

Public Abstract

First Name:Olayinka

Middle Name:

Last Name:Oshokoya

Adviser's First Name:Renee

Adviser's Last Name:JiJi

Co-Adviser's First Name:

Co-Adviser's Last Name:

Graduation Term:SP 2015

Department:Chemistry

Degree:PhD

Title:Development of advanced chemometric methods for analysis of deep-ultraviolet resonance Raman and circular dichroism spectroscopic data for protein secondary structure determination

Determination of protein secondary structure has become an area of great importance in biochemistry and biophysics as protein secondary structure is directly related to protein function and protein related diseases. While NMR and x-ray crystallography can predict placement of each atom in proteins to within an angstrom, optical methods are the preferred techniques for rapid evaluation of protein secondary structure content. Such techniques require calibration data to predict unknown protein secondary structure content where accuracy may be improved with the application of multivariate analysis. We compare protein secondary structure predictions obtained from multivariate analysis of ultraviolet resonance Raman (UVRR) and circular dichroism (CD) spectroscopic data using classical and partial least squares, and multivariate curve resolution-alternating least squares is made. Based on this analysis, the suggested best approach to rapid and accurate secondary structure determination is a combination of both CD and UVRR spectroscopy. While initial studies suggest that a complementary use of spectroscopic data from optical methods such as circular dichroism (CD), infrared (IR) and ultraviolet resonance Raman (UVRR) coupled with multivariate calibration techniques like multivariate curve resolution-alternating least squares (MCR-ALS) is the preferred route for real-time and accurate evaluation of protein secondary structure, further study presents a new strategy for the improvement of secondary structure determination of proteins by fusing CD and UVRR spectroscopic data. Also, a new method for determining the structural composition of each protein is employed, which is based on the relative abundance of the ( $\phi$ , $\psi$ ) dihedral angles of the peptide backbone as they correspond to each type of secondary structure. Comparison of the predicted protein secondary structures from MCR-ALS analysis of CD, UVRR and fused data with definitions obtained from dihedral angles of the peptide backbone, yields lower overall root mean squared errors of calibration for helical, sheet, poly-proline II type and total unfolded secondary structures with fused data. Considering that a disadvantage of multivariate calibration methods is the requirement of known concentration or spectral profiles, and second-order calibration methods, such as parallel factor analysis (PARAFAC), do not have such a requirement due to the "second-order advantage", PARAFAC was employed for analysis of UVRR data. An exceptional feature of UVRR spectroscopy is that UVRR spectra are also dependent on excitation wavelength as they are on secondary structure composition. Thus, higher order data can be created by combining protein UVRR spectra of several proteins collected at multiple excitation wavelengths. PARAFAC has been used to analyze UVRR data collected at multiple excitation wavelengths on several proteins to determine secondary structure content.