# Use of DLS/Raman to study the unfolding thermodynamics of proteins

**A Malvern Instruments' Bioscience Development Initiative**

## Executive summary

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

Protein aggregation is an extremely complicated process. It can be induced by changes in pH, ionic strength, or temperature, induced via chemical degradation, or alternatively triggered by external factors during manufacturing, i.e. air-water interface, agitation, external surfaces (silicone oil droplets, stainless steel, glass surfaces), freeze-thaw cycles, or light. Effective mechanisms to control against these issues for one protein may not be readily transferred to another. Although multiple mathematic models have been developed to describe protein aggregation mechanisms and pathways, the prediction of aggregation propensity under different conditions and therefore also of long term stability is challenging.

Here, we demonstrate the unique capabilities of DLS/Raman spectroscopy to characterize protein denaturation under thermally - and chemically-induced stress conditions by providing additional information that is not readily obtainable via traditional approaches.

## Introduction

The ability to understand the mechanisms of protein aggregation is gaining tremendous attention in the biopharmaceutical industry because of reduced efficacy and possible immunogenicity issues arising from the presence of protein aggregates. Partially unfolded species are proposed to play a crucial role in aggregation. Thus, a comprehensive understanding of protein unfolding, in tems of both conformation and size, will greatly enhance the capability to formulate or design stable protein therapeutics. Here, we demonstrate the utility of a combined Dynamic Light Scattering (DLS) and Raman spectroscopy platform to provide new insights into the unfolding free energy ($\Delta G$) by thermal stress or chemical denaturation (Guanidine-HCl or urea).

## Materials and methods

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 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 a series of Raman and DLS data over a pre-defined time interval at a desired temperature set-point.

## Results and Discussion

### Protein folding and unfolding - Basic theory

The simplest approximation of protein unfolding is to assume a two state model in which the protein exists as either folded or unfolded, with no intermediate state, which is described by the following expression:
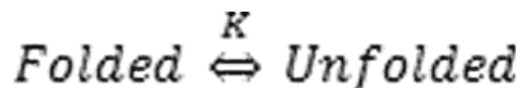
$$Folded \overset{K}{\Leftrightarrow} Unfolded$$

where K is the equilibrium constant, and Folded and Unfolded indicate the native and denatured states respectively. Therefore, at a specific temperature, K is defined as:

$$K = \frac{[F]}{[U]} = \frac{f_F}{f_U} = \frac{f_F}{1 - f_F}.$$

Here [F] and [U] indicate the concentration of folded and unfolded fractions of the tested protein. If we assume that the sample is purely monomeric, then the concentration term can be dropped and replaced by the fraction of folded, $f_F$, with the unfolded fraction given by (1- $f_F$). It has to be noted that for thermal denaturation, the process needs to be confirmed to be fully reversible and with no apparent aggregation, for this substitution to be valid. Once reversibility is proven, the goal then is to relate a specific Raman parameter to the fraction of folded protein:

$$f_F = \frac{v_i - v_f}{v_u - v_f}.$$

Here, v represents any Raman spectroscopic property, and $v_u$ and $v_f$ refer to the value of that property at unfolded and folded states, and vi is the value at a particular measurement point. The Raman marker can be a measure of a number of secondary

or tertiary structural markers that characterize α-helical or β-sheet content, or the local environment (hydrophobic/hydrogen bonding) of tyrosine and tryptophan residues.

Once the equilibrium constant K has been determined, then the folding free energy, ΔG, is found via the expression:

$$\Delta G = -RT \ln K.$$

Here, R is the gas constant and T is the absolute temperature.

For chemical denaturation investigations, the sample and a series of denaturant [i.e. Guanidine hydrochloride (GdnHCl)] concentrations are incubated together. The resulting data is fitted to the following equation:

$$\Delta G_i = \Delta G_o + m[GdnHCl].$$

Here, $\Delta G_0$ and $\Delta G_i$ refer to the folding free energy in absence of the denaturant and that in presence of the $i^{th}$ denaturant concentration, respectively. The slope is defined as m, as for many protein unfolding processes, the free energy is a linear function of the denaturant concentration added.

For thermal denaturation, temperature ramping is usually applied to the protein sample and a measurement is conducted at each temperature point. All other previously derived equations are equivalent, but the free energy is described by "the modified Gibbs-Helmholtz equation":

$$\Delta G = \Delta H \left(1 - \frac{T}{T_m}\right) - \Delta C_p \left((T_m - T) + T\ln\left(\frac{T}{T_m}\right)\right).$$

Here, T is the measurement temperature and $T_m$ is the melting temperature, where K=1. The enthalpy at T is represented by ΔH, whereas $\Delta C_p$ is the heat capacity change before and after the unfolding process. For a monomeric protein unfolding, this value is set to 0. A direct measurement of $\Delta C_p$ can be obtained using differential scanning calorimetry (DSC), if needed. If we assume that the enthalpy does not change over the studied temperature range, then the Van't Hoff enthalpy, $\Delta H_v$, can be obtained by the following equation:

$$\ln\left(\frac{K_1}{K_2}\right) = \frac{-\Delta H_v}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right).$$

## Chemical denaturation

To demonstrate that Raman spectroscopy can characterize the thermodynamics of a protein folding-unfolding process, we used GdnHCl to denature a protein sample. A series of GdnHCl concentrations (from 0 M to 4.5 M) was added to 20 mg/mL human serum albumin (HSA) solutions at pH 7.0 in PBS buffer. The protein plus GdnHCl mixtures were equilibrated at room temperature overnight before measurement. Only the Amide I region, which provides protein secondary structural information, is used for the analysis that follows.
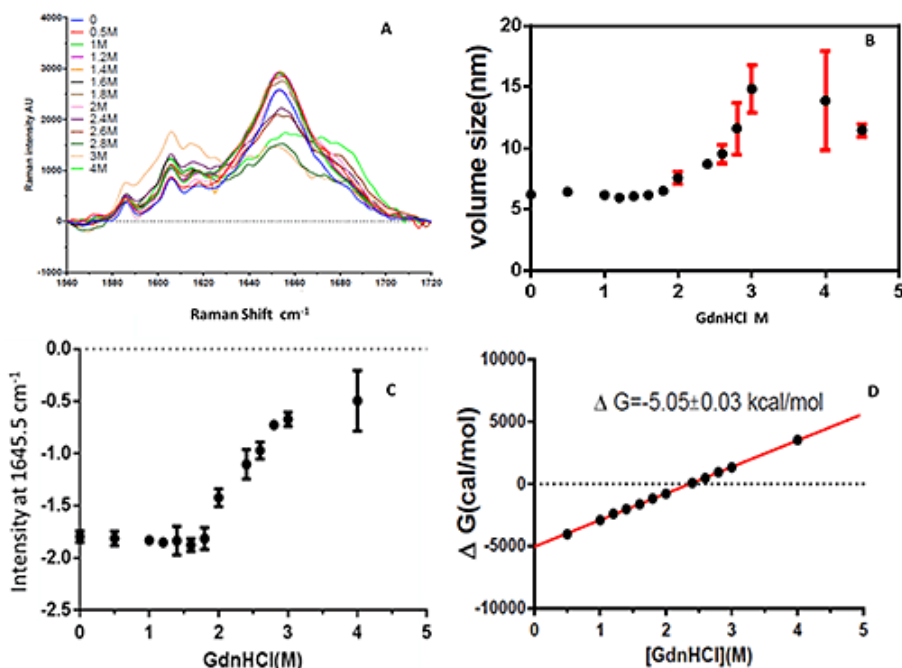
Figure 1: Buffer-subtracted Raman spectra of HSA in presence of varied concentrations of GdnHCl at pH 7.0 with PBS buffer, 0.15 M ion strength at 25 °C (A). The size trend from volume percentage measured from DLS and corrected by viscosity and refractive index for corresponding GdnHCl concentration (B). The intensity trend of secondary derivative of HSA Raman spectra at 1645.5 cm-1 with midpoint = 2.31 M ± 0.09 M GdnHCl (C). Plot of free energy against GdnHCl concentration, fitted to give a free energy of -5.05 kcal/mol ± 0.03 kcal/mol (D)

Figure 1 shows data obtained for HSA chemical denaturation experiments. Panel A shows the spectra from solutions with varying concentrations of GdnHCl, where the evolution of the Amide I peak profile indicates that helical content is reduced and β-sheet structure increases as a function of increased GdnHCl concentration. Panel B reveals a consistent trend of increasing size as unfolding progresses. Panel C shows how the secondary structure changes as a function of GdnHCl concentration, and panel D presents the data fit that results in the determination of the free energy for this specific sample. The values are generally consistent with literature reports[1,2].

## Thermal denaturation

For the thermal denaturation study, lysozyme from chick egg white was prepared in 20 mM phosphate-citrate buffer, pH 4.0, at 30 mg/mL. The sample went through a typical thermal ramping experiment with Raman and DLS data collected on the sample sequentially at a single temperature. Lysozyme was studied previously to confirm the complete reversibility of the unfolding (Use of DLS/Raman to Study the Thermal Unfolding of Lysozyme). This is necessary to validate simplifying assumptions used to obtain the thermodynamic parameters, i.e. that concentration can be replaced by "fraction folded" values.
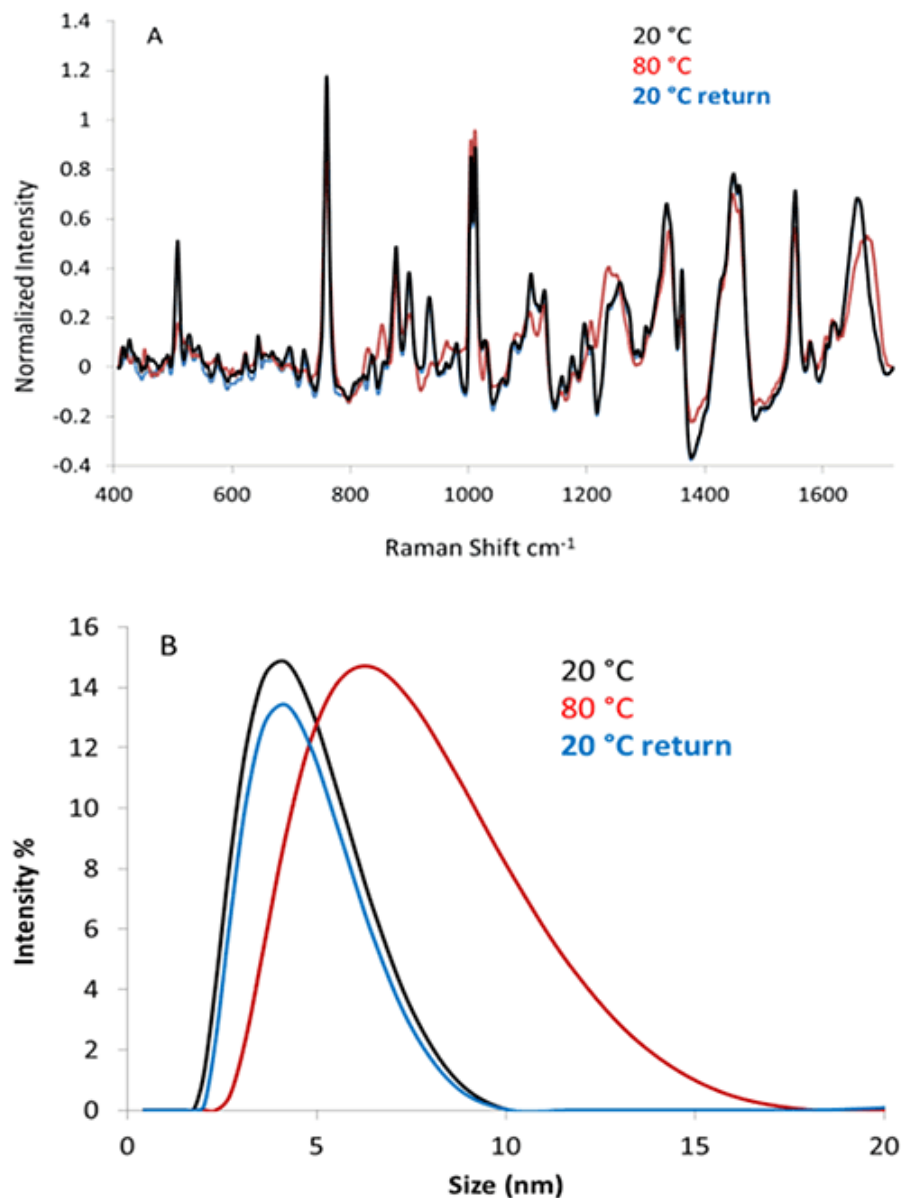
Figure 2: The reversibility test of lysozyme at pH 4.0. A. Representative Raman spectra of the sample at 20°C, 80°C and retuning to 20°C. B. The corresponding size transition data measured for the same sample. Both Raman and DLS indicate reversibility

As shown in Figure 2A, the Raman spectrum taken at 20°C before heating (black curve) overlays almost exactly with that (blue curve) taken after returning to 20°C post heating to 80°C. However, the spectrum taken at 80°C (red curve) shows clear variation in the Amide I and Amide III peaks (1580 cm$^{-1}$ to 1720 cm$^{-1}$ and 1230 cm$^{-1}$ to 1300 cm$^{-1}$), indicative of increased β-sheet structure. Additional changes are seen in the tyrosine region (820 cm$^{-1}$ to 860 cm$^{-1}$, indicative of unfolding, and lower intensity for disulfide bond region (500 cm$^{-1}$ to 550 cm$^{-1}$) that could be indicative of disulfide bond conformational changes. DLS data from the same sample/experiment (shown in Figure 2B) reveal a corresponding increase in size on heating to 80°C, with a corresponding reversion to the original size and polydispersity on return to 20°C. Both the Raman and DLS data indicate clear reversibility.

Having established the validity of reversibility for the sample, a complete thermal ramping experiment was conducted with 1°C increase per step from 20°C to 90°C. A PLS model was applied to obtain the secondary structure information for the sample across the temperature range. As shown in Figure 3, by plotting the helical percentage against temperature, and assuming a typical two state model, we obtain the following values: $T_m$, $T_{onset}$, and the Van't Hoff energy, $\Delta H_v$, which represents the dependence of free energy on temperature. The values are fairly close to those reported in the literature under similar conditions[3,4]. The free energy for a specific temperature can be calculated once the equilibrium constant, K, is determined.
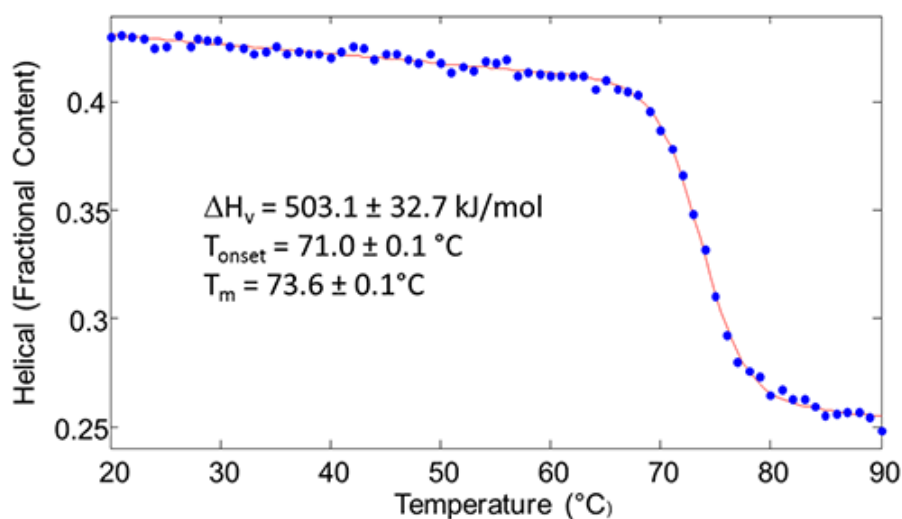


Figure 3: The thermal transition curve of lysozyme sample at pH 4.0 with the fitted parameters, Tm, Tonset, and the Van't Hoff energy, ΔHv

The above determined values are based on a specific secondary structure component, e.g. the helical percentage. However, due to the abundant structural markers offered by Raman, and the generalizability of the approach, it is possible to determine and compare these values for a variety of structural and size parameters.

Table 1: The thermal transition parameters fitted for lysozyme unfolding

| Lysozyme | $T_{onset}$ °C | $T_m$ °C | $\Delta H_V$ kJ/mol |
|---|---|---|---|
| α-Helix | 71.0 ± 0.1 | 73.6 ± 0.1 | 503.1 ± 32.7 |
| β-sheet | 72.0 ± 0.3 | 73.6 ± 0.3 | 674.5 ± 44.5 |
| Random coil | 71.0 ± 0.1 | 73.6 ± 0.1 | 477.7 ± 11.5 |
| $Trp_{com}$ 1550 $cm^{-1}$ | 72.0 ± 0.3 | 74.0 ± 0.3 | 642.0 ± 149.7 |
| $Trp_{com}$ 870 $cm^{-1}$ | 72.0 ± 0.2 | 74.0 ± 0.4 | 503.1 ± 17.2 |
| S-S $I_{507.5}$ | 70.0 ± 0.3 | 73.4 ± 0.2 | 338.3 ± 35.1 |
| Z-averaged size | 72.0 ± 0.4 | 74.0 ± 0.4 | 624.9 ± 29.7 |

In Table 1, the thermal parameters derived from a variety of structural markers are summarized and compared. It is interesting to note that although $T_m$ and $T_{onset}$ are relatively consistent, the Van't Hoff energy values derived from the different parameters vary. For instance, the $\Delta H_v$ value derived from the disulfide bond confirmation transition is the lowest; values derived from helical and random coil content, as well

as the hydrogen bonding environment of tryptophan ( $Trp_{870}$) are comparable to each other; while the values for β-sheet, tryptophan dihedral angle ($Trp_{1550}$) and size are similar. This indicates that the technique may provide a more thorough characterization of the unfolding process than is available by just monitoring a single structural parameter, which is the current "gold standard".

Table 2: The thermal transition parameters fitted for HSA at different pHs.

| HSA | $T_{onset}$ °C | $T_m$ °C | $\Delta H_v$ kJ/mol |
|---|---|---|---|
| pH 3 | 53.0 ± 0.3 | 59.5 ± 0.3 | 134.7 ± 2.5 |
| pH 5 | 63.0 ± 0.2 | 66.3 ± 0.2 | 313.8 ± 14.2 |
| pH 7 | 68.0 ± 0.2 | 71.6 ± 0.2 | 297.1 ± 15.9 |
| pH 8 | 64.0 ± 0.2 | 69.3 ± 0.3 | 225.9 ± 7.1 |

The technique can also be used to compare the thermal stability of one protein sample under varied formulation conditions, such as: pH, ionic strength, buffer components. As an example of this capability, 37 mg/mL HSA samples in 20 mM phosphate-citrate buffer are compared at different pHs. As summarized in Table 2, by fitting the transition curve for β-sheet structure and comparing the different pH values, the sample of HSA at pH 7 is determined to be the most stable. The $\Delta H_v$ values derived from β-sheet content also show pH dependence.

The Van't Hoff energy obtained using this method may not directly comparable to that obtained from DSC. Comparisons would be more valid under certain circumstances: if the protein unfolding process is highly cooperative, the experimental/sampling conditions are comparable, and no aggregation or self-association is taking place. Interestingly though, as the Van't Hoff energy is a fitted model parameter, differences between it and DSC derived values could be used to judge deviations from the assumption of a real two-state transition; i.e. the greater the variation between these values, the less likely the unfolding follows a two-state model.

## What combined DLS/Raman measurements add to thermal stability measurements

As has been stressed throughout this note, protein folding-unfolding needs to be investigated under well-defined conditions to be valid, i.e. low concentration, fully reversible, etc. However, in practice, especially during thermal stability or formulation screening studies, experiments might need to be conducted at formulation conditions, i.e. in the presence of excipients, at relatively high concentrations, etc. Moving away from ideal conditions, unfolding and aggregation compete with each other as shown here with a very simplified model:

$$Folded \underset{k_f}{\overset{k_u}{\longleftrightarrow}} Unfolded \overset{k_{agg}}{\longrightarrow} Aggregates$$
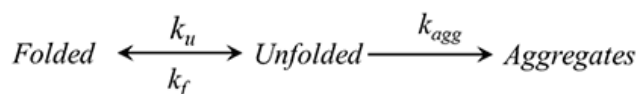
Here, the term 'aggregates' includes all kinds of protein aggregates, such as soluble oligomers, fibrils, precipitates, etc. $k_u$, $k_f$, and $k_{agg}$ refer to the rate constant of unfolding, folding and aggregation, respectively. In the scenario,

$k_{agg}$ >> $k_f$ , a very small fraction of unfolded species can trigger significant aggregation. As increasing temperature and concentration will increase the aggregation rate

constant, it is reasonable to assume that typical thermal stability measurements reflect both unfolding and aggregation processes.

The current work horse techniques for thermal stability testing, circular dichroism (CD) and differential scanning calorimetry (DSC), operate under the assumption that only unfolding is being monitored, and as such provide an incomplete picture of what is happening. Specifically, CD provides a window into conformation changes, while DSC characterizes the thermal behavior of a sample. In order to thoroughly understand protein aggregation, these separate pieces of information may not be sufficient. Recently, as shown in Figure 4, Sahin et al. point out that "DSC alone is not sufficient to determine the onset of soluble aggregate formation".



RESEARCH ARTICLES
*Biotechnology*

**Aggregation and pH–Temperature Phase Behavior for Aggregates of an IgG2 Antibody**

ERINC SAHIN,[1] WILLIAM F. WEISS IV,[1] ANDREW M. KROETSCH,[1] KEVIN R. KING,[2] R. KENDALL KESSLER,[2] TAPAN K. DAS,[2] CHRISTOPHER J. ROBERTS[1]

[1]Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716
[2]Pfizer Biotherapeutics Research and Development, Chesterfield, Missouri 63017

Received 28 August 2011; revised 19 December 2011; accepted 27 December 2011
Published online 13 January 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.23056

As a result, DSC alone is not sufficient to determine the onset of soluble aggregate formation during thermal scanning, and therefore which endothermic event correlates most strongly with aggregation. Monomer

Figure 4: Paper from Sahin et al. 2012 in J. Pharm. Sci. 101, 1678-1687. The figure on the left is a representation of DSC (line, right axis) and monomer loss (circle, left axis) as functions of temperature at three pHs. The sample concentration is 1 mg/mL IgG2 samples in 10 mM citrate buffer with 1°C/min scanning rate

Furthermore, aggregation and unfolding are convolved in the endotherm peak during DSC thermal ramping. The DSC profile in Figure 3 (left column) shows a typical 3-peak endotherm profile for a mAb, with size exclusion chromatography (SEC) results shown as open circles. The SEC monomer fraction values are shown on the left y-axis, and indicate that significant monomer loss, upwards of 50% at the transition midpoint, occurs during the course of the DSC measurement. With >50% monomer loss, the DSC $\Delta C_p$ results cannot be interpreted strictly as unfolding energy, although this is the typical assumption made. The sample concentration for this data was typical for DSC experiments, at approximately 1 mg/mL. Higher sample concentrations (typical for formulated biopharmaceuticals) would be expected to exhibit even higher levels of aggregation. This discussion points to the importance of the technique presented here, in which protein conformation change (via Raman) and size distribution evolution (via DLS) is determined on the same sample on a single analytical platform. The ability to couple these two analytical techniques has the potential to enhance our understanding of the protein change upon thermal stress. Because of the richness of structural data contained within a Raman spectrum, the combined approach is not limited to only thermal stress, but is capable of characterizing the impact of a broad range of induced sample stresses.
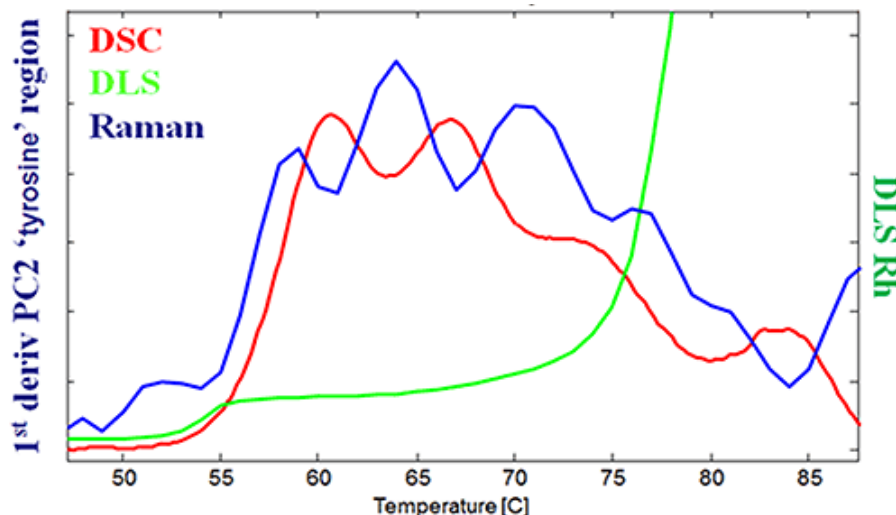
Figure 5: Thermal ramping profile of a monoclonal antibody with DSC, and DLS-Raman.

Figure 5 overlays transition profiles of DSC, Raman and DLS, highlighting the benefits of combining DLS-Raman to characterize a mAb sample via thermal ramping. The Raman and DLS data were acquired on the same sample at the same time, whereas the DSC data was acquired separately. DSC data shows multiple transitions, as do the Raman and DLS results. The DLS and Raman data correlate well with the first DSC transition at ~ 53°C. The high temperature transition at ~75°C is also seen also by all three modalities. The addition of the DLS data to the DSC trace though, reveals significant aggregation at this temperature. Between these two temperature extremes, transitions are seen in both the DSC and Raman data, but are missing from the DLS data. The fact that multiple Raman transitions appear to directly precede those seen in DSC provides a tantalizing glimpse into the capability provided by the correlation of conformation, size and thermal events during a temperature ramp experiment. Additionally, with the ability to work with high concentration samples, and the abundance of protein markers (secondary and tertiary structure), coupled with the insights from DLS (e.g., size distribution, polydispersity), make the DLS-Raman system a powerful tool to characterize protein thermal stability. Although it might not completely decouple unfolding from aggregation process, the system provides a real-time picture of the sample.

## Conclusions

The combination of DLS and Raman spectroscopy in a single analytical platform provides a variety of benefits:

1.  It is compatible with traditional protein unfolding studies, using either thermal or chemical denaturants, and provides data that correlate well with the literature values for model proteins.
2.  The DLS-Raman system provides a more complete characterization of protein unfolding, given the ability to probe abundant protein structural makers (secondary and tertiary) and the simultaneous measurement of size distributions.
3.  It provides a complete real-time picture of a protein under thermal or other stresses, and begins to decouple unfolding and aggregation processes, which other analytical approaches do not.

## 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 markets.

## References

1.  Beatriz Farruggia and Guillermo A. Pico. International Journal of Biological Macromolecules. 26 (1999) 317-323;
2.  Basir Ahmad, Md Z. Ahmed, Soghra K. Haq, Rizwan H. Khan. Biochimica et Biophysica Acta 1750 (2005) 93-102;
3.  N. Kishore and B. Sabulal. Pure& Appl. Chem. 70 (1998) 665-670;
4.  Life Sciences Application Note. Calorimetry Sciences Corporation. Document No. 20211021306. Feb. 2006. http://www.bioch.ox.ac.uk/aspsite/services/equipmentbooking/biophysics/CSCnote3.pdf

Malvern Instruments Limited

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

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