Mass Spectral and Fluorescent Identification of Tryptophan Residuals in Unknown Proteins

Kit Canny, Simran Gill, Dorian Schwartz

Chemical Frontiers | CHEM 3500-1

3/19/2023

Mass Spectral and Fluorescent Identification of Tryptophan Residuals in Unknown Proteins

Kit Canny*, Simran Gill, Dorian Schwartz

University of Denver

Introduction

Tryptophan is one of 20 naturally occurring alpha amino acids. Thought to play a rate limiting role during protein synthesis, it is found in the lowest concentrations among these amino acids despite exerting a vital role in several vital metabolic functions. Structurally, tryptophan is composed of an alpha-amino group, alpha carboxylic acid group, and an indol-3-yl substituent at position 3 on the aromatic R group. It has a pK_1 of 2.38, pK_2 of 9.39 and a Hydropathy Index of -0.9. Where the hydropathy index is a scale combining hydrophobicity and hydrophilicity of R groups. This value is reflective of the Gibbs free energy (ΔG) transfer of the amino acid side chain from a hydrophobic solvent to a nearby water molecule. This transfer is favorable for charged or polar amino acid side chains, and unfavorable for amino acids with nonpolar or more hydrophobic side chains.

Because of the aromatic R group, tryptophan is a fluorophore which absorbs light at an excitation wavelength of 280 nm and ultraviolet light. Measurement with a spectrophotometer can be used to detect molecules and give their concentration in solution. The fraction of the incident light absorbed by a solution at a given wavelength is related to the thickness of the absorbing layer (path length) and the concentration of the absorbing species. This is the Lambert-Beer law ($\log (I_0/I) = \varepsilon c \ell$).

Given this, fluorescence can be used for tryptophan determination, where the position of tryptophan within the protein has an effect on its fluorescent emission wavelength. If the aromatic tryptophan residue is present on the surface of the protein, the fluorescence is much higher at approximately 350 nm compared to the fluorescence that can be captured when tryptophan is located inside the protein of approximately 330 nm.

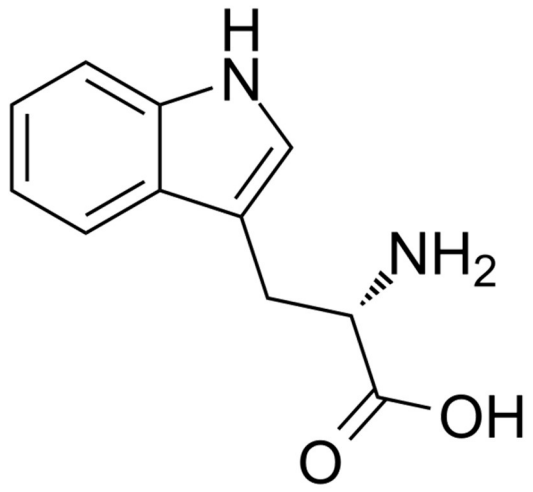


Figure 1: L-Tryptophan Diagrams

Mass Spectrometry is an instrumental analysis tool which is used to determine the mass to charge ratio, often denoted as m/z, of specified molecules present within a presented sample. Using this ratio, the molecular weight can then be determined of the component, allowing for the identification of unknown compounds, concentration of a specified compound, determine structure, and other various chemical properties of molecules within the sample[4]. Within the proposed reaction, an arylation modification is proposed at the tryptophan residuals in order to tag and determine the number of residuals based on the molecular weight modifications of the native protein in comparison[1]. This is done by denaturing a sample of the native protein with a gentle detergent, urea, which solubilizes the protein at high concentrations (8M) and disrupts the noncovalent bonds, and then undergoing the reaction demonstrated below in Figure 2.

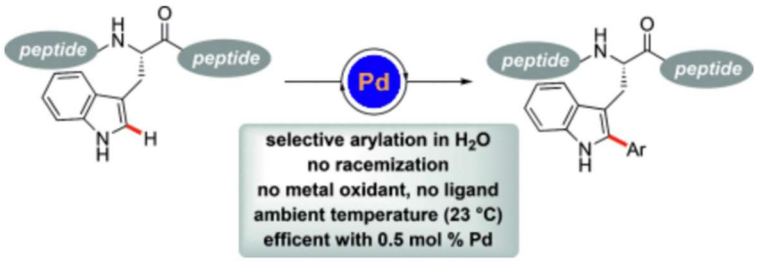
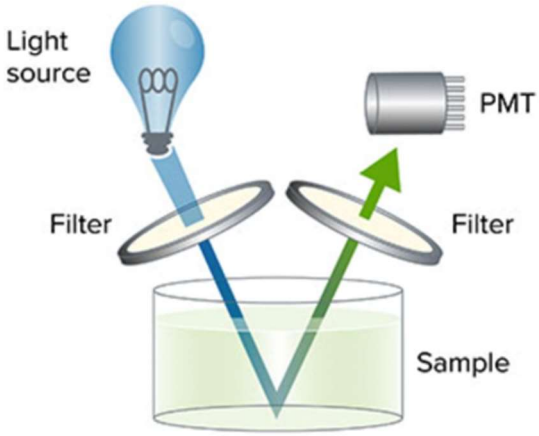


Figure 2: Shows an aqueous arylation reaction as proposed to be done to Samples B and C in the Mass Spectral Method.

Fluorescence is the chemical phenomenon of luminescence via the absorption of a photon in the ground state which results in the promotion to an excited state, and then emitting this photon at a lower energy than the absorbed photon. Fluorescence is a quality that is specific to a selective species of chemicals, denoted as fluorophores: usually composed of networks of delocalized pi bonds. Aromatic molecules, for this reason, are fluorophores. Fluorescence is also specific, with the excitation and emission wavelength corresponding to the conjugation of this pi bond network. Using this, the emission wavelength given in UV-Vis spectroscopy or the analysis in fluorimetry can be used to narrow down the structure, components, and or analyze the compound provided[2,5]. Within the proposed method, the analysis of the excitation and emission of arylated Tryptophan peptides will be analyzed in a phosphate buffer saline. The establishment of a serial dilution calibration curve will allow for the arylated and arylated and denatured samples (Samples D and E in the proposition) to be compared at a percentage and the number of residuals approximated internally and surface from both samples. The Sample D, arylated and nondenatured, will give the number of



surface residuals as compared to the calibration curve, and the samples E, denatured and arylated, will give the total number of residuals, from which the number from sample D can be subtracted giving the number of internal residuals.

Figure 3: Illustrates fluorimetry setup for the L-Tryptophan and residual analysis. Light source indicated is the excitation wavelength, which then passes through a filter, into the sample, where the emission wavelength is emitted and passed through a secondary filter and analyzed.

Limitations of the procedure are that, with fluorescence of the Tryptophan residuals, there needs to be caution in reading the results with error of similar aromatic residuals such as L-Tyrosine. Amino acids with similar fluorescent emission have the potential to cause confusion in the actual number of residuals located within the unknown protein, and can cause possible conflation of number of residuals recorded.

Limitations of the procedure within instrumentation are that, ESI-TOF analyzers for Mass Spectrometry send the sample through an electrospray nebulizer, from there it travels through a small capillary opening to a pulsar into the flight tube. As such, these instruments cannot handle continuous flow nor continuous ionization of samples. Given this, the solution is to maintain a flow rate low enough for ESI at 10 µl/min while not holding a continuous flow rate for TOF. This procedure also is dependent on the belief that the unknown protein is water soluble, and will not proceed if otherwise.

Limitation of the procedure otherwise are that this arylation reaction has been specified for L-tryptophan. For other amino acids, the reaction mechanism must be modified and specified to each independent structure.

Proposed Procedure

Proposed Mass Spectrometry Methodology

Prepare protein samples for mass spectral analysis. Sample A will contain the native version of an unknown protein, Sample B will contain the arylated version of the unknown protein, and Sample C will contain the denatured and arylated version of the protein, all in a solution of 0.1% formic acid and LCMS-grade water.

First, Sample C should be denatured. This will occur under chemical denaturation. Add 8 M urea to solution of the native protein in 0.1% formic acid and LCMS-grade water. This addition of a detergent will cause a disruption in the noncovalent interactions in the native protein. The removal of the urea prior to running the sample through the LCMS system will be done through running the sample through a Pierce™ spin column (25 x 125 µL)[9]. The spin column contains resin beads that will trap urea molecules as they traverse the column due to the relatively small size of urea molecules. The unknown protein will be too large to get trapped by the beads and will therefore elute out of the column prior to the urea. Allow the sample to run through the column for 2-5 minutes and then centrifuge the eluent to obtain a urea-free solution. Both Sample B and Sample C will then be subjected to selective arylation of Trp residues of the unknown protein. To perform the arylation reaction, combine $Pd(NO_3)_2$ in HNO_3 0.5 mol/l 1000 mg/l Pd with the given protein sample in an aqueous solution and allow the reaction to run for 24 hours at room temperature.

Run the samples using LC-ESI-TOF-MS (Liquid Chromatography- Electron Spray Ionization- Time of Flight- Mass Spectrometry)[3]. The mobile phase will be 0.1% formic acid in LCMS-grade water with an organic modifier of acetonitrile, and the stationary phase will be a C-18 column. Use a flow rate of 0.2 L/min and injection volume will be 5-10 µL. The ESI-TOF-MS will be operated in positive ion mode. The

obtained spectra will need to be individually analyzed to determine each sample's molecular weight. Sample A's calculated molecular weight will be that of the native protein, Sample B the molecular weight of the arylated, non-denatured protein, and Sample C will be the molecular weight of the arylated, denatured protein. Subtracting the molecular weight of Sample A from the molecular weight of Sample C and dividing the resulting value by the molecular weight of an arylated Trp residue will give an estimate of the total number of Trp residues in the given unknown protein. Then, following the same process of comparison between Sample A and Sample B, an estimate of the amount of exterior Trp residues in the unknown protein can be obtained by subtracting the amount of exterior residues from the total number of Trp residues giving an estimate of the number of interior Trp residues.

Proposed Fluorimetry Methodology

A 500 mL solution of 1x PBS, giving an approximate pH of 7.4, will be prepared by dissolving 4g of NaCl, 0.1 g of KCl, 0.71 g of Na_2HPO_4, and 0.12 KH_2PO_4 in 500 mL[6] of water. After dissolving the salts, a pH probe will be used to measure the pH of the buffer. If the pH is below 7.4, add NaOH in a dropwise manner until the pH reading reaches the desired 7.4. If the pH is below 7.4, proceed similarly with HCl until the pH reading reaches the desired 7.4.

An arylated Trp sample will be prepared using the same reaction described in above section. However, a notable difference between this reaction and the aforementioned one is that the starting material should be L-tryptophan, a known amino acid. To examine the excitation and emission characteristics of the arylated Trp, simultaneously scanning of the absorbance spectrum and the emission spectrum of a

sample of arylated Trp in PBS will need to take place.

To do this, first establish a calibration curve, using the concentration of arylated tryptophan on the x-axis and fluorescence intensity on the y-axis. After the arylation is completed, a serial dilution of the arylated Trp product in 1x PBS will be performed. The concentration range that the calibration curve data will bracket will be 10 mM to $1x10^{-5}$ mM. Sample D will contain the arylated, non-denatured version of the unknown protein in PBS, and Sample E will contain the arylated, denatured version of the unknown protein in PBS. Arylation and denaturation will be performed using the same process as described in the above section.

Load 200 µL aliquots both Samples D and E into a 96-well opaque black microplate. In addition to Sample D and Sample E, also include the calibration samples containing known concentrations of arylated tryptophan in PBS and three blank samples containing only 1x PBS in water. The blank samples will be used to obtain background fluorescence measurements and calculate the limit of detection (avg. blank signal + $3\sigma_{blank}$) and the limit of quantification (avg. blank signal + $10\sigma_{blank}$). The microplate will be loaded into the fluorimeter and each sample will be subjected to a fluorescence excitation light source. Any resulting fluorescence from the samples will be detected and recorded by the instrument. Both the excitation wavelength and the wavelength will be recorded for the fluorescence data and will selected based on the results of analysis of the excitation and emission characteristics of arylated Trp. Effort should be made to avoid choosing a detection wavelength that overlaps with the emission spectrum of other fluorescent amino acids such as L-Tyrosine.

The concentration of arylated Trp in both Sample D, the non-denatured sample,

and Sample E, the denatured sample, will be estimated by plugging the fluorescence intensity signal measured for each sample into the equation of the calibration curve. The molecular weight determined from the arylated protein mass spectral results and the concentration determined from the fluorescence data for the non-denatured arylated sample will be used to calculate the amount of exterior arylated Trp residues. The total amount of arylated Trp residues will be calculated using the same molecular weight value and the concentration determined from the fluorescence data for the denatured arylated sample. The amount of interior residues will be estimated by subtracting the number of exterior Trp residues from the total number of Trp residues. These results will be compared with the results from the mass spectral data.

References

1 Chemical modifications of tryptophan residues in ... - Wiley Online Library https://onlinelibrary.wiley.com/doi/full/10.1002/psc.3286 (accessed Mar 20, 2023).

2 Libretexts. Fluorescence https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Electronic_Spectroscopy/Radiative_Decay/Fluorescence (accessed Mar 19, 2023).

3 Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) https://www.creative-biolabs.com/glycoprotein/liquid-chromatography-electrospray-ionization-mass-spectrometry-lc-esi-ms.htm#:~:text=LC%2DESI%2DMS%20is%20a%20hyphenated%20mass%20spectrometry%20technique.,accuracy%20of%20the%20mass%20spectrometer. (accessed Mar 19, 2023).

4 Mass spectrometry applications areas https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/mass-spectrometry-learning-center/mass-spectrometry-applications-area.html#:~:text=Specific%20applications%20of%20mass%20spectrometry,protein%20identification%2C%20and%20carbon%20dating. (accessed Mar 19, 2023).

5 Omary, M. A. Luminescence, theory https://www.sciencedirect.com/science/article/pii/B978012803224400193X (accessed Mar 19, 2023).

6 Palladium standard solution traceable to SRM from NIST $pd(no_3)_2$ in HNO_3 0.5 mol/L 1000 mg/L PD certipur®: Sigma-aldrich https://www.sigmaaldrich.com/US/en/product/mm/114282 (accessed Mar 19, 2023).

7 Peptide and amino acid quantification using UV fluorescence in synergy HT multi-mode Microplate Reader: April 18, 2003 https://www.biotek.com/resources/application-notes/peptide-and-amino-acid-quantification-using-uv-fluorescence-in-synergy-ht-multi-mode-microplate-reader/ (accessed Mar 19, 2023).

8 Phase-dependent long persistent phosphorescence in coumarin-phosphine ... https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202300497 (accessed Mar 20, 2023).

9 Urea, ACS grade, 500 g https://agscientific.com/products/detergents/additional-detergents/urea-acs-grade-500-g.html#:~:text=Urea%20can%20be%20used%20as,concentrations%20by%20disrupting%20noncovalent%20bonds. (accessed Mar 19, 2023).