Journal of Undergraduate Chemistry Research, **2021**, 20 (3), 43

A NOVEL PROTEIN STRUCTURE ELUCIDATION TECHNIQUE BY CIRCULAR DICHROISM AND NEAR INFRARED SPECTROSCOPY

Stephen C. Layng*, Jackson Harris*, Rebecca Miller*, Maria Leising*, Hannah Schroeder*, Alex Nguyen*, Meaghan Raab*, Cole Phelps* and Barry Hicks†

United States Air Force Academy Biochemistry Laboratory, 2355 Fairchild Drive, Suite 2N-225, United States Air Force Academy, CO 80840

Abstract

Identification and quantification of known or unknown proteins is essential in the biochemical and medical fields. Most importantly, the ability to elucidate a protein structure necessarily lends to an understanding of that protein's function. This is because structure determines function. In this experiment, near infrared (NIR) spectroscopy was employed to support and confirm circular dichroism data for determining secondary structure of proteins – specifically the bovine IgG antibody. Due to the little time required to perform NIR analyses, this research has shown the potential for a new application of NIR in order to significantly reduce the elucidation time for a protein structure in the laboratory.

†corresponding author: barry.hicks@usafa.edu

Keywords: Circular dichroism, NIR spectroscopy, SDS-PAGE, alpha helix, beta sheet, PBS buffer.

Introduction

Historically, protein isolation and elucidation has been a time-consuming and tedious process.[1] Current identification strategies include UV-Vis spectroscopy, SDS-PAGE, fluorescence spectroscopy, and circular dichroism (CD) spectroscopy.[2] Individually, these techniques are often insufficient to identify a protein. However, when applied in unison, they can be used as a reliable and effective means of elucidation.[2]

Previously, circular dichroism has been the primary technique for determining the secondary structure of proteins. Circular dichroism is a strategy that applies circularly polarized light to a protein resulting in different responses from the alpha helices, beta sheets, and random coils. Referencing the circular dichroism data, the NIR readings of several spectroscopic bands were analyzed and compared to CD results. The laboratory experiment discussed in this paper was aimed at accelerating the process of protein identification by using CD and NIR spectroscopy.

The analytical techniques of CD and NIR spectroscopy are relatively new in the realm of protein structure elucidation and both are used when attempting to determine the secondary structure of a protein. Currently, CD spectroscopy is considered the more consistent and reliable source of secondary structure determination[3]. Further analysis with NIR spectroscopy is required before conclusions can be made regarding its potential application in protein identification. However, based on an analysis of the second derivative plots of several bands in the NIR spectra, it is promising that this information will provide revealing data on the secondary structure of proteins.

Plotting the second derivative of the NIR data against the known percent alpha helix and beta sheet values provided evidence for a promising correlation between these variables. With R^2 values of 0.869 and 0.861 for some NIR band figures, this appears to be a possible new strategy for approximating secondary structure of proteins.

Experimental

CD Spectroscopy

For this experiment, a JASCO J-1500 CD spectrometer was used to determine the secondary structure of the unknown protein. The values for the unknown's molecular weight, number of amino acid residues, and concentration had to be completed prior to analyzing the sample with the CD spectrometer. This required information came from UV-Vis spectroscopy, SDS-PAGE, and fluorescence spectroscopy[4]. One mL of a 1:2 dilution of various protein stock solutions were prepared using a PBS buffer. A 1.0 mm quartz cuvette was filled with the protein solution. A separate cuvette was filled with buffer to act as the blank background spectrum for the NIR reading. The UV CD spectrum was collected from 250 – 190 nm.

NIR Spectroscopy

Near-Infrared scans were taken with the CARY Agilent Series UV-Vis-NIR spectrophotometer. Stock solutions of several proteins were used for the NIR analysis. 300 mg of protein sample were added to a 15-mL plastic tube.1.5 mL of PBS buffer were added to the same 15-mL plastic tube. 300 µL of the solution were transferred to a 1.0 mm cuvette. Several regions were analyzed: 1550 – 1850nm and 2130-2300nm. These regions were selected to avoid the NIR signals from water at 1400nm and 1900nm.

Results

CD Spectroscopy

Circular dichroism was implemented to characterize the secondary structure of the unknown protein. CD spectroscopy matched the known secondary structures in the protein database to the unknown spectrum (Figure 1).

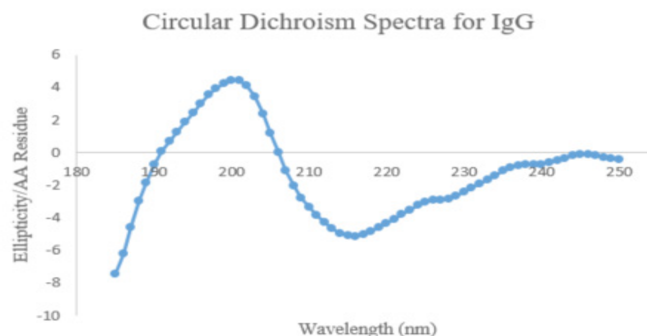

Figure 1. Circular dichroism spectrum for the IgG antibody.

The alpha helix and beta sheet calculation was completed after IgG's molecular weight, amino acid count, and concentration were input. The CD software labeled 10.3% of the IgG protein as alpha helix and 34.6% as beta sheet. The protein ribbon structure from the protein database also confirmed this finding (Figure 2).

NIR Spectroscopy

Near-infrared spectroscopy was experimentally used to test whether secondary structure could be determined with IR radiation[5]. The second derivative of the absorbance was plotted for various regions (Figures 2 and 3). One region was from 1550 – 1850nm[5-6]. This avoids the peak at 1400nm corresponding to water (Figure 3).

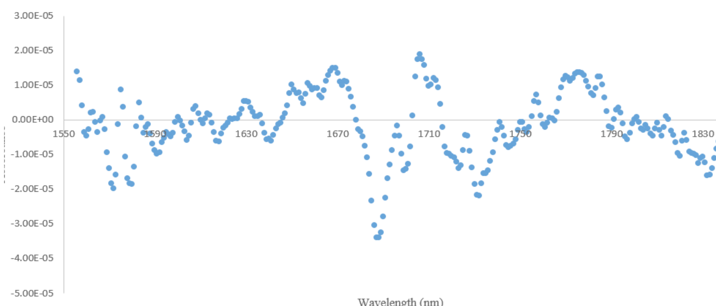

Figure 2. Second derivative of NIR absorbance from 1550 – 1850nm.

Another region, from 2130-2300nm, was examined (Figure 3). This avoids another peak at 1900nm, which is associated with water absorbance. It was hypothesized that these absorbance values of certain regions could be used to determine secondary structure as well.[5-7]

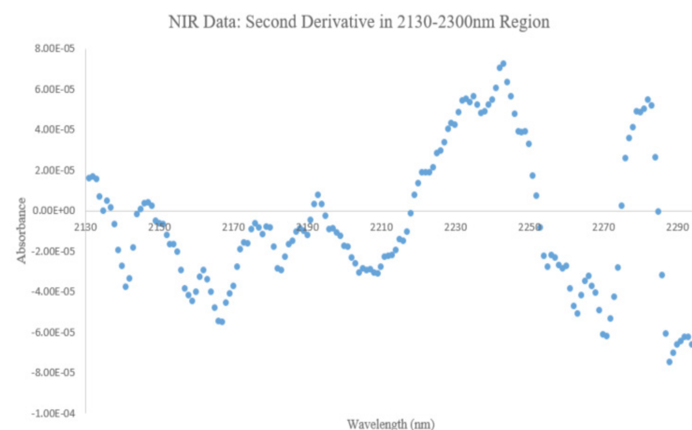

Figure 3. Second derivative of NIR absorbance from 2130 – 2300nm.

Eight separate proteins were examined via NIR spectroscopy with the same method as IgG: Bovine Serum Albumin (BSA), Bovine Hemoglobin (Hb), Bovine Cytochrome C (CytC), Equine Myoglobin (My), Gallus Lysozyme (Lys), Bovine Gamma Globulin (IgG), Bovine Beta-Lactoglobulin (bLac). The goal of the experiment was to determine whether or not certain NIR peaks corresponded to the secondary structure of the proteins. To test this hypothesis, the second derivative information from the above figures were plotted against the known percent alpha helix and beta sheet values of each of the eight proteins.

Only certain ranges were tested for this experiment. However, there was promising evidence of a correlation between the known secondary structure elements of proteins and certain bands in the NIR spectra. One of these successful trials was observed at the 2148.8-nm band (Figure 4).

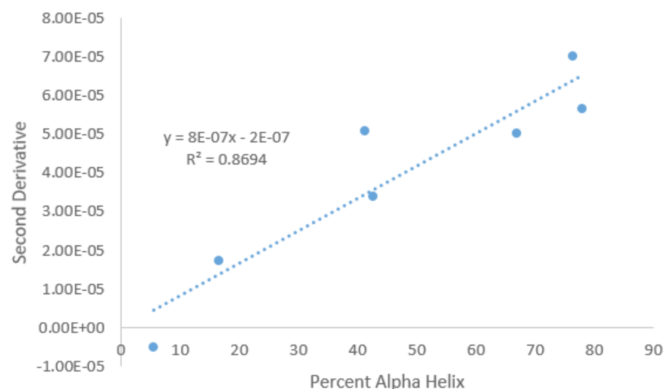

Figure 4. Second derivative data plotted against known percent alpha helix values for eight different proteins.

The correlation was noted in the R^2 value as 0.869. This indicates that the analyzed NIR band was possibly successful in correlating secondary structure with NIR frequencies. A similarly large correlation was discovered for the beta sheets at this same NIR band frequency (Figure 5) In this case, the R^2 value was 0.861. This correlation was one of the highest observed correlations in the experiment (not all of the bands were tested).

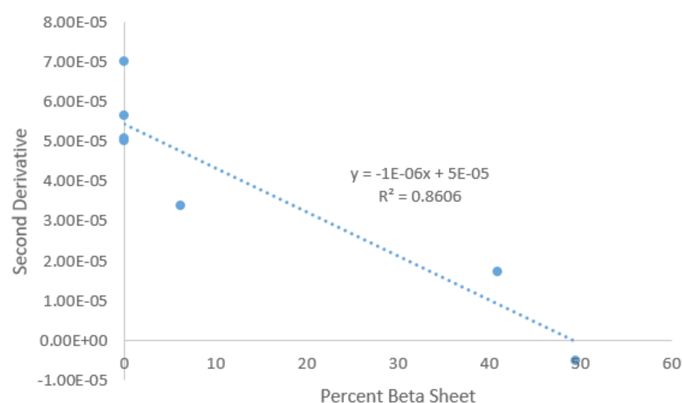

Figure 5. Second derivative data plotted against known percent beta sheet values for eight different proteins.

Discussion

CD Spectroscopy

Circular dichroism takes advantage of the optical activity of protein structures. CD involves sending circularly polarized light through a sample of protein. The light interacts differently with alpha helices, beta sheets, and random coils within the protein of interest. The reading began at 250nm and ended at 190nm because that is the region where differences are observed between second structures. The most important variable to control is the flow of

nitrogen into the system. Because the wavelength is so low and the energy is so high in the spectrophotometer, any oxygen could result in photo-oxidation. Incorrect readings and damaged instrumentation result from photo-oxidation.

Components of the secondary structure were approximated by comparing the IgG protein CD spectrum to numerous known spectra within the protein database. The secondary components of the IgG antibody were calculated to be 10.3% alpha helix and 34.6% beta sheet. The true value for most IgG antibodies is approximately 8.98% alpha helix and 42.1% beta sheet. The CD spectrum for the IgG was relatively accurate when compared to the database. Rigidity is an important element of antibodies in order to maintain the appropriate structure for binding to an antigen.[8-9]

NIR Spectroscopy

Near-Infrared spectroscopy was employed in this experiment to test the potential applications of NIR. The figures that resulted from plotting the second derivative data of the NIR spectra against the known secondary structure components served as significant evidence of a correlation between variables. The primary improvement that will be incorporated in follow-up experiments is increasing the number of analyzed proteins. An experiment with only eight references results in limited data output. However, these were the available proteins in the laboratory. Regardless, the purpose of this experiment was to demonstrate the potential of analyzing NIR data to deduce secondary structure components. With the current results and data, it appears as though this is purpose was achieved.

Conclusion

Protein elucidation is an area of unwavering and rising importance. Over 150,000 proteins have already been crystallized and placed into the RCSB protein database.[10] However, the demand for quick and reliable protein identification is only increasing. As protein and antibody elucidation techniques are refined and improved, so will our understanding of their vast capabilities. This experiment confirmed the hypothesis that some NIR band frequencies appear to be correlated with a protein's secondary structure. In future experiments, more band frequencies will be analyzed for a greater number of proteins. However, the scope of this experiment was to simply gather preliminary data from NIR spectra to determine the potential of elucidating secondary structures of proteins. Overall, these efforts were successful and were confirmed and supported by several analytical techniques in the laboratory.

References

(1) Lacroix E, Viguera AR and Serrano L, *J Mol Biol*, **1998**, 284, 173–191.

(2) Pan et al., 2011. X. Pan, P. Qin, R. Liu, J. Wang. *J. Agric. Food Chem*., 59, **2011**, pp. 6650-6656.

(3) Li, C. Arakawa T. *Int. J. Biol. Macromol*. **2019**, 132, pp. 1290-1295

(4) Li, C. & Arakawa, T. *Int. J. Biol. Macromol*. **2019**, 140, 668–671.

(5) Dousseau, F. & Pezolet M. *Biochemistry*. **1990**, 29, 8771–8779.

(6) Miyazawa, M., Sonoyama, M. *J. Near Infrared Spectrosc*. **1998**, 6, pp. 253-257.

(7) Hsu LN, Lin TP, Sane SU. *Journal of Near Infrared Spectroscopy.* **2008**, 16(5):437-444.

(8) Da Vela, S.; Roosen-Runge, F.; Skoda, M. W. A.; Jacobs, R. M.J.; Seydel, T.; Frielinghaus, H.; Sztucki, M.; Schweins, R.; Zhang, F.;Schreiber, F. *J. Phys. Chem.* **2017**, 121, 5759−5769.

(9) Berg J. M.; Tymoczko J. L.; Stryer L. The Immunoglobulin Fold Consists of a Beta-Sandwich Framework with Hypervariable Loops. Biochemistry. 5th edition. New York: W. H. Freeman; **2002**. Section 33.2.

(10) https://www.rcsb.org/.