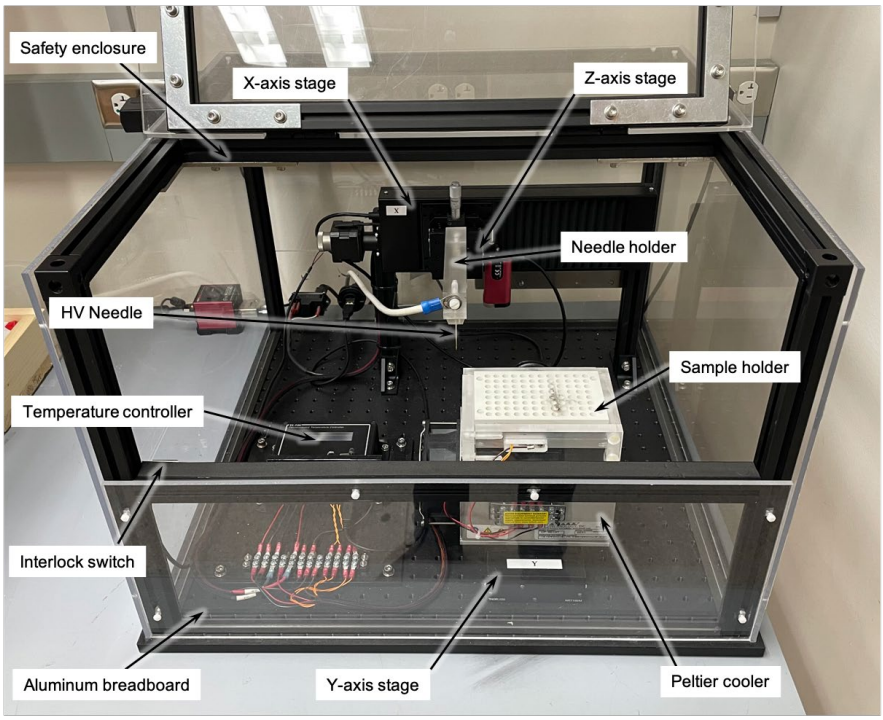


Figure 3-15: Photograph of multiplexing system

3.7 Plasma current and voltage measurements

In order to gain further insight into the nature of the plasma being used to generate hydroxyl radicals, it is required to have a system to measure the plasma current and/or the electric current flowing into the plasma. Figure 3-16 shows a diagram of the experimental setup that was used to measure plasma current. It includes a Pearson Model 110 current monitor to measure the current and a Tektronix Model P6015A high-voltage probe to simultaneously measure the applied voltage, each of which is displayed on an oscilloscope. An example of a current and voltage measurement can be seen in Figure 3-17, which was produced by Joshua Blatz. [14] These results show that over a single 10 kHz voltage cycle, plasma current flows in ~20 short pulses, each having a width of around 1 nanosecond and a separation of about 10 microseconds. Therefore, the number of pulses in a plasma exposure is roughly 2×10^5 per second.

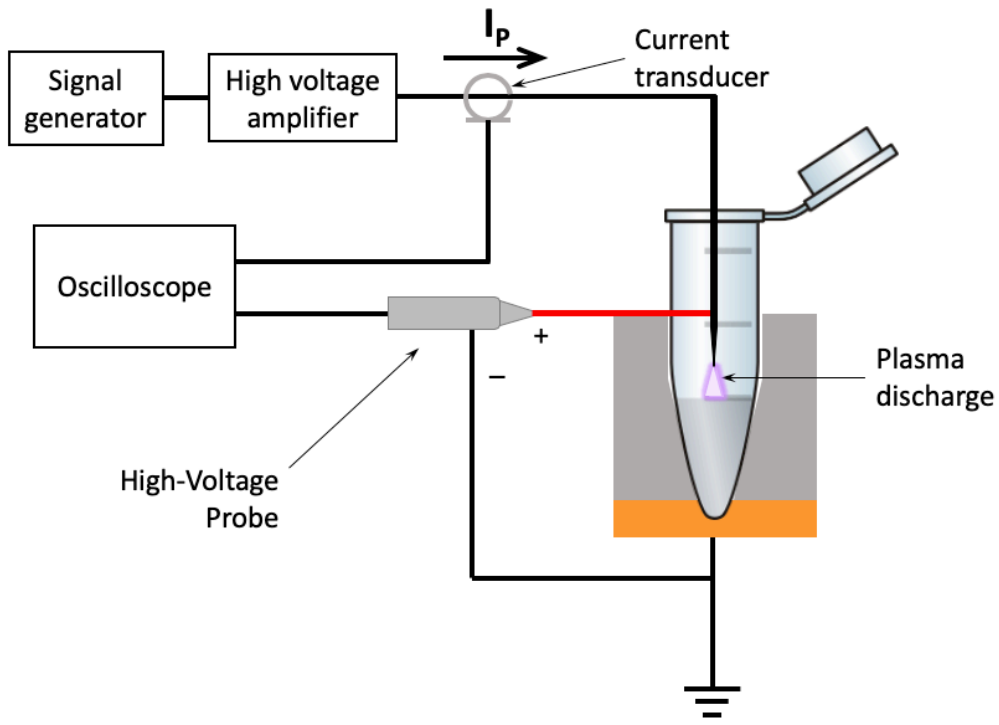


Figure 3-16: Plasma current and voltage measurement setup

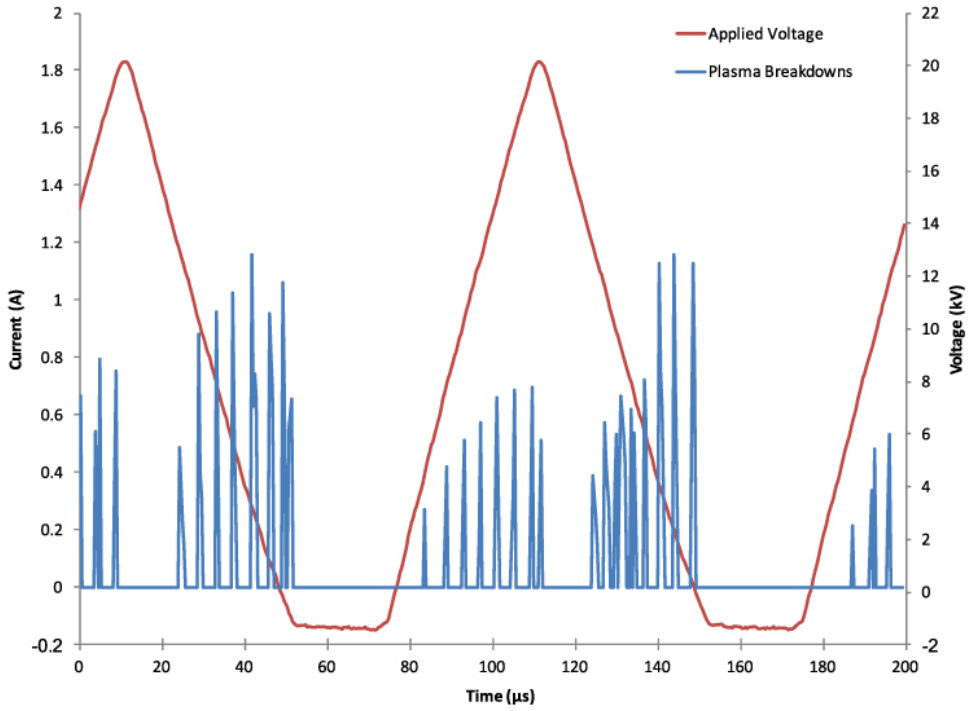


Figure 3-17: Plasma current and voltage vs. time plot [14]

This suggests that plasma discharges are generated over sub-microsecond bursts, with spacings between bursts that are variable, but are on the order of about 10 microseconds. As we will show in Section 3.8, each plasma discharge is the source of a small amount of hydroxyl-radical generation. Therefore, the timescale over which hydroxyl radicals are produced is similar to that of a plasma current pulse.

Because hydroxyl radicals dissipate or recombine within microseconds, [29] short bursts of radicals with spacing of at least 1 microsecond are ideal for minimizing structural deformation during labelling. This is because protein structural changes typically take place on a timescale greater than 1 microsecond. [30] It is widely hypothesized in the field of hydroxyl radical protein footprinting that any structural changes due to hydroxyl radicals themselves will be minimized when the radicals interact with the protein in short bursts that are less than 1 microsecond in duration. To demonstrate that this is the case with PLIMB, it was shown by Benjamin Minkoff

that exposure to PLIMB does not significantly disrupt structure, by exposing protein bovine serum albumin (BSA) to PLIMB for up to 60 seconds and analyzing with mass spectrometry. [14]

It should also be noted that the current pulses in Figure 3-17 vary in magnitude. It has been reported that, in a similar plasma setup to PLIMB, narrow spikes in plasma current directly correlate with bursts of hydroxyl-radical generation as measured with UV spectroscopy. [31] Because of this, we believe that the variations of amplitude and spacing in the current pulses directly correlate with the variability in hydroxyl-radical generation. This hypothesis has been tested, and the results are shown in Chapter 5.

3.8 Photoemission spectroscopy

A significant amount of preliminary work has been done by Joshua Blatz to measure the hydroxyl-radical concentration during plasma exposure and highlight the relationship between plasma current and hydroxyl generation. UV-spectroscopy measurements were made using the experimental setup shown in Figure 3-18, which can be set up for either absorption or emission spectroscopy. Light emitted from the PLIMB plasma for optical emission, and/or light from a UV source placed behind the plasma system for optical absorption, passes into a monochromator. The appropriate use of optical components allows for spatial resolution of hydroxyl-radical production along the plasma discharge as well as within the protein solution. Light is directed into a McPherson 234-302 monochromator and is diffracted with a 1200 grooves/mm grating blazed for 140 nm. A Hamamatsu R928 photomultiplier tube (PMT) at the exit of the monochromator is used to measure the intensity of light coming through the monochromator as a function of wavelength.

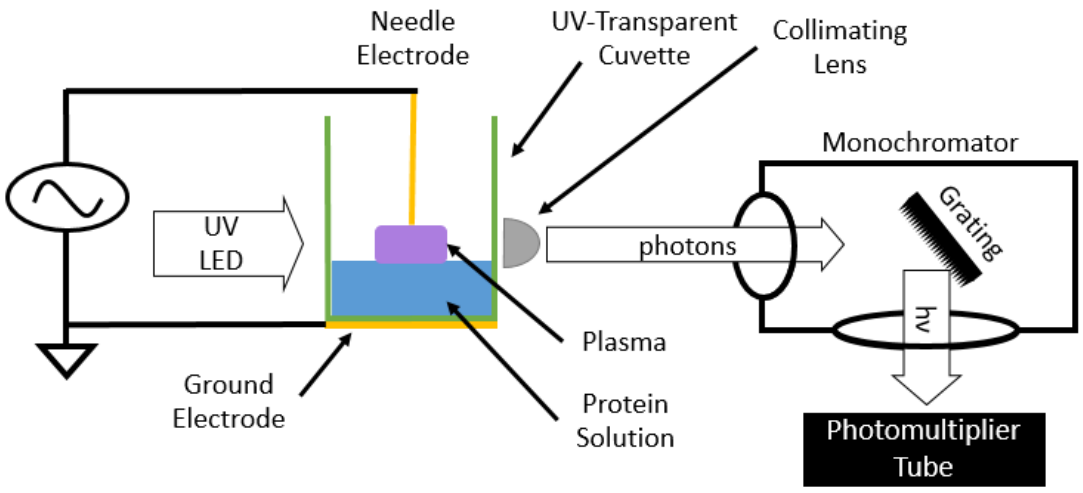


Figure 3-18: Photoemission spectroscopy setup

This diagram was created by Joshua Blatz.

In order to investigate the relationship between plasma discharges and hydroxyl generation, emission spectroscopy was performed while simultaneously measuring plasma current. The monochromator was set to 309 nm, a wavelength of light known to be emitted from hydroxyl radicals. [29, 31-35] The photomultiplier current was measured along with the plasma-current pulses. Figure 3-19 shows a spike in hydroxyl generation lasting about 800 nanoseconds (ns) measured from the beginning of a sharp increase in PMT current, that was preceded by a plasma discharge. These results show that hydroxyl-radical generation occurs in quick succession with the individual plasma discharges and suggests that plasma current may be an effective parameter for indirectly measuring hydroxyl-radical generation. The ~200 ns delay between the plasma discharge and the spike in PMT current is an electrical delay caused by the response time of the photomultiplier amplifier circuitry (see Appendix F for details).

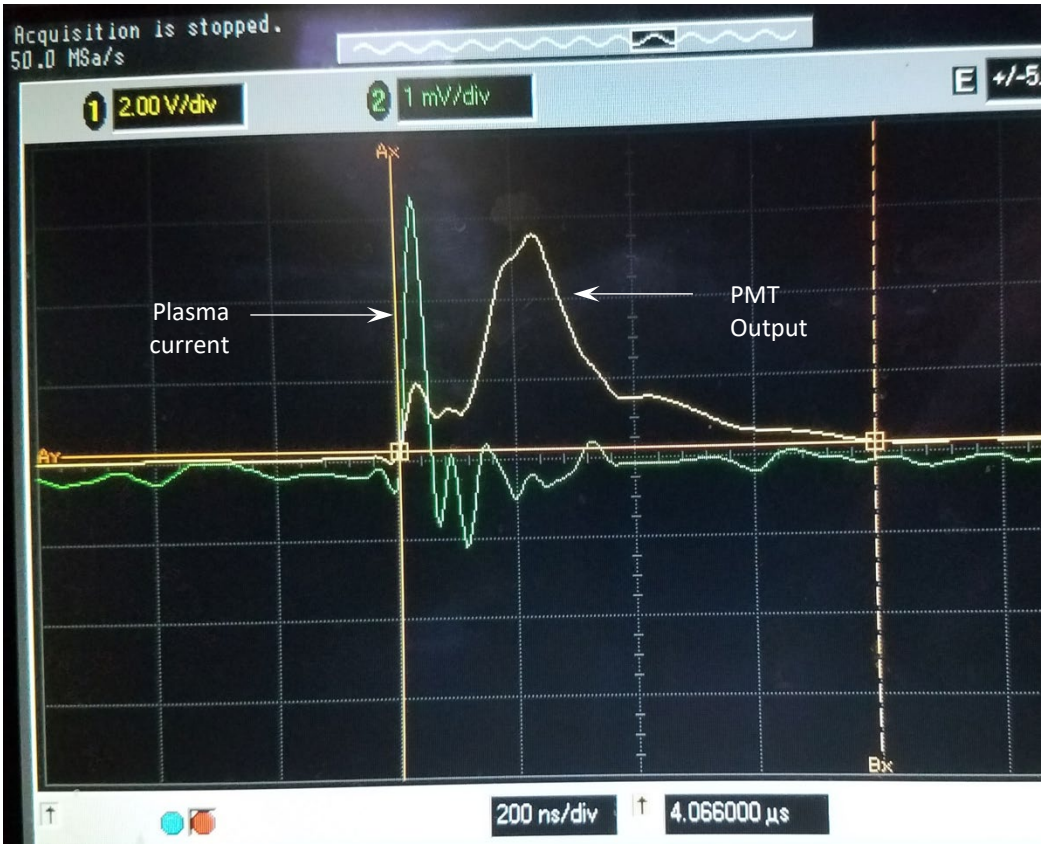


Figure 3-19: Plasma current and 309nm light measurement

3.9 References

[1] K. S. Li, L. Shi, and M. Gross, "Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization," in *Accounts Chem. Res.* vol. 51, ed, 2018, pp. 736-744.

[2] X. R. Liu, M. M. Zhang, and M. L. Gross, "Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications," *Chemical Reviews,* vol. 120, no. 10, pp. 4355-4454, 2020/05/27 2020, doi: 10.1021/acs.chemrev.9b00815.

[3] B. Niu and M. L. Gross, "MS-based hydroxyl radical footprinting: Methodology and application of fast photochemical oxidation of proteins (FPOP)," *Mass Spectrometry-Based Chemical Proteomics,* pp. 363-416, 2019.

[4] K. S. Li, L. Shi, and M. L. Gross, "Mass spectrometry-based fast photochemical oxidation of proteins (FPOP) for higher order structure characterization," *Accounts of chemical research,* vol. 51, no. 3, pp. 736-744, 2018.

[5] F. A. Choudhury, G. Sabat, M. R. Sussman, Y. Nishi, and J. L. Shohet, "Fluorophore-based sensor for oxygen radicals in processing plasmas," *Journal of Vacuum Science & Technology A,* vol. 33, no. 6, p. 061305, 2015, doi: 10.1116/1.4930315.

[6] H. Ren, G. Jiang, G. A. Antonelli, Y. Nishi, and J. L. Shohet, "The nature of the defects generated from plasma exposure in pristine and ultraviolet-cured low-k organosilicate glass," *Applied Physics Letters,* vol. 98, no. 25, p. 252902, 2011, doi: 10.1063/1.3601922.

[7] H. J. Ahn, K. I. Kim, G. Kim, E. Moon, S. S. Yang, and J.-S. Lee, "Atmospheric-pressure plasma jet induces apoptosis involving mitochondria via generation of free radicals," *PLoS One,* vol. 6, no. 11, p. e28154, 2011.

[8] S. Uchida *et al.*, "Plasma damage mechanisms for low-k porous SiOCH films due to radiation, radicals, and ions in the plasma etching process," *Journal of Applied Physics,* vol. 103, no. 7, p. 073303, 2008.

[9] E. Stoffels, Y. A. Gonzalvo, T. Whitmore, D. Seymour, and J. Rees, "Mass spectrometric detection of short-living radicals produced by a plasma needle," *Plasma Sources Science and Technology,* vol. 16, no. 3, p. 549, 2007.

[10] S. Yonemori and R. Ono, "Flux of OH and O radicals onto a surface by an atmospheric-pressure helium plasma jet measured by laser-induced fluorescence," *Journal of Physics D: Applied Physics,* vol. 47, no. 12, p. 125401, 2014.

[11] B. Locke, M. Sato, P. Sunka, M. Hoffmann, and J.-S. Chang, "Electrohydraulic discharge and nonthermal plasma for water treatment," *Industrial & engineering chemistry research,* vol. 45, no. 3, pp. 882-905, 2006.

[12] L. O. de Brito Benetoli, B. M. Cadorin, V. Z. Baldissarelli, R. Geremias, I. G. de Souza, and N. A. Debacher, "Pyrite-enhanced methylene blue degradation in non-thermal plasma water treatment reactor," *Journal of hazardous materials,* vol. 237, pp. 55-62, 2012.

[13] G. R. Stratton, C. L. Bellona, F. Dai, T. M. Holsen, and S. M. Thagard, "Plasma-based water treatment: conception and application of a new general principle for reactor design," *Chemical Engineering Journal,* vol. 273, pp. 543-550, 2015.

[14] B. Minkoff, J. Blatz, F. Choudhury, D. Benjamin, J. Shohet, and M. Sussman, "Plasma-Generated OH Radical Production for Analyzing Three- Dimensional Structure in Protein Therapeutics," *Sci Rep,* vol. 7, no. 1, pp. 12946-12946, 2017, doi: 10.1038/s41598-017-13371-7.

[15] M. R. Sussman, J. L. Shohet, F. A. Choudhury, J. M. Blatz, B. B. Minkoff, and D. I. Benjamin, "Methods, systems, and compositions for studying solvent accessibility and three-dimensional structure of biological molecules," Patent 10,571,460, February 25, 2020.

[16] J. C. Barreto, G. S. Smith, N. H. Strobel, P. A. McQuillin, and T. A. Miller, "Terephthalic acid: a dosimeter for the detection of hydroxyl radicals in vitro," *Life sciences,* vol. 56, no. 4, pp. PL89-PL96, 1994.

[17] Y. Manevich, K. D. Held, and J. E. Biaglow, "Coumarin-3-carboxylic acid as a detector for hydroxyl radicals generated chemically and by gamma radiation," *Radiation research,* vol. 148, no. 6, pp. 580-591, 1997.

[18] G. Xu and M. R. Chance, "Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting," *Analytical chemistry,* vol. 77, no. 14, pp. 4549-4555, 2005.

[19] G. Kim, S. J. Weiss, and R. L. Levine, "Methionine oxidation and reduction in proteins," *Biochimica et Biophysica Acta (BBA)-General Subjects,* vol. 1840, no. 2, pp. 901-905, 2014.

[20] X. Xu, "Dielectric barrier discharge—properties and applications," *Thin solid films,* vol. 390, no. 1-2, pp. 237-242, 2001.

[21] F. Massines, N. Gherardi, N. Naude, and P. Segur, "Glow and Townsend dielectric barrier discharge in various atmosphere," *Plasma physics and controlled fusion,* vol. 47, no. 12B, p. B577, 2005.

[22] U. Kogelschatz, B. Eliasson, and W. Egli, "From ozone generators to flat television screens: history and future potential of dielectric-barrier discharges," *Pure and Applied Chemistry,* vol. 71, no. 10, pp. 1819-1828, 1999.

[23] S. Eletr and G. Inesi, "Phase changes in the lipid moieties of sarcoplasmic reticulum membranes induced by temperature and protein conformation changes," *Biochimica et Biophysica Acta (BBA)-Biomembranes,* vol. 290, pp. 178-185, 1972.

[24] H. M. Levy, N. Sharon, E. M. Ryan, and D. Koshland Jr, "Effect of temperature on the rate of hydrolysis of adenosine triphosphate and inosine triphosphate by myosin with and without modifiers. Evidence for a change in protein conformation," *Biochimica et Biophysica Acta,* vol. 56, pp. 118-126, 1962.

[25] O. V. Koroleva, E. V. Stepanova, V. I. Binukov, V. P. Timofeev, and W. Pfeil, "Temperature-induced changes in copper centers and protein conformation of two fungal laccases from Coriolus hirsutus and Coriolus zonatus," *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology,* vol. 1547, no. 2, pp. 397-407, 2001.

[26] M. J. Campbell, R. P. Walter, R. McLoughlin, and C. J. Knowles, "Effect of temperature on protein conformation and activity during ultrafiltration," *Journal of membrane science,* vol. 78, no. 1-2, pp. 35-43, 1993.

[27] F. J. DiSalvo, "Thermoelectric Cooling and Power Generation," *Science,* vol. 285, no. 5428, pp. 703-706, 1999, doi: 10.1126/science.285.5428.703.

[28] D. Zhao and G. Tan, "A review of thermoelectric cooling: Materials, modeling and applications," *Applied Thermal Engineering,* vol. 66, no. 1, pp. 15-24, 2014/05/01/ 2014, doi: https://doi.org/10.1016/j.applthermaleng.2014.01.074.

[29] A. Pankaj *et al.*, "Generation mechanism of hydroxyl radical species and its lifetime prediction during the plasma- initiated ultraviolet (UV) photolysis," *Scientific Reports,* vol. 5, no. 1, 2015, doi: 10.1038/srep09332.

[30] Z. Bu and D. J. E. Callaway, "Chapter 5 - Proteins MOVE! Protein dynamics and long-range allostery in cell signaling," in *Advances in Protein Chemistry and Structural Biology*, vol. 83, R. Donev Ed.: Academic Press, 2011, pp. 163-221.

[31] Y. Hong *et al.*, "Measurement of hydroxyl radical density generated from the atmospheric pressure bioplasma jet," *Journal of Instrumentation,* vol. 7, no. 03, p. C03046, 2012.

[32] M. Sun *et al.*, "Emission spectra of hydroxyl radical generated in air corona discharge," *Guang pu xue yu Guang pu fen xi= Guang pu,* vol. 25, no. 1, pp. 108-112, 2005.

[33] H.-P. Dorn, R. Neuroth, and A. Hofzumahaus, "Investigation of OH absorption cross sections of rotational transitions in the A 2∑+,>(υ'= 0>)← X 2 Π,>(υ ″= 0>) band under atmospheric conditions: Implications for tropospheric long-path absorption measurements," *JGR,* vol. 100, no. D4, pp. 7397-7409, 1995.

[34] G. Dieke and H. Crosswhite, "The ultraviolet bands of OH fundamental data," *Journal of Quantitative Spectroscopy and Radiative Transfer,* vol. 2, no. 2, pp. 97-199, 1962.

[35] Y. H. Kim *et al.*, "Measurement of reactive hydroxyl radical species inside the biosolutions during non-thermal atmospheric pressure plasma jet bombardment onto the solution," *Plasma Chemistry and Plasma Processing,* vol. 34, no. 3, pp. 457-472, 2014.

Chapter 4: Epitope Mapping with PLIMB

Hydroxyl-radical footprinting has been previously shown to be an effective tool for epitope mapping. [1-3] Although different methods for producing hydroxyl radicals are used, the experimental process is generally similar. It involves preparing two sets of samples: one containing the antigen alone, and one containing the antibody bound to the antigen. All variables besides the presence or absence of the antibody and the plasma exposure time are kept constant between these sets of samples when comparing the degrees of oxidation in order to isolate the effect of antibody binding.[40]

The two sets of samples are then exposed to hydroxyl radicals, and then prepared for mass spectrometry analysis. This preparation is generally done as a bottom-up proteomics approach, [4] whereby the protein is digested with a protease, and the resulting peptides are analyzed with mass spectrometry. The extent of oxidation can then be calculated on a peptide level for each set of samples. Data analysis can then be performed to identify regions of the antigen which undergo protection, or a decrease in solvent accessibility as indicated by a decrease in oxidation, when bound to the antibody. These identified regions can then be considered as part of the epitope.

This Chapter outlines the process by which the binding location of an antibody to its target antigen is mapped using PLIMB-induced modification, followed by mass-spectrometry analysis. The process includes preparing samples of the antigen and antibody-bound antigen, exposing the samples to PLIMB to induce modification, preparing the samples for mass spectrometry analysis by digestion and cleanup, analyzing the digested samples via liquid-chromatography mass

[40] Note that a complete multi-factorial analysis should be performed so that all other variables such as the plasma parameters, sample concentration, buffer conditions, etc. in order to find the optimal combination for the epitope-mapping samples.

spectrometry (LC/MS) for peptide identification, followed by data analysis for quantifying the degree of modification and identifying the epitope regions, and/or locating the regions of the antigen involved in binding.

The target protein that was used for this experiment is the blood-clotting agent thrombin (human), with a murine (mouse) monoclonal-antibody (mAb) binder (Haematologic Technologies), referred to here as "antithrombin." This antibody/antigen pair was chosen in order to compare these results with previously published results generated by other protein-footprinting techniques, namely HDX and hydroxyl-radical footprinting using FPOP. [1, 5]

4.1 Experiment

The following experiment was conducted by Faraz Choudhury, Benjamin Minkoff, and me.

4.1.1 Sample preparation

For the epitope-mapping experiment, thrombin and anti-thrombin mAb at stock concentrations (8.9 mg/ml thrombin and 7.8 mg/ml antithrombin) were mixed together at a 1:1 molar ratio and incubated for 1 hour at room temperature. After incubation, phosphate-buffered solution (PBS) (0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7.2) was added to both the thrombin and antibody-bound thrombin before PLIMB treatment. The final concentration of thrombin in all samples was 1.1 μM in 50 μL.

4.1.2 PLIMB Exposure

The 50 μL samples of thrombin alone and thrombin complex (thrombin and antibody) were placed in 0.25 ml PCR tubes and were exposed to PLIMB (see Section 3.3 for plasma parameters) for two different times: 1 and 3 seconds. The two plasma exposure times will be used to analyze the extent of thrombin oxidation based on treatment with two doses of hydroxyl radicals. The temperature of

the sample holder cooling block was set to 2° C during the plasma exposure. The experiment was performed with two replicates, yielding a total of eight samples: 2 sample conditions (±antibody) × 2 plasma exposure times × 2 replicates. Immediately following plasma exposure, 50 µL of 25mM methionine in PBS was added to each sample in order to quench any longer-lived reactive species that may be present in solution after exposure to plasma, [6] which would result in an unwanted increase in background modification. Methionine quenching after hydroxyl labelling has been shown to be effective for reducing background modification for other protein-footprinting techniques. [6-9] After PLIMB exposure and quenching, samples were immediately prepared for mass spectrometry analysis as described below.

4.1.3 Post-PLIMB sample handling

Sample handling in this experiment was performed by Faraz Choudhury and me following a procedure developed by both of us with guidance and instruction from Benjamin Minkoff and Gregory Barret-Wilt of UW – Madison. Following PLIMB treatment described aboveo, the samples were precipitated by adding 12 µL of trichloroacetic acid (TCA) and incubated on ice for 30 minutes. The protein was then pelleted by centrifugation at 16,000×g, 4 °C for 10 minutes. The supernatant[41] was removed, and the protein was washed by adding 100 µL of ice-cold acetone, vortexing, and incubating on ice for 10 minutes. The samples were then re-pelleted by centrifugation at 4 °C and 16,000×g for 10 minutes. The supernatant was removed, and the pellet was resolubilized into 8M urea in 50 mM ammonium bicarbonate and then diluted to 4 M urea with 50 mM ammonium bicarbonate. After resolubilization, the proteins' disulfide bonds[42] were reduced by adding 5 mM dithiothreitol (DTT) and incubating at 50 °C 30 minutes. Disulfide bonds

[41] The supernatant here refers to the liquid lying above a solid residue after centrifugation.
[42] Disulfide bonds between pairs the side chains of cystines found in a protein's amino-acid sequence. Reducing and alkylating disulfide bonds allows a more complete denaturation prior to digestion.

were then alkylated by adding 15 mM iodoacetamide (IAA) and incubating at room temperature in the dark for 30 minutes. Samples were then diluted to 1M urea with 50 mM ammonium bicarbonate, and a mixture of trypsin and lys-C proteases were added at a protease/protein mass ratio of 1:20. The samples were digested overnight for 16 hours at 37° C. Following digestion, samples were acidified with neat formic acid to a final concentration of 1%. The samples were then desalted and concentrated using Omix C18 pipette tips (Agilent) using the following protocol: The tips were rinsed three times with 100 µL of neat acetonitrile, and then equilibrated by rinsing three times with 100 µL of 0.1% formic acid. The sample was then bonded to the tip by slowly running the sample through the tip 15 times, and then the tip was rinsed three times with 0.1% formic acid. The sample was then eluted with 100 µL of elution buffer (70% acetonitrile, 1% formic acid) into a fresh tube. Following elution, the samples were dried down via a Speed Vac vacuum concentrator. Dried-down peptides were resuspended into 0.1% formic acid at a concentration of 200 µg/ml, and 2 µL of peptides were injected into the LC/MS system for spectral analysis (details below).

4.1.4 Mass-spectrometry analysis

Samples were analyzed by Benjamin Minkoff using an Orbitrap Fusion Lumos mass spectrometer using the following protocol. A 20-minute chromatographic gradient from 2 to 30% acetonitrile with 0.1% formic acid was used for separation over a 3 µM, 50 cm Acclaim PepMap C18 column from ThermoFisher Scientific. Data-dependent acquisition was performed with a cycle time of 1 second using MS1 parameters of 240K resolving power in the Orbitrap, a scan range of 350-2000 m/z, a GC target of 10^6, and MS2 parameters of charge state 2-7 selection, a quadrupole isolation window of 0.7 Da, HCD collision energy of 28%, an AGC value of 3×10^5, and an automatic scan

range starting at 110 m/z in the ion trap with the scan rate set to "turbo." Dynamic exclusion of 10 seconds was used with a duration exclusion of 10 seconds after seeing an ion once.

4.2 Data analysis

While there are many different computer programs that can be used for peptide identification from mass spectrometry data, a company called Protein Metrics has developed an integrated workflow for their Byos® mass-spectrometry search software specifically designed for oxidative-footprinting experiments. It has many features that simplify data the analysis for footprinting analysis much easier, including a function for automatically calculating the percent oxidation of a given peptide. For these reasons, Protein Metrics software was used for mass-spectrometry searching and quantifying oxidation. Data analysis for this experiment was performed by Faraz Choudhury and me, with the help and guidance from Benjamin Minkoff of UW – Madison and St. John Skilton, Ilker Sen, and Claire Bramwell from Protein Metrics. The mass-spectrometry data was analyzed using Protein-Metrics software. The '.raw' data files were submitted to Byos® (Protein Metrics) for a database search using the sequence of Thrombin, and automated generation of extracted-ion chromatograms ('XIC') of the precursor ions. The peptides were identified using MS and MS/MS criteria, [10] and the following modifications, noted by their molecular mass shift in Daltons (Da) and the amino acids which are modified noted by their single-letter abbreviation, were used in our search. Here, standard modifications are those that occur naturally, and are searched for by the software, but are not quantified as a footprinting modification during data analysis. PLIMB modifications are those that can be induced by plasma exposure and are quantified during data analysis.

Standard modifications:

Carbamidomethyl: +57.021464 Da @ C (fixed)

Gln->pyro-Glu: -17.026549 Da @ NTerm Q

Glu->pyro-Glu: -18.010565 Da @ NTerm E

Phospho: +79.966331 Da @ S, T

PLIMB modifications:

Oxidation: +15.994915 Da @ G, C, D, F, H, K, M, N, P, R, W, Y

Dioxidation: +31.989829 Da @ C, F, K, M, P, R, W, Y

Cys-Oxidation: +15.994915-57.021464 Da @ C

Nitro: +44.985078 Da @ W, Y

The peptides were identified through automated analysis of the Byos software using MS and MS/MS analysis using a 1% false discovery rate (FDR), and the proportion of plasma-modified species (marked as PLIMB modifications above) were calculated based on extracted-ion chromatogram (XIC) relative peak areas of modified versus unmodified peptides. The sequence coverage of thrombin was about 95% as shown in Figure 4-1. The percent modification of each thrombin peptide was compared with samples containing thrombin alone and with samples containing thrombin bound to anti-thrombin mAb.

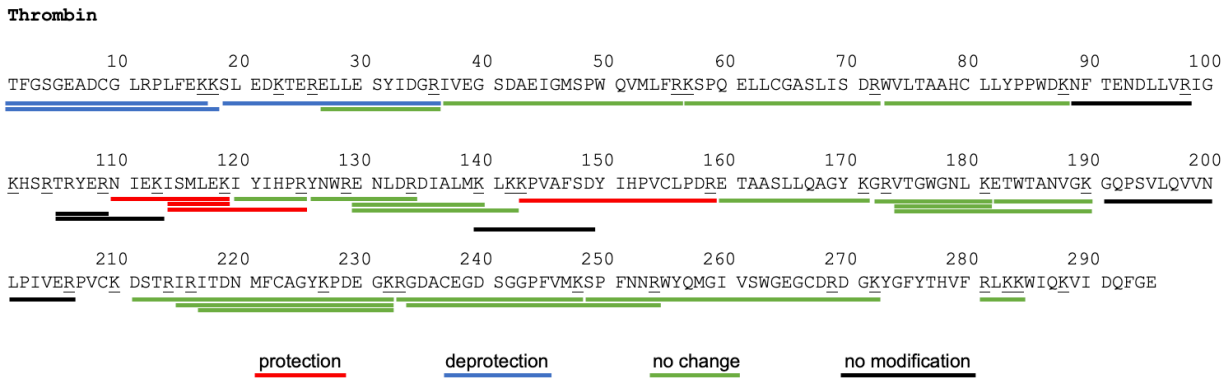


Figure 4-1: Coverage map of the tryptic (trypsin-produced) digest of thrombin exposed to 3 seconds of PLIMB

This figure shows: 1) peptides that were not labeled (black lines), 2) no change between antibody-bound thrombin and thrombin alone (green lines), 3) protection in the antibody-bound form (red lines, modified residue in bold), and 4) de-protection in the antibody-bound form (blue lines).

4.3 Results

A t-test[43] with $p \leq 0.05$ was performed to identify statistically significant differences in modification. Six peptides of thrombin that were identified underwent a statistically significant change in PLIMB-induced modification when comparing 1) samples of thrombin alone and 2) thrombin bound to the antithrombin mAb. These six peptides are shown in Figure 4-2. In addition to observing peptides that experienced a statistically significant change in percent modification in the antibody bound thrombin samples, many peptides did not exhibit a significant change in modification, as shown in Figure 4-3. These peptides not only indicate regions that are not affected by the binding interactions, but also suggest that the overall oxidative potential of the bound and unbound samples is equivalent. This is because, if the overall oxidative potential between the two sets of samples had changed, we would expect to see a corresponding increase or decrease in percent oxidation across many of the peptides in the protein between the sets of samples.

[43] A t-test is a statistical hypothesis test in which the test statistic follows the Student's t-distribution [11] under the null hypothesis. The p-value is the probability that the results of the sample data occur by chance.

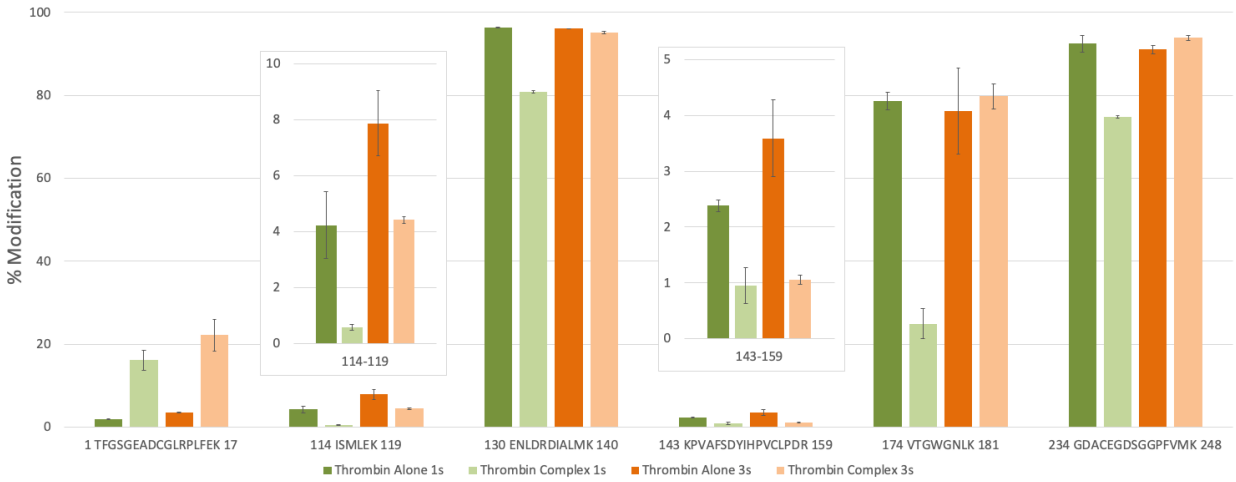


Figure 4-3: Oxidation of Thrombin and Thrombin Complex (antibody bound thrombin) for peptides that were modified by 1 and 3 seconds of PLIMB treatment.

Here, the peptides of thrombin which showed a statistically significant change in oxidation upon thrombin binding are shown. The error bars represent the ± standard deviation in percent modification between replicate samples. The inlaid charts show the oxidation of peptides 114-119 and 143-159.

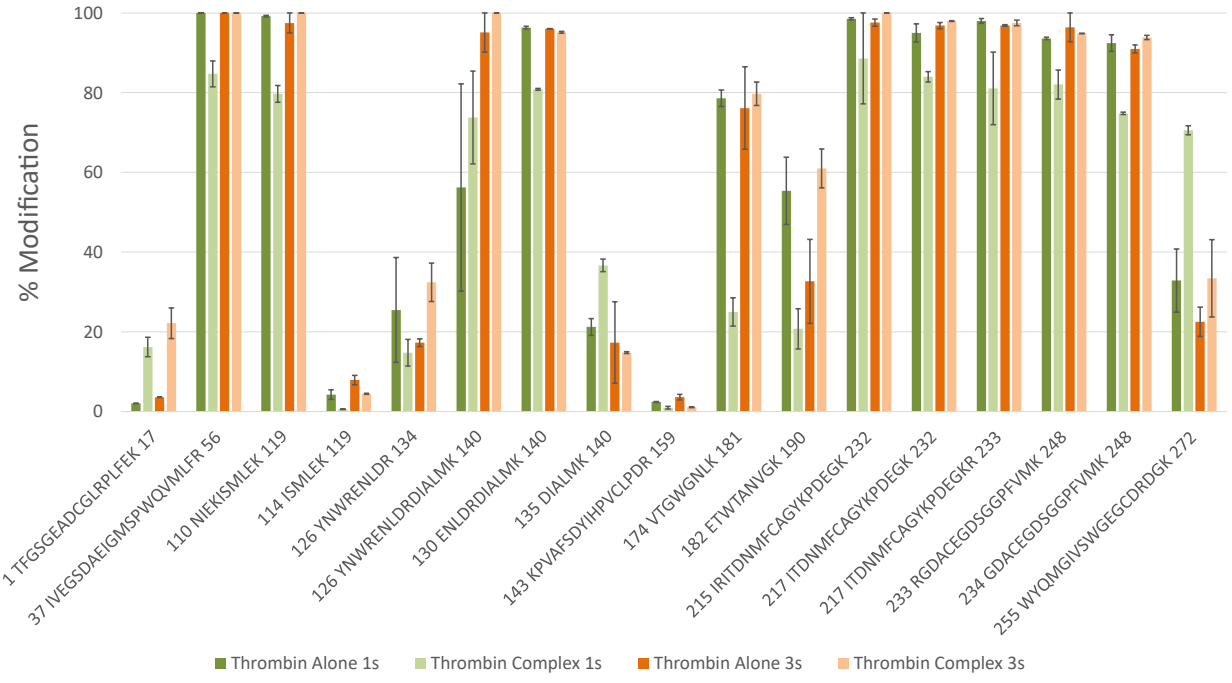


Figure 4-2: Full sequence oxidation data for thrombin and thrombin complex

Here, all peptides of thrombin that showed significant levels of modification after 1 or 3 seconds of PLIMB exposure are shown.

4.4 Discussion

Out of the six peptides that showed statistically significant changes, three peptides (peptides 130-140, 174-181 and 234-248) showed decreases in modification upon binding after a one-second exposure, but interestingly, an insignificant change in modification after a three-second exposure. Two peptides (peptides 114-119 and 143-159) experienced decreased levels of modification for both 1- and 3-second PLIMB exposure times. These results suggest a stronger level of protection upon binding compared to peptides that showed only a decrease in oxidation for 1-second exposure. The regions which show the strongest levels of protection are likely part of the epitope, as opposed to regions that experience protection due to a conformational change, which is likely a weaker interaction than the epitope interaction. These regions are also spatially clustered when mapped to the thrombin crystal structure [12] as shown in Figure 4-4. Because of the protection seen in both the 1 and 3 second PLIMB exposure data, the data suggests that peptides 114-119 and 143-159 comprise the epitope region of thrombin and form a discontinuous, conformational epitope.

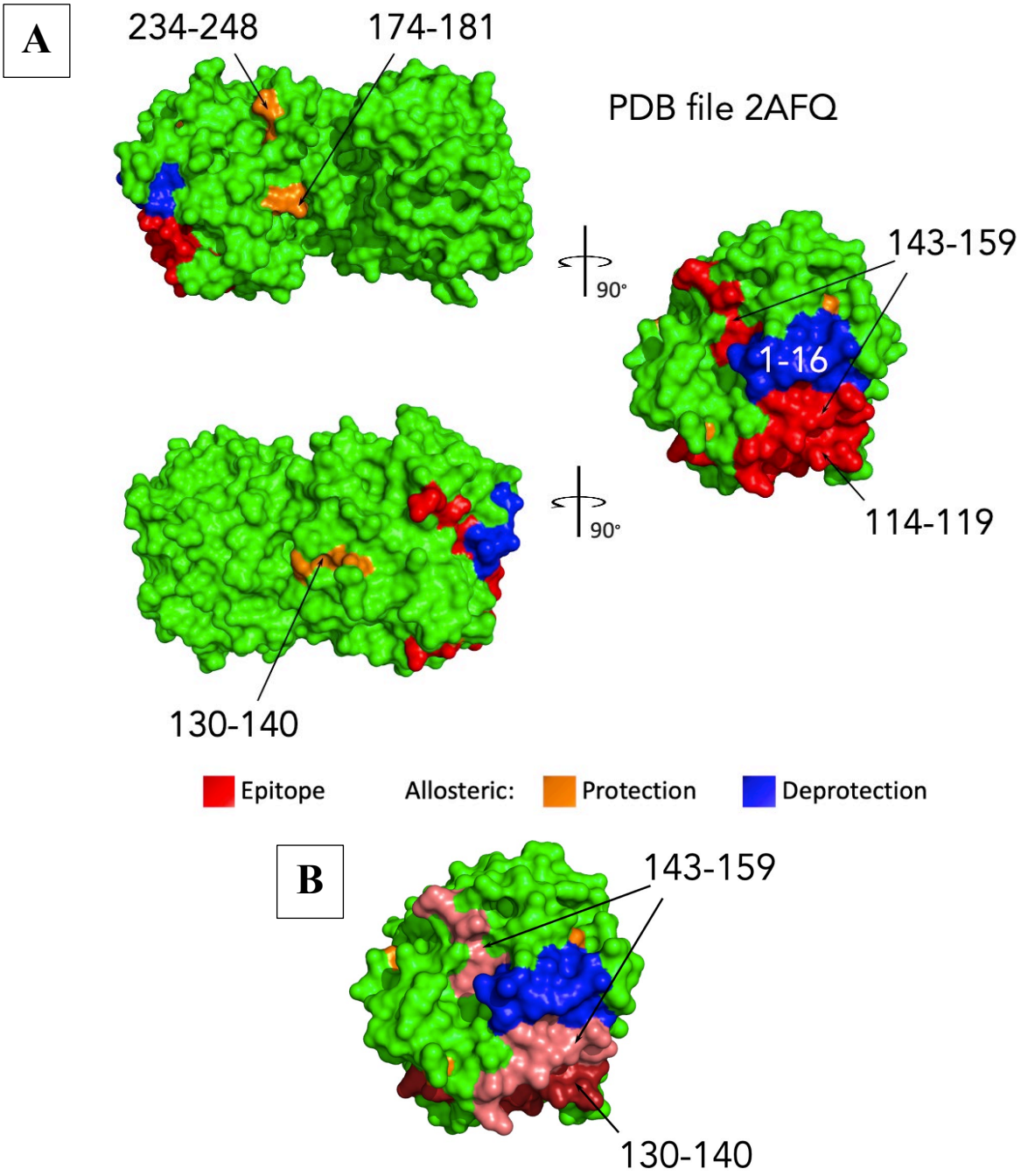


Figure 4-4: PLIMB epitope results mapped to thrombin crystal structure

Figure 4-4A shows PLIMB data mapped to a three-dimensional crystal structure (Protein Data Bank file 2AQF [12]) of a thrombin dimer. Note that the sequence numbers shown here do not exactly align with those of the crystal structure. The various regions of thrombin were colored as follows: 1) no change in modification upon binding shown in **green**, 2) protection due to the epitope shown in **red**, 3) projection due to allosteric conformational changes upon binding, shown in **orange** and 4) deprotection upon binding shown in **blue**. Each of the three images of the thrombin structure is rotated 90° from the image above it.

Figure 4-4B shows the two peptides that make up the epitope in different shades of red. Peptide 143-159 is shown in **light red**, and peptide 130-140 is shown in **dark red**.

Comparing these results to those that were generated using HDX and FPOP, we see a notable amount of consistency, as can be seen in Figure 4-5. Komives *et al*. saw significant protection upon antibody binding as measured by HDX in two peptides: 97-117 and 139-149. [5] After seeing protection in these two peptides, a competition assay was performed to test the effect each of these peptides of thrombin had on antithrombin mAb binding. Peptides 97-117 and 139-149 were each synthesized, and then introduced to the solution in increasing concentrations while measuring the binding of antithrombin mAb to thrombin. They found that peptide 97-117 did not affect antithrombin binding even at 10,000-fold molar excess to antithrombin, while peptide 139-149 competed for binding to antithrombin at 2000-fold molar excess. Based on these results, they concluded that peptide 139-149 comprises the primary epitope of thrombin. PLIMB results also show protection in regions close to both of these peptides, namely in peptides 114-119 and 143-159. Jones *et al*. show protection over peptide 114-119 and several peptides over amino acids 129-171. Note that Komives *et al.* experiments were conducted using the protease pepsin, as is typically done with HDX experiments, while both the FPOP experiment by Jones *et al.* and the PLIMB experiment shown here were conducted using trypsin. Therefore, the peptides for both the PLIMB and FPOP experiments have an n-terminal K or R, while the peptides for the HDX experiment do not.

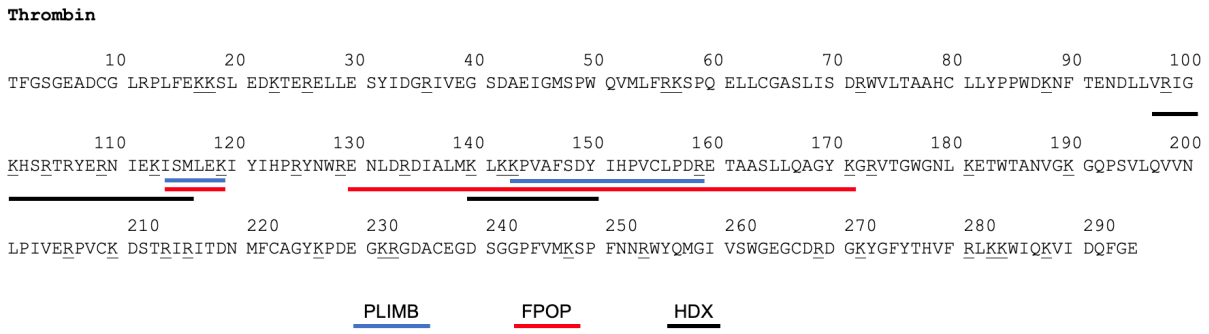


Figure 4-5: Comparison of thrombin epitope data generated with PLIMB, FPOP, and HDX

PLIMB epitope data, as detected by statistically significant decreases in oxidation after 3 seconds of PLIMB exposure, is indicated in **Blue** in the above sequence map of thrombin. Similarly, epitope results gerated from FPOP are shown in **Red** [1] and from HDX are shown in **Black**. [5]

One of the challenges of epitope mapping with footprinting techniques has been differentiating allosteric effects*, i.e.,* conformational changes to antigen protein upon binding, with epitope regions. This is because footprinting methods look for protection, or decreases in solvent accessibility, to locate epitope regions. Allosteric effects can also show up in data analysis, as they will also appear as changes in solvent accessibility. When two or more distinct regions of a protein show protection from an epitope-mapping experiment, there often isn't a straightforward method for differentiating allosteric effects from epitope regions. One strategy for differentiating epitopes from allosteric effects is by utilizing one or more additional experimental techniques such as HDX, site-specific mutagenesis[44], or chemical cross-linking mass spectrometry.[45] [13] While using multiple experimental techniques, although time-consuming, can certainly provide a more detailed

[44] Site-specific mutagenesis is a method that involves making mutations, or changes, to a protein's primary sequence in order to study the effect that individual amino acids have on a protein's function.

[45] Chemical cross-linking mass spectrometry is a method that involves creating a chemical bond between two chains of proteins, in which the chemical bond is highly dependent on the physical distance between regions of proteins and analyzing the regions that become bonded using mass spectrometry. Cross-linking data can be informative for epitope mapping, as it can be used to identify regions of a protein which are spatially close to another protein, as is the case with an antibody bound to its target. The regions of these proteins will become bound chemically during a crosslinking experiment, and the resulting bound regions can be identified with mass spectrometry.

and thorough analysis, and have been shown to be effective for differentiating epitopes from allosteric effects, a method for differentiating using HRPF alone has not yet been established.

Three peptides, 130-140, 174-181 and 234-248, all show protection after 1 second of PLIMB exposure, indicating decreased solvent accessibility, but not at 3 seconds. Interestingly, Figure 4-4 shows that these peptides are located in the crevices of the protein and are not located near the proposed epitope region. It has been previously reported by Komives that the structure of thrombin tightens upon binding with thrombomodulin, a protein that acts as an anticoagulant when bound to thrombin. This was detected with HDX showing a decrease in deuterium uptake in several regions of thrombin in addition to the binding site. [14-16] Specifically, Komives et al. report decreased solvent accessibility of residues 167-180, a region where a decreased level of modification was seen in our PLIMB results (residues 174-181). [14] Therefore, based on solvent accessibility measurements of thrombin using PLIMB and previously reported HDX data, it is hypothesized that the structure of thrombin tightens up upon binding to the mAB, resulting in a decrease in solvent accessibility in regions located in the crevices of the protein. The decrease in solvent accessibility associated with these regions (peptides 130-140, 174-181 and 234-248) provides a level of protection which blocks plasma-induced modification for a 1-second exposure time yet was insufficient at blocking modification from the larger radical dose generated by the longer 3-second exposure. While more experiments will need to be performed to either validate or invalidate this technique, using the response of two or more PLIMB doses may prove to be an effective method for differentiating epitopes from allosteric effects using an HRPF technique alone.

In addition to the peptides that showed protection at 1 second of PLIMB exposure, an increase in oxidation on peptide 1-16 was observed, which is located on the surface of the protein

next to the proposed epitope region (see Figure 4-4). This may be attributed to allosteric conformational changes near the epitope region which unfold and expose the amino-acid residues in that region, thereby increasing their solvent accessibility.

Gross *et. al.* also reported allosteric conformational changes in thrombin due to antibody-binding measured with FPOP. However, all of the allosteric effects observed by FPOP in thrombin showed deprotection, *i.e*. an increase in solvent accessibility. [1] This may be an indication of unfolding of the protein caused by the addition of hydrogen peroxide, which is the precursor chemical of hydroxyl radicals used in the FPOP system. Another possible reason for this observed difference may be due to the hydroxyl-radical generation rate and total radical dose produced by the two systems. FPOP produces a burst of many radicals while PLIMB generates a series of relatively smaller sub-microsecond bursts of radicals, and the total dose can be controlled by the total plasma exposure time.

4.5 Conclusions

Here we have shown PLIMB's utility for mapping the discontinuous epitope of a monoclonal antibody binder to thrombin and characterized the allosteric conformational changes in the protein upon binding. The epitope data shown here agrees with the epitope region previously determined using other well-established techniques. In the context of other epitope mapping techniques, the use of PLIMB for epitope mapping is significant because it can be used to map conformational epitopes relatively quickly and does not require: 1) large amounts of ultrapure proteins or 2) the use of any chemical additives, which can potentially disrupt the native structure or binding.

4.6 References

[1] L. M. Jones, J. B Sperry, J. A Carroll, and M. L. Gross, "Fast photochemical oxidation of proteins for epitope mapping," *Analytical chemistry,* vol. 83, no. 20, p. 7657, 2011, doi: 10.1021/ac2007366.

[2] Y. Yan, "Fast photochemical oxidation of proteins (FPOP) maps the epitope of EGFR binding to adnectin," *Journal of the American Society for Mass Spectrometry.,* vol. 25, no. 12, p. 2084, 2014, doi: 10.1007/s13361-014-0993-x.

[3] K. S. Li *et al.*, "Orthogonal Mass Spectrometry-Based Footprinting for Epitope Mapping and Structural Characterization: The IL-6 Receptor upon Binding of Protein Therapeutics," *Analytical Chemistry,* vol. 89, no. 14, pp. 7742-7749, 2017/07/18 2017, doi: 10.1021/acs.analchem.7b01748.

[4] R. L. Gundry *et al.*, "Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow," *Current Protocols in Molecular Biology,* vol. 90, no. 1, pp. 10.25.1-10.25.23, 2010, doi: https://doi.org/10.1002/0471142727.mb1025s88.

[5] A. Baerga-Ortiz, C. A. Hughes, J. G. Mandell, and E. A. Komives, "Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein," *Protein Science,* vol. 11, no. 6, pp. 1300-1308, 2002, doi: https://doi.org/10.1110/ps.4670102.

[6] J. Saladino, M. Liu, D. Live, and J. S. Sharp, "Aliphatic peptidyl hydroperoxides as a source of secondary oxidation in hydroxyl radical protein footprinting," *Journal of the American Society for Mass Spectrometry,* vol. 20, no. 6, pp. 1123-1126, 2009/06/01 2009, doi: 10.1021/jasms.8b03489.

[7] B. Zhang, M. Cheng, D. Rempel, and M. L. Gross, "Implementing fast photochemical oxidation of proteins (FPOP) as a footprinting approach to solve diverse problems in structural biology," *Methods,* vol. 144, pp. 94-103, 2018/07/15/ 2018, doi: https://doi.org/10.1016/j.ymeth.2018.05.016.

[8] G. Xu, J. Kiselar, Q. He, and M. R. Chance, "Secondary Reactions and Strategies To Improve Quantitative Protein Footprinting," *Analytical Chemistry,* vol. 77, no. 10, pp. 3029-3037, 2005/05/01 2005, doi: 10.1021/ac048282z.

[9] J. G. Kiselar, R. Mahaffy, T. D. Pollard, S. C. Almo, and M. R. Chance, "Visualizing Arp2/3 complex activation mediated by binding of ATP and WASp using structural mass spectrometry," *Proceedings of the National Academy of Sciences,* vol. 104, no. 5, pp. 1552-1557, 2007.

[10] B. Minkoff, J. Blatz, F. Choudhury, D. Benjamin, J. Shohet, and M. Sussman, "Plasma-Generated OH Radical Production for Analyzing Three- Dimensional Structure in Protein Therapeutics," *Sci Rep,* vol. 7, no. 1, pp. 12946-12946, 2017, doi: 10.1038/s41598-017-13371-7.

[11] W. contributors. "Student's t-distribution." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Student%27s_t-distribution&oldid=1011045143 (accessed March, 26, 2021).

[12] D. J. Johnson, T. E. Adams, W. Li, and J. A. Huntington, "Crystal structure of wild-type human thrombin in the Na+-free state," *Biochemical Journal,* vol. 392, no. 1, pp. 21-28, 2005.

[13] J. Li *et al.*, "Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis," *Analytical Chemistry,* vol. 89, no. 4, pp. 2250-2258, 2017/02/21 2017, doi: 10.1021/acs.analchem.6b03058.

[14] C. H. Croy, J. R. Koeppe, S. Bergqvist, and E. A. Komives, "Allosteric Changes in Solvent Accessibility Observed in Thrombin upon Active Site Occupation," *Biochemistry,* vol. 43, no. 18, pp. 5246-5255, 2004/05/01 2004, doi: 10.1021/bi0499718.

[15] J. R. Koeppe, A. Seitova, T. Mather, and E. A. Komives, "Thrombomodulin Tightens the Thrombin Active Site Loops To Promote Protein C Activation," *Biochemistry,* vol. 44, no. 45, pp. 14784-14791, 2005/11/01 2005, doi: 10.1021/bi0510577.

[16] J. G. Mandell, A. Baerga-Ortiz, S. Akashi, K. Takio, and E. A. Komives, "Solvent accessibility of the thrombin-thrombomodulin interface11Edited by M. F. Moody," *Journal of Molecular Biology,* vol. 306, no. 3, pp. 575-589, 2001/02/23/ 2001, doi: https://doi.org/10.1006/jmbi.2000.4416.

Chapter 5: Feedback-Control System

The aim of a hydroxyl radical protein-footprinting (HRPF) experiment, whether it be using PLIMB or another method for generating hydroxyl radicals, is to use levels of oxidation as a proxy for solvent accessibility and to infer structural characteristics from these results. However, using PLIMB exposure time as a measure of hydroxyl dose often does not yield sufficiently repeatable data. It is hypothesized that this is due in part to variability in radical generation.

As an example of data which shows significant variability, Figure 5-1 shows results from a PLIMB experiment conducted by Benjamin Minkoff of UW – Madison. Peptide-level oxidation of epidermal growth factor receptor (EGFR) is shown under two conditions: (1) EGFR alone, and (2) EGFR bound to epidermal growth factor EGF. The goal of this experiment is to detect changes in solvent accessibility of EGFR through detection of statistically significant changes in peptide-level oxidation when EGFR is bound to EGF. The error bars here are quite large relative to averaged data points, making it difficult to assign statistical significance to what may be true differences between experimental conditions. For this reason, protein-footprinting experiments often need to be repeated to obtain reliable data, and even then, as shown in Figure 5-1, data points that represent a noticeable trend may not be statistically significant. This can limit the conclusions that can be drawn from a dataset and thus its utility. Therefore, in order to increase the reliability of the PLIMB instrument for generating data for epitope mapping, the setup requires specific improvements to reduce variability in labeling.

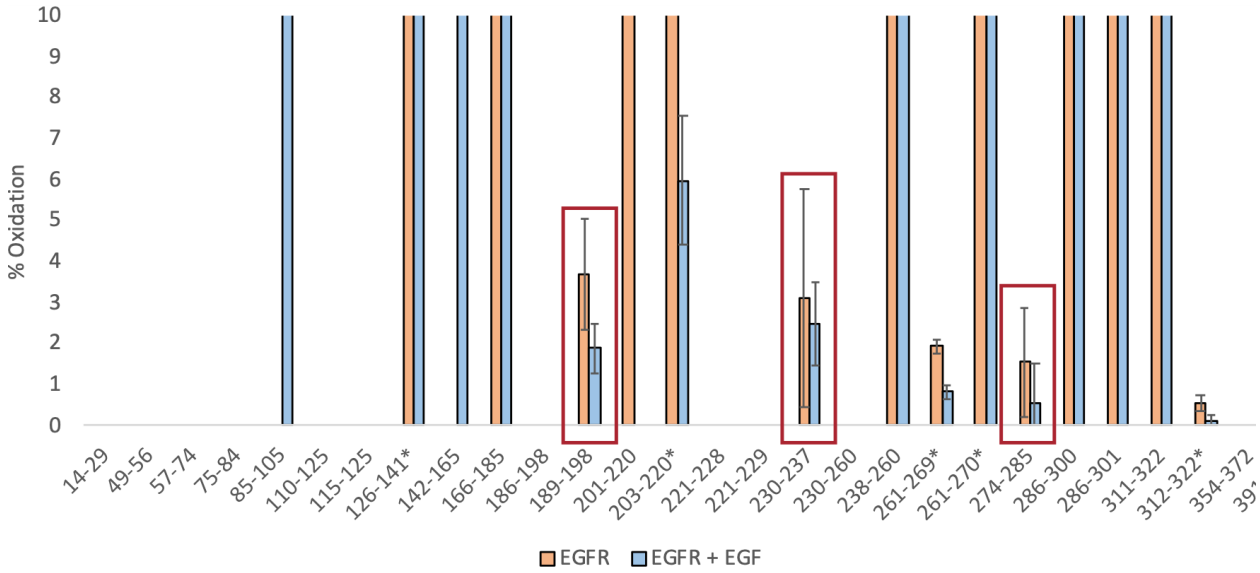


Figure 5-1: Example of Protein Footprinting Data

Protein footprinting data typically looks at relative changes in solvent accessibility under different conditions. This figure shows peptide-level oxidation, with variability shown as error bars, of the protein EGFR under two conditions: "EGFR" (alone) and "EGFR + EGF." High variability in oxidation often leads to ambiguous experimental results.

The goal of this chapter is to augment the PLIMB instrument with an additional system that will allow for more reproducibly modified replicate samples. The approach will be able to directly measure the rate of hydroxyl radicals being generated during exposure and use this measurement in a feedback-control system that will expose a sample to a specific radical dose, in order to reduce variability in oxidation of protein samples.

5.1 Design

The basic approach that was used for detecting hydroxyl-radical generation during plasma exposure, was to measure the intensity of 309-nm UV light that emanates from the plasma. 309-nm light is known to be generated specifically by hydroxyl radicals [1-6], and is therefore a direct measurement of their presence. The following sections will describe the design of the UV detection system, the feedback control system, and how each of these are integrated together.

5.1.1 Light detection

In order to detect and measure the amount light being generated by the plasma at a particular wavelength, a McPherson 234-302 monochromator with a 1200 grooves/mm diffraction grating blazed for 140 nm was used and was set up as described in Section 3.8. A diagram of the setup can be seen in Figure 5-2. A photomultiplier tube (Hamamatsu R928a) was chosen for the purposes of the feedback-control system, as it has several characteristics which are necessary for this application.

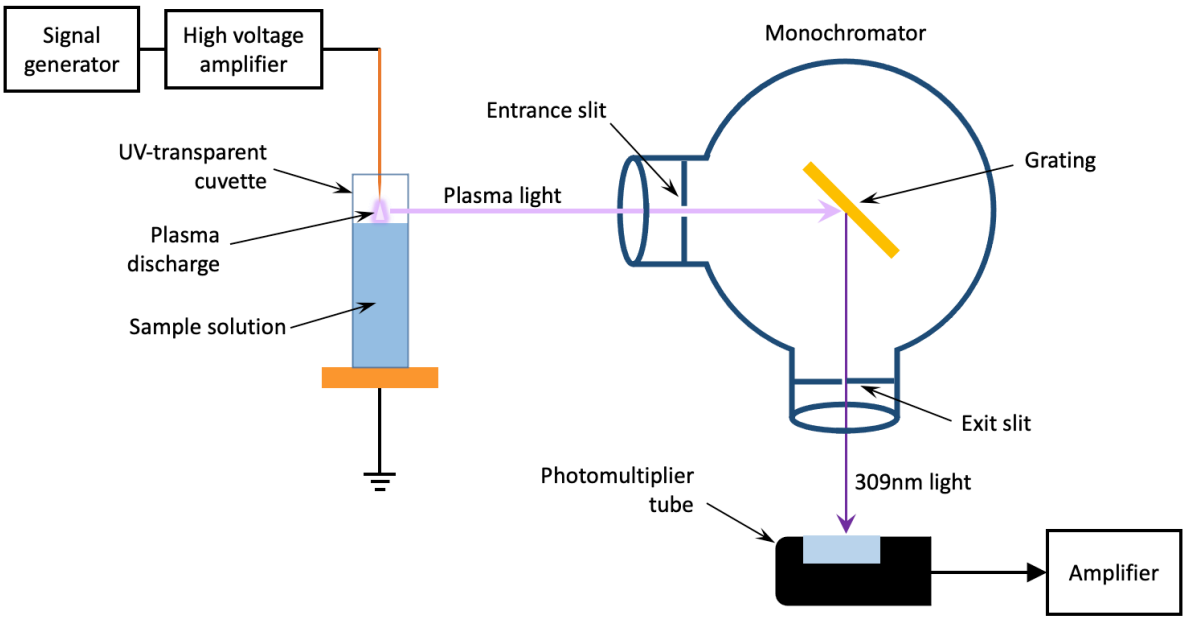


Figure 5-2: Single wavelength measurements with monochromator with photomultiplier tube

The primary characteristic of a photomultiplier tube that is important for this purpose, is its sensitivity. Sensitivity here refers to the ability to generate a large output signal given a very small amount of incident light. The Hamamatsu R928a has a luminous sensitivity of around 2500 amperes per lumen (A/lm). When fed into a transimpedance amplifier with a gain of 0.15V/µA, this yields an overall sensitivity of 375 MV/lm. Given that a suitable range of output voltages

should be 0.1–1V, this allows for the detection of approximately 10^{-9} lumens of incident light. Assuming there is about a 1% efficiency of the monochromator system (given the path length from the plasma to the detector and the efficiency of the detraction grating), this would require the total amount of light to be about 10^{-7} lumens. Given that the plasma is visible to the human eye when viewed at a distance of about 1 meter in a well-lit room, the amount of visible light given off by the plasma should certainly be equal to or greater than 10^{-7} lumens. Therefore, assuming the UV light given off by the plasma is of similar intensity or higher to that of the visible light, the sensitivity of the photomultiplier tube should be high enough for the experimental setup.

The other key feature of a photomultiplier tube which lends itself to this experimental setup is its response time. The response time refers to the amount of time it takes the photomultiplier to change its output signal in response to a change in incident light intensity. Because it was established that the plasma exhibits short, ~0.1-μs discharges, as seen by the current measurements described in Section 3.7, the timescale of the light-detector response should be on a similar or shorter timescale. The typical response time for the R928 is 2.2 ns, which is appropriate for measuring bursts of light on the timescale of the current pulse for each plasma discharge.

5.1.2 Analog-to-digital conversion

Once the experimental setup for single-wavelength light measurement had been established, it was then necessary to augment the system in order to measure the overall hydroxyl-radical dose over an entire plasma-exposure time, which, for the purposes of epitope mapping, is typically 1-10 seconds. The overall goal would be to integrate the 309-nm light over each 1-10 ns plasma discharge and keep a real-time running sum over the course of an entire 1-10 second exposure time. The resulting sum should then give the total amount of 309-nm light emitted over the entire exposure time and can be related to the hydroxyl-radical dose.

If we take R to be the total radical dose for a given exposure, and $r(t)$ to be the radical generation rate, then integrating $r(t)$ over the entire exposure time, from time 0 to the end of the exposure, T, would give us the entire radical dose. Note that $r(t)$ is zero when each discharge has stopped and before the next begins.

$$R = \int_0^T r(t)dt \tag{5-1}$$

If we then consider that the radical generation rate is proportional to the total amount of 309-nm light emitted by the plasma, Φ_{total} (see equation (5-2)), then one could simply integrate the amount of 309-nm light being generated by the plasma to generate a quantity that is proportional to the radical dose, R, (equation (5-3)).

$$r(t) \sim \Phi_{total} \tag{5-2}$$

$$R \sim \int_0^T \Phi_{total}(t)dt \tag{5-3}$$

If we also consider that the voltage measured at the output of the photomultiplier tube amplifier, V_{out}, is proportional to the photomultiplier-tube current, I_{PMT}, which itself is proportional to the intensity of 309-nm light incident on the PMT window, Φ_{PMT}. Although the light incident on the PMT is certainly less than the total the total amount of 309-nm light emitted Φ_{total} due to the fact that light emanates from the plasma in all directions and only a small amount is incident the monochromator and the insertion loss of the monochromator. Nonetheless, Φ_{PMT} will be proportional to the total amount of light emitted, Φ_{total} (equation (5-4). Therefore, by the transitive property, we can say that the radical dose should be proportional to the PMT amplifier voltage, V_{out}, integrated over the exposure time (equation (5-5)).

$$V_{out} \sim I_{PMT} \sim \Phi_{PMT} \sim \Phi_{total} \sim r \quad \textbf{(5-4)}$$

$$R \sim \int_0^T V_{out}(t)dt \quad \textbf{(5-5)}$$

The first approach to accomplish this integration was to use an analog-integrator circuit using operational amplifiers (op-amps), like the one shown in Figure 5-3. The input of the integrator circuit could be connected to the output of the PMT amplifier, and the output voltage would represent the integral of the input voltage from the time that the circuit was connected to an active input signal. This output voltage would represent a quantity proportional to the hydroxyl-radical dose and could be sampled as frequently as was required by a computer program that controlled the plasma.

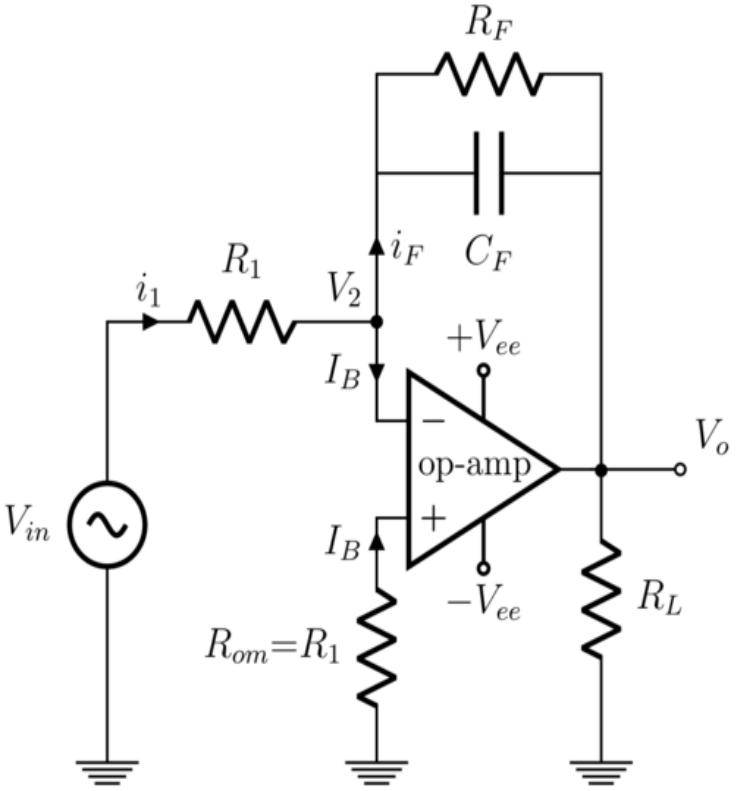


Figure 5-3: Op-amp analog integrator circuit [7]

An OPA657 was used, along with an R1 of 100Ω, and depending on the particular configuration that was tested, an R_F of 1.5 to 48kΩ and a C_F of 1 to 47pF.

For such a circuit to accurately work for this purpose, the bandwidth would have to be large enough to respond to rapid changes in PMT output voltage. As can be seen in the following oscilloscope capture (Figure 5-4), a single burst of 309nm light (channel 1, yellow trace) that follows a current pulse (channel 2, green trace), rises from 0 to ~5V in about 200ns. If we take the signal frequency to be about equal to the inverse of the rise time, then the signal frequency would be equal to 5 MHz and would require an amplifier with a bandwidth of at least that large to track the rapid changes in voltage.

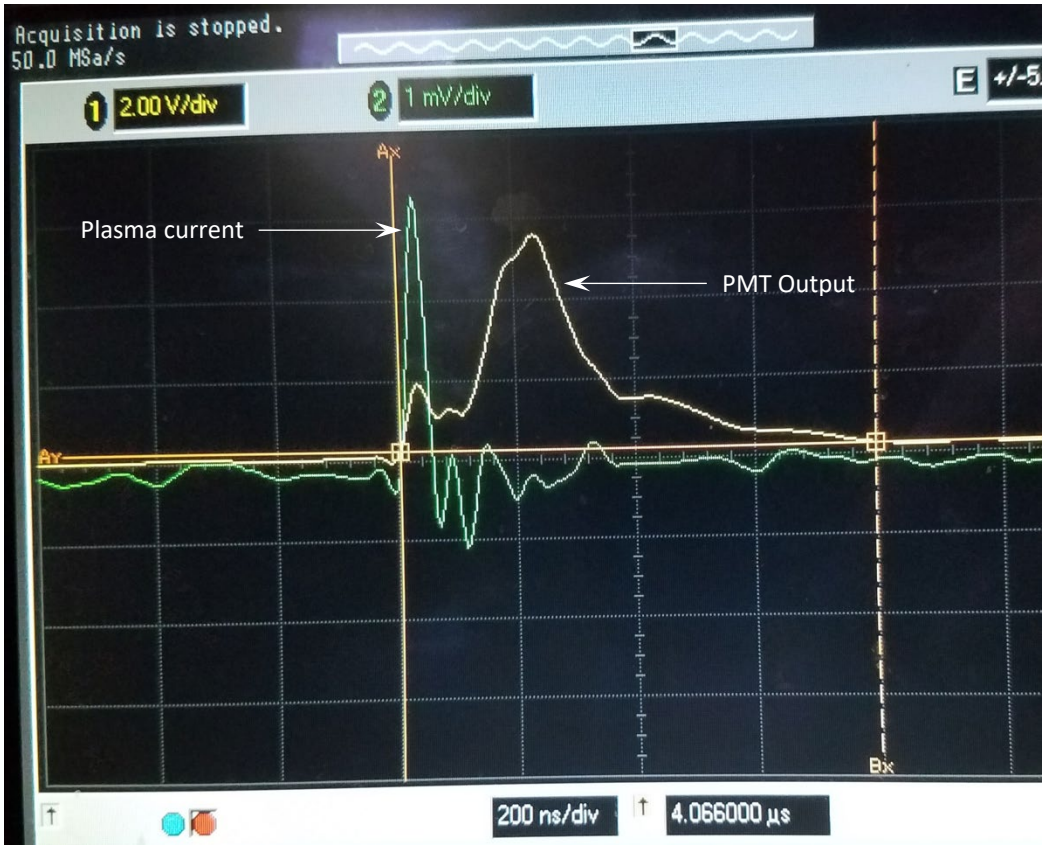


Figure 5-4: Plasma current and 309nm light measurement

While it should in principle be possible to create such a circuit, it was found to be quite challenging in practice. A modified form of the circuit shown in Figure 5-3 was assembled using an OPA657 op-amp, which has a bandwidth of 350 MHz. The resulting circuit had an overall

bandwidth of 10 MHz, which was sufficient for tracking the rapid changes in voltage produced by the PMT amplifier circuit. After testing the circuit with simulated AC input signals, it was found that the circuit was quite sensitive to electrical noise. It was concluded that the only way to significantly reduce all of the noise present in the circuit would be to design a fully integrated printed-circuit board (PCB) with on-board power supplies.

Following the realization that an analog integrator circuit would be quite challenging to implement for this project, an alternative approach was taken. This approach was to use real-time analog to digital conversion (ADC), generate digital samples of the output voltage, $S(V_{out})$, and sum the resulting digital signals to record a real-time measure of the summed PMT output voltage. The resulting quantity, shown in Equation (5-6), should then also be proportional to the radical dose, given a small enough sampling period, *dt*, and digital resolution to sufficiently approximate the integral shown in Equation (5-5).

$$R \sim \sum_0^T S(V_{out})dt \tag{5-6}$$

If the sampling period, *dt*, is held constant, then the radical dose should also be proportional to the sum of each output voltage sample, $S(V_{out})$.

$$R \sim \sum_0^T S(V_{out}) \tag{5-7}$$

The sampling rate for this experimental setup would need to be ~ $N \times f$, where the frequency, $f = 1/T$. T is the period of a single pulse of 309-nm light given off by a plasma discharge, and N is a constant for the number of samples per pulse that should be taken. As shown in Figure 5-4, a single pulse of 309-nm light is about 200 ns in duration. Assuming N should be about 100 samples/pulse to accurately measure waveshape and intensity of each pulse, the required sampling

rate for this system should be at least 500 million samples per second (MS/s). The ADC that was chosen to sample the analog voltage signal from the photomultiplier tube amplifier circuit was an Analog Devices model AD9680, which has a maximum sampling rate of 2.8 GS/s and a 14-bit ADC resolution.

5.1.3 Real-time signal summation

Next, an additional piece of equipment is needed to perform the real-time summation of the digital samples generated by the ADC. The block diagram shown in Figure 5-5 describes what needs to be accomplished by the additional electronic equipment. A running sum of all sampled data would need to be kept, which would represent the hydroxyl radical dose over the plasma exposure time, as described in Equation (5-7). A new 14-bit sample would need to be taken in and added to the running sum, which is kept as a number stored in a register, at a speed of $\geq 5\times10^{8}$ s^{-1} in order to keep up with the sequential stream of samples that would be outputted by the ADC at 500 MS/s.

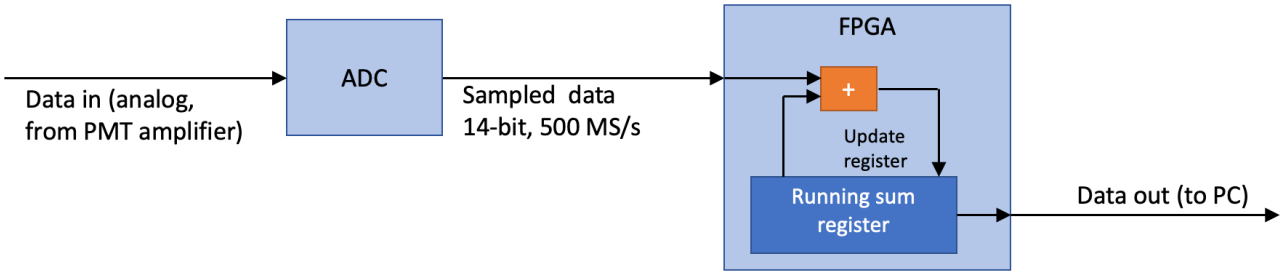


Figure 5-5: ADC/FPGA Overview Block Diagram

Because of the significant speed at which this these operations would need to be performed at, the device that was chosen to accomplish this task was a field-programmable gate array (FPGA) controller. The FPGA board that was chosen was an iWave Systems Intel Arria10 SoC/FPGA Development kit, which was supplied by Arrow Electronics. This product included an Intel Arria-10 System-On-a-Chip (SOC), and an iWave carrier board for interfacing with ADC evaluation

boards. This model was chosen because it is available as a demonstration board and could be paired with the AD9680 ADC in the form of an AD-FMCDAQ2-EBZ evaluation board. The FPGA and ADC module were chosen as a pair because they had been documented by Arrow Electronics as an example project with a platform-user guide, which made for easy development of a proof-of-concept setup. The ADC had sufficient specifications including a maximum sampling rate of 2.8 GS/s and a resolution of 14 bits per sample, and the FPGA development board had the necessary features and specification for this project including an onboard SOC with a Dual Arm Cortex A9 processor for running a Linux operating system, 2 GB of RAM, which would be sufficient for the operating system and programs that needed to be run during operation, a floating-point performance of 1.5 TeraFLOPs, which was more than enough for implementing the summation functionality, a 12.5 Gbps backplane data rate, which is fast enough to accept the 500 MS/s sent by the ADC. Programming of the ADC/FPGA controller board was done by Dean Kostan at Logic Tectonics according to the specifications shown in the final block diagram shown in Figure 5-6.

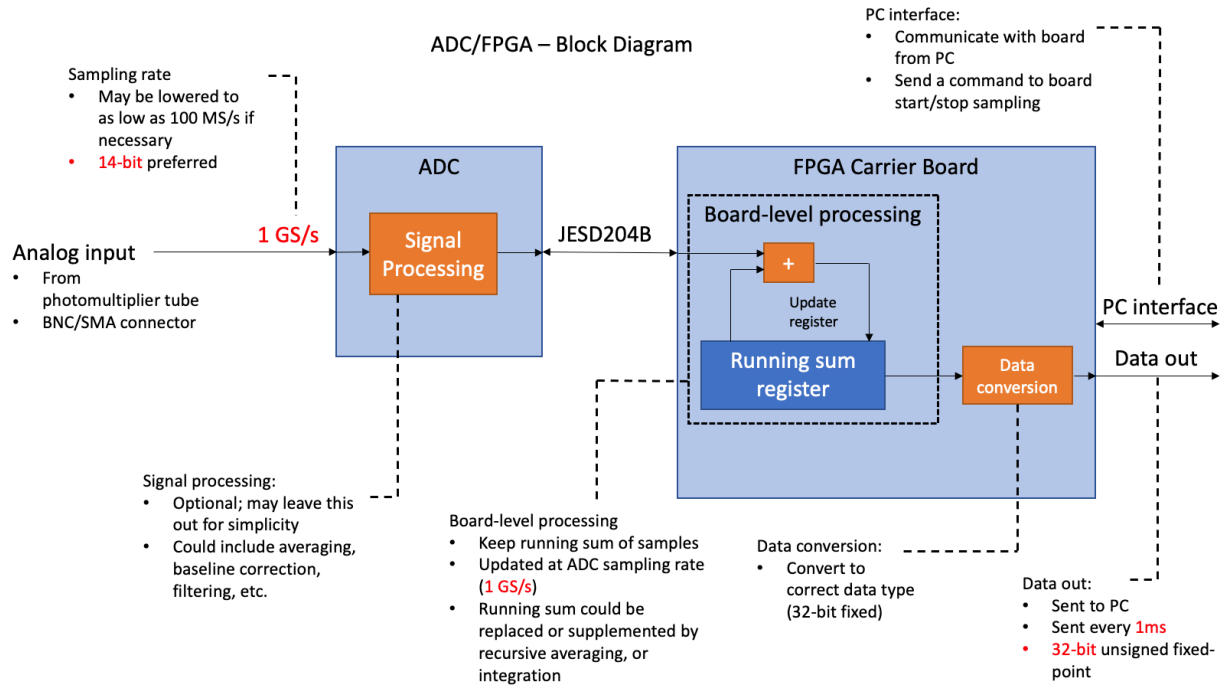


Figure 5-6: ADC/FPGA Specification Block Diagram

The final ADC/FPGA board setup is shown in Figure 5-7. It communicates with the computer via a USB serial connection and has a separate output-only serial transmission (UART over 5V TTL serial), which connects to a computer with a 5V serial-to-USB adapter, and outputs the "accumulator value," or the current value of the running ADC sum, every 1 ms.[46] The USB serial connection is used to communicate with the board via Linux command line. Functionality was added to the board in the form of programs that could be run through the Linux operating system, which controls various functions on the FPGA. They include the following:

1. accum_on: This command starts the collection of data from the ADC to the FPGA, and starts the "accumulator" function, which generates the running sum of sampled data.
2. accum_off: This command stops the accumulator function.

[46] The 1 ms transmission time was chosen to be approximately the fastest time at which the LabView program could reasonably keep up with the data stream and respond by turning off the plasma once a particular ADC value was reached.

3. accum_reset: This command resets the accumulator value and restarts the accumulator function.

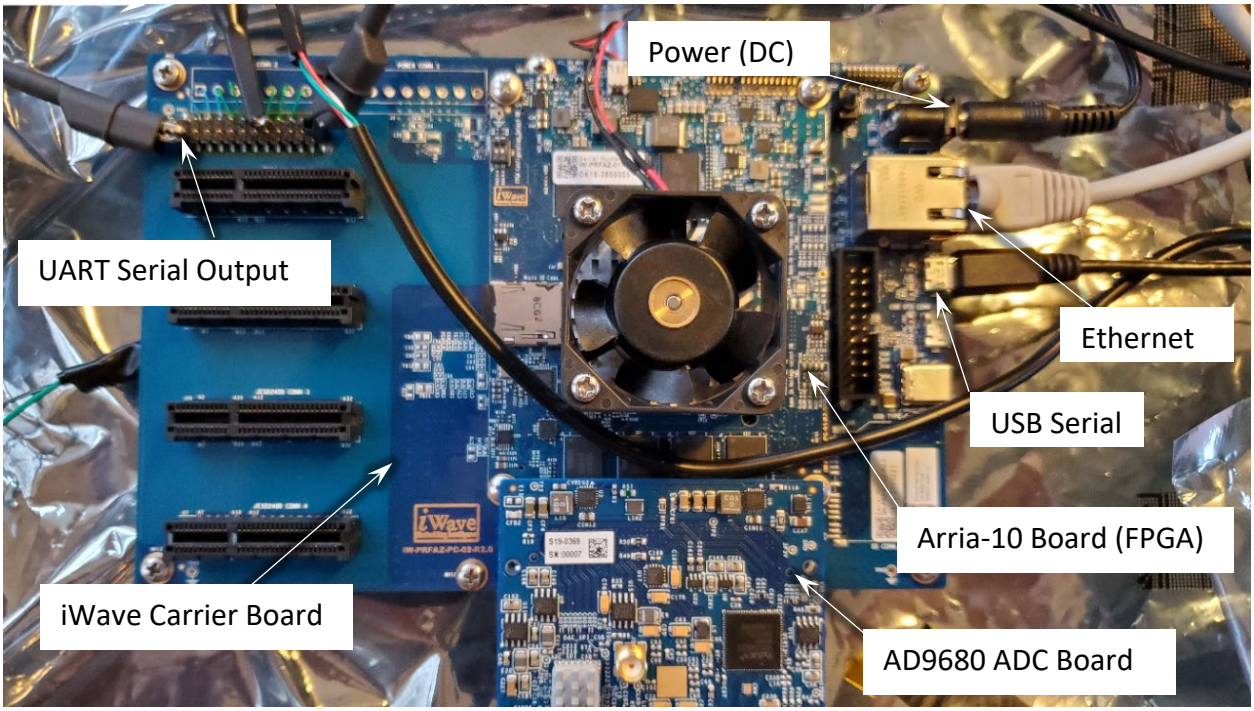


Figure 5-7: ADC/FPGA Board

5.1.4 LabView program

A computer running LabView was used to interface with the FPGA/ADC board, and with the hardware that controlled the plasma generation. Two primary LabView codes were written: “ADC Timed” and “ADC Control”. The “ADC Timed” program was used to expose a sample to a user-inputted plasma-exposure time and records the ADC accumulator value at the end of an exposure. This program was primarily used to find the relationship between plasma-exposure time and the corresponding ADC accumulator values. The “ADC Control” program would alternatively expose a sample to plasma until a user-inputted ADC accumulator value was reached or exceeded, at which point the plasma is stopped by the program, and the final ADC accumulator value is recorded. These programs worked by sending the Linux command-line instructions to the board

in the following sequence: reset and start the accumulator with the "accum_reset" command, start the plasma by enabling the function generator output, and then record the accumulator values sent by the FPGA's UART connection. Once the desired plasma-exposure time or accumulator value was reached, the program would stop the plasma by disabling the function generator.

The first test that was performed after assembling the system was to test its response to plasma. 300 μL samples of PBS were prepared in UV-transparent cuvettes, and were exposed to 0, 1, 2.5, and 5 seconds of plasma in triplicates using the "ADC Timed" LabView program to record the ADC accumulator value for each sample. As shown in Figure 5-8, the resulting ADC values follow a linear relationship with exposure time, with a coefficient of determination (R^2) value of 0.9996, and coefficient of variation (C.V.)[47] values below 5% for all replicates. This experiment was then repeated, but with the high-voltage needle of the PLIMB instrument raised high enough above the surface of the solution so that a plasma discharge did not form. The resulting ADC accumulator values also followed a linear relationship with time, but with a slope that is 5% of the slope resulting from the values obtained with the plasma on. The increasing ADC accumulator values here likely result from small amounts of noise being generated from the PMT amplifier output. Note that when the power supply of the amplifier is turned off and the ADC accumulator is turned on for comparable amounts of time, the magnitude of resulting ADC values, relative to that of the data shown here, is negligible.

[47] Coefficient of variation is the standard deviation of a set of data divided by the average.

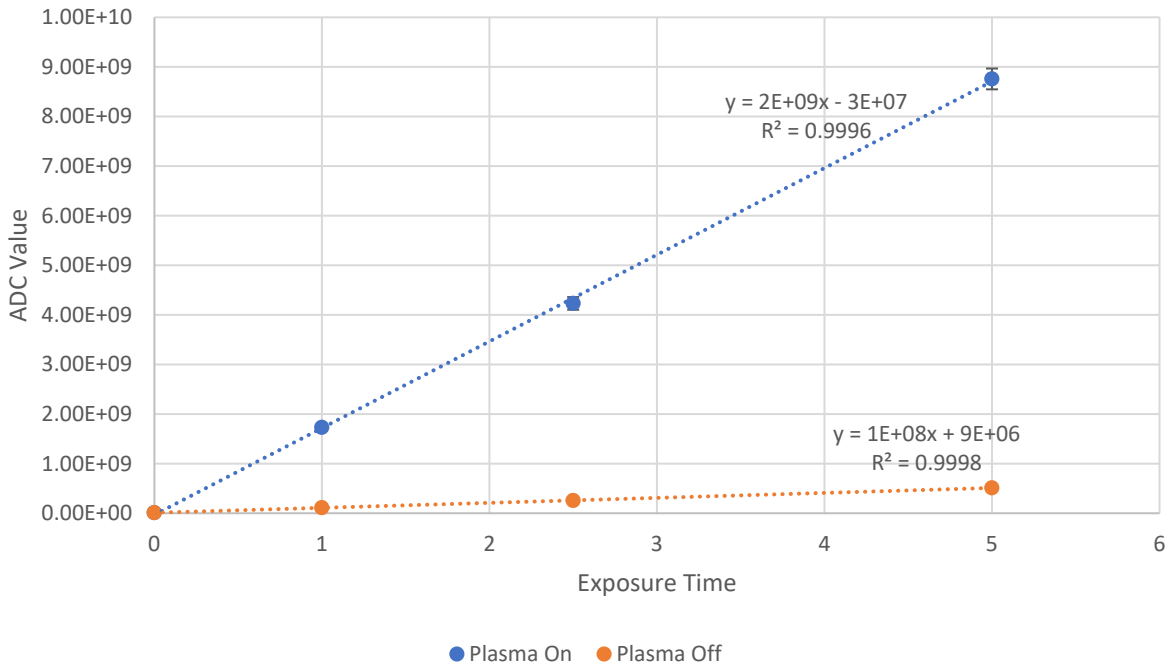


Figure 5-8: ADC Accumulator values vs. time with plasma off and on

5.2 Coumarin exposure and fluorescent measurement

Coumarin was first used to test the function of the UV spectroscopy setup, and to begin to develop a relationship between the measured UV signal and the hydroxyl-radical dose. A 10-mM coumarin solution was prepared in phosphate buffered solution (PBS), and samples of 300-mL volume were prepared in UV-transparent cuvettes. These samples were exposed to 0, 10, 20, and 30 seconds of plasma (see Section 3.3 for plasma parameters) in triplicate (12 samples) using the "ADC Timed" LabView program to record the amount of 309-nm UV light detected by the photomultiplier tube that was sampled and summed using the ADC/FPGA board. The coumarin samples were then analyzed for hydroxyl modification with a Tecan SPECTRAFluor Plus fluorometer, using an excitation wavelength of 360 nm and a detection wavelength of 465 nm. The results can be seen in Figure 5-9.

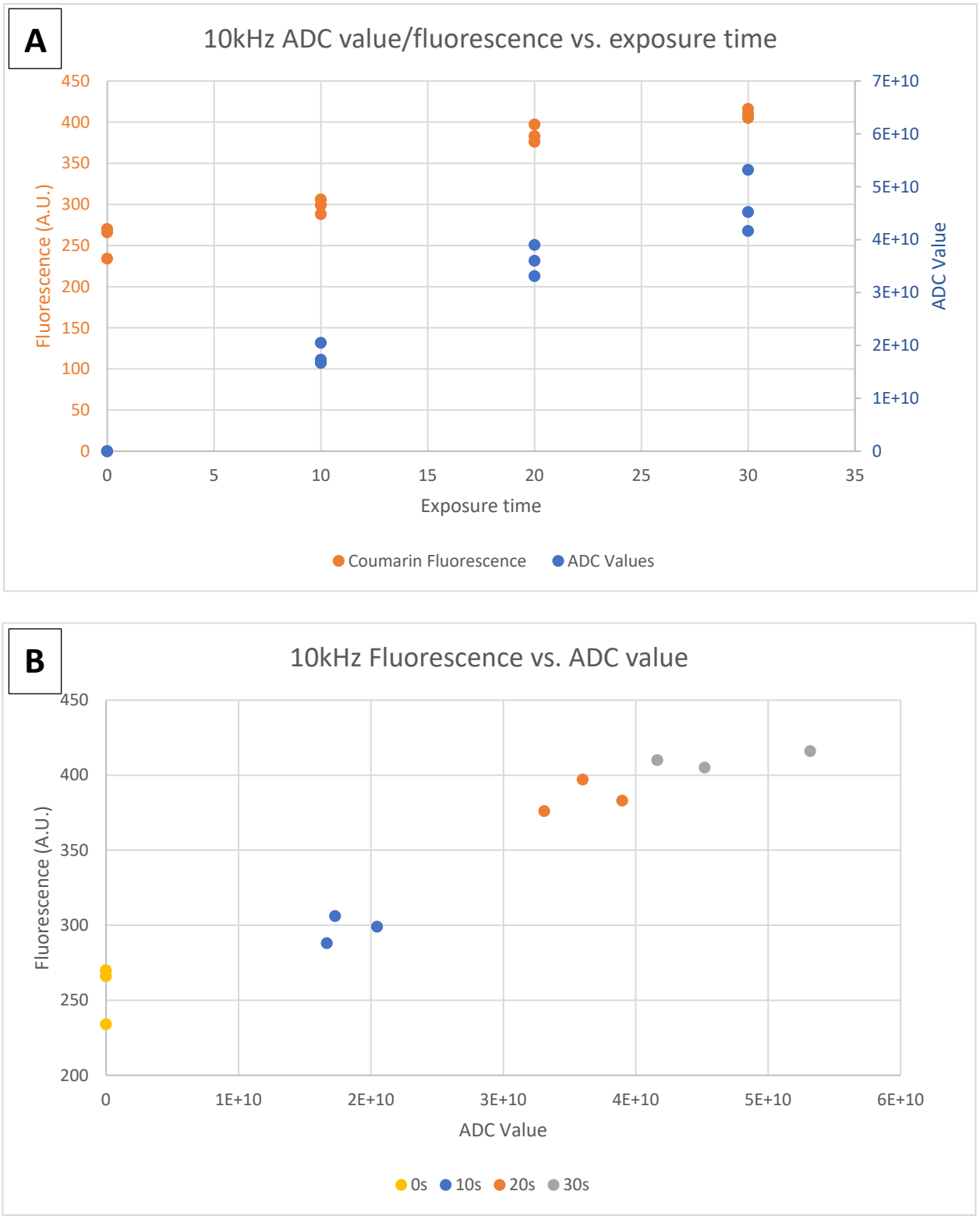


Figure 5-9: Coumarin/ADC accumulator values for 10kHz plasma

As can be seen in Figure 5-9A, both the coumarin fluorescence and ADC accumulator values increase with longer exposure times. Figure 5-9B shows the relationship between coumarin fluorescence and ADC values. This graph shows that there is a roughly linear relationship between coumarin fluorescence and ADC values. However, it should be noted that the relationship between coumarin fluorescence and hydroxyl-radical dose is not completely linear, especially beyond about 15 seconds of 10kHz plasma exposure [8] (see Section 3.1). Therefore, a perfect-linear relationship between coumarin fluorescence and ADC values cannot reasonably be expected.

This experiment was then repeated using a high voltage frequency of 2kHz rather than the standard 10kHz. Because it is known that a 2kHz frequency plasma will generate a smaller dose of hydroxyl radicals over the same exposure time, [8] it was reasonable to assume that a smaller measure of coumarin fluorescence alongside smaller ADC values would result. The results can be seen in Figure 5-10.

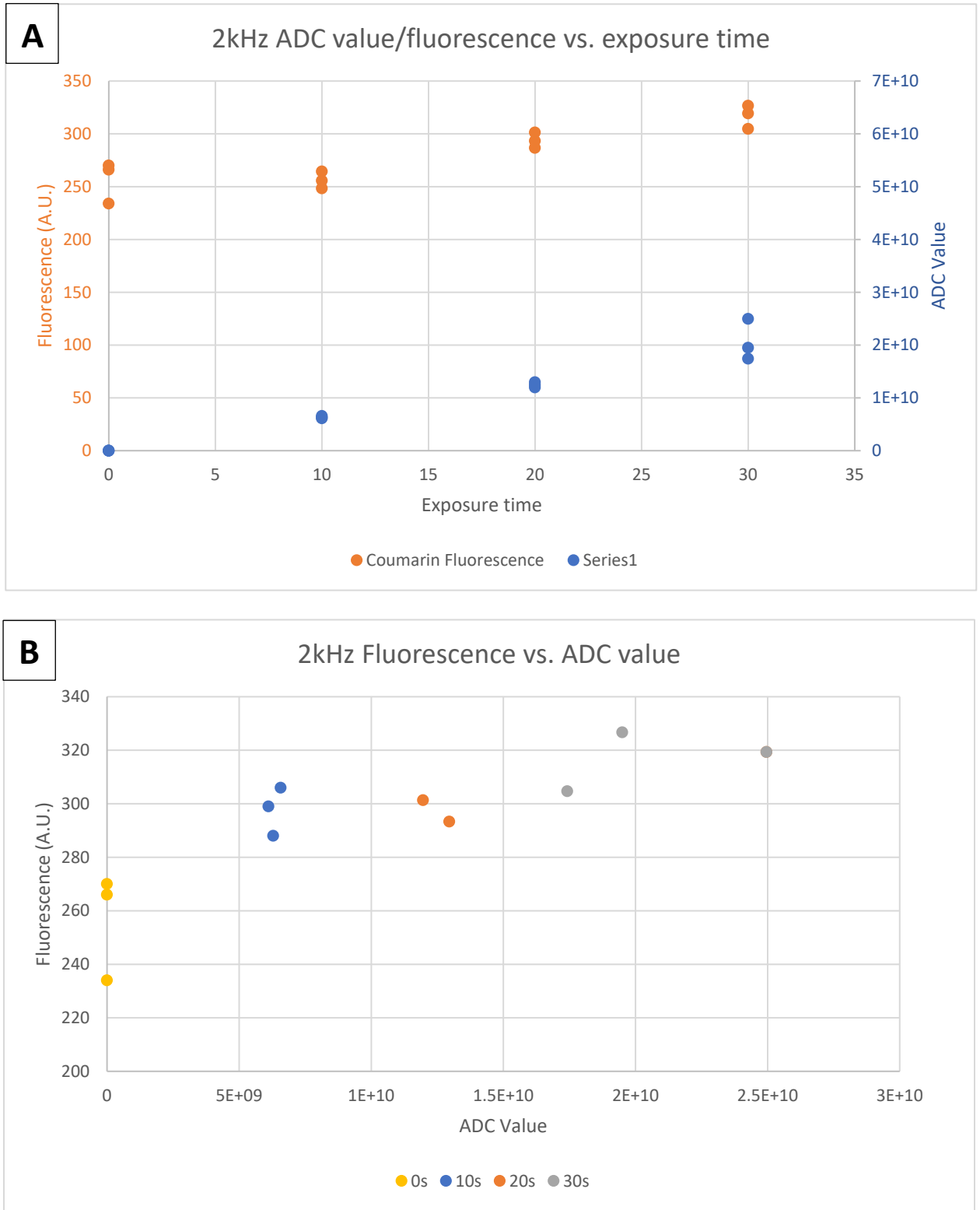


Figure 5-10: Coumarin/ADC accumulator values for 2kHz plasma

Here, we see similar results to the 10kHz plasma exposure, but with coumarin fluorescence and ADC accumulator values significantly lower, as would be expected for a plasma that generates a lower dose of hydroxyl radicals in the same amount of exposure time relative to 10 kHz. Plotting the coumarin fluorescence vs. ADC value data for both 10 kHz and 2 kHz plasmas on the same graph (Figure 5-11) shows an interesting result. Here we see that the linear relationship between coumarin fluorescence and ADC values follows the same trajectory for both 2 and 10 kHz plasmas. This suggests that the relationship between the ADC accumulator value for a given plasma exposure and the measured coumarin fluorescence value, and thus the hydroxyl dose, is linearly related regardless of the radical generation rate.

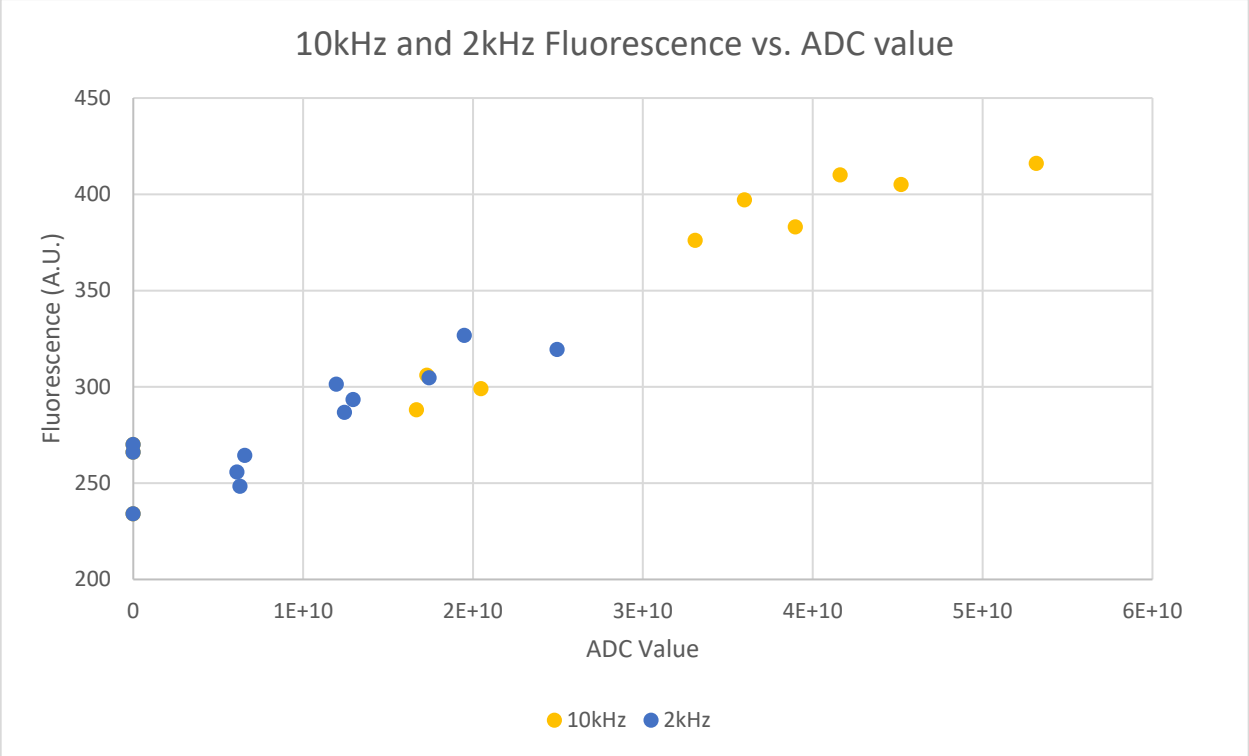


Figure 5-11: Coumarin Fluorescence vs. ADC accumulator value for 2 and 10kHz plasma

5.3 Methionine exposure

After showing promise with coumarin exposure, the UV-spectroscopy setup was tested further using samples of methionine. As shown in Section 3.1, exposed methionine samples show a dose-dependent increase in percent oxidation when exposed to PLIMB when measured with mass spectrometry, and the relationship is linear over a wider range of plasma exposure times than coumarin fluorescence. The data also typically exhibits smaller error bars when compared to coumarin-fluorescence measurements. For these reasons, methionine is considered a better standard for measuring plasma-induced oxidation than coumarin.

A 50 μM solution of methionine was prepared in 25mM, pH 7 ammonium acetate, and samples of 100-μL volume were prepared in UV-transparent cuvettes. Three samples were exposed to PLIMB for 20 seconds using the "ADC Timed" LabView program to record the amount 309-nm UV light detected by the photomultiplier tube and sampled and summed using the ADC/FPGA board. Then, the average value of the recorded ADC accumulator value was calculated, and three additional samples were exposed to PLIMB using the "ADC Control" LabView program, using the average ADC accumulator value for the first three samples as the control value. This program recorded both the exposure time and the final ADC accumulator value when the plasma had stopped. These samples were then analyzed with an electrospray ionization time-of-flight (ESI-TOF) mass spectrometer soon after exposure. Percent oxidation was calculated by taking the signal produced by ions with a mass of 165 Da (methionine sulfoxide) and dividing by the sum of the signals produced by ions with mass of 149 Da (unmodified methionine) and ions of methionine sulfoxide, as shown in equation (5-8). The results can be seen in Table 5-1, Table 5-2 and Figure 5-12.

Table 5-1: Methionine exposure time, percent oxidation, and ADC accumulator values

$$\%\ oxidation = \frac{(methionine\ sulfoxide\ signal)}{(unmodified\ methionine\ signal) + (methionine\ sulfoxide\ signal)} \quad \textbf{(5-8)}$$

	Timed			ADC Control		
Sample:	T1	T2	T3	C1	C2	C3
Exposure time (s)	20	20	20	20.688	21.096	21.465
Percent Oxidation	6.16%	12.99%	16.38%	11.95%	12.78%	15.40%
ADC Value (Hex)	9A873D27F	870F28441	8D41827D3	8FBF2B721	8FC0EF4EB	8FBDEA6DD
ADC Value (Dec)	4.15E+10	3.63E+10	3.79E+10	3.86E+10	3.86E+10	3.86E+10

Table 5-2: Average percent oxidation and ADC accumulator values

		Timed	ADC Control
% Oxidation:	Average	11.84%	13.37%
	Standard Dev.	4.25%	1.47%
	Coefficient of Variation	35.90%	11.01%
ADC Value:	Average	3.86E+10	3.86E+10
	Standard Dev.	2.18E+09	1.30E+06
	Coefficient of Variation	5.65%	0.00%

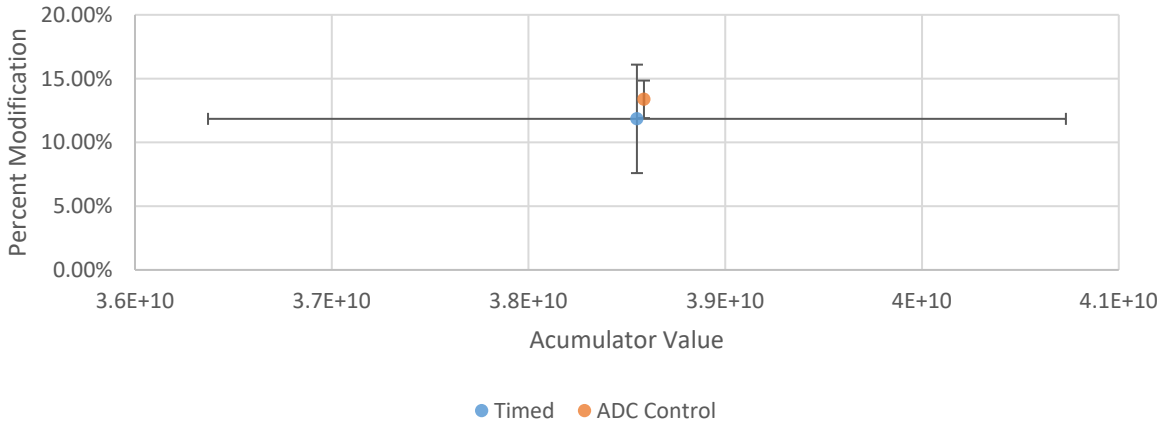


Figure 5-12: Average percent oxidation vs. ADC accumulator value with standard deviation error bars

As can be seen in Table 5-2, the coefficient of variation (C.V.) of percent oxidation and ADC values for the timed samples are 35.9% and 5.65% respectively. When exposing to a particular ADC value rather than a set exposure time, the C.V. of the percent-oxidation shrank to 11% as the C.V. of the ADC value was set to 0% by the feedback-control system. This experiment demonstrates that using the feedback-control system for exposing samples to a particular dose, as measured by summed ADC values, can significantly reduce variability in percent oxidation of replicate samples exposed to PLIMB, when compared to samples exposed to a particular exposure time.

5.4 Bandpass filter

While it has been demonstrated that the feedback-control system using a monochromator for UV spectroscopy is effective in achieving the goal of reducing variability in oxidation between replicate PLIMB-exposed samples, there are some drawbacks in its use for this role. The first is that the monochromator is quite bulky, requiring the entire PLIMB/feedback-control setup to take up an entire workbench. Second, monochromators can be expensive. Monochromators like the one used for the experiments described in this chapter, a McPherson 234-302, can cost tens of thousands of dollars. While these drawbacks may not necessarily be prohibitive, they would surely present challenges for the commercial adoption of such a system.

Another drawback is that, because of the monochromator's design, it exhibits a significant amount of signal loss between the intensity of incoming and outgoing light. This is for several reasons: 1) The entrance and exit slits are only about 0.5 mm in width. While this is important for achieving high spectral resolution, it only allows for a small amount of light through the monochromator. 2) The path length between the entrance and exit of the monochromator is about 0.5 meters. [9] Given that light intensity is inversely proportional to the square of the distance it

travels, this long of a path between the light source, here being the plasma, and the photomultiplier-tube detector, results in a significant reduction in signal than if the detector was placed very close to the source. Finally, 3) monochromators require a diffraction grating, which also results in an additional loss of signal. The 1200g/mm grating that was used for the experiments in this chapter, which was blazed for 140 nm, has only about a 15% theoretical efficiency at 309nm. [9]

The combination of these factors results in a system which has significant signal loss. However, using the monochromator setup, the peak output signal of the photomultiplier tube amplifier was about 750 mV during the plasma discharge, as can be seen in Figure 5-13. The magnitude of this signal falls within the 1.5V input range of the ADC. As a result, the system is able to measure the light intensity with reasonable resolution. However, the plasma and its associated power supplies also generate a significant amount of electrical noise, which gets introduced into the output voltage. As can be seen in Figure 5-13, the magnitude of the electrical noise is about 250-mV peak (the ADC only measures positive voltages). Therefore, the signal-to-noise ratio (SNR) that was achieved by this system is estimated to be about 3:1. It is hypothesized that, by increasing the amount of light signal incident to the photomultiplier tube while keeping the electrical noise constant, the SNR could be reduced, thus improving the performance of the system for accurately measuring intensities of light corresponding to a hydroxyl-radical dose.

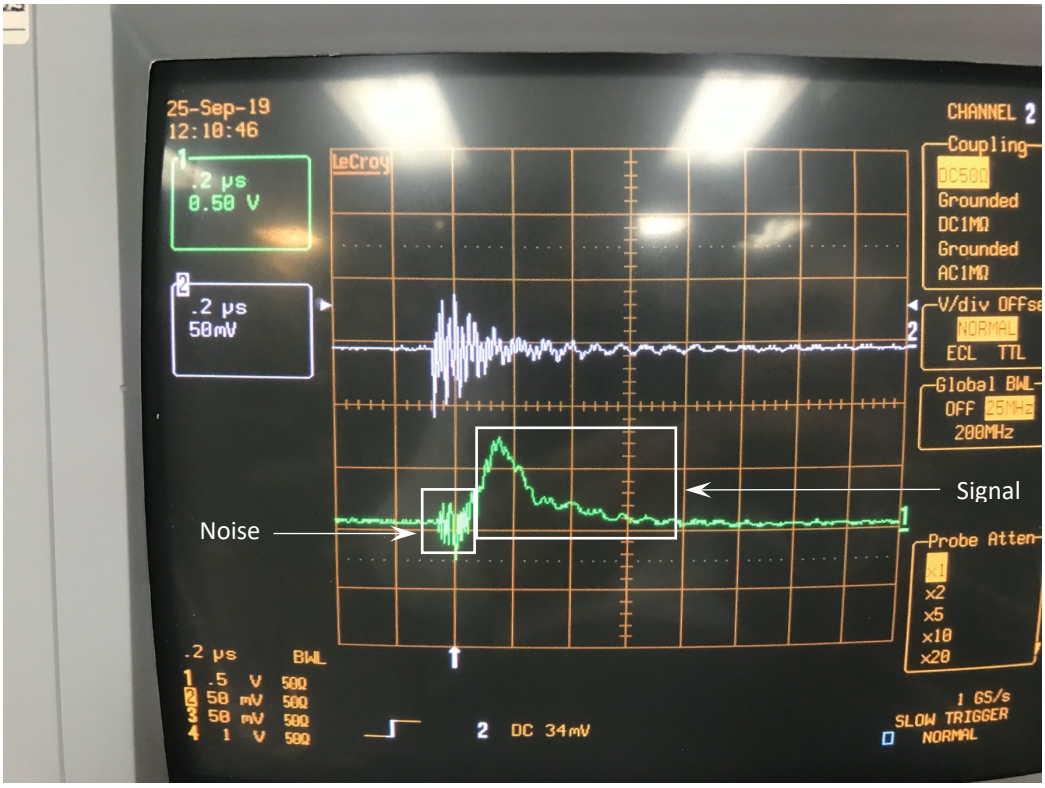


Figure 5-13: PMT output voltage with monochromator

The oscilliscope capture here shows two channels: channel 1 (green) shows the response of the photomultiplier tube (PMT) amplifier output voltage, and channel 2 (white) shows the plasma current over a single plasma discharge. On the PMT output signal, we see two distinct waveshapes overlayed on one another. The first is the high frequency noise, which is noticeably similar to the shape of the plasma discharge current. Here, we see that the noise has a magnitude of about 250mV. The other distinct waveshape is the lower frequency rise and fall of the PMT output, here labelled as the "signal," which has a magnitude of about 750mV.

The approach that was used here to increase the light signal at the input of the PMT, was to replace the monochromator with a UV-bandpass filter. Although a bandpass filter would have some disadvantages, such as a wider, non-adjustable bandwidth, when compared to a monochromator for this purpose, it would allow for a much stronger light signal. This is because: 1) the theoretical efficiency of many UV-bandwidth filters can be as high as 80% or more, 2) the setup would no longer need to include an input or output slit, and 3) the light path could be significantly shortened, allowing for a distance of just a few centimeters between the plasma-light

source and the photomultiplier tube detector. Furthermore, the overall setup would be much smaller and cheaper than could be achieved with a monochromator.

The bandpass filter that was chosen was an Edmund Optics model 34-974, a 25-mm diameter filter with a center wavelength of 310 ± 2.0-nm, a full width-half maximum wavelength of 10 ± 2-nm, and a maximum transmission ≥70% at center wavelength. [10] This filter would allow the detection of 309-nm light, while blocking out most other bands emitted by the plasma.

By observing the output of the photomultiplier-tube amplifier with an oscilloscope after switching to the UV bandpass filter, it was found that the peak voltage after a plasma discharge was now about 10 V, while the noise signal remained at around 250 mV. This indicated that the SNR had indeed been increased from about 3 to about 40. What now needed to be done, was to scale the output voltage down to about 1.5 V peak, in order to achieve the maximum resolution with the ADC, which has an input voltage range of 1.5V. The output needed to be scaled down in such a way as to maintain the high SNR that was now achieved. This was done by including a simple voltage divider circuit using a potentiometer as seen in Figure 5-14. This would scale down the voltage of both the output signal and the noise proportionally, so long as the noise was induced upstream (towards the photomultiplier tube) in the circuit. While running plasma and monitoring the output voltage with an oscilloscope, the potentiometer was adjusted until the peak voltage was around 1.5V. Along with the PMT output voltage signal, the noise signal was also reduced significantly.

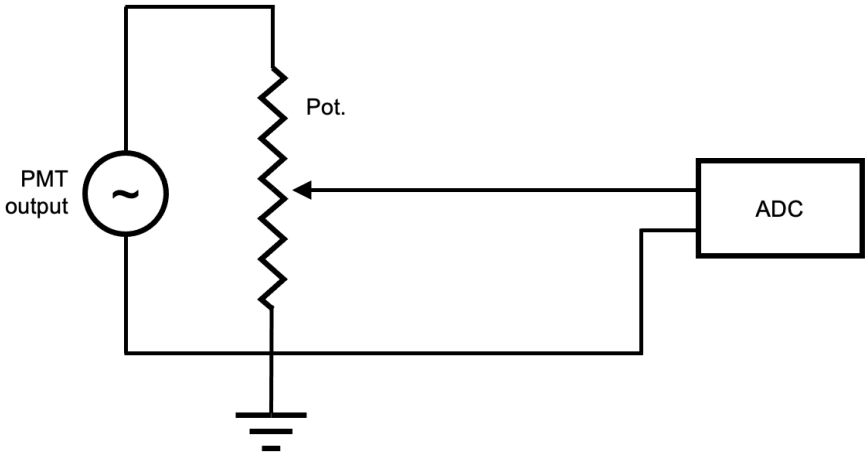


Figure 5-14: Potentiometer circuit used to scale PMT output voltage

5.5 Methionine exposure with bandpass filter

To test the new experimental setup, six samples of methionine were again prepared and exposed to PLIMB. The first set of three samples was exposed to PLIMB for 15 seconds using the "ADC Timed" LabView program, and the ADC accumulator values for each were recorded. Then, a second set of three samples was exposed using the "ADC Control" LabView program, which exposed each sample to PLIMB for the average value of the ADC of the first set of samples. The exact exposure time for each sample was recorded. Following PLIMB exposure, the samples were analyzed by mass spectrometry in order to measure the percent concentration of oxidized methionine in each sample. The results can be seen below in Table 5-3, Table 5-4, and Figure 5-15.

Table 5-3: Methionine exposure results with UV bandpass filter

	Timed			ADC Control		
Sample:	T1	T2	T3	C1	C2	C3
Exposure time (s)	15	15	15	12.759	13.414	15.177
Percent Oxidation	3.63%	4.94%	5.19%	3.19%	3.59%	3.60%
ADC Value (Hex)	9A873D27F	870F28441	8D41827D3	8FBF2B721	8FC0EF4EB	8FBDEA6DD
ADC Value (Dec)	4.15E+10	3.63E+10	3.79E+10	3.86E+10	3.86E+10	3.86E+10

Table 5-4: Average percent oxidation and ADC accumulator values for UV bandpass filter

		Timed	ADC Control
% Oxidation:	Average	4.59%	3.46%
	Standard Dev.	0.68%	0.19%
	Coefficient of Variation	14.90%	5.53%
ADC Value:	Average	3.48E+11	3.48E+11
	Standard Dev.	2.90E+10	2.92E+07
	Coefficient of Variation	8.32%	0.01%

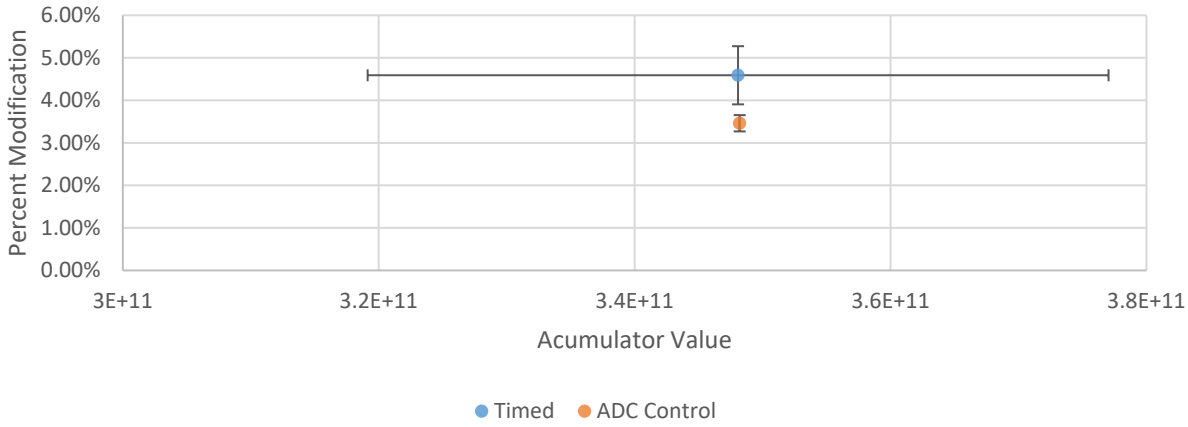


Figure 5-15: Average percent oxidation vs. ADC accumulator value with standard deviation error bars for UV-bandpass filter

As can be seen in Table 5-4, the coefficient of variation (C.V.) of percent oxidation and ADC values for the timed samples are 14.9% and 8.32% respectively. When exposing up to a particular ADC value rather than for a set exposure time, the C.V. of percent oxidation and final ADC accumulator value shrank to 5.53% as the C.V. of the ADC value was set to 0% by the feedback-control system. This result shows that by using the UV-bandpass filter as a replacement for the monochromator in the feedback-control system, it is effective for reducing variability in percent oxidation of replicate samples exposed to PLIMB, when compared to samples exposed for a particular exposure time. Furthermore, the system is shown to have improved performance, as the percent oxidation C.V. between replicate ADC-controlled samples decreased from 11.01% to

5.53%, when replacing the monochromator with the UV-bandpass filter. This is likely due to the increased SNR that was achieved with the UV-bandpass filter when compared to that using the monochromator.

5.6 Self-oxidation

While the feedback-control system addresses the primary cause of variability in oxidation between replicate samples exposed to PLIMB, *i.e.,* the variability in hydroxyl-radical dose caused by varying hydroxyl-generation rates over an entire exposure, there are other factors which could cause variability. One in particular is oxidation that occurs *after* PLIMB exposure, but before analysis by mass spectrometry. Oxidation could be introduced throughout any of the sample handling procedures that are followed to prepare protein samples for mass spectrometry analysis. Optimizing *each* of these steps to minimize the potential for oxidation is something that should be done to further reduce variability in oxidation. However, this would be quite an involved process, and is beyond the scope of this thesis (see Section 7.4.2).

Another potential cause of oxidation after PLIMB exposure is in sample storage. It is often not feasible to perform all of the sample-handling procedures immediately followed by mass spectrometry analysis after exposing samples to PLIMB, especially if there is a queue of samples that need to be analyzed on a mass spectrometer. Therefore, sample storage is often required.

Throughout our experience with PLIMB, it was only assumed that the potential for self-oxidation should be minimized by minimizing the storage time as much as possible and by storing samples in as cold a temperature as possible. This section aims to develop a method for quantifying the self-oxidation of samples in order to determine how samples should be stored, and for how long before self-oxidation starts to cause concern in data variability.

To test self-oxidation of samples over time, coumarin was prepared at a 10 mM concentration in a 25 mM, pH 7 ammonium acetate buffer. Two sets of 96 samples were aliquoted into 250 µL PCR tubes at 50 µL volumes, and one set was exposed to PLIMB for 15 seconds, while the other was not exposed to PLIMB. Half of each set of samples was quenched with a 50 µL solution of 10 mM methionine in 25 mM ammonium acetate buffer after exposure while the other half was quenched with 50 µL of 25 mM ammonium acetate buffer. This was done to test whether methionine quenching affects the rate of self-oxidation during sample storage. Each of these sets of 12 samples were then separated and stored at four different temperatures: 25, 4, -20, and -80° C. Three samples, for triplicate measurements, from each set were removed from storage, and the fluorescence was measured after 0, 3, 6, and 9 days of storage.

In order to analyze the effect of storage at various temperatures, a change in coumarin fluorescence was calculated for set of samples by linear regression, taking the storage time (0 through 9 days) as the independent variable values, and the measured coumarin fluorescence as the dependent values. The data was grouped based on the experimental condition: 0 or 15s PLIMB exposure, and methionine vs. buffer quenching after exposure. The change in fluorescence versus storage temperature was plotted, and the results can be seen in Figure 5-16.

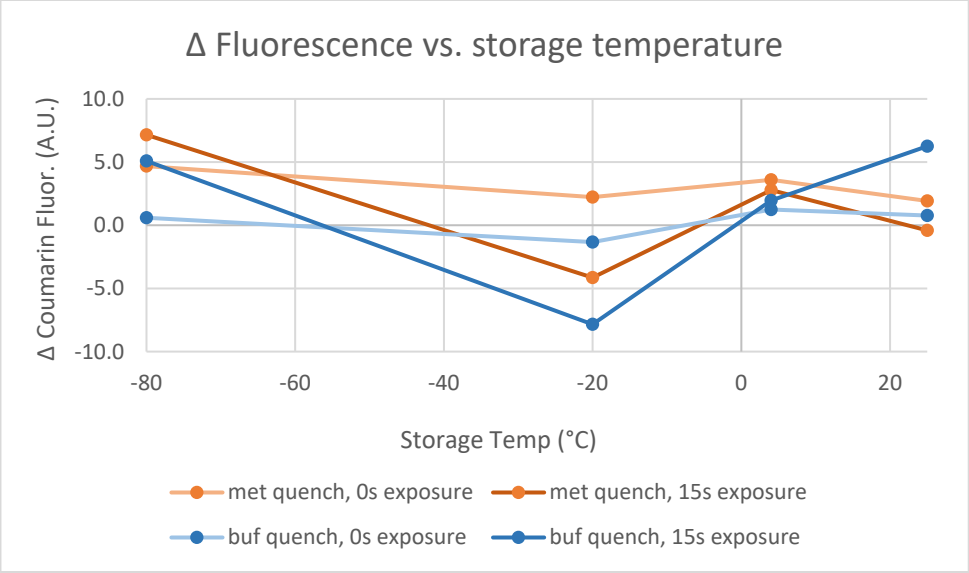


Figure 5-16: Change in coumarin fluorescence vs. temperature storage

As we can see from these results, samples stored at -20° C experienced the lowest amount of self-oxidation among samples of all conditions. Interestingly, we even see a decreasing amount of fluorescence over the 9-day period over which these rates were calculated, indicated by a negative value for Δ Coumarin Fluor. It is not yet known what might cause a decrease in fluorescence, especially since oxidation is generally considered a permanent modification under most circumstances. However, assuming the coumarin samples only experienced an increase in oxidation after storage, and that any decrease in fluorescence is caused by an unrelated reaction, storage at -20° seems to result in the least amount of self-oxidation.

Another important consideration for storing samples is the expected percent change in oxidation over time at various storage temperatures. The measured coumarin fluorescence values after storage were normalized by the values measured immediately after PLIMB exposure and are shown in the following graphs. Figure 5-17A shows methionine-quenched samples, and Figure 5-17B shows buffer-quenched samples.

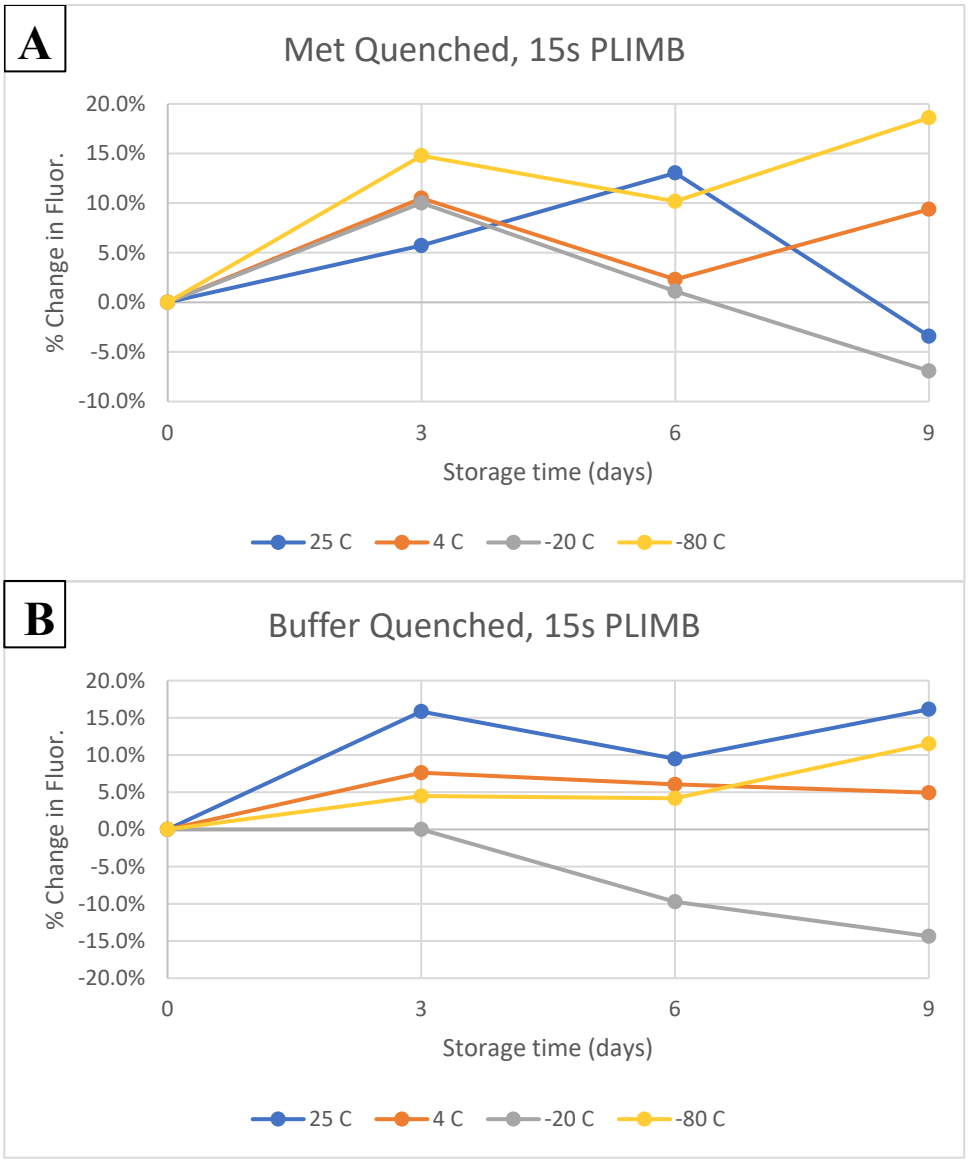


Figure 5-17: Percent change in coumarin fluorescence vs. storage time for methionine (A) and buffer-quenched (B) samples

Here again we see that storage at -20° (shown in grey) shows the smallest increase in fluorescence for samples under all conditions and over all storage times, with the sole exception of methionine-quenched coumarin (Figure 5-17A) stored for three days at room temperature (25° C, shown in blue). In fact, storage of buffer-quenched samples at -20° even shows a decrease in fluorescence over time, which again, may be due to a separate reaction that is unrelated to an increase in fluorescence due to oxidation.

While we do see significant amounts of change in coumarin fluorescence after storage, ranging from 0 to 20%, a rise in oxidation among all samples is not as concerning as a differential increase in oxidation between replicate samples. As is evident by the analysis shown in Chapter 4, a rise in the standard deviation between replicate samples is detrimental to the detection of statistically significant differences between samples. To assess the potential for differences in the rate of self-oxidation between replicate samples, the coefficient of variation (C.V.) in fluorescence between replicate samples as a function of storage time is shown in Figure 5-18 for samples stored at all temperatures.

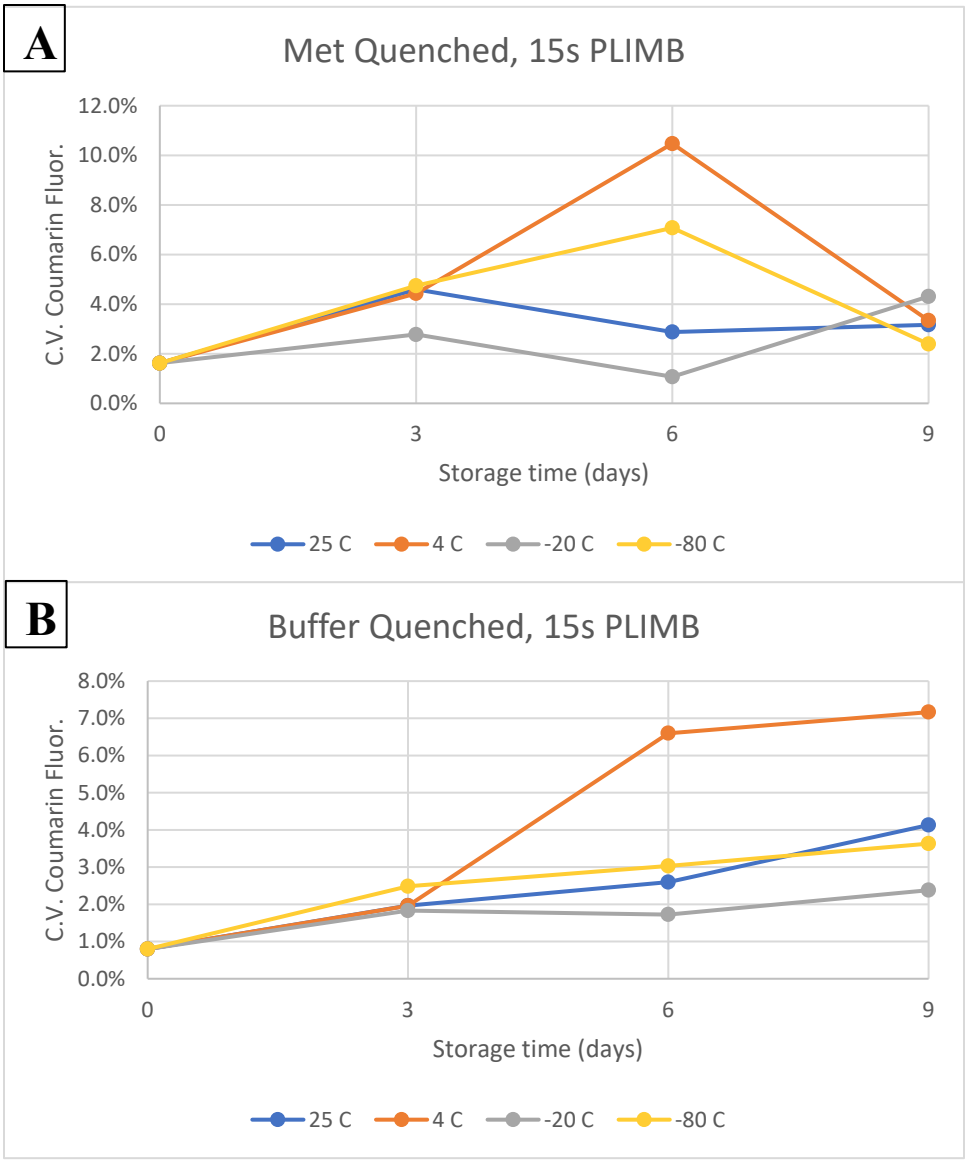


Figure 5-18: Coefficient of variation (C.V.) between replicate samples vs. storage time

Here we see that the C.V. generally increases slightly over time, but the values are somewhat smaller in magnitude than those of the percent change in oxidation shown in Figure 5-17. We also see that, for both methionine and buffer-quenched samples, three days of storage or less keeps the C.V. below 5%. We also see that -20° storage (shown in grey) shows a minimal amount of variation between replicates for all storage times, both methionine and buffer quenched, with the sole exception of methionine-quenched coumarin (Figure 5-18B) stored for 9 days at -20°

(grey) which shows a higher C.V. than the other storage temperatures. This again suggests that -20° may be the ideal storage temperature.

While the data presented here is informative for assessing rates of self-oxidation after PLIMB exposure, it may be necessary to perform a similar experiment using samples of methionine or protein for increased accuracy. As discussed in Section 3.1, coumarin is a less-accurate reporter of oxidation when compared to methionine or proteins. Furthermore, it is not yet known what causes the decrease in coumarin fluorescence that can be seen in these results, or whether this decrease is: 1) independent of storage temperature, 2) constant over time, or 3) is indicative of any chemical reaction that can also be observed in proteins. For these reasons, a similar experiment may even need to be performed if, for example, the precise rate of self-oxidation of a protein of interest needs to be established for a given PLIMB experiment before it is carried out. The resulting information will be informative for determining how long samples may be stored, what temperature is ideal, and how much variability is likely to result from storage prior to mass-spectrometry analysis.

5.7 References

[1] A. Pankaj *et al.*, "Generation mechanism of hydroxyl radical species and its lifetime prediction during the plasma- initiated ultraviolet (UV) photolysis," *Scientific Reports,* vol. 5, no. 1, 2015, doi: 10.1038/srep09332.

[2] M. Sun *et al.*, "Emission spectra of hydroxyl radical generated in air corona discharge," *Guang pu xue yu Guang pu fen xi= Guang pu,* vol. 25, no. 1, pp. 108-112, 2005.

[3] H.-P. Dorn, R. Neuroth, and A. Hofzumahaus, "Investigation of OH absorption cross sections of rotational transitions in the A 2∑+,>(υ'= 0>)← X 2 Π,>(υ ″= 0>) band under atmospheric conditions: Implications for tropospheric long-path absorption measurements," *JGR,* vol. 100, no. D4, pp. 7397-7409, 1995.

[4] G. Dieke and H. Crosswhite, "The ultraviolet bands of OH fundamental data," *Journal of Quantitative Spectroscopy and Radiative Transfer,* vol. 2, no. 2, pp. 97-199, 1962.

[5] Y. H. Kim *et al.*, "Measurement of reactive hydroxyl radical species inside the biosolutions during non-thermal atmospheric pressure plasma jet bombardment onto the solution," *Plasma Chemistry and Plasma Processing,* vol. 34, no. 3, pp. 457-472, 2014.

[6] Y. Hong *et al.*, "Measurement of hydroxyl radical density generated from the atmospheric pressure bioplasma jet," *Journal of Instrumentation,* vol. 7, no. 03, p. C03046, 2012.

[7] W. contributors. "Op amp integrator." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Op_amp_integrator&oldid=984122996 (accessed December 9, 2020).

[8] J. M. Blatz, D. Benjamin, F. A. Choudhury, B. B. Minkoff, M. R. Sussman, and J. L. Shohet, "Effect of frequency and applied voltage of an atmospheric-pressure dielectric-barrier discharge on breakdown and hydroxyl-radical generation with a liquid electrode," *Journal of Vacuum Science & Technology A,* vol. 38, no. 4, p. 043001, 2020, doi: 10.1116/6.0000125.

[9] McPherson. "VUV and Deep Ultraviolet Spectrometer | McPherson." https://mcphersoninc.com/spectrometers/vuvuvvis/model234302.html (accessed January 17, 2021).

[10] E. O. Inc. "310nm CWL, 25mm Dia., Hard Coated OD 4.0 10nm Bandpass Filter." http://www.edmundoptics.com/p/310nm-cwl-25mm-dia-hard-coated-od-4-10nm-bandpass-filter/33102/ (accessed January 17, 2021).

Chapter 6: Extracting Solvent Accessibility

Solvent accessibility (SA) is a protein higher-order structural measurement that describes the degree to which a given amino-acid residue, peptide, or region of a protein is exposed to its surrounding solvent (highly solvent accessible) or buried in the structure of the protein (not solvent accessible). This information is quite useful for many different protein analyses because it can elucidate how different regions of a protein may interact with other biomolecules. In the case of epitope mapping, differences in solvent accessibility are used to determine which region of a target protein can be interacting with a protein therapeutic binder.

The topic of solvent accessibility is notably significant for understanding hydroxyl-radical footprinting techniques such as PLIMB, because footprinting data is a measurement of solvent accessibility when other extraneous variables, such as the varying inherent reactivity rates of amino acids to hydroxyl radicals, are controlled for, as will be described here.

All of the previous work done by PLIMB has employed a technique that allows for differential analysis of solvent accessibility. By comparing the extent of modification to a particular residue or peptide of a protein sample that was exposed to PLIMB under different conditions, a direct comparison between solvent accessibilities can be achieved. This is because all of the other variables that effect PLIMB-induced modification are kept constant, as described in the table below. However, it should be noted that *many* of the experimental conditions and variables could have a significant effect on the results of a PLIMB experiment. Therefore, the experimental conditions should be chosen based on a multi-factorial analysis in order to optimize all experimental conditions based on the desired outcome for a particular experiment (see Section 7.4.1 for details).

Table 6-1: Variables controlled for in differential-PLIMB analysis

Variable:	Controlled by:
Plasma parameters	Holding all plasma parameters constant between replicates
Radical dose	Constant hydroxyl-radical dose controlled by exposure time or feedback-control system
Temperature	Constant sample temperature set by Peltier cooler
Inherent reactivity	Comparing only identical residues or peptides
Sequence context	Comparing only the residues or peptides with identical protein sequence positions

For epitope mapping experiments with PLIMB, the variable that is changed between two sets of samples is the presence or absence of an antibody binder. This allows for the detection of epitope regions, as they are indicated by a decrease in modification upon binding. Therefore, if differences in modification between these sets of samples can be detective, they are most likely reflective of differences in solvent accessibility.

However, this approach results in a limited amount of structural information that can be gained through analysis of PLIMB data. There may be information other than differences in solvent accessibility due to binding that can be extracted from PLIMB data that would also be relevant for epitope mapping. For example, one might wish to determine whether a protein sample became denatured prior to or resulting from PLIMB exposure. Protein denaturation would call into question whether the data from an epitope mapping or other experiment performed by PLIMB is accurate. Therefore, testing for denaturation would be an excellent effect to test for using PLIMB itself. Testing for denaturation with PLIMB, could involve measuring relative solvent accessibility over many regions of the protein, and comparing these to a known, three-dimensional structure of the protein, where the structure was generated with x-ray crystallography or another high-

resolution technique. Then, it could be determined whether any regions show significant discrepancies in their relative solvent accessibility. For example, if a region of the protein was shown to be completely buried in the crystal structure, but showed to be completely solvent exposed by PLIMB, this may be an indication of denaturation. An alternative explanation that would also need to be considered in this case, is that the data could indicate that the crystal structure itself does not accurately reflect the protein's in-solution conformation, and that PLIMB *is* accurately measuring its structure, without causing denaturation. In either case, such a discrepancy in data would certainly be significant.

The goal of this chapter is to develop a method for analysis of PLIMB data to characterize the solvent accessibility of proteins in solution. The approach will be to come up with a measurement which can be calculated directly from protein-footprinting data that correlates with solvent accessibility and then compare this measurement with solvent accessibility extracted from high-resolution protein structures generated by x-ray crystallography or cryo-EM. It should however be noted that these high-resolution protein structures are an imperfect comparison for these purposes, due to the following reasons:

1. High-resolution structures may represent a pure protein without ligand or antibody binders and may include a truncated sequence of a full-length protein in the case of crystallography. This is done in order to aid in crystallization. Alternatively, cryo-EM structures often represent a protein in a larger ensemble or protein complex, involving ligand/receptors, cellular structures such as membranes, etc. Capturing larger-order structures can be used to show the interplay between these molecules in order to gain insight into their biological roles, but also is done in order to meet the requirement of a minimum molecular weight for high resolution cryo-EM structures. For this reason, the ideal structure should be one that

represents the full-length target protein alone, without added proteins or other biomolecules.

2. High-resolution structures represent the structure of a pure protein crystal, or protein frozen in solution, rather than the structure of a dynamic protein, which exhibits breathing, binding and dissociation, allosteric conformation changes, etc., which affect solvent accessibility of the protein and higher-order structure due to allosteric effects.

Despite these drawbacks of using high-resolution structures as a comparison for solvent-accessibility measurements, crystal and cryo-EM structures are still quite valuable for this purpose, mainly because they can provide structural information at an atomic-level resolution. When comparing a crystal or cryo-EM structure to that of the protein in biologically relevant conditions (nanomolar to millimolar concentrations, ~37° C, pH 7.4, salty conditions, etc.), the structural details can be thought of as overdetermined at the level of individual atoms or amino-acid residues, but can accurately represent secondary structure, and depending on the context, higher-order structure as well.

What we aim to do is to use crystallography data to compare data generated from PLIMB, in order to gain insight into the following:

1. Whether PLIMB exposure denatures a protein, thus exposing regions of the protein to solvent and thus hydroxyl radicals for labelling, where they would otherwise be buried under normal conditions.
2. Whether PLIMB-generated hydroxyl-radical protein footprinting (HRPF) data be used to measure relative solvent accessibility between different regions of a protein. One might think that comparing levels of PLIMB modification between peptide A and peptide B of a

protein may be suitable for a relative comparison of solvent accessibility. However, the reactivity rates of amino acids to hydroxyl or other radical species must be considered. Since the relative reactivity rates between individual amino acids can differ by several orders of magnitude, the relative reactivity rates of peptide A and peptide B may be substantially different as, thus rendering a direct comparison of percent modification as a measure of solvent accessibility invalid. This is because the peptide that is more reactive would show a much higher percent oxidation after a given hydroxyl dose than the less reactive peptide, all other things being equal. As described in Chapter 4, our epitope-mapping approach compares PLIMB-induced modification of a particular region of a protein. The modification is measured when the protein is in two distinct states: bound or unbound to an antibody. This approach allows us to compare percent modification of a peptide directly, as the inherent reactivity between the peptides is identical. The question thus remains: can we use PLIMB data to compare solvent accessibility between regions of a protein?

3. Determine whether PLIMB can accurately generate a measurement which is equivalent or similar to solvent accessibility, and if so, determine how it compares to a measurement of solvent accessibility generated by crystallography.

Work to extract solvent accessibility from hydroxyl-radical footprinting data has been published, including data generated with FPOP, published by Sharp *et. al.* [1], and with synchrotron radiation, published by Chance *et. al.* [2] Both involved exposing proteins to a range of hydroxyl-radical doses in order to generate a rate of oxidation on an amino-acid level, and then normalizing these oxidation rates by the intrinsic reactivity rates of the amino acid, resulting in a

calculated value called the normalized protection factor, which can then be compared directly with solvent-accessibility measurements taken from a three-dimensional model of the protein. Each of these steps will be described in detail in the following sections.

6.1 Normalized protection factor (NPF)

In order to answer the questions listed above, a measurement must be obtained from both PLIMB and crystal-structure data that can be directly compared. Sharp *et. al.* described a process they used to make these measurements and comparisons between protein crystal structure and FPOP-generated HRPF data. One of the measurements they calculate from HRPF data is the Normalized protection NPF). The NPF incorporates both the degree of FPOP-induced oxidation of a particular amino-acid residue or peptide, and the relative reactivity rate of that residue or peptide, in order to generate a value which is designed to correspond to the solvent accessibility of that residue or peptide. The formula for the NPF of a single amino acid is listed below:

$$NPF = \frac{R_i}{K_{R_i}/K_{R_P}} \quad \textbf{(6-1)}$$

Here, R_i is the oxidation rate of a given amino acid i measured with hydroxyl-radical protein footprinting, K_{Ri} is the intrinsic reactivity rate to hydroxyl radicals of that amino acid, and K_{Rp} is the intrinsic reactivity rate of the amino-acid proline. [1] This formula normalizes the rate of modification by the relative reactivity rate of the amino acid being modified, thus resulting in a measurement of solvent accessibility that can be compared between amino acids of varying reactivity rates.

6.2 Reactivity rates

The relative reactivity rates that were used by Sharp *et al.* [1] and by other researchers in the field of HRPF [3-6] were measured by Xu and Chance using synchrotron radiation [7, 8] (see Section 2.3.2 for details on synchrotron-generated hydroxyl radicals). These reactivity rates were generated by exposing samples of free amino acids in solution to a set amount of synchrotron radiation to induce oxidation, and then measuring the molar concentration of unmodified amino acid versus PLIMB exposure time. An additional amino acid, proline, was added in solution as an internal control, and thus, a relative reaction rate, Kr/Kp (see Equation (6-1)), was calculated. The relative reactivity rates, as measured by Xu and Chance, [7] are listed in the table below.

Table 6-2: Relative hydroxyl radical relative reactivity rates of amino acids with respect to proline [7]

Side chain	k/k_P
Cys*	30
Met*	21
Trp	17
Tyr	12
Phe	11
Cystine†	10
His	9.3
Leu**	4.4
Ile**	4.3
Arg	2.9
Lys	2.2
Val	1.9
Thr	1.6
Ser	1.4
Pro	1
Glu	0.69
Gln	0.66
Asn	0.44
Asp	0.42
Ala	0.14
Gly	0.04

* The reactivity was measured with Gly-X-Gly peptide using Gly-Trp-Gly as an internal control, resulting in a relative reactivity rate of Met and Cys with respect to Trp (k/k_W). The k/k_P was calculated by multiplying k/k_W by the reactivity of Trp to proline (k_W/k_P), measured with free tryptophan.

** The reactivity of Leu and Ile were measured with phenylalanine (Phe) as an internal control, resulting in k/k_F. The k/k_p was calculated by multiplying k/k_F by the reactivity of Phe to proline (k_F/k_P).

† Cystine here refers to the reactivity of free cystine, as opposed to Cys, which refers to the reactivity of cystine in the tripeptide Gly-Cys-Gly. Free cystine exists as a dimer in solution, which is representative of disulfide bonded cystine residues in a protein, while the peptide Gly-Cys-Gly, is a monomer, and is representative of free cystine residues in a protein.

One of the crucial considerations that must be taken when using the reactivity rates of free amino acids as shown above for calculating NPFs for analysis of proteins, is that the inherent reactivity of an amino acid found in a protein or peptide may not be identical to that of the free amino acid. In other words, in order to generate an accurate measurement of solvent accessibility of a particular amino-acid residue in a full protein or peptide, the amino acid in its sequence context must be accurately measured or estimated. [1] In other words, the reactivity of a particular amino-acid residue depends on which amino-acid it is *and* on which amino acids are near it in the protein sequence. The effects of sequence context on the reactivity of amino-acid residues in a polypeptide structure are quite complex and not fully known. [9] Nonetheless, NPFs calculated with amino-acid reactivity rates have been shown to be correlated with the solvent accessibility calculated from a protein-crystal structure. [1, 2]

6.3 Solvent accessibility

Solvent accessibility (SA) can be calculated from a static, three-dimensional protein structure generated using computer simulations, or experimentally, using NMR, cryo-EM, or crystallography. [10] The calculation typically involves a "ball-rolling" method, whereby a test sphere of a particular radius, usually set to 1.4Å, which is the approximate radius of a water molecule, is used as a probe to calculate the surface area of the protein molecule. [10] After calculating a surface area for a particular residue, the surface area is then normalized by the relative surface area (RSA). Normalization by RSA here is analogous to the normalization by the relative reactivity rate of each amino acid that is used in calculating the NPF, as described in Section 6.1. RSA values are typically derived from Gly-X-Gly peptides [10-13], where X is the amino acid of interest. After normalization, the resulting value is an absolute measure of solvent accessibility of a given residue in the static, three-dimensional protein structure.

6.4 Analysis of myoglobin

Now that a method has been established for calculating NPFs and comparing them to SA measurements from a known three-dimensional structure, testing was performed to determine whether this process could be successfully used with PLIMB-generated data—ultimately to show if PLIMB can be used to effectively measure SA. As was used by Sharp *et. al.*, [1] the protein myoglobin was used as a model protein for PLIMB analysis. Myoglobin is a relatively small protein, with primary sequence of about 150 amino acids, and a molecular weight of 17 kDa. [14] It is well characterized, and has a full-length crystal structure. [15] Benjamin Minkoff performed the following experimentation on myoglobin including sample preparation, PLIMB treatment, digestion, and mass spectrometry analysis.

6.4.1 PLIMB Treatment

A sample of equus (horse) heart myoglobin was prepared in 50 mM phosphate buffered solution (PBS) at a concentration of 30 μM. This was aliquoted into 50 μL volumes in 250 μL PCR tubes and exposed to PLIMB (see Section 3.3 for plasma parameters) for 0, 2.5, 5, and 10 seconds in duplicates. Immediately after exposure, 50 μL of 25 mM methionine in PBS was added to each sample to quench excess radicals.

6.4.2 Post-PLIMB sample handling

Immediately following PLIMB treatment and quenching, the samples were precipitated by adding 12 μL of trichloroacetic acid (TCA) and incubated on ice for 30 minutes. The protein was then pelleted by centrifugation at 16,000×g, 4 °C for 10 minutes. The supernatant[48] was removed, and the protein was washed by adding 100 μL of ice-cold acetone, vortexing, and incubating on ice for

[48] The supernatant here refers to the liquid lying above a solid residue after centrifugation.

10 minutes. The samples were again pelleted by centrifugation at 4 °C and 16,000×g for 10 minutes. The supernatant was removed, and the pellet was resolubilized into 8 M urea in 50 mM ammonium bicarbonate and then diluted to 1 M urea with 50 mM ammonium bicarbonate. A mixture of trypsin and lys-C proteases were added at a protease/protein mass ratio of 1:20. The samples were digested overnight for 16 hours at 37 °C. Following digestion, samples were acidified with neat formic acid to a final concentration of 1%. The samples were then desalted and concentrated using Omix C18 pipette tips (Agilent) using the following protocol: The tips were rinsed three times with 100 µL of neat acetonitrile, and then equilibrated by rinsing three times with 100 µL of 0.1% formic acid. The sample was then bonded to the tip by slowly running the sample through the tip 15 times, and then the tip was rinsed three times with 0.1% formic acid. The sample was then eluted with 100 µL of elution buffer (70% acetonitrile, 1% formic acid) into a fresh tube. Following elution, the samples were dried down via a SpeedVac vacuum concentrator. Dried-down peptides were resuspended into 0.1% formic acid at a concentration of 200 µg/ml, and 2 µL of peptides was injected into the LC/MS system for mass-spectral analysis (details below).

6.4.3 Mass-spectrometry analysis

Samples were analyzed using an Orbitrap Elite mass spectrometer with the following settings. A 60-minute chromatographic gradient from 2% to 50% acetonitrile with 0.1% formic acid was used for separation over a 3µM, 25cm Acclaim PepMap C18 column from ThermoFisher Scientific. Data-dependent acquisition was performed using MS1 parameters of 120K resolving power in the Orbitrap, a scan range of 350-1800 m/z, AGC target of 1×10^6. A top 20 method was used for data-dependent ion selection. MS2 parameters of charge state ≥ 2+ selection, a CID normalized collision energy of 35%, an activation Q of 25, an activation time of 10ms, an AGC value of 3×10^5, and an automatic scan range starting at 110 m/z in the ion trap with the scan rate set to “turbo.”

Dynamic exclusion of 12 seconds was used with a duration exclusion of 12 seconds after seeing an ion twice.

6.4.4 Data analysis

As was done in Chapter 4, the mass spectrometry data was analyzed using Protein Metrics software. The '.raw' data files were submitted to Byos® (Protein Metrics) for a database search using the sequence of Myoglobin, and automated generation of extracted ion chromatograms ('XIC') of the precursor ions. The peptides were identified using MS and MS/MS analysis, [16] and the following modifications were used in our search, noted by their molecular mass shift in Daltons (Da) and the amino acids which are modified noted by their single-letter abbreviation, were used in our search. Here, standard modifications are those that occur naturally, and are searched for by the software, but are not quantified as a footprinting modification during data analysis. PLIMB modifications are those that can be induced by plasma exposure and are quantified during data analysis.

Standard modifications:

Carbamidomethyl / +57.021464 Da @ C (fixed)

Gln->pyro-Glu / -17.026549 Da @ NTerm Q

Glu->pyro-Glu / -18.010565 Da @ NTerm E

Phospho / +79.966331 Da @ S, T

PLIMB modifications:

Oxidation / +15.994915 Da @ G, C, D, F, H, K, M, N, P, R, W, Y

Dioxidation / +31.989829 Da @ C, F, K, M, P, R, W, Y

Cys-Oxidation / +15.994915-57.021464 Da @ C

Nitro / +44.985078 Da @ W, Y

The peptides were identified using MS and MS/MS criteria using a 1% false discovery rate (FDR), and the proportion of plasma-modified species (marked as PLIMB-modifications above) were calculated based on extracted ion chromatogram (XIC) relative peak areas of modified versus unmodified peptides.

After peptide identification was completed, Protein Metrics Byologic software was used to identify and quantify PLIMB-induced modification at the amino-acid level. For each MS/MS hit, or spectrum that matches a particular peptide in the target protein, that is identified as a *modified* peptide, Byologic will automatically assign a specific residue in that peptide that is the most likely to contain the modification. This is often referred to as the "localization." It will also assign a measure of the confidence in the localization, called the Delta Mod Score. Based on the Delta Mod Scores, as well as manual analysis of the MS/MS spectrum, one can determine which MS/MS hits correspond to which specific residues in the peptide that are being modified. Modifications to a specific residue can then be matched to individual peaks on the extracted ion chromatogram (XIC) plot, allowing for quantification. As with peptide-level analysis, percent modification on a residue level is done by calculating the relative peak areas of modified peptides versus the unmodified peptides.

Fifteen unique amino-acid residues of myoglobin were identified to show significant levels of PLIMB-induced modification with high confidence in localization. Table 6-3 below shows a list of these identified residues, along with the following: 1) the sequence of the peptide containing the residue, 2) the start (labelled Start AA) and end (labelled End AA) position in the myoglobin sequence, 3) the PLIMB-induced modification and mass shift (labelled Mod. Name/Mass) that was identified, 4) the amino-acid species (labelled Mod. AA) of the residue, and 5) the position of the residue along the protein sequence (labelled Var. Pos. Protein). Note that some residues are

listed more than once, as they may have been identified both with both oxidation and dioxidation (doubly oxidized), and/or on different unique peptides. The total percent modification for each unique amino acid was calculated by summing the total percent modification over each peptide and for each modification, weighting dioxidation by twice that of oxidation, since it represents two oxidation events on a single residue.

Table 6-3: Amino acid residues of myoglobin identified and quantified for PLIMB-induced modification

Peptide Sequence	Start AA	End AA	Mod. Name/Mass (Da)	Mod. AAs	Var. Pos. Protein
GLSDGEWQQVLNVWGK	1	16	Dioxidation/31.9898	W	7
GLSDGEWQQVLNVWGK	1	16	Oxidation/15.9949	W	7
LSDGEWQQVLNVWGK	2	16	Dioxidation/31.9898	W	7
GLSDGEWQQVLNVWGK	1	16	Dioxidation/31.9898	W	14
GLSDGEWQQVLNVWGK	1	16	Oxidation/15.9949	W	14
LSDGEWQQVLNVWGK	2	16	Dioxidation/31.9898	W	14
VEADIAGHGQEVLIR	17	31	Oxidation/15.9949	D	20
VEADIAGHGQEVLIR	17	31	Oxidation/15.9949	H	24
VEADIAGHGQEVLIR	17	31	Dioxidation/31.9898	R	31
VEADIAGHGQEVLIR	17	31	Oxidation/15.9949	R	31
LFTGHPETLEK	32	42	Oxidation/15.9949	H	36
KHGTVVLTALGGILK	63	77	Dioxidation/31.9898	K	63
KHGTVVLTALGGILK	63	77	Oxidation/15.9949	K	63
HGTVVLTALGGILK	64	77	Oxidation/15.9949	H	64
HGTVVLTALGGILKK	64	78	Oxidation/15.9949	H	64
GHHEAELKPLAQSHATK	80	96	Oxidation/15.9949	H	82
GHHEAELKPLAQSHATK	80	96	Dioxidation/31.9898	K	87
GHHEAELKPLAQSHATK	80	96	Oxidation/15.9949	K	87
YLEFISDAIIHVLHSK	103	118	Oxidation/15.9949	F	106
HPGDFGADAQGAMTK	119	133	Dioxidation/31.9898	H	119
HPGDFGADAQGAMTK	119	133	Dioxidation/31.9898	M	131
HPGDFGADAQGAMTK	119	133	Oxidation/15.9949	M	131
ALELFR	134	139	Oxidation/15.9949	F	138
ALELFR	134	139	Oxidation/15.9949	R	139

6.4.5 NPF calculation

After generating percent modification values for fifteen individual amino acids of myoglobin for each sample, a linear-regression analysis was used to generate a percentage modification per unit of exposure time for each individual amino acid. This was done by taking the exposure times, 0, 2.5, 5, and 10 seconds, as the independent variable values, the average of the percent oxidation between replicates as the dependent variable values and using the "LINEST" function of Microsoft Excel to generate a linear regression using the least squares method. The slope of the linear function for each amino acid was taken to be the percentage modification per unit of exposure time. The slope was then divided by the relative reactivity (Table 6-2) of the particular amino acid being oxidized, as described in Equation (6-1), to calculate the NPF value. Note that the slope, *i.e.*, rate of oxidation, for each amino-acid residue was calculated using all plasma exposure times, as there did not appear to be any signs of protein unfolding or degradation at longer exposure times, as would be indicated by a sharp increase in modification after a given amount of plasma exposure. Rather, the rates of oxidation per unit of exposure time all appeared to be consistent throughout the 10-second range.

6.4.6 Solvent-accessibility analysis

The crystal structure of wild-type myoglobin [15] (Protein Data Bank file 1WLA) was used to generate solvent accessibility for each of myoglobin's amino acids. This was done using a program called GetArea, which is available for free as an online resource. [17] GetArea calculates information about the solvent accessibility of a protein using a three-dimensional crystal structure. [18] The program outputs the following information on each amino acid of a protein's sequence:

1. **Total area** – the total solvent accessible surface area of a particular amino-acid residue.

2. **Apolar area** – the total accessible area of the apolar atoms (those that have no electrical polarity) of a particular residue. [19]
3. **Backbone area** – the total solvent accessible area of a residue's amide[49] backbone. [17]
4. **Sidechain area** – the total solvent accessible area of a residue's functional group (side chain). [17]
5. **Ratio** – The ratio of a residue's side chain solvent accessible surface area to a "random coil" value. [17] The "random coil" value, also referred to as the residue surface area (RSA), [11] of a residue X is the average solvent-accessible surface area of X in the tripeptide Gly-X-Gly in an ensemble of 30 random conformations. [17] Hence, the "Ratio" is a dimensionless value of solvent accessibility of a particular residue normalized by its inherent solvent accessibility.
6. **In/Out** – Residues are marked as solvent exposed (out), if the "Ratio" value exceeds 50%, or as buried (in), if the ratio value is less than 20%. [17]

Because the "Ratio" value normalizes the solvent accessible surface area by the inherent accessibility, or RSA, of the amino acid side chain, it is analogous to the NPF calculation, because it is a measure of plasma modification per unit of time normalized by the inherent reactivity rate of the amino-acid side chain. The Ratio value was therefore used as the value for comparing solvent accessibility of the crystal structure with that generated by PLIMB data.

[49] Amide refers to a group of compounds characterized by the N-C=O moiety. A chain of amides makes up a protein's "backbone."

6.4.7 Results

The Ratio value versus the calculated NPF value for each of the fifteen amino acids that were quantified for PLIMB-induced modification, were graphed, and can be seen in Figure 6-1. The amino acids are grouped by their relative reactivity rates as follows:

Table 6-4: Amino acid groupings based on reactivity rates relative to proline

Group name	Relative reactivity range	Amino acids included
M & W	> 15	M, W
High	5 – 15	Y, F, C, H, L
Med.	1 – 5	I, R, K, V, S, T, P
Low	0.25 – 1	E, Q, N, D
Not included	< 0.25	A, G

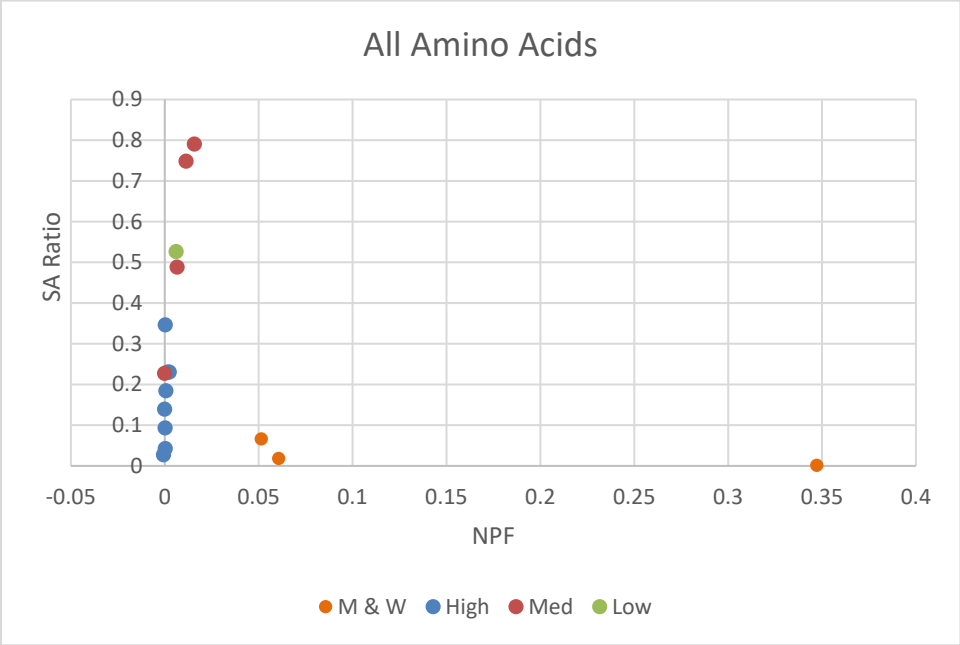


Figure 6-1: Solvent Accessibility Ratio vs. calculated NPF for all amino acids

As can be seen in Figure 6-1, there does not appear to be a distinct correlation between the calculated SA and NPF values. However, all but the methionine and tryptophan (M & W) amino

acids appear to fall into a distinct region of the graph with linearity. Sharp. *et al*. also noticed in their work that the amino acids with the highest reactivity rate, notably the sulfur-containing residues (methionine and cysteine), did not show correlation between their calculated NPF and SA. [1] This because their extremely high reactivity rates resulted in high levels of background oxidation before, during, and after hydroxyl exposure, which results in artificially high levels of oxidation after normalization when compared to other amino acids. [1] For this reason, they excluded them from their solvent accessibility analysis. Here, we see a similar trend. Both methionine and tryptophan seem to show artificially high levels of oxidation which are not correlated with their solvent accessibility, presumably due to their extremely high relative reactivity rates. Note that a similar result may have been observed with cysteine as well, but no cystines were identified and quantified in this experiment. Figure 6-2 shows a similar graph of calculated SA ratio versus NPF, but with the "M & W" residues excluded. Here, a much better correlation can be seen, as the linear fit with the equation shown below, shows an R^2 value of 0.856.

$$y = 60.14x \tag{6-2}$$

The linear relationship shown here can be used to calculate an estimated value of SA based on the calculated NPF value.

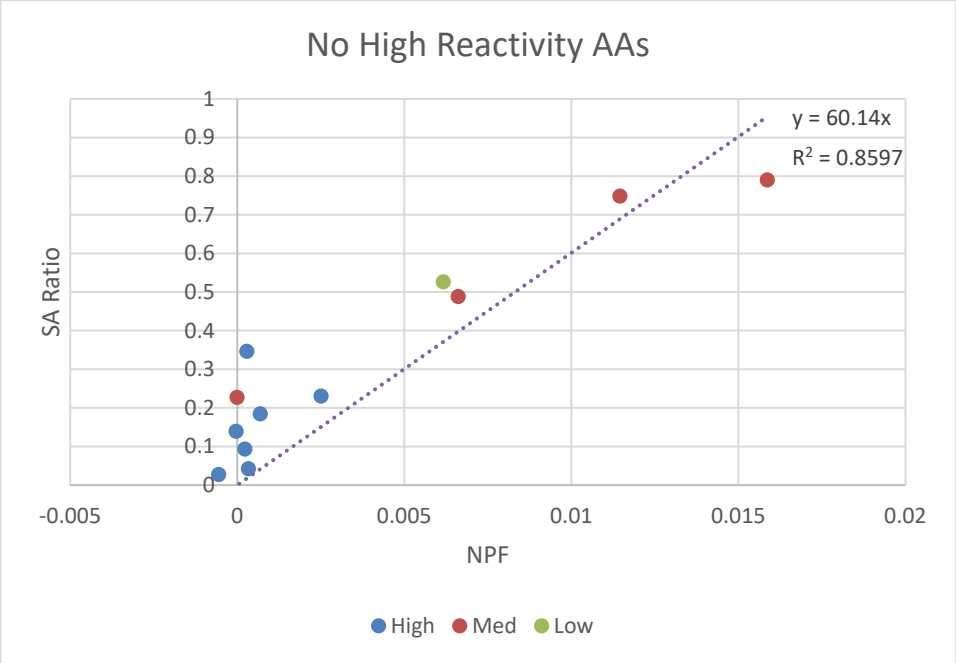


Figure 6-2: Solvent Accessibility Ratio vs. calculated NPF for all amino acids except methionine (M) and Tryptophan (W)

After the linear relationship between NPF and SA ratio was established, solvent accessible surface areas (SASA) were calculated for all of the twelve unique amino acids that were used in the model. This was done by first calculating a SA ratio from the NPF value for each amino-acid residue using Equation (6-2). Then, an SASA was calculated for each residue by multiplying the SA ratio by the corresponding RSA value (see Section 6.3) and reassigning negative SASA values to 0. The PLIMB-generated SASA values were then compared to those generated from the crystal structure, and the results can be seen below in Table 5. A total root-mean-squared (RMS) error was calculated to be 24.98 $Å^2$, which is comparable to the errors reported by Sharp *et al.* for their solvent accessibility models. [1]

Table 5: Solvent accessible surface area (SASA) comparison

Residue	Crystal SA (Å2)	PLIMB SA (Å²)	Error (Å²)
D20	59.44	42.28	-17.16
H24	6.49	3.17	-3.32
R31	95.40	78.43	-16.97
H36	53.49	2.72	-50.77
K63	129.96	158.24	28.29
H64	35.56	23.55	-12.01
H82	4.17	0.00	-4.17
K87	123.05	114.26	-8.78
F106	25.03	0.00	-25.03
H119	28.45	6.50	-21.95
F138	16.75	2.52	-14.23
R139	44.38	0.00	-44.38
		Total RMS Error (Å²):	**24.98**

6.5 Conclusion

The work described in this chapter establishes a method for extracting information on a protein's solvent accessibility using PLIMB data alone. In summary, the process is as follows:

1. **PLIMB exposure**: Expose samples of the target protein to PLIMB at a range of hydroxyl-radical doses. Samples should be prepared in replicates of at least two, with a higher number of replicates likely increasing the confidence in data by accounting for variability in oxidation due both to PLIMB exposure itself, and to oxidation induced in the sample-handling processes or storage.
2. **Sample handling and MS analysis**: Perform a protein digestion which yields a mass spectrometry coverage that includes all regions of interest, with peptides that are of

reasonable length (5 to 15 amino acids) so as to be conducive to single amino-acid level analysis.

3. **Data analysis**: Identify and quantify levels of oxidation on an amino-acid level for each replicate sample exposed to each PLIMB dose. Only amino acids that have high confidence in both localization (which amino acid of a peptide contains a specific modification) and quantification (identifiable XIC peaks associated with a specific amino acid modification) should be included in further analysis.
4. **NPF calculation**: Calculate a linear regression slope for each amino acid, taking the PLIMB doses (exposure times) as the independent variable values, and the percent oxidation at each dose as the dependent variable values. Then, normalize the slope by the relative reactivity rate of that particular amino acid.
5. **SA calculation**: Calculate an estimated solvent accessibility ratio value from the calculated NPF values using a linear regression formula.
6. **SASA calculation**: Calculate the SASA from the SA ratio by multiplying by the RSA (see Section 6.3) of each amino acid.

SASA values can then be used as a measure of solvent accessibility that is independent of the chemical composition of the region of protein being analyzed or of the method being used to generate solvent accessibility data.

6.6 References

[1] B. Xie, A. Sood, R. Woods, and J. Sharp, "Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection," *Sci Rep,* vol. 7, no. 1, pp. 4552-4552, 2017, doi: 10.1038/s41598-017-04689-3.

[2] W. Huang, Krishnakumar M. Ravikumar, Mark R. Chance, and S. Yang, "Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis," *Biophysical Journal,* vol. 108, no. 1, pp. 107-115, 2015/01/06/ 2015, doi: https://doi.org/10.1016/j.bpj.2014.11.013.

[3] X. R. Liu, M. M. Zhang, and M. L. Gross, "Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications," *Chemical Reviews,* vol. 120, no. 10, pp. 4355-4454, 2020/05/27 2020, doi: 10.1021/acs.chemrev.9b00815.

[4] D. Hambly and M. Gross, "Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin," *International Journal of Mass Spectrometry,* vol. 259, no. 1, pp. 124-129, 2007/01/01/ 2007, doi: https://doi.org/10.1016/j.ijms.2006.08.018.

[5] W. Shi *et al.*, "Structural Characterization of HIV gp41 with the Membrane-proximal External Region *," *Journal of Biological Chemistry,* vol. 285, no. 31, pp. 24290-24298, 2010, doi: 10.1074/jbc.M110.111351.

[6] Y. Sheng, J. Capri, A. Waring, J. S. Valentine, and J. Whitelegge, "Exposure of Solvent-Inaccessible Regions in the Amyloidogenic Protein Human SOD1 Determined by Hydroxyl Radical Footprinting," *Journal of the American Society for Mass Spectrometry,* vol. 30, no. 2, pp. 218-226, 2019/02/01 2019, doi: 10.1021/jasms.8b05938.

[7] G. Xu and M. R. Chance, "Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting," *Analytical chemistry,* vol. 77, no. 14, pp. 4549-4555, 2005.

[8] G. Xu and M. R. Chance, "Hydroxyl radical-mediated modification of proteins as probes for structural proteomics," *Chemical reviews,* vol. 107, no. 8, pp. 3514-3543, 2007.

[9] J. S. Sharp and K. B. Tomer, "Effects of Anion Proximity in Peptide Primary Sequence on the Rate and Mechanism of Leucine Oxidation," *Analytical Chemistry,* vol. 78, no. 14, pp. 4885-4893, 2006/07/01 2006, doi: 10.1021/ac060329o.

[10] A. Shrake and J. A. Rupley, "Environment and exposure to solvent of protein atoms. Lysozyme and insulin," *Journal of Molecular Biology,* vol. 79, no. 2, pp. 351-371, 1973/09/15/ 1973, doi: https://doi.org/10.1016/0022-2836(73)90011-9.

[11] M. Z. Tien, A. G. Meyer, D. K. Sydykova, S. J. Spielman, and C. O. Wilke, "Maximum allowed solvent accessibilites of residues in proteins," (in eng), *PLoS One,* vol. 8, no. 11, pp. e80635-e80635, 2013, doi: 10.1371/journal.pone.0080635.

[12] G. Rose, A. Geselowitz, G. Lesser, R. Lee, and M. Zehfus, "Hydrophobicity of amino acid residues in globular proteins," *Science,* vol. 229, no. 4716, pp. 834-838, 1985, doi: 10.1126/science.4023714.

[13] S. Miller, J. Janin, A. M. Lesk, and C. Chothia, "Interior and surface of monomeric proteins," *Journal of Molecular Biology,* vol. 196, no. 3, pp. 641-656, 1987/08/05/ 1987, doi: https://doi.org/10.1016/0022-2836(87)90038-6.

[14] J. Zaia, R. S. Annan, and K. Biemann, "The correct molecular weight of myoglobin, a common calibrant for mass spectrometry," (in eng), *Rapid Commun Mass Spectrom,* vol. 6, no. 1, pp. 32-6, Jan 1992, doi: 10.1002/rcm.1290060108.

[15] R. Maurus *et al.*, "A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket," *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology,* vol. 1341, no. 1, pp. 1-13, 1997/08/15/ 1997, doi: https://doi.org/10.1016/S0167-4838(97)00064-2.

[16] B. Minkoff, J. Blatz, F. Choudhury, D. Benjamin, J. Shohet, and M. Sussman, "Plasma-Generated OH Radical Production for Analyzing Three- Dimensional Structure in Protein Therapeutics," *Sci Rep,* vol. 7, no. 1, pp. 12946-12946, 2017, doi: 10.1038/s41598-017-13371-7.

[17] S. Negi, H. Zhu, R. Fraczkiewicz, and W. Braun. "Solvent Accessible Surface Areas, Atomic Solvation Energies, and Their Gradients for Macromolecules." Sealy Center for Structural Biology. http://curie.utmb.edu/getarea.html (accessed February 5, 2021).

[18] R. Fraczkiewicz and W. Braun, "Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules," *Journal of Computational Chemistry,* vol. 19, no. 3, pp. 319-333, 1998, doi: https://doi.org/10.1002/(SICI)1096-987X(199802)19:3<319::AID-JCC6>3.0.CO;2-W.

[19] B. von Freyberg, T. J. Richmond, and W. Braun, "Surface Area Included in Energy Refinement of Proteins: A Comparative Study on Atomic Solvation Parameters," *Journal of Molecular Biology,* vol. 233, no. 2, pp. 275-292, 1993/09/20/ 1993, doi: https://doi.org/10.1006/jmbi.1993.1506.

Chapter 7: Conclusions and Future Directions

PLIMB is a new technology for performing a protein structural technique known as hydroxyl-radical protein footprinting (HRPF) that has some distinct advantages over currently available HRPF technologies. Because of the characteristics of the technology and the data it can generate, PLIMB may prove to be uniquely qualified to fill the specific role of mapping the epitopes, or binding locations, of protein therapeutics to their target proteins. Epitope mapping is a critical step in the development of protein therapeutics; however, current methods for performing epitope mapping have significant drawbacks. PLIMB addresses many of the drawbacks of technologies currently available for epitope mapping.

The hypothesis of this research is to develop methods and systems to allow the use of PLIMB to map the epitopes of protein therapeutics. The tasks needed to prove this hypothesis include 1) developing an experimental workflow that can be followed for using PLIMB technology to map epitopes, 2) augmenting the PLIMB technology to include a system for reducing data variability between replicate samples, and 3) developing a method for extracting solvent accessibility measurements from PLIMB data. This chapter provides a summary of the significant results and findings that came from this research and recommendations for future work.

7.1 Epitope mapping with PLIMB

The epitope of a monoclonal antibody (mAb) binder to human thrombin was mapped using PLIMB following a protocol that was developed for other HRPF techniques [1] as a guide. The process involved preparing samples of thrombin with and without a bound antibody and exposing each to PLIMB for 1 and 3 seconds. Following PLIMB exposure, samples were quenched with methionine and prepared for mass spectrometry analysis. Mass-spectrometry preparation involved

precipitation, digestion, and solid-phase extraction. Mass-spectrometry analysis was performed, and the resulting data was analyzed with Protein Metrics Byos software for identification of peptides and quantification of percentage oxidation. A statistical t-test was then performed to determine statistically significant differences in oxidation between samples containing thrombin alone and those containing antibody-bound thrombin for both 1 and 3-second exposed sample sets. The results identified six peptides that showed statistically significant changes in oxidation after 1 second of PLIMB exposure. Interestingly, only three showed statistically significant changes after 3 seconds of exposure. These results suggest that the epitope region was located in two peptides that showed protection, or a decrease in oxidation upon antibody binding, at both 1 and 3 seconds of exposure. These results agree with previously published epitope results that were found using hydrogen deuterium exchange (HDX) [2] and fast-photochemical oxidation of proteins (FPOP). [1]

7.2 Feedback-control system

In order to address issues of variability in oxidation between replicate PLIMB-exposed samples, a feedback-control system was developed. First, a method for measuring hydroxyl-radical generation was established. It involved using a monochromator to measure 309-nm light, which is known to be emitted by hydroxyl radicals. [3-8] The emitted light was measured with a photomultiplier tube (PMT). Measurements of the plasma current and the PMT signal showed that plasma discharges, indicated by a narrow (0.1 μs) burst of plasma current, proceeded a corresponding burst of 309-nm light.

An analog-to-digital converter (ADC) was then used to sample the PMT signal at 500 MS/s to capture the overall intensity and waveshape of each individual 309-nm light burst. A field-programmable gate array (FPGA) controller was used to perform a summation of the sampled

signal over an entire PLIMB exposure time in order to obtain a value which is, in theory, proportional to the hydroxyl-radical dose. The FPGA was programmed to send a summed signal to an external computer running a LabView program that controlled the output of the high-voltage signal driving the plasma. The LabView program was configured so that samples could either be exposed to a particular plasma exposure time, or a particular ADC value corresponding to a given hydroxyl dose. This system was thus configured as a feedback-control system, that would take in hydroxyl-radical generation data and control the overall hydroxyl dose to which a sample was exposed.

Testing of this system was first done by exposing samples of coumarin to both 2 and 10-kHz generated plasma. A linear relationship between oxidized coumarin, measured by fluorescence, and the FPGA-summed ADC value was seen for both frequencies. The relationship also held after combining data from both the 2 and 10-kHz plasmas, suggesting a relationship between ADC values and hydroxyl dose that was independent of plasma parameters. Further testing was done by exposing samples of methionine and analyzing the degree of oxidation with mass spectrometry. Two sets of samples were prepared: one set was exposed to PLIMB for 20 seconds, and the other was exposed up to a given ADC value. For each exposure, the exact ADC value and exposure time were recorded. The resulting data showed a decrease in the coefficient of variation (C.V.) of percent oxidation between replicate samples from 35.9% to 11.1% when exposed to ADC values rather than plasma exposure times.

Although these results are promising, the system was improved to further reduce variability. This was done by replacing the monochromator with a 310-nm UV-bandpass filter, allowing for a much stronger signal at the PMT which reduced the signal-to-noise ratio. A test like that of the previous setup was performed. Two sets of coumarin samples were prepared. One set

was exposed to a 15-second plasma exposure time, and the second to a fixed ADC value. The degree of oxidation was measured with mass spectrometry and the resulting data showed a decrease in percent oxidation C.V. from 14.9% to 5.53% This is a significant improvement over the original monochromator setup.

Lastly, self-oxidation due to sample storage was tested as another possible source of data variability. Samples of coumarin were prepared and exposed to PLIMB for 0 and 15 seconds and then stored at 25, -4, -20, and -80° C. Triplicate samples were analyzed for fluorescence after 0, 3, 6, and 9 days of storage. It was found that samples should ideally be stored at -20° C for less than 3 days to reduce the potential for variability due to self-oxidation after PLIMB exposure.

7.3 Extracting solvent accessibility

Although the raw data from PLIMB, *i.e.,* percent oxidation of regions of a protein, has significant utility for assessing structural differences of a protein that is exposed to different environments, such as the presence or absence of a bound antibody in the case of epitope mapping, these results cannot be directly compared to different regions of the same or a different protein. This is because the relative reactivity of the various regions can vary over orders of magnitude and will therefore show a markedly different oxidative response to a particular dose of exposed hydroxyl radicals. To overcome this issue, a method for extracting solvent accessibility from PLIMB data was established. Solvent accessibility, unlike percentage oxidation, is a measurement that is independent of the chemical composition of the region being analyzed or the technology/method of analysis. [9]

Samples of the protein myoglobin were prepared and exposed to PLIMB for 0, 2.5, 5, and 10 seconds. The samples were then prepared and analyzed by mass spectrometry. The mass-spectrometry data was analyzed with Protein Metrics Byos software for identification of peptides

and quantification of percentage oxidation on an amino-acid level. A rate of oxidation per unit of PLIMB exposure time was established for each of the 15 individually identified amino acids by linear regression. The oxidation rates were normalized by the relative reactivity of the amino acid, resulting in a calculated normalized protection factor (NPF). The NPF values for each amino acid were compared to solvent accessibility measurements generated from a published three-dimensional structure of myoglobin, [10] and the results showed a linear relationship between the solvent accessibility measurements calculated from PLIMB data and from the crystal structure when including all amino acids except the most reactive, namely methionine and tryptophan. This linear relationship provides a mathematical model for converting NPF values found through PLIMB experimentation into solvent accessibility measurements. Using this model, solvent accessible surface areas (SASAs) were calculated from PLIMB-generated NPF values and were compared to SASA values calculated from the crystal structure. Results showed a total root-mean-squared (RMS) error of 24.98 $Å^2$, which is comparable to the errors reported SA models reported by other HRPF researchers. [9]

7.4 Suggestions for future work

7.4.1 Optimization of experimental conditions

As mentioned in Chapter 4, the experimental conditions that were varied to map the epitope of antithrombin mAb to thrombin were 1) the presence or absence of a bound antibody, and 2) the plasma exposure time. Varying these two factors proved to be sufficient for experimentally mapping the epitope using PLIMB. However, other experimental conditions may have been varied as well. In other words, rather than a two-factor analysis design, a multi-factor or fractional-factorial design could have been used to achieve higher confidence in the statistical significance

in the differences in oxidation between the two conditions of primary interest: the presence or absence of a bound antibody.

It is also not known whether the protocol described in Chapter 4 would be effective for mapping the epitopes of all antibody/antigen pairs. The mechanism of antibody binding as well as the various binding properties such as affinity can vary greatly between antibodies. [11-13] Therefore, detecting statistically significant levels of protection for all antibody/antigen pairs using the same protocol may not be possible. A multi-factorial or fractional-factorial design may be beneficial for improving the overall resolution that is achievable with a PLIMB epitope mapping experiment. The variables that may be included in a fractional-factorial design are as follows:

1. **Plasma parameters**: The parameters associated with the PLIMB plasma are known to influence the production of hydroxyl radicals, [14] and are likely to influence structural aspects of the protein in solution as well. These parameters include but are not necessarily limited to 1) the magnitude of the applied voltage, 2) the DC offset of the applied voltage, 3) the waveshape of the applied voltage, 4) the frequency of the applied voltage, and 5) the spacing between the high voltage electrode and the surface of the liquid solution.
2. **Sample conditions**: Although there are some technical limitations on the sample conditions that may be used in a PLIMB experiment, there are a significant number of factors which may have a significant effect on the outcome of a PLIMB experiment. These include, but are not limited to: 1) the protein concentration, 2) the buffer conditions including concentration and composition, 3) the inclusion of a radical scavenger, as is done with many FPOP experiments, [15, 16] which absorbs hydroxyl radicals in solution, thus changing the dynamics of their interactions with the sample

protein molecules, 4) the solution temperature during exposure, 5) the sample volume, and 6) the molar ratio between antibody and antigen.

3. **Sample-handling procedures**: The sample handling procedures used to prepare PLIMB-exposed samples for mass-spectrometry analysis is a critical step in the epitope-mapping process. It involves several steps, as outlined in Section 4.1.3, and each of these steps may be varied in order to achieve improved results from a PLIMB epitope-mapping experiment. This is explored in further detail in the following section.

7.4.2 Optimization of sample-handling procedures

The protocol that was used to prepare PLIMB-exposed samples for mass-spectrometry analysis described in Chapter 4 proved to be effective based on the following metrics: 1) The peptide samples resulting from the sample-handling protocol were indeed compatible with a liquid-chromatography mass spectrometer and did not show signs of contamination or impurities which would degrade the quality of the mass spectrometry data. 2) The resulting mass-spectrometry data showed 95% sequence coverage. This is a suitable level of coverage for a protocol that only included a single protease for protein digestion. 3) Significant levels of peptide-level oxidation were seen, and statistically significant changes in oxidation between samples of different conditions were able to be detected.

However, the degree to which the sampling protocol itself induces oxidation in the samples is not yet known. Oxidation produced throughout these steps is unwanted, as it is indistinguishable from PLIMB-induced oxidation, and thus has the potential to decrease the effective signal-to-noise ratio of the PLIMB data. In addition, because oxidation resulting from sample handling occurs after exposure, the feedback-control system and minimization of storage self-oxidation described

in Chapter 5 would not be effective in reducing variability in the oxidation between replicate samples caused in these steps.

A reasonable extension of the work described in Chapter 4 would be to analyze the degree to which each of the procedures used throughout the mass-spectrometry preparation protocol induces a significant level of oxidation in the protein samples. For example, the first step after PLIMB exposure is precipitation. Precipitation of the protein sample involves crashing the protein out of solution and removing the supernatant[50] solvent. This procedure thus exposes the protein directly to oxygen-rich air, which may cause a significant amount of protein oxidation. If this step were isolated, tested, and was found to indeed contribute a significant amount of oxidation, an alternative approach would need to be explored. Alternatives to any of the sample handling procedures would need to be developed to reduce unwanted oxidation, while also maintaining the performance of the overall protocol, as described by the metrics in the previous paragraph.

7.5 Closing statement

In closing, PLIMB is a technology that shows significant potential to fill the role of epitope mapping of protein therapeutic candidates. The PLIMB instrument has been shown to generate hydroxyl radicals to oxidize solvent-accessible regions of proteins in solution. A process has been developed to use PLIMB to map epitopes, data variability has been reduced through a feedback-control system, and a data analysis protocol has been established for measuring solvent accessibility directly with PLIMB. This thesis and its associated work are strong evidence that PLIMB can be successfully used for epitope mapping.

[50] The supernatant here refers to the liquid lying above a solid residue after centrifugation.

7.6 References

[1] L. M. Jones, J. B Sperry, J. A Carroll, and M. L. Gross, "Fast photochemical oxidation of proteins for epitope mapping," *Analytical chemistry,* vol. 83, no. 20, p. 7657, 2011, doi: 10.1021/ac2007366.

[2] A. Baerga-Ortiz, C. A. Hughes, J. G. Mandell, and E. A. Komives, "Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein," *Protein Science,* vol. 11, no. 6, pp. 1300-1308, 2002, doi: https://doi.org/10.1110/ps.4670102.

[3] A. Pankaj *et al.*, "Generation mechanism of hydroxyl radical species and its lifetime prediction during the plasma- initiated ultraviolet (UV) photolysis," *Scientific Reports,* vol. 5, no. 1, 2015, doi: 10.1038/srep09332.

[4] M. Sun *et al.*, "Emission spectra of hydroxyl radical generated in air corona discharge," *Guang pu xue yu Guang pu fen xi= Guang pu,* vol. 25, no. 1, pp. 108-112, 2005.

[5] H.-P. Dorn, R. Neuroth, and A. Hofzumahaus, "Investigation of OH absorption cross sections of rotational transitions in the A 2∑+,>(υ'= 0>)← X 2 Π,>(υ ″= 0>) band under atmospheric conditions: Implications for tropospheric long-path absorption measurements," *JGR,* vol. 100, no. D4, pp. 7397-7409, 1995.

[6] G. Dieke and H. Crosswhite, "The ultraviolet bands of OH fundamental data," *Journal of Quantitative Spectroscopy and Radiative Transfer,* vol. 2, no. 2, pp. 97-199, 1962.

[7] Y. H. Kim *et al.*, "Measurement of reactive hydroxyl radical species inside the biosolutions during non-thermal atmospheric pressure plasma jet bombardment onto the solution," *Plasma Chemistry and Plasma Processing,* vol. 34, no. 3, pp. 457-472, 2014.

[8] Y. Hong *et al.*, "Measurement of hydroxyl radical density generated from the atmospheric pressure bioplasma jet," *Journal of Instrumentation,* vol. 7, no. 03, p. C03046, 2012.

[9] B. Xie, A. Sood, R. Woods, and J. Sharp, "Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection," *Sci Rep,* vol. 7, no. 1, pp. 4552-4552, 2017, doi: 10.1038/s41598-017-04689-3.

[10] R. Maurus *et al.*, "A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket," *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology,* vol. 1341, no. 1, pp. 1-13, 1997/08/15/ 1997, doi: https://doi.org/10.1016/S0167-4838(97)00064-2.

[11] I. A. Wilson and R. L. Stanfield, "Antibody-antigen interactions," *Current opinion in structural biology,* vol. 3, no. 1, pp. 113-118, 1993.

[12] N. Sinha, S. Mohan, C. A. Lipschultz, and S. J. Smith-Gill, "Differences in electrostatic properties at antibody–antigen binding sites: implications for specificity and cross-reactivity," *Biophysical journal,* vol. 83, no. 6, pp. 2946-2968, 2002.

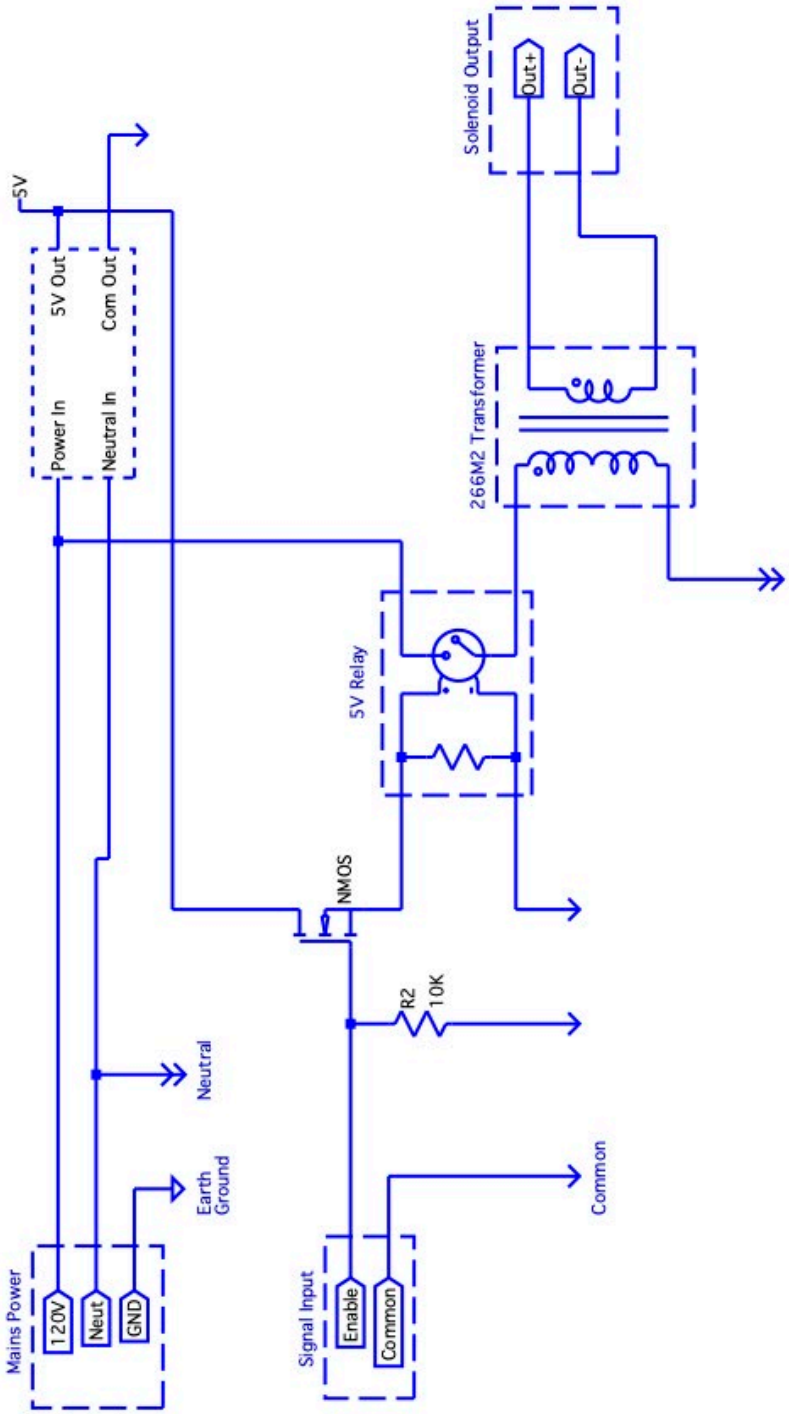
[13] J. Maynard and G. Georgiou, "Antibody engineering," *Annual review of biomedical engineering,* vol. 2, no. 1, pp. 339-376, 2000.

[14] J. M. Blatz, D. Benjamin, F. A. Choudhury, B. B. Minkoff, M. R. Sussman, and J. L. Shohet, "Effect of frequency and applied voltage of an atmospheric-pressure dielectric-barrier discharge on breakdown and hydroxyl-radical generation with a liquid electrode," *Journal of Vacuum Science & Technology A,* vol. 38, no. 4, p. 043001, 2020, doi: 10.1116/6.0000125.

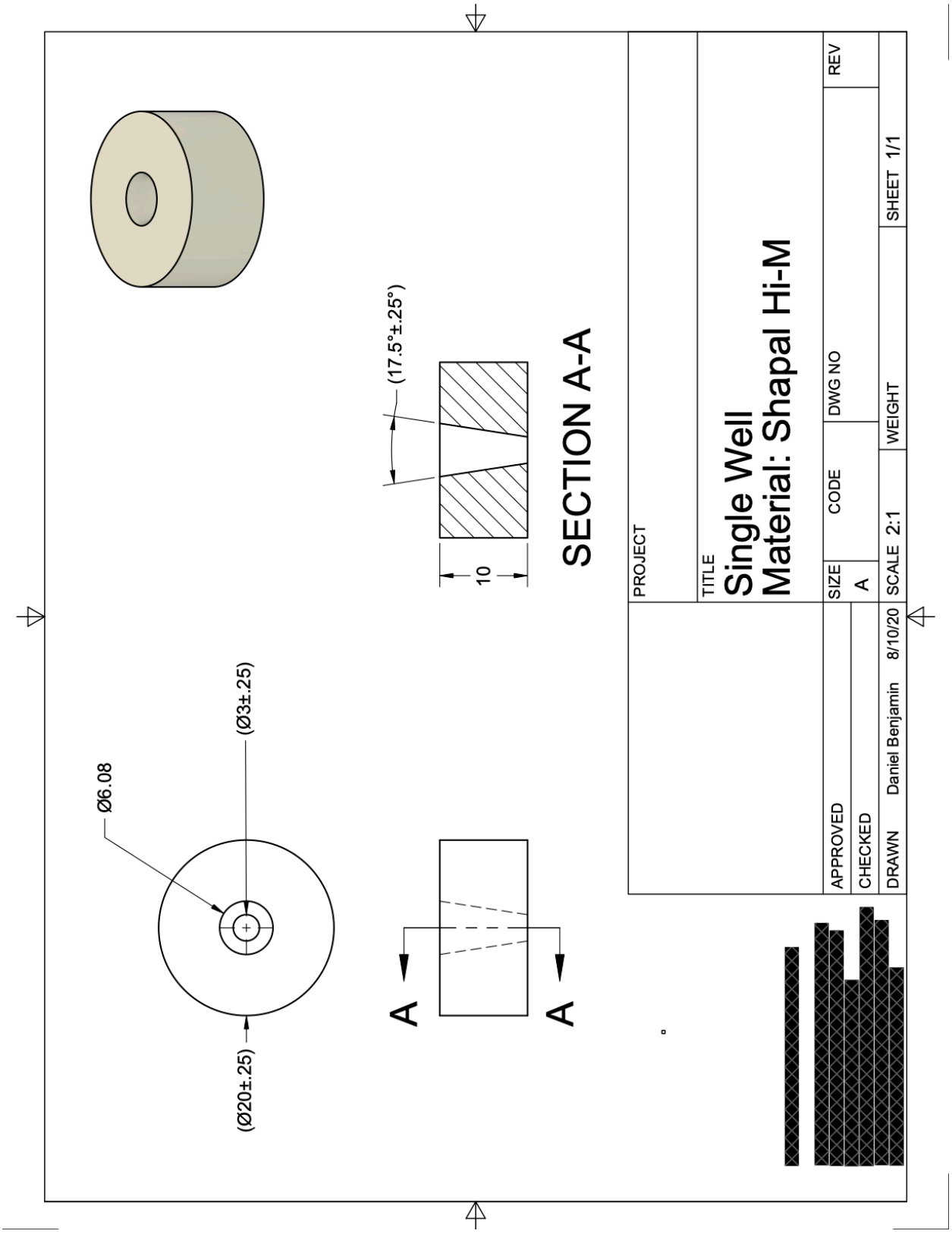
[15] X. R. Liu, M. M. Zhang, and M. L. Gross, "Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications," *Chemical Reviews,* vol. 120, no. 10, pp. 4355-4454, 2020/05/27 2020, doi: 10.1021/acs.chemrev.9b00815.

[16] B. Niu *et al.*, "Incorporation of a Reporter Peptide in FPOP Compensates for Adventitious Scavengers and Permits Time-Dependent Measurements," *Journal of The American Society for Mass Spectrometry,* vol. 28, no. 2, pp. 389-392, 2017/02/01 2017, doi: 10.1007/s13361-016-1552-4.

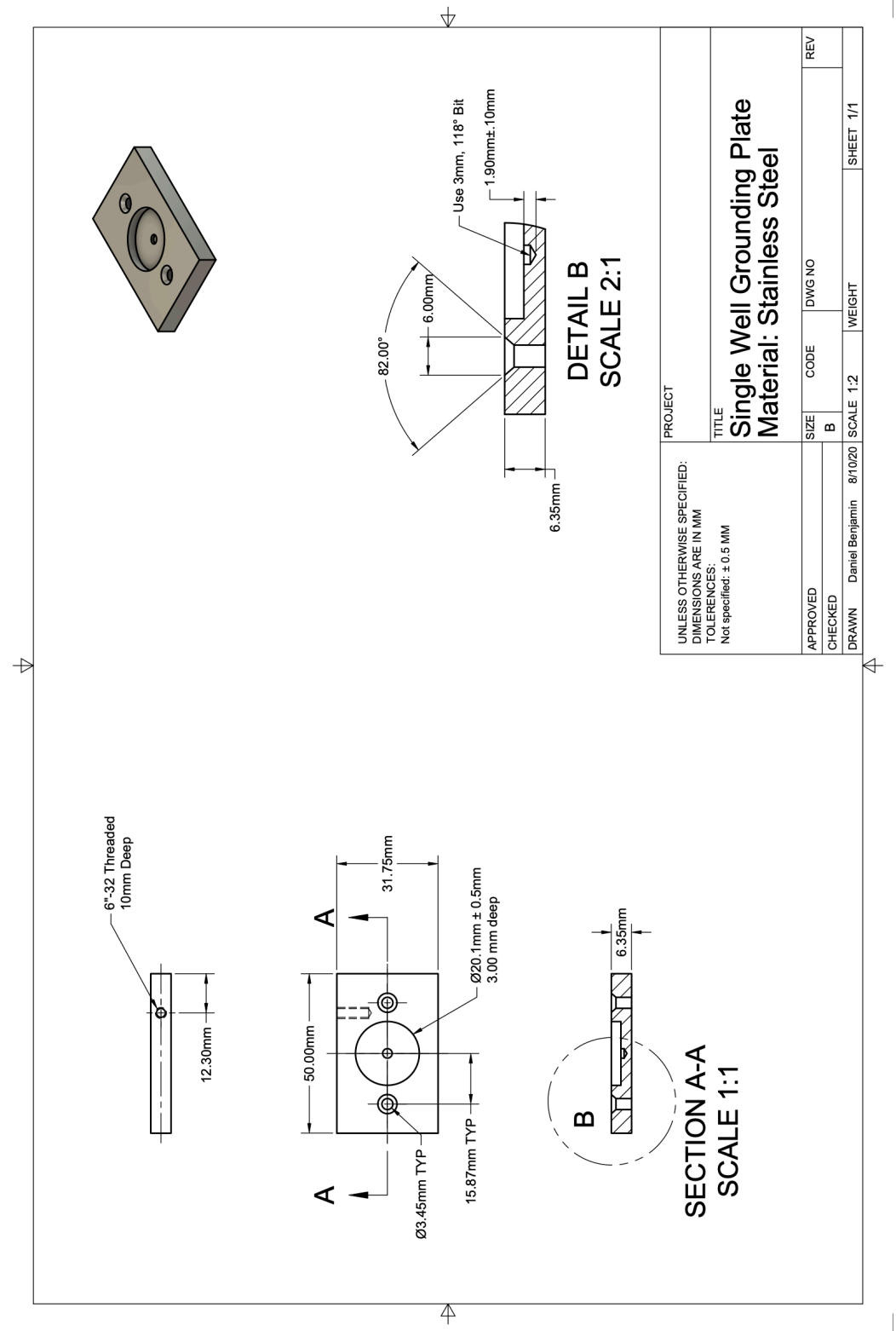
Appendix A: Solenoid switch circuit diagram



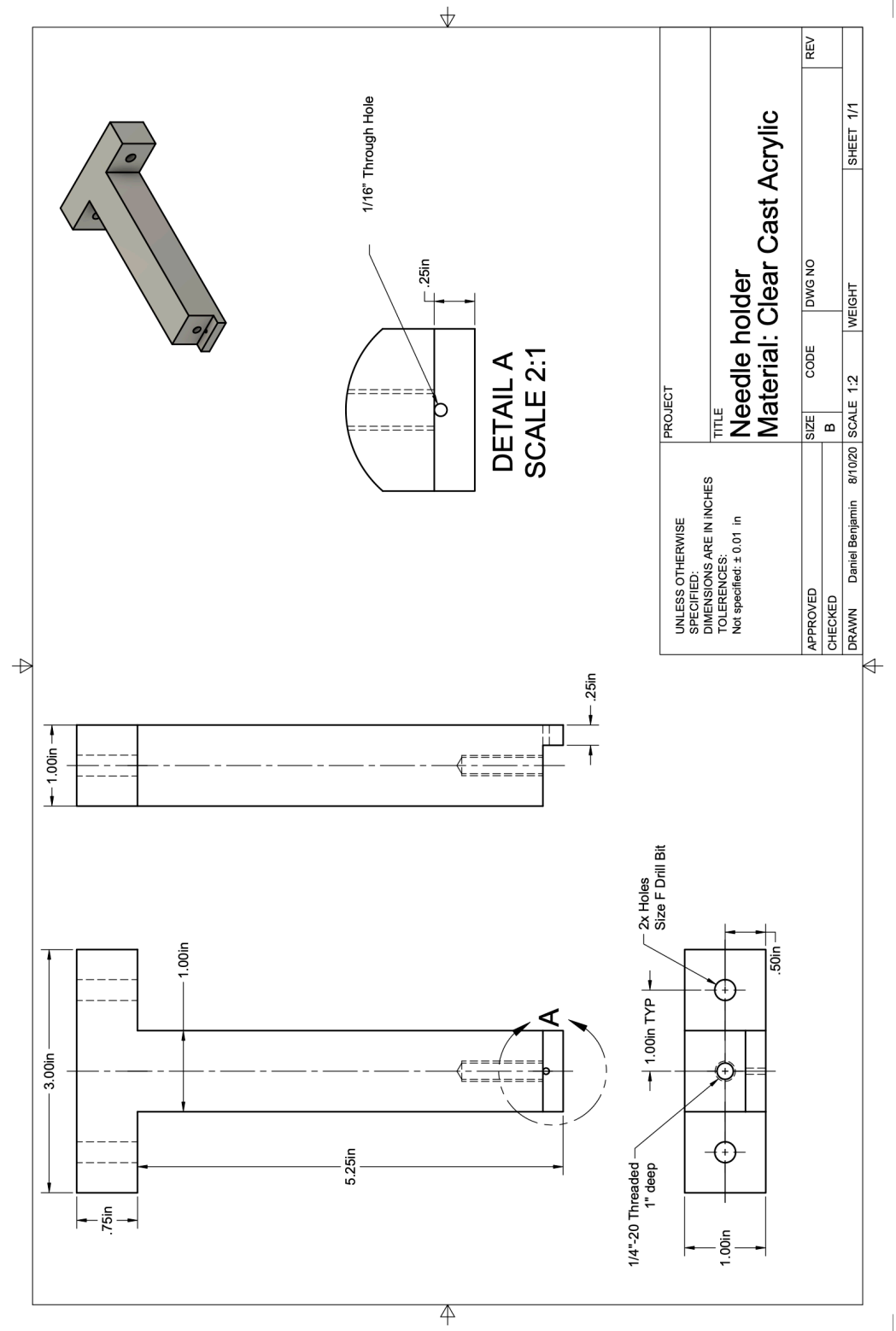
Appendix B: Single-well ceramic sample holder technical drawing



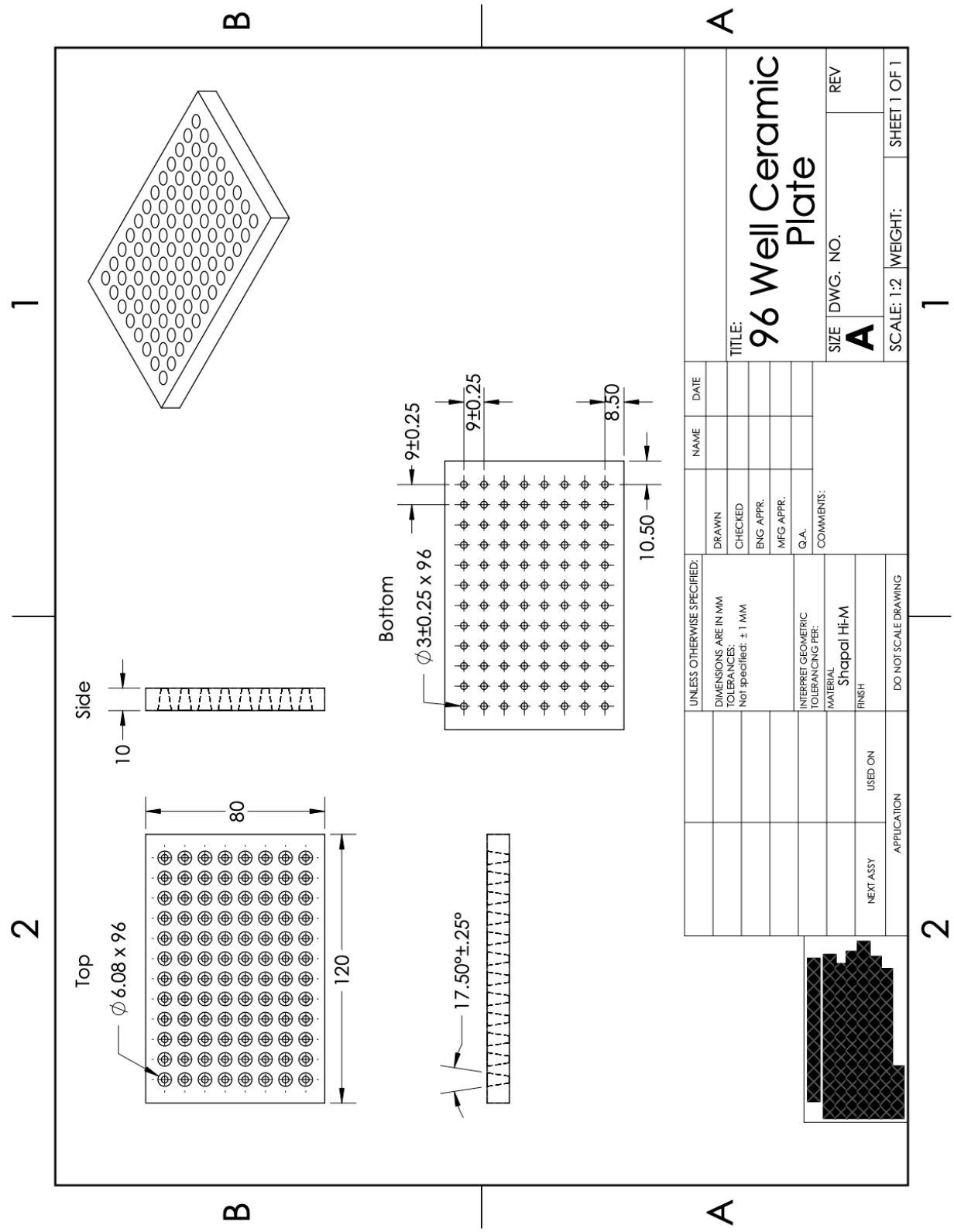
Appendix C: Single-well grounding plate technical drawing



Appendix D: Needle holder technical drawing



Appendix E: Ninety-six well ceramic sample holder



Appendix F: Photomultiplier tube amplifier delay

As was shown in Figure 3-20 and Figure 5-4, there appears to be a ~200 ns delay between a plasma discharge, as measured by a spike in plasma current, and the corresponding rise in photomultiplier tube (PMT) current induced by 309-nm light. This phenomenon was hypothesized to be caused by a delay in the PMT amplifier circuit, which is used to convert current from the PMT into an output voltage. This hypothesis was tested by removing the PMT from the amplifier circuit and applying a 1 V, 500 kHz square wave pulse with a signal generator through a 1 MΩ resistor to the input of the amplifier as shown in Figure F-1. The output of the amplifier circuit, as well as the voltage across the 1 MΩ resistor, were both measured with an oscilloscope.

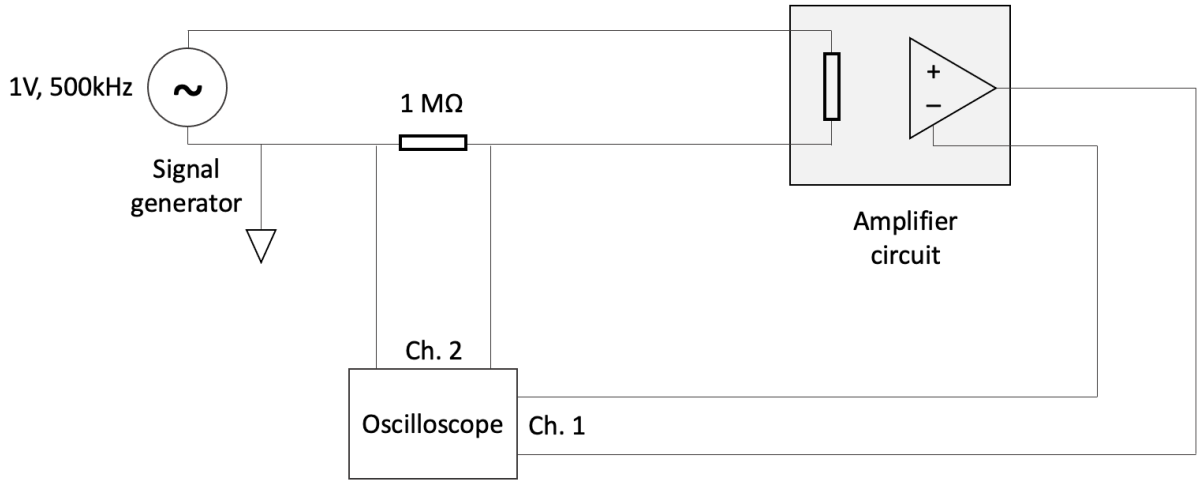


Figure F-1: Photomultiplier tube amplifier delay experimental setup

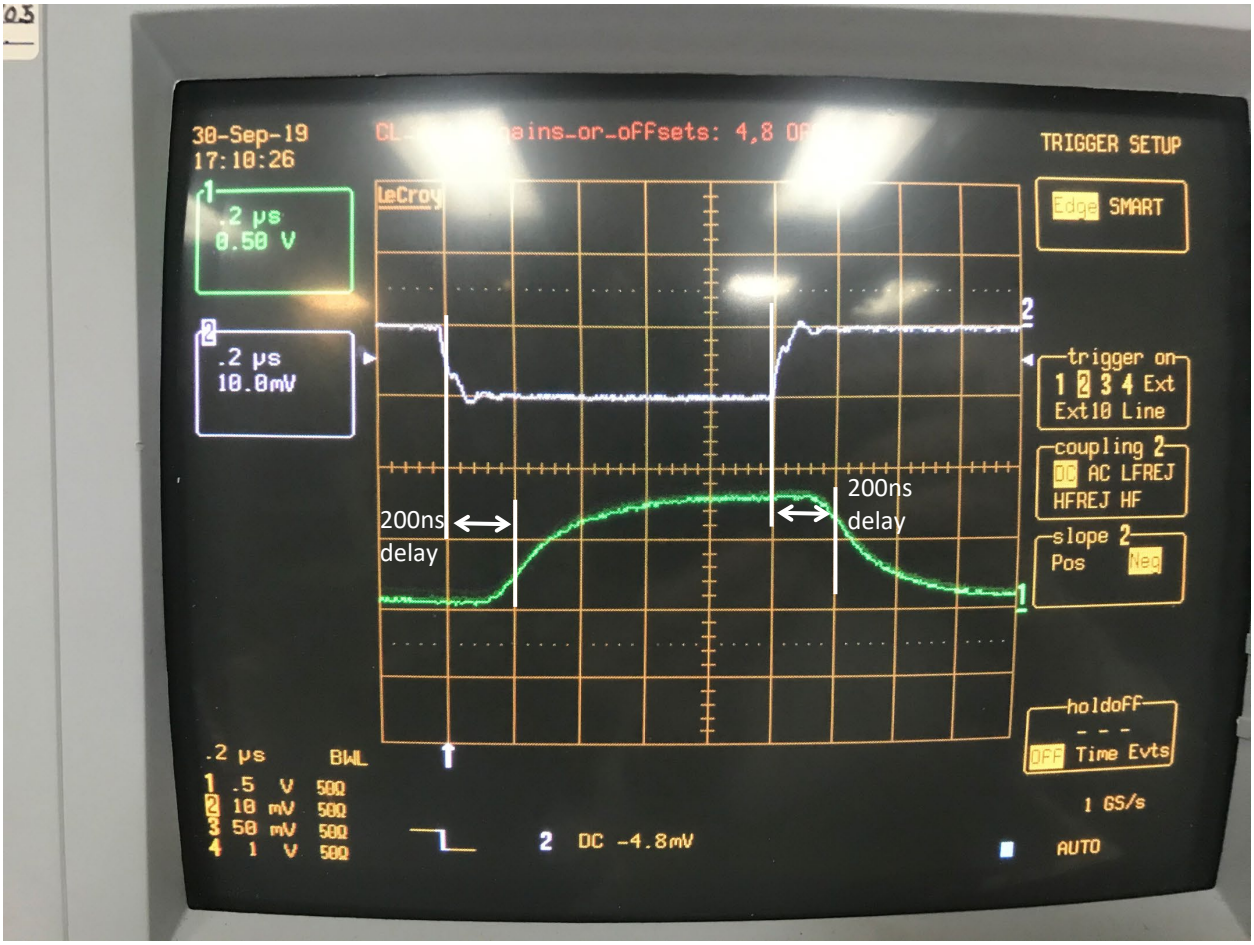


Figure F-2: Photomultiplier tube amplifier delay oscilloscope capture

This oscilloscope capture shows the current flowing into the PMT amplifier circuit by monitoring the voltage across a 1MΩ resistor connected from the signal generator, which is outputing a 1V, 500kHz square wave signal, to the input of the PMT amplifier. This is shown in Channel 2 as a **purple** trace. The output of the PMT amplifier is shown in Channel 1 as a **green** trace. The delay between the falling edge of the input current and the rise of the output signal and between the rising edge of the input current and the fall of the output signal can be seen here.

As can be seen in Figure F-2, the falling edge of the square-wave corresponds with a rise in output voltage. This is because the overall gain of the amplifier has a negative value. Presumably, this is due to the fact that a light signal captured by the PMT produces a negative current at its output because of the photoelectric effect, so in order to create a positive voltage at the output of amplifier circuit, the overall gain would be negative.

Figure F-2 also shows a delay between the falling edge of the square wave input and the rise of the amplifier output signal of about 200 ns. Similarly, we see a delay between the rising edge of the square wave input and the fall of the amplifier's output signal of about 200 ns.