Real-time assessment of thermal perturbation of therapeutic proteins: Aggregation

With thanks to Wei Li[1], Xianglei Liu[2] and Dimiter Dimitrov[1,2]
[1]National Cancer Institute, MD, USA and [2]Center for Antibody Therapeutics, University of Pittsburgh, PA, USA

Summary

- This Application Note illustrates the use of the ProteinMentor platform for a comparative assessment of the stability of an array of proteins in solution.
- The unique combination of microscopy and spectroscopy, along with the ability to apply a thermal stress across the array of samples, provides for real-time stability assessment in a single experiment.
- Visual and spectral analysis of the data provides detailed insight into the identity, region prone to aggregation, extent and mechanism for aggregation.

Introduction

Recent years have seen an unprecedented growth in the development of biological molecules and biological products as therapeutic treatments for a wide variety of diseases. The challenges of formulating these products to optimize stability, preserve therapeutic efficacy and minimize immunogenicity are well-known to be far more complicated than traditional pharmaceutical products. One class of biological therapies that have been heavily focused on are therapeutic proteins, including monoclonal antibodies (mAb). According to the Antibody Society[1,2] there are approximately 100 approved antibody therapeutics in the EU and US, with an additional 18 in regulatory review.

Analytical methodologies to assess critical quality attributes (CQAs) of therapeutic proteins are varied, time-consuming and can involve complicated sample preparation and analysis. Multiple analytical tools are routinely deployed for each CQA assessment, and the results provide limited understanding of how, and to what extent, each CQA impacts safety and efficacy. This fragmented approach contributes to the high cost, long timelines and increased risk of safety and efficacy of the therapeutic proteins.

Understanding stability and aggregation at the molecular level is critical for the development and commercialization of therapeutic proteins. Comprehensive characterization and efficient monitoring of these two CQAs can effectively reduce the risk of immunogenicity.

In this application note, we describe the use of the ProteinMentor platform for the characterization and monitoring of model therapeutic proteins as they undergo thermal stress. The model proteins used were constructs of the C_H2 domain of IgG1, which comprises of three helical segments, several different β-turns and three- and four-strand β-sheets pinned by an intrachain disulfide bond.

ProteinMentor is a unique combination of Quantum Cascade Laser (QCL) microscopy and spectroscopy which allows for the direct visualization of aggregates and the determination of the regions, amino acids and mechanisms that lead to aggregation. We demonstrate the applicability of the platform for the assessment of stability and aggregation. A detailed description of the full study was previously published by Pastrana et al., 2021, see reference 3 for details.

Methods

Samples

Briefly, a model system using three constructs of an unglycosylated C_H2 domain of human IgG1 was used for the study. The native construct (residues 231-342) consisted of two anti-parallel β-sheets and a native disulfide bond between residues 261 and 321, which stabilizes the domain. The second construct (C_H2s) was a shortened construct (residues 238-342), which has been shown to have increased thermal stability[4]. The third construct was C_H2s modified to introduce a second disulfide (between residues 242 and 334). For a full description of the three constructs, refer to the Methods in reference 3.

Construct Name	IgG1 Heavy Chain Residues	Description	Melting Temperature
C_H2	231-342	Native C_H2 from IgG1; C-terminal extension (16 amino acids)	~57 °C
C_H2s	238-342	N-terminal truncation of C_H2	C_H2 + 5 °C
m01s	238-342	N-terminal truncation and site directed mutation of 242 and 334 to create second disulfide	C_H2s + 20 °C

Instrumentation

The ProteinMentor™ platform consists of a Quantum Cascade Laser Microscope (QCLM) with a fully 21 CFR part 11 compliant computer data management system (CDMS). For full experimental details see reference 3.

The stability study was performed within the temperature range of 28-60 °C. These temperatures allowed for the investigation of early events that cause destabilization of the proteins prior to their unfolding.

Data Processing

Correlational analysis was performed using both two-dimensional infrared correlational spectroscopy (2D-COS) and two-dimensional co-distribution correlation spectroscopy (2-D CDS) in the Correlation Dynamics™ software (Protein Dynamic Solutions). The temperature range of 28 – 48 °C was selected to compare the stability of the different constructs. For a comparative analysis we chose a temperature range well below the onset of thermal transition to establish differences in stability that would not involve thermal unfolding to the point of denaturation.

Results and Discussion

Hyperspectral Images – Visual Inspection for Aggregation

ProteinMentor provides a visual indicator of aggregates greater than 4.25 µm, the spatial resolution of the system. For the study here, no visible aggregates were apparent, with clear, homogenous solutions shown by the hyperspectral (HS) images. Representative images are shown in Figure 1 for each of the proteins at 28 °C and 60 °C. Each of the images are clear, with no visible sign of aggregation.

Even in the absence of visual aggregation, however, aggregates smaller than 4.25 µm may be present. Evaluation of the spectral data from ProteinMentor is able to determine if any aggregation was induced by the thermal stress imparted on the proteins. The spectral evaluation can also detect changes in the protein structure that may lead to the onset of aggregation or other issues that impact therapeutic efficacy or immunogenicity.

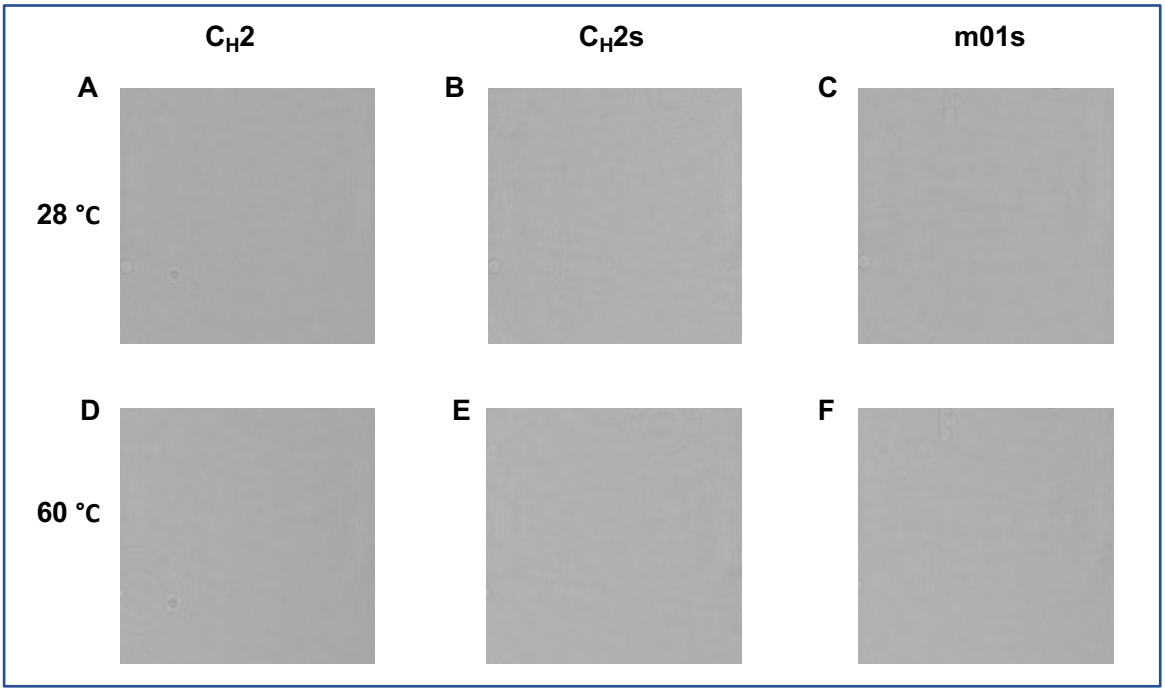

Figure 1. Representative HS images of the three constructs of the mAb fragments. Images shown for the initial temperature of 28 °C (A-C) and final temperature of 60 °C (D-F). HSIs for each of the constructs are displayed in columns - C_H2 (A, D), C_H2s (B, E) and m01s (C, F) No aggregates were observed within the spatial resolution of 4.25 μm. These HS images were generated from 223,000 QCL IRM transmission spectra.

Spectral Analysis and 2-Dimensional Correlation Spectroscopy

Each HS image is comprised of 223,000 QCL IRM spectra and takes less than 30 seconds to acquire. To understand how the platform uses these high signal-to-noise HS images, evaluation of the spectral data focuses on the assignment and inspection of well-known peaks within IR spectra of protein samples. The peak intensities are then compared as the samples are heated. Figure 2 shows overlays of the IR spectral region from 1780-1485 cm^{-1} for each protein at the 6 different temperatures. This selected spectral region includes two of the signature IR peaks used in protein analysis (the amide I band at 1700-1600 cm^{-1} and the amide II band at 1600-1500 cm^{-1}). The absorbance of water overlaps the region of the amide I band, however, even with the high absorptivity of water, the enhanced signal-to-noise (S/N) provided by the QCLM, enables visualization of the incremental intensity changes in the amide I band. Furthermore, the design of the ProteinMentor platform, with the heated Slide Cell Array (SCA), provides a high S/N with accurate thermal control and allows for the difference spectra to be used.

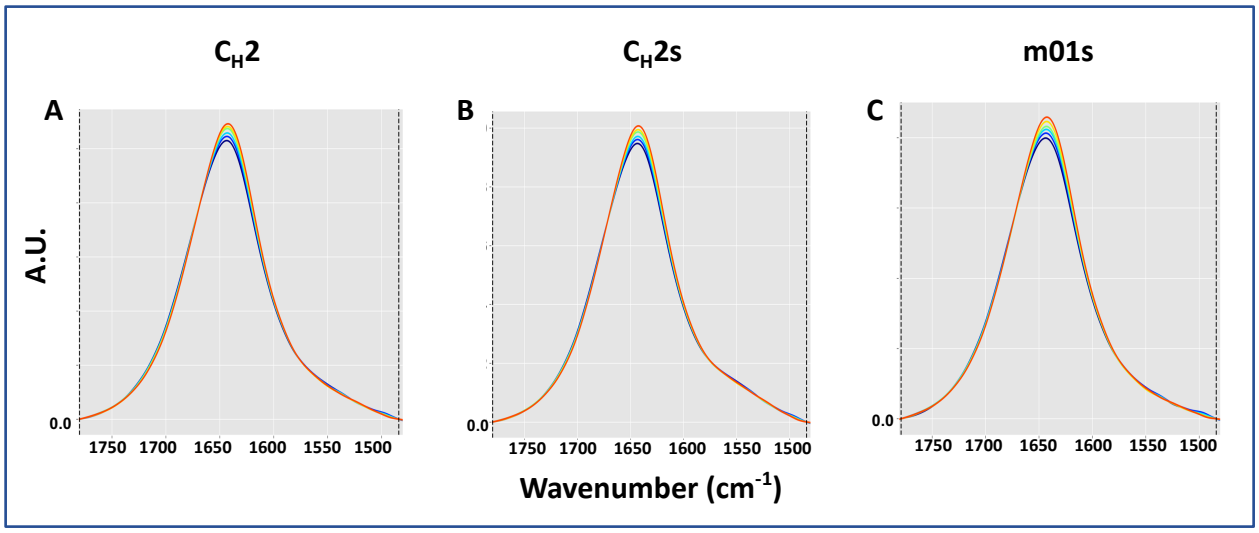

Figure 2: QCL IRM baseline corrected spectral overlays within the spectral region of 1780 – 1485 cm^{-1} and temperature range of 28 - 48 °C to address comparability. (A) C_H2, (B) C_H2s and (C) m01s. The vertical dashed lines represent the exact point selected for the linear baseline correction. The color temperature code for each spectral line used was: black, 28 °C; blue, 32 °C; green, 36 °C; yellow, 40 °C; orange, 44 °C; and red, 48 °C.

Figure 3 displays the difference spectra for the same data as shown in Figure 2. The difference spectra are generated by simply subtracting the first spectrum from all subsequent spectra. This allows for the subtraction of the H_2O contribution and also removes any spectral contributions that did not change over the course of the thermal stress study. As can be seen in Figure 3, both positive and negative peaks are apparent in the difference spectra. It is these changes that have occurred during thermal stress. Correlation Dynamics uses the difference spectra of key regions of the IR spectrum for comprehensive analysis and reporting of the thermal stress study.

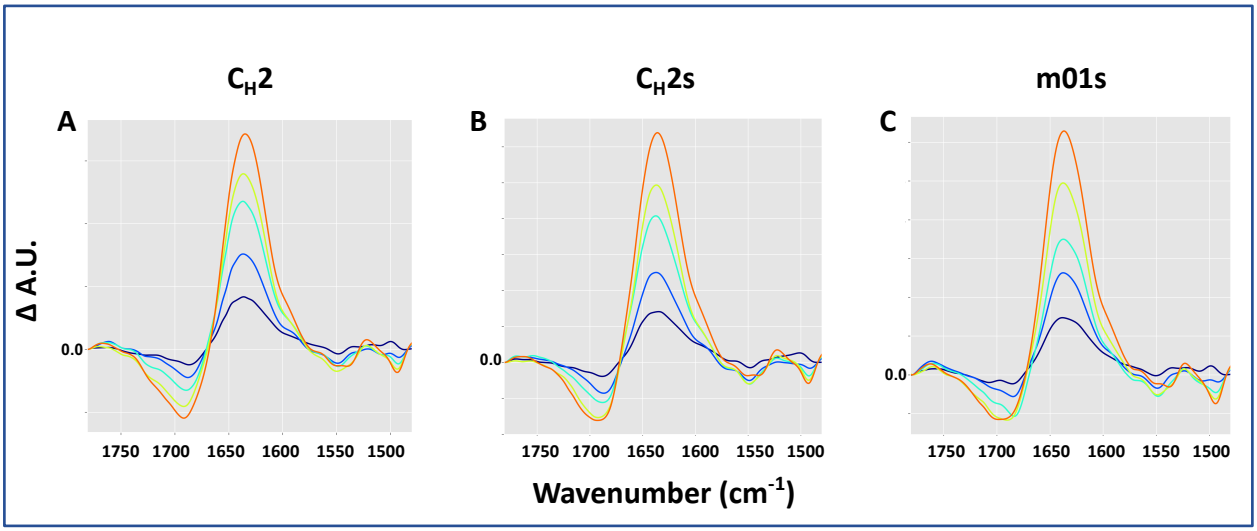

Figure 3: QCL IRM difference spectral overlay within the spectral region of 1780 – 1485 cm^{-1} and temperature range of 28 - 48 °C to address comparability. (A) C_H2, (B) C_H2s and (C) m01s. The QCL IRM difference spectra were generated by subtracting the first spectra from all subsequent spectra. The color temperature code for each spectral line used was: black, 28 °C; blue, 32 °C; green, 36 °C; yellow, 40 °C; orange, 44 °C; and red, 48 °C.

2D-COS, introduced by Noda[4], allows for the study of changes in the secondary structure of a protein and also side chain interactions as a function of external factors. In this work, the external factor was increasing temperature to introduce thermal stress to the proteins in the array. The ProteinMentor array specifically allows for the determination the temperature at which the changes occur. Using this technique, a typical IR spectrum is used as the basis for peak assignments, as described in detail in reference 3. Based on the known protein sequence, the location of the changes within the protein are also determined and protein aggregation can be assigned to specific region(s). The use of correlation plots provides enhanced resolution of the peak intensity changes within the spectral region of interest that are due to the temperature increase. Furthermore, the order of events can also be determined from the correlation plots[3,4,5].

Typical 2D-COS plots, shown in Figure 4, consist of two correlation spectral sets – the synchronous plot (Figure 4A, shown with a blue background) and the asynchronous plot (Figure 4B, shown with a green background). The synchronous plot can be considered as a similarity plot, showing the events that occur in-phase with one another[5]. The peaks on the diagonal of this plot ("auto peaks") provide information on changes in domain stability as the sample undergoes thermal stress. The asynchronous plot represents dissimilarity and identifies changes that occur out of phase with one another. Comparing the corresponding peaks in each of the synchronous and asynchronous plots allows for the determination of the order in which the structural changes occur. The asynchronous plot contains only peaks

off the diagonal ("cross peaks"). The position of cross peaks in the asynchronous plots makes it possible to map the region of the protein that is prone to aggregation, providing the amino acid level resolution.

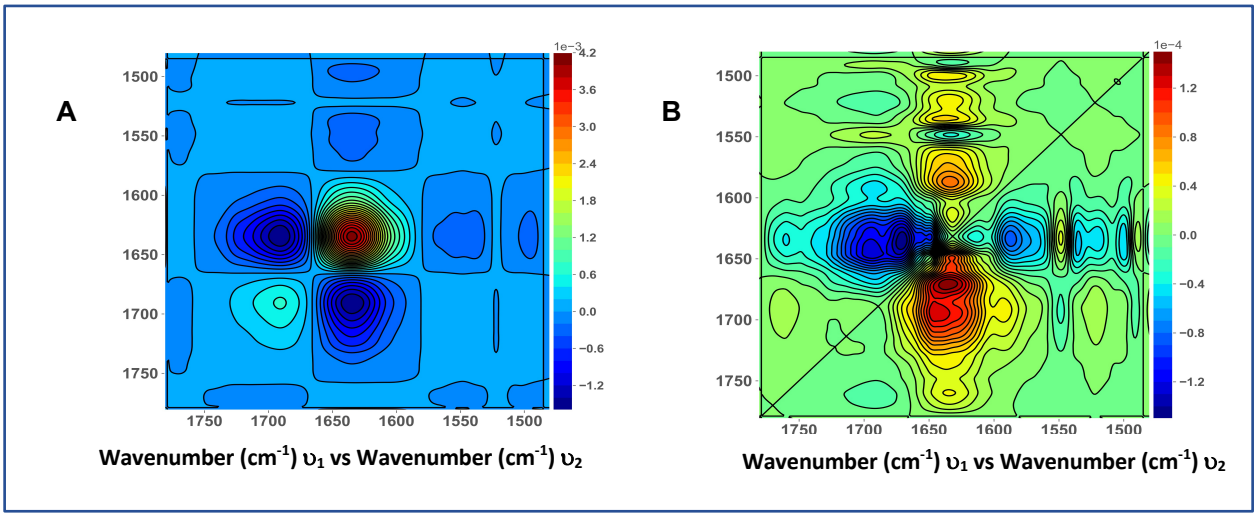

Figure 4: 2D-COS plots within the spectral region of 1780-1485 cm^{-1} and temperature range of 28 - 48 °C for the C_H2 fragment. The x- and y-axes represent wavenumber υ_1^{-1} and wavenumber υ_2^{-1}, respectively. Peaks within the plot represent changes that correlate between υ_1^{-1} and υ_2^{-1}. The magnitude of the intensity changes is shown by color, with the color scale shown on the right-hand side of each contour plot. Peaks that change in-phase with one another are represented in the synchronous plot (A, blue) and peaks that change out-of-phase with one another are represented in the asynchronous plot (B, green). The plots are symmetrical and the top triangle (left) is used to describe the features. The asynchronous plot provides the enhanced resolution needed for the molecular understanding of the mechanism of aggregation.

A second algorithm used for deeper understanding of the HS images is co-distribution correlation spectroscopy (2D-CDS). 2D-CDS uses the same dataset to highlight the changes that occur to the highest populations of proteins within the sample. The resulting data from 2D-CDS enables rapid determination of the extent of aggregation.

Using the combination of the ProteinMentor platform with correlation spectroscopy algorithms enables the determination of the relative stability of different regions of a protein, comparative analysis between proteins and identification of the regions of highest risk of thermal instability. This could be the same protein from different lots or under different formulation conditions, or, as in the case of this study, similar proteins with modifications introduced to improve the stability.

Determination of Aggregation in the Absence of Visual Evidence

The simplest way to interpret the correlation plots from a thermal stress study is to evaluate the data set for the samples that have the greatest number of peaks and the most intense color changes. It is those samples and conditions that indicate lower stability. When different proteins are analyzed on the same array, comparisons between thermal stability of different proteins can be assessed in real time.

In addition to determining which samples and regions show the greatest change, a key feature visual indicator of aggregation that is apparent in the 2D-CDS analysis, is the presence of a clover-shaped feature in the center of the plot. Figure 5 shows the 2D-CDS plot for the three constructs.

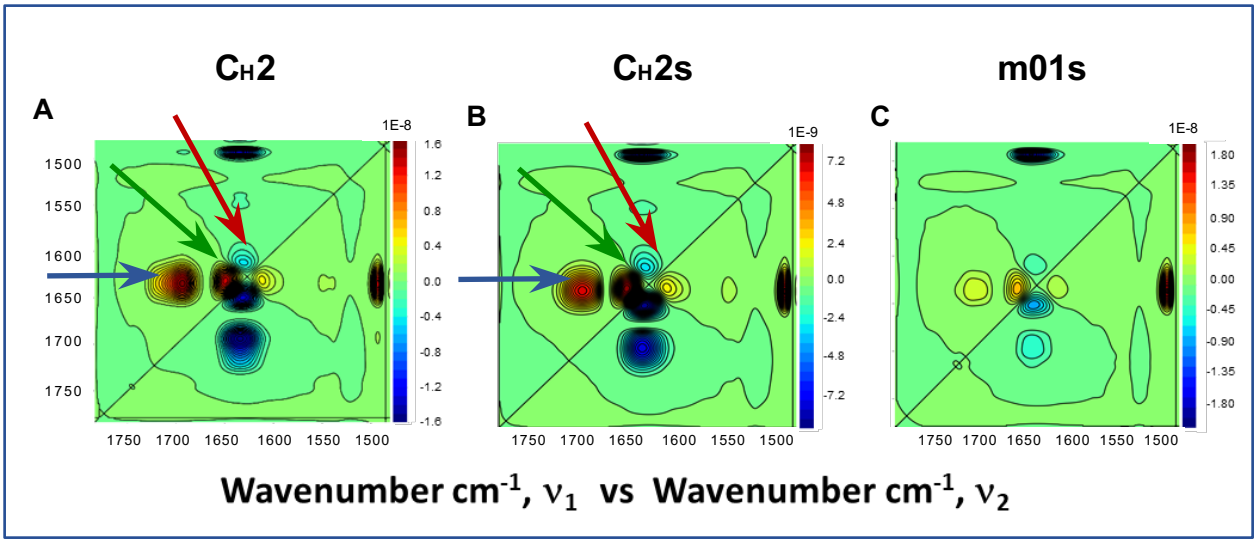

Figure 5: Co-distribution asynchronous correlation plots within the spectral region of 1780-1485 cm^{-1} and temperature range of 28 - 48 °C for: (A) C_H2, (B) C_H2s and (C) m01s. These plots represent the distribution population of the mAb fragments in solution. The color bar that defines the intensity changes is shown on the right-hand side of each plot. An overall evaluation of stability assessment can be performed. For C_H2 and C_H2s evidence of aggregation is observed within the cloverleaf set of peaks near the center of the plot, indicated by the red and green arrows in A and B. Each of the three arrows indicate the features associated with the β–sheet (υ_2 1636 cm^{-1}) that underwent the greatest changes over the temperature ramp. The red arrow indicates the self-association (υ_1 1624 cm^{-1}), the green arrow indicates the α-helix (υ_1 1653 cm^{-1}), and the intense response at 1680 cm^{-1}, 1636 cm^{-1} (υ_1, υ_2) indicated by the blue arrow is from the β-turn associated with the β-sheet. The m01s construct is the most stable of the three mAb fragments due to the greatly attenuated intensity changes observed.

Adapted with permission from Pastrana, B. et al., *Anal. Chem.*, 93, 3, 1342–1351. Copyright 2021 American Chemical Society.

As can be seen in Figure 5A and B, there is strong evidence of the clover-shaped feature for C_H2 and C_H2s, respectively. By contrast, Figure 5C shows a far lower intensity change for

m01s, indicating that m01s was much more stable when heated. For details of the underlying peak assignments and explanation of the correlation plots, refer to Pastrana et al., 2020[3].

Determination of the Order of Events Leading to Aggregation

In order to understand the increased stability of the m01s construct, 2D-COS was used to elucidate the molecular events leading to the aggregation process of C_H2 and C_H2s. Individual amino acids located within different secondary structures (namely tryptophans, arginines and glutamines) can be used as internal probes of the structural changes of the three protein constructs. For example, of the two tryptophans (1496 cm^{-1}) in the sequence of the constructs, one is buried within a β sheet motif, while the other is in a helical segment. Monitoring the changes in spectral features associated with these internal probes provides insight into changes in the protein's secondary structure as the temperature increased.

A full description of the specific changes for each of the constructs has been previously described[3,4]. The correlational analysis provided the molecular understanding for the increased stability of the m01s construct and the events leading to aggregation of C_H2 and C_H2s. The presence of a 3_{10}-helix and β-turn type III within the N-terminal end of the C_H2s and m01s constructs provided improved stability of both those constructs over the C_H2 construct, however, the C_H2s construct did undergo aggregation even with these two structural elements present. The additional disulfide bridge within the m01s construct provided greater structural stability and reduced the susceptibility for aggregation.

Comparison of Stability Between Different Proteins

The 2-D correlational algorithms not only provide information on the molecular order of events leading to aggregation and the extent of aggregation, they also identify key signature features within the HS images that can be monitored to quantitatively compare the stability between the proteins in an array.

For the three antibody constructs in this study, the key features that showed changes over increasing temperature included the β-sheets and β-turns, as indicated in Figure 5. Furthermore, stability evaluation should always include side chains involved in weak interactions such as hydrophobic, hydrogen bonding and salt-bridges. The lysine, glutamate and tryptophan residues identified by 2D-COS to be important, were therefore used to monitor the disruption of weak interactions.

A linear incremental change in β-sheet composition was observed with increasing temperature (Figure 6A) for all three constructs. A likely explanation for this observation is that the neighboring random coil segments adjacent to the fourth β-strand can convert to the β-strand. The instability of the β-turns is evident for C_H2 with the linear decrease in intensity over the course of the temperature increase (Figure 6B). C_H2s and m01s plateau at 40°C, suggesting this structure is more stable for those constructs. For C_H2s, however, the β-turn intensity starts to decrease again at 44°C (Figure 6B). For the lysine, glutamate and tryptophan residues, the changes observed were very similar for all three protein constructs (data not shown, refer to Reference 3 for details).

Overall, the results suggest that m01s has the potential for generating anti-parallel β-sheets, which was not observed for C_H2 nor C_H2s. In addition, the β-turns in m01s stabilized at 40°C compared to C_H2 and C_H2s. The presence of the additional disulfide stabilized the structure and eliminated flexibility in the C-terminal end. Even though the additional disulfide did result in some destabilization of the glutamate residues of m01s (through loss of interactions available in the native form), the combination of N-terminal truncation and the additional disulfide prevented the aggregation that was apparent with C_H2 and C_H2s constructs.

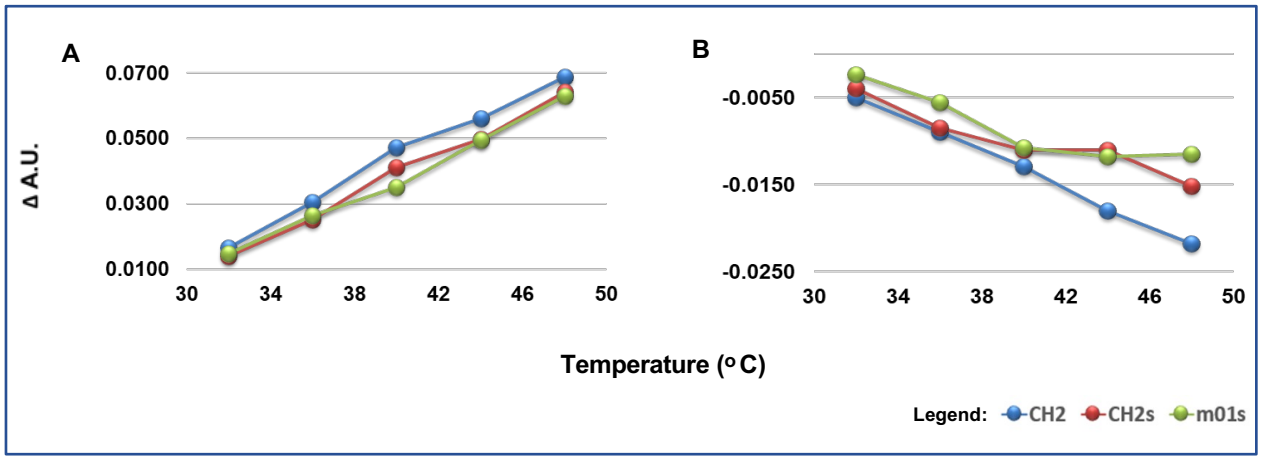

Figure 6: Change in intensity using QCL IRM difference spectra associated with (A) β-sheets (1636 cm^{-1}), and (B) β-turns (1680 cm^{-1}) over the temperature range of 32 - 48 °C for C_H2 (blue), C_H2s (red) and m01s (green). The smaller peak intensity change is indicative of lesser conformational change, therefore greater stability. The results show confirmation of the greater stability observed for m01s when comparing conformational changes of the main structural components of theses mAb fragments.

Conclusions

The ProteinMentor platform allows for the real-time acquisition of hyperspectral images for an array of therapeutic proteins in solution under controlled thermal perturbation.

2D-COS analysis of the difference spectra revealed the regions that were prone to aggregation and the difference in stability between the three C_H2 domain constructs.

The experiment described here, (from 28-60 °C) took less than 2 hours and used only 1 µL sample per well.

The ProteinMentor platform provides key advantages for the analysis of aggregation and stability of therapeutic proteins, namely:

- Minimal sample requirements
- No sample preparation
- Drug substance or drug product analyzed within the formulation
- Aggregation identity, mechanism and extent determined in a single assay
- Protein stability assessment in solution
- Stability comparison between different samples in a single assay
- Predictive results for protein developability
- Determination of the potential for immunogenicity

Protein Dynamic Solutions
Wakefield, Massachusetts
USA
www.pdsbio.com
info@pdsbio.com

References:

1. Antibody Society, Inc. https://www.antibodysociety.org/

2. Kaplon, H. and Reichart, J.M., "Antibodies to watch in 2021" *MABS,* 13:e1860476 (34 pages) **2021**. https://DOI.org/10.1080/19420862.2020.1860476.

3. Pastrana, B., Nieves, S., Li, W., Liu, X., and Dimitrov, D.S. "Developability Assessment of an Isolated CH2 Immunoglobulin Domain" *Anal. Chem.*, 2021, **93**, 3, 1342–1351. https://doi.org/10.1021/acs.analchem.0c02663

4. Gong, R.; Wang, Y.; Ying, T.; Feng, Y., Streaker, E. Prakaran, P., and Dimitrov, D.S. "N-terminal truncation of an isolated human IgG1 CH2 domain significantly increases its stability and aggregation resistance" Mol. Pharm. 10:2642−2652, 2013.

5. Noda, I. Dowrey, A.E. Marcott, C. Story, G.M. Ozaki, Y. "Generalized two-dimensional correlation spectroscopy" *Appl. Spectrosc*., 2000, **54**, 236A-248A. https://doi.org/10.1366/0003702001950454

6. Noda, I. "Techniques useful in two-dimensional correlation and co-distribution spectroscopy (2DCOS and 2DCDS) analyses" *J. Mol. Struct*., 2016, **1124**, 29e41. https://doi.org/10.1016/j.molstruc.2016.01.089