Understanding the conformational stability of protein therapeutics using Raman spectroscopy

A Malvern Instruments' Bioscience Development Initiative

Executive summary

The combination of Dynamic Light Scattering (DLS) with Raman Spectroscopy provides the ability to extract a wealth of chemical, structural, and physical information about biotherapeutic proteins under formulation conditions.

Raman spectroscopy simultaneously derives protein secondary structure (Amide I and III) and tertiary structure markers (aromatic side chains, disulfide bonds, hydrogen bonding, local hydrophobicity). These higher order structural determinations are performed at actual formulation concentrations (50 mg/mL or greater for mAbs), rather than at the lower concentrations required by conventional methods, i.e. less than a few mg/mL for Circular Dichroism (CD).

By use of Raman spectroscopy, protein secondary and tertiary structure perturbation/ unfolding, melting temperature, onset temperature of aggregation, and van 't Hoff enthalpy values can all be derived, leading to improved understanding of competing pathways of unfolding/structure change and aggregation, and ultimately, unique insights into the mechanism(s) of aggregation to help improve formulation stability.

The unique coupling of Raman spectroscopy and DLS provides the ability to simultaneously correlate protein structure with colloidal parameters to enhance the understanding of protein therapeutic formulations under a variety of stress conditions, e.g., thermal, formulation, chemical degradation, extrinsic particulates.

Introduction

Raman spectroscopy derives information about protein secondary and tertiary structure by monitoring molecular vibrations. Dynamic light scattering (DLS) is a workhorse technique for determining the hydrodynamic diameter of biotherapeutic proteins in solution, in addition to their polydispersity and interactions. By combining these two analytical approaches into a single system, a wealth of chemical, structural and physical parameters can be determined. Protein secondary and tertiary structure,

melting temperature, onset temperature of aggregation and transition enthalpy values can all be derived, as well as aggregation propensity, protein solubility, and the potential for high viscosity at formulated concentrations. The complementary nature of the results obtained from DLS and Raman spectroscopy on the same sample may provide unique insights into its unique mechanisms of aggregation and unfolding. Here we describe the role of Raman spectroscopy in understanding the conformational stability of protein therapeutics.

The approach

Malvern Instruments' Zetasizer Helix (ZS Helix) integrates a fiber-coupled Raman spectrometer with a Zetasizer Nano ZSP to provide DLS (colloidal stability) and Raman (conformational stability) data sequentially on a single sample. The Zetasizer Nano system integrates proprietary non-invasive backscatter (NIBS) detector technology with dynamic (DLS), static (SLS) and electrophoretic (ELS) light scattering to measure the hydrodynamic radius of proteins from 0.15 nm - 5 µm, from 0.1 mg/mL to ≥ 100 mg/mL. Raman spectra are collected using 785 nm excitation (~280 mW) from 150 cm^{-1} - 1925 cm^{-1} at 4 cm^{-1} resolution. Sample aliquots (~120 µL) are placed into a 3 mm quartz cuvette and positioned in a temperature-controlled compartment that provides temperature control from 0°C - 90°C ± 0.1°C. Thermal ramp studies are conducted by collecting Raman and DLS data over a series of pre-defined 0.1°C - 5°C step-wise increments. Isothermal incubation studies are conducted by collecting as series of Raman and DLS data over a pre-defined time interval at a desired temperature set-point.

Results and Discussion

Higher order structure

Secondary structure

Raman spectroscopy is an effective approach to determine protein secondary structure characteristics. A variety of secondary structural determination algorithms have been suggested, including 1) band deconvolution of a particularly information-rich spectral feature called the Amide I band and 2) a multivariate modeling approach that incorporates a broad spectral range (~990 cm^{-1} to 1730 cm^{-1}).

The Amide I band, ~1600 cm^{-1} to 1700 cm^{-1}, results mainly from the C=O stretch combined with a small amount of C-N stretch. The Amide III feature, ~1200 cm^{-1} to 1340 cm^{-1}, results from coupling of the C-N stretch to the N-H bend. The overall shapes of these bands change with secondary structure. A band at ~930 cm^{-1} to 950 cm^{-1} is based on N-C_α-C skeletal stretching mode, and its intensity is indicative of α-helix content. The above features are the major characteristic bands of the protein secondary structure and are summarized in Table 1.

Table 1: Correlation between Raman shift and protein secondary structure.

Raman Shift (cm^{-1})	Protein Backbone	Secondary Structure Motif
1670-1680	Amide I	β-Turn/random β-space
1660-1670	Amide I	β-sheet
1650-1655	Amide I	α-helix
1330-1340	Amide III	α-helix
1235-1250	Amide III	β-sheet
930-950	Skeletal stretch	α-helix

Figure 1 presents representative Raman spectra of BSA (Bovine Serum Albumin) before and after thermal treatment, highlighting the major secondary structure features described above. In this example, the clear transition from an α-helix rich structure into a β-sheet abundant structure is evident. A variety of analyses (e.g. spectral deconvolution, second derivative) can be applied to the Amide I, III, and skeletal stretching regions for detailed analysis of secondary structure. These approaches will be covered in future technical notes.

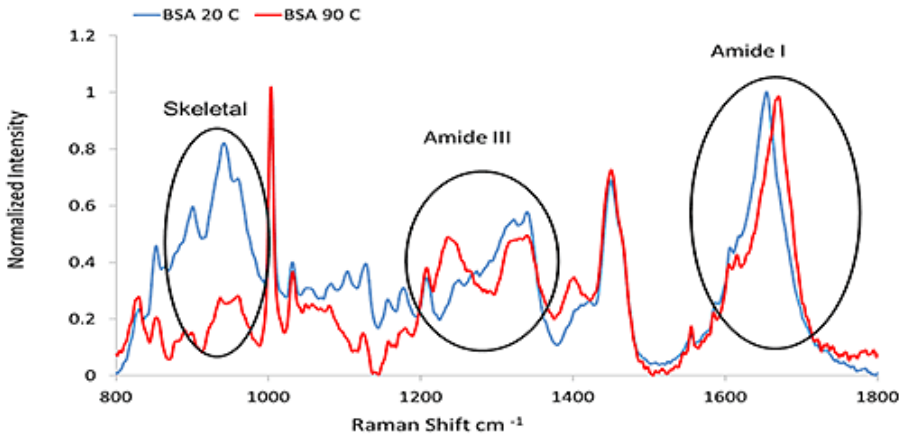

Figure 1: The Raman spectra of BSA (50 mg/mL at pH 7.4, PBS buffer), collected at 20°C and 90°C, respectively. Spectra are normalized to the phenylalanine intensity at 1004 cm^{-1}.

A multivariate approach, on the other hand, considers the entire spectral profile, not just a single Raman band or frequency. In this approach, spectral and structural data from a representative number of proteins is used to create a PLS model that can predict protein structures on unknown proteins, i.e. those not contained within the library. The data presented in Table 2 demonstrate the type of structural information that can be derived from this model automatically, along with comparison to reported values from www.pdb.org.

Table 2: Secondary structural motifs derived from a multivariate analysis of Raman spectra of proteins, as compared to PDB reported values.

Protein	PDB File	Multivariate Model Value		PDB Value	
		α-helix	β-sheet	α-helix	β-sheet
Concanavalin A	3CNA	0	46	0	40
Trypsin	3RN3	18	33	20	35
Lysozyme	7LYZ	28	8	39	10
Human Serum Albumin	1YSX	69	0	67	0

Figure 2 compares α-helical and β-sheet content predicted via the multivariate model as a function of temperature for the same set of BSA data shown in Figure 1. It clearly shows the secondary structure changing as the temperature increases; this is chiefly an increase in β-sheet and random content, while α-helix drops significantly.

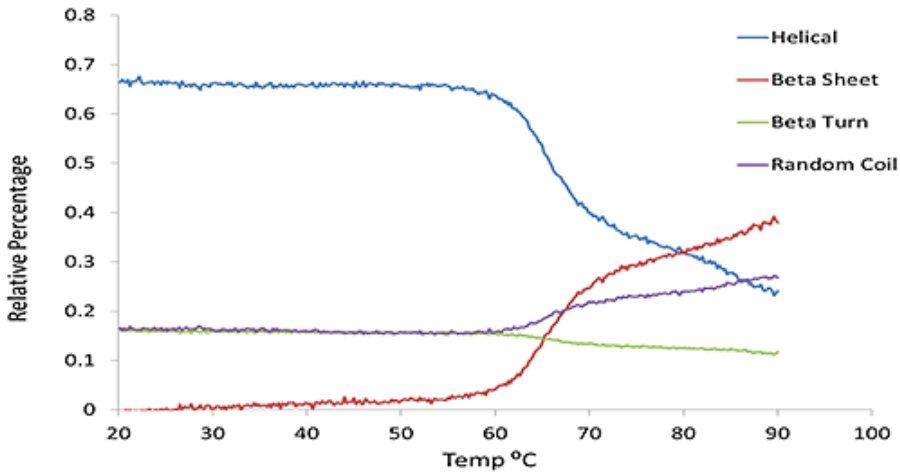

Figure 2: The predicted secondary structural changes derived from multivariate analysis for the same BSA sample shown in Figure 1.

Tertiary structure

Raman spectroscopy is also highly sensitive to the symmetrical vibrational modes of the aromatic side chains, i.e. phenylalanine (Phe, F), tyrosine (Tyr, Y) and tryptophan (Trp, W), and as such has been widely used as a probe of protein tertiary structure. There are several well-established and frequently-used Raman markers that provide insights into the environment of these aromatic side chains. Table 3 summarizes the dominant tertiary structure band assignments, and their corresponding structural indications.

Figure 3 clearly illustrates that as the temperature increases, the secondary structure (green circles), changes from α-helix abundant to β-sheet rich structure: α-helix drops from 40% to 18%, and β-sheet increases from 6% to 28%. Concurrent with changes in secondary structure are changes in the aromatic side chains. Take Trp for example: a decreasing ratio of 1360 cm^{-1} /1340 cm^{-1} indicates probable exposure to an aqueous environment. A 2 cm^{-1} shift in the W17 region (870 cm^{-1} -885 cm^{-1}) indicates changes in the hydrogen bonding structure of these side chains. The W3 region (~1550 cm^{-1}) shifted 3 cm^{-1}, indicating a change in the dihedral angle from 100° to close to 90°. The

tyrosine region also displays interesting behavior, with an initial peak near 836 cm^{-1} at 20 °C, transitioning to two peaks at 828 cm^{-1} and 853 cm^{-1} at 80 °C. It should be noted that not all of the aromatic side chains will show changes, or even change at the same temperature. More importantly, each protein is unique and these markers should be carefully assigned to the specific protein.

Table 3: Summary of the common aromatic markers used for protein tertiary structure.

Raman Shift (cm^{-1})	Source	Structure Indication
1360/1340 >1	Trp Fermi Doublet (W7)	Indole ring in hydrophobic environment or in contact with aliphatic chains
1360/1340 <1	Trp Fermi Doublet (W7)	Indole ring in hydrophilic environment or exposed to aqueous medium
1550	Trp dihedral angle (W3)	The angle between the indole ring and the peptide bond plane
870-885	Trp hydrogen bonding (W17)	883 non H-bonded 877 medium strength 871 strong H-bonded
760	Trp (W18)	Cation-π interaction
850/830 ~ 0.25	H-bonding of Tyrosine	Donor of strong H-bond
850/830 ~ 1.25	H-bonding of Tyrosine	Donor and acceptor of strong H-bond/exposed to aqueous environment
850/830 ~ 1.25	H-bonding of Tyrosine	Acceptor of strong H-bond/buried tyrosine
1002-1004	Phe ring breathing mode	Normalize the Raman intensity

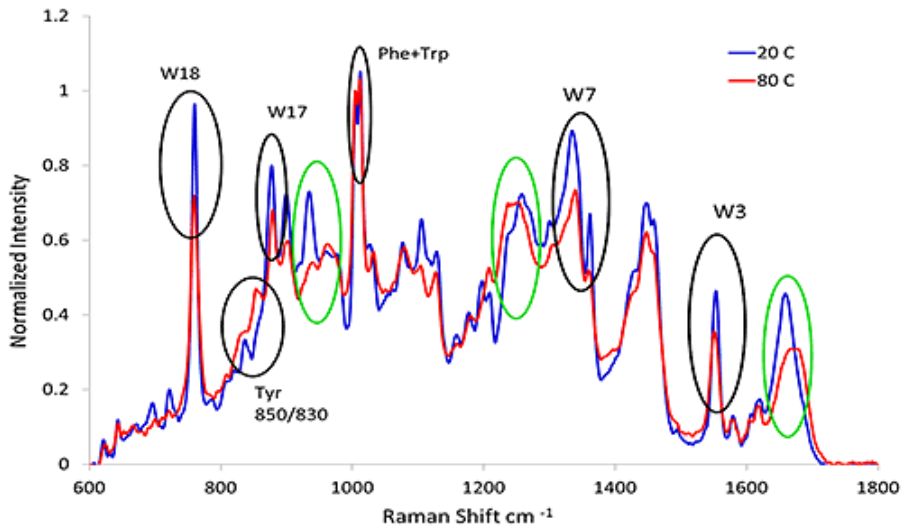

Figure 3: Representative spectra from lysozyme at pH 7.4 before and after heat treatment. The black circles indicate the aromatic markers as shown in Table 3 and the green circles refer to the secondary structure described in Table 1.

Disulfide bond

The disulfide bond is a covalent bond between sulfur units on cysteine residues, and is critical for maintaining the native protein structure. Three unique disulfide bond conformers are summarized in Table 4. More specific definitions of the different conformers will be covered in a future application note.

Table 4: Summary of disulfide bond conformers in Raman spectra.

Raman Shift (cm^{-1})	Source	Structure Indication
508-512	S-S stretch	GGG conformer
523-528	S-S stretch	GGT conformer
540-545	S-S stretch	TGT conformer

Figure 4 shows the disulfide region of lysozyme (four total disulfide bonds) before and after thermal treatment. At 20 °C, the 508 cm^{-1} , 528 cm^{-1}, and 545 cm^{-1} bands indicate that the lysozyme disulfide bonds have character assigned to the gauche-gauche-gauche (GGG), gauche-gauche-trans (GGT) and trans-gauche-trans (TGT) conformers. At 80 °C, the peaks at 528 cm^{-1} and 545 cm^{-1} have disappeared, indicating that only the GGG conformer still remains. In the presence of a reducing agent (dithiothreitol - DTT) the disulfide bond intensity decreases with time, especially the GGG conformation, indicating the loss of these conformers.

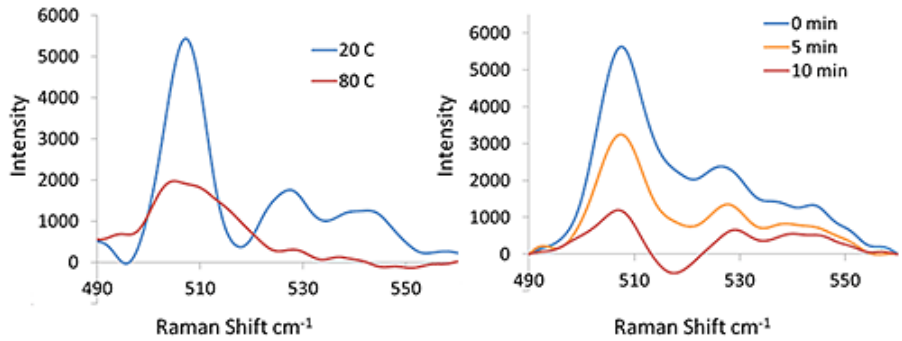

Figure 4: Raman spectra (disulfide spectral region) of lysozyme at 20°C and 80°C (Left) and the lysozyme sample at 40°C (Right) with DTT (50mg/mL protein with 50 mM DTT).

Additional Raman spectra - Protein structure applications

Raman spectroscopy has significant promise as a workhorse technique in the rapidly expanding realm of protein therapeutics. Its strengths in sampling, ruggedness and selectivity promise a host of future applications including forensic particle identification (silicone oil droplets, glass, etc); protein structural determination at interfaces (water-oil, gold nanoparticles); protein dynamics (H/D exchange); accelerated degradation (thermal, reducing agent, agitation and oxidization), and many others. DLS and Raman are complementary techniques, and when data from both are taken as a whole, unique insights into the mechanisms of protein aggregation may be elucidated. Future application notes will explore the relationship between size and protein structure, investigating questions such as: 1) does a change in protein conformation trigger an aggregation event, or vice versa?; and 2) is there a marker that can predict protein stability efficiently and robustly?

Conclusions

The combination and integration of DLS and Raman spectroscopy into a single instrumental platform provides unique analytical capabilities to determine protein secondary and tertiary structure, melting temperature, size and onset of aggregation for the same sample without altering the testing conditions. Here we have described the unique spectra-structure properties that may be correlated using Raman Spectroscopy. Subsequently, we describe the utility of DLS to derive the bulk viscosity/ restricted diffusion interaction parameter (k_D), particle interaction parameter (B_{22}), melting temperature, onset temperature of aggregation and transition enthalpies as indices of formulation stability (Understanding the Colloidal Stability of Protein Therapeutics with Dynamic Light Scattering). Future studies will report on the unique correlations that may be derived from the combination of DLS and Raman spectroscopy to improve the understanding of the kinetics, thermodynamics and mechanism(s) of aggregation and the correlation of specific protein structural motifs to aggregation to improve product knowledge and formulation stability.

About Malvern's Bioscience Development Initiative

Malvern Instruments' Bioscience Development Initiative (BDI) was established to accelerate innovation, development and the promotion of new technologies, products and capabilities to address unmet measurement needs in the biosciences market.

Malvern Instruments Limited
Grovewood Road, Malvern,
Worcestershire, UK. WR14
1XZ

Tel: +44 1684 892456
Fax: +44 1684 892789
www.malvern.com