Two-dimensional Correlation and Two-dimensional Co-distribution Spectroscopies of Proteins

Belinda Pastrana PhD,

Protein Dynamic Solutions, Inc., Wakefield, MA, USA

Curtis W. Meuse PhD,

Institute for Bioscience and Biotechnology Research, National Institute of Standards and Technology, Rockville, MD, USA

Given the complexity of proteins, the broad nature of infrared bands, and the overlap of the HOH bending mode of water within the 1700–1600 cm^{-1} spectral region, it was often considered too difficult to include side-chain vibrational modes in a description of a protein's dynamic behavior in solution. Yet, it is the roles of side-chains that allow a description of a protein's dynamics to become a site-specific understanding of stability and provide a basis for the development of protein drugs.

This article describes how, in certain cases, site-specific analysis of protein infrared spectra can be performed using two-dimensional infrared correlation spectroscopy

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

(2D-COS) developed by Dr. Isao Noda. The 2D-COS algorithm spreads overlapped peaks within a spectral region of interest into a second dimension that allows for a complete evaluation of all of the peaks that change during a perturbation. 2D-COS analysis has been implemented successfully for the study of proteins under different perturbations such as pressure, chemical, pH, concentration, and most commonly temperature. Initially, the interest was in determining conformational changes during thermal unfolding, which then led to evaluating solvent accessibility through hydrogen/deuterium exchange, leading to the use of isotope editing to probe local structural changes and interactions.

Herein, we will explore the extent that both 2D-COS and two-dimensional co-distribution spectroscopy (2D-CDS) have been implemented in the study of proteins. However, we were also keen to include recent advances to site-specifically monitor protein deamidation in aqueous solution. Protein deamidation, like oxidation and glycosylation, is one of the processes that can initiate protein degradation pathways which could lead to decreased target binding, decreased stability, and possibly aggregation. These concerns have become part of the critical quality attributes (CQAs) evaluated during therapeutic protein development because they can potentially impact drug safety and efficacy. To date, tandem liquid chromatography–mass spectrometry (LC-MS) techniques have been the primary method used to ascertain the presence and sites of deamidation in protein samples. However, the LC-MS process requires multiple steps, including enzymatic digestion and peptide mapping. 2D-COS and 2D-CDS analyses provide an alternative that allows the determination of the deamidation sites and the evaluation of the stability of the protein populations within a sample without any sample preparation steps.

The 2D-COS method utilizes an innovative platform comprised of a quantum cascade laser microscope (QCLM), slide cell array, and software. The QCLM provides an enhanced signal/noise ratio and real-time acquisition of hyperspectral images of an array of proteins in solution under accurate thermal control. The method monitors the deamidation process during thermal stress. Spectral data is then subject to 2D-COS analysis to determine the extent of and site at which both asparagine and glutamine deamidation occur. In addition, 2D-CDS analysis allows assessment of changes in stability in the distributions of the protein populations during the perturbation. The results presented here for the deamidation of the NISTmAb have been validated by LC-MS to evaluate this new analytical tool. We foresee this approach as a required step in a therapeutic protein's developability assessment and release testing.

1 INTRODUCTION TO INFRARED EVALUATION OF PROTEINS

The infrared (IR) spectra of proteins are complex, and for many years the focus has been the study of the conformationally sensitive peptide bond carbonyl stretching mode ν(C═O) and N—H deformation mode δ(N—H) within the amide I band (1700–1600 cm^{-1}).[1–5] Proteins were studied in both H_2O and D_2O for validation of the secondary structure determination.[2,3] Also, the amide hydrogens of proteins were fully exchanged to deuterium to allow for the evaluation of the amide I′ and amide II′ bands (the prime nomenclature denotes the resulting peak shifts due to the change in reduced mass upon exchange from δ(N—H) to δ(N—D)). This treatment of the protein sample would effectively simplify the amide I band to contain only the peptide bond carbonyl stretching mode ν(C═O), while effectively eliminating the highly absorbing and overlapping HOH bending contribution.[2,6] The use of isotope editing was exploited to study the structural dynamics of a protein or peptide under thermal stress, which proved effective.[7] The isotope editing of the peptide bond involved $\nu(^{13}C{=}O)$ with respect to ($\nu(^{12}C{=}O)$ or $\nu(C{=}^{18}O)$ with respect to $\nu(C{=}^{16}O)$ to evaluate defined regions of a protein or peptide, taking advantage of the reduced mass effect causing a shift to a lower wavenumber for the stretching vibration of the secondary structure of interest. Although isotope editing is an effective tool, it is an expensive alternative.

2 INTRODUCTION TO 2D CORRELATION AND CO-DISTRIBUTION SPECTROSCOPIES

Isao Noda developed two-dimensional correlation spectroscopy (2D-COS) to enable spectral interpretation.[8–13] The essential component of the method is a perturbation (chemical, pressure, thermal, concentration, pH, among others) of the sample. Given this minimal experimental requirement, 2D-COS is independent of any particular analytical tool and has been applied to many types of spectroscopies. The purpose of the 2D-COS analysis is to understand the causes of peak changes based on insights from their phase and intensity changes. 2D-COS has many advantages over traditional spectroscopy for the study of proteins, including (i) enhanced spectral resolution allowing for the evaluation of side-chain modes critical to the stability of proteins; (ii) confirm peak assignments within the spectral region of interest; (iii) temporal resolution allowing for the description of the sequential order of molecular events, providing information to support

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

the molecular basis of cooperativity, which had eluded investigators for years; and (iv) the ability to monitor signature peak intensity changes or peak shifts by separating them into two categories based on their phase: (a) *in-phase,* observed in the synchronous plot and (b) *out-of-phase,* observed in the asynchronous plot. This last advantage has allowed for definitions and descriptions of dynamic molecular events within the protein. We will describe the logical implementation and progression of the 2D-COS algorithm by scientists worldwide for the analysis of proteins using a wide range of perturbations.

The value of 2D-COS[8–13] and two-dimensional co-distribution spectroscopy (2D-CDS)[14,15] analyses for the study of proteins cannot be understated. These algorithms do not require dedicated instrumentation, just a series of spectra as a function of a sample perturbation. As a basis for discussion, we include below the mathematical expressions that define the functionality of the algorithms.

2.1 Two-dimensional Infrared Correlation Spectroscopy Algorithm

2D-COS has proven helpful in that it provides a detailed molecular description of the effects of the perturbation on both the side chains and, in turn, the conformational stability of the protein. The 2D-COS algorithm[8] is defined as:

$$\tilde{A}(\nu_j, t_k) = \begin{cases} A\ (\nu_j, t_k) - \bar{A}(\upsilon_j) \text{ for } 1 \le k \le m \\ 0 \text{ otherwise} \end{cases} \quad (1)$$

Where, $\bar{A}(\upsilon_j)$ is the initial spectrum of the dataset. Synchronous 2D correlation intensities of the covariance spectral data are defined by:

$$\Phi(\nu_1, \nu_2) = \tilde{A}(\nu_1, t_j) \cdot \tilde{A}(\nu_2, t_j) \quad (2)$$

The resulting correlation intensity $\Phi(\nu_1, \nu_2)$ as a function of two independent wavenumber axes, ν_1 and ν_2, is the synchronous plot.

Asynchronous 2D correlation intensities of the covariance spectral data are defined by:

$$\Psi(\nu_1, \nu_2) = \tilde{A}(\nu_1, t_j) \cdot N_{ij}\tilde{A}(\nu_2, t_i) \quad (3)$$

The term N_{ij} is the element of the so-called Hilbert-Noda transformation matrix and is given by:

$$N_{ij} = \begin{cases} 0 \ \text{ for } i = j \\ \frac{1}{\pi(j-i)} \ \text{ otherwise} \end{cases} \quad (4)$$

The cross-correlation function is applied to the difference spectral dataset, which results in two separate, yet symmetrical 2D plots. The first plot is referred to as the ***synchronous plot***. It contains positive peaks on the diagonal, known as the *auto peaks*, and provides the overall changes observed in the spectral dataset. The relationship established in this synchronous plot relates to the spectral intensity changes that occur synchronously, hence the name. The second 2D plot is known as the ***asynchronous plot***. This plot relates the asynchronous intensity changes, resulting in enhanced spectral resolution, which can be used to correlate out-of-phase peak intensity changes, such as those defined for the deamidation process (e.g. changes observed for key signature peaks due to the deamidation of a protein). Both plots contain off-diagonal peaks, which are referred to as *cross peaks*; these peaks correlate with the spectral changes observed at the frequencies assigned to the various bands (e.g. backbone and side-chain structural elements). Spectral intensity changes observed are due to the incremental thermal perturbation applied to the monoclonal antibody (mAb) sample. No a priori knowledge of the system is required for interpretation of the results. The information in both plots allows for determining the sequential order of molecular events that occur during the perturbation by following Noda's rules.[8–10] These plots are symmetrical, and for discussion purposes, we will always refer to the top triangular portion of the plot for analysis. The application of Noda's rules begins with the plot that has the most significant resolution enhancement (i.e. the asynchronous plot):

I Asynchronous cross peak ν_2: if positive, then ν_2 is perturbed prior to ν_1 ($\boldsymbol{\nu_2 \rightarrow \nu_1}$).
II Asynchronous cross peak ν_2: if negative, then ν_2 is perturbed after ν_1 ($\boldsymbol{\nu_2 \leftarrow \nu_1}$).
III If the corresponding synchronous *cross peak is* positive, then the order of the events is established using the asynchronous plot (rules I and II).
IV However, if the corresponding synchronous *cross peak* is negative and the asynchronous *cross peak* is positive, the order is reversed.

Using this process an order of events can be established for each peak observed on the ν_2 axis.

2.2 Two-dimensional Co-distribution Spectroscopy Algorithm

Whereas 2D-COS analysis provides detailed molecular events that the protein undergoes upon thermal (or other) stress, 2D-CDS plots reflect the distribution of the protein ensemble in solution and reveal the main structural events common to the majority of

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

the proteins in the sample. The spectral correlations observed can be compared directly between protein samples, allowing for determinations of relative stability and aggregation. Used together, the methods can provide an understanding of the mechanisms of aggregation and stability within the different structural elements of a protein.

2D-CDS is especially suited for evaluating the dynamic elements common to an ensemble of protein structures in a solution subjected to a perturbation.[14,15] For a set of m spectra, $\widetilde{A}(\upsilon_j, t_k)$ are obtained as a function of the *spectral variable* (υ_j) with $j = 1, 2, \ldots, n$ and some *perturbation variable* t_1, with $i = 1, 2, \ldots, k$ at defined intervals between t_1 and t_k. In our data, the *spectral variable* υ is the QCLM (see Section 11) wavenumber, and the *perturbation variable* is temperature.

The dynamic spectrum $\widetilde{A}(\upsilon_j, t_k)$ is defined as:

$$\widetilde{A}(\upsilon_j t_k) = A(\upsilon_j, t_k) - \overline{A}(\upsilon_j) \tag{5}$$

with the average $\overline{A}(\upsilon_j)$ as the reference spectrum given by:

$$\overline{A}(\upsilon_j) = \frac{1}{m} \sum_{k=1}^{m} A(\upsilon_j, t_k) \tag{6}$$

The asynchronous 2D-CDS spectrum $\Delta(\upsilon_1, \upsilon_2)$ is defined as:

$$\Delta(\upsilon_1, \upsilon_2) = \frac{\mathrm{T}(\upsilon_1, \upsilon_2)}{m(m-1)} \sum_{k=1}^{m} k \left\{ \frac{\widetilde{A}(\upsilon_2, t_k)}{\overline{A}\,(\upsilon_2)} - \frac{\widetilde{A}(\upsilon_1 t_k)}{\overline{A}(\upsilon_1)} \right\} \tag{7}$$

The total joint variance $\mathrm{T}(\upsilon_1, \upsilon_2)$ is given by:

$$\mathrm{T}(\upsilon_1, \upsilon_2) = \sqrt{\Phi(\upsilon_1, \upsilon_1) \cdot \Phi(\upsilon_2, \upsilon_2)} \tag{8}$$

By convention, the value of $\Delta(\upsilon_1, \upsilon_2)$ is set to be zero, when $\overline{A}(\upsilon_1) = 0$ or $\overline{A}(\upsilon_2) = 0$ is encountered, indicating a lack of spectral intensity at either QCLM wavenumber. The asynchronous 2D-CDS intensity is a measure of the difference in the distribution of two spectral signals along the perturbation variable axis based on the mathematical tool known as moment analysis.

The interpretation of the 2D-CDS asynchronous plot is direct. For a positive cross peak $\Delta(\upsilon_1, \upsilon_2) > 0$, the presence of the spectral signal at υ_1 is distributed predominantly at an earlier stage, prior to υ_2. In the case of a negative cross peak $\Delta(\upsilon_1, \upsilon_2) < 0$, the order is reversed. Finally, when $\Delta(\upsilon_1 \upsilon_2) \approx 0$, the average distributions of the spectral signals are similar, and therefore they co-exist.

3 TWO-DIMENSIONAL INFRARED CORRELATION SPECTROSCOPY ANALYSIS FOR SECONDARY STRUCTURE DETERMINATION

2D-COS enhances resolution by separating the spectral data into a second dimension, providing both spatial and temporal resolution. The only requirement is a gradual perturbation during spectral acquisition. This has allowed for the evaluation of proteins under different perturbation conditions such as pressure,[16] ligand binding,[17] hydrogen/deuterium exchange,[18–23] pH,[23–25] and concentration,[23–26] with the most commonly used perturbation being thermal.[24–49] Some of these 2D-COS experiments determined the minimum number of carbonyl stretching peaks that contributed to the amide I and II bands, eliminating the subjectivity associated with curve-fitting analysis.[17,24,25,27,29,31,32,34–37,40–42] The minimum number of peaks required for curve fitting analysis of the amide I and II bands is based on the number of auto peaks and their position on the diagonal of the synchronous plots, which provides all of the peaks changed by the perturbation. Spectral simulations were employed for comparison with the empirical spectral data to verify the information gleaned from the interpretation of the synchronous and asynchronous plots while applying Noda's rules. The selectivity of the 2D-COS method was validated using simulation spectra of the amide I and II bands, providing links to various types of protein structures.[39–41]

4 SOLVENT ACCESSIBILITY AND PROTEIN FLEXIBILITY

Several groups have applied 2D-COS to the kinetics of hydrogen/deuterium exchange of amide bonds to describe protein solvent accessibility.[18–22,24] The approach involved attenuated total reflectance (ATR) IR spectroscopy whereby a thin layer of protein film was deposited onto a high refractive index ATR crystal (germanium, ZnSe). The film was then purged with D_2O vapor to allow for the exchange of the ionizable protons of the protein to exchange while spectral data were acquired. Different classes of proteins have been subject to the analysis of the kinetics of hydrogen/deuterium exchange, allowing for the rate constants to be determined and classified. The first of such studies was reported by Nabet and Pézolet.[18] 2D-COS extended the analysis to allow the determination of regions of the proteins that were most solvent exposed and their extent of exchange. The use of hydrogen/deuterium exchange

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

has also been employed in the study of protein complexes to evaluate the changes in solvent accessibility upon complex formation.

One 2D-COS example is the evaluation of human serum albumin kinetics of exchange to allow for the separation of the rates of amide proton exchange based on secondary structure.[20] Another example is *Chlamydomonas reinhardtii (Cr)* centrin, a highly soluble calcium-binding protein, where both its full-length form and terminal end fragments were studied.[21] This combinatorial approach allowed for the distinction of the different rates of exchange for a highly symmetrical protein. This definitive assessment showed a greater solvent accessibility for the N-terminal fragment (residues 1–94), which exhibited faster kinetics of exchange when compared to the C-terminal fragment (residues 99–169), while the full-length protein presented a combination of the two rates of exchange as a function of time.[21] A third example is the comparative analysis of hydrogen/deuterium exchange involving human epidermal growth factor receptor ligands upon binding to the receptor. These included epiregulin, EGF, TGF-α, amphiregulin, and heparin-binding-epidermal growth factor receptor which were found to have differences in both their exchange rates and their aqueous accessible surface regions. For amphiregulin and heparin-binding-epidermal growth factor receptor, the secondary structure composition was also determined for the first time.[22]

Hydrogen/deuterium exchange rates associated with the different structural elements of a pure protein and for its complex are essential for the determination of a binding interface. One example is streptavidin and its complex with biotin, which was used to monitor the exchanging of amide protons. The slowest exchange rate was observed for the β-sheets, which occurred on the order of hours. The binding interface was determined by the slower rates of exchange when compared to the pure components' exchange rates for the same secondary structural elements.[23,24]

5 SIDE CHAIN VIBRATIONAL MODES AS INTERNAL PROBES IN PROTEINS

The enhanced spectral resolution and peak correlations from 2D-COS allow side-chain vibrational mode analysis to improve both spatial and temporal resolution.[17,25,27,29,31,35–37,40,41,45–49] The peaks associated with side-chain vibrational modes can serve as probes of their peptide backbone locations by correlating with the conformationally sensitive ν(C═O) in the asynchronous plot.

5.1 Advantage of H_2O Over D_2O in Studying Proteins

A fully hydrogen/deuterium exchanged protein requires the use of D_2O at 99.9 atom % D, the number of side-chain modes within the amide I′ and II′ bands is limited to four: glutamate $\nu(COO^-)$, aspartate $\nu(COO^-)$, arginine $\nu_s(CN_3H_5^+)$, and the tyrosine ν(C═C) ring.[1–5] The side-chain vibrational modes not observed are due to decreased absorptivity upon hydrogen/deuterium exchange of the exchangeable protons and the shifts to lower wavenumbers that fall outside the spectral region of 1700–1500 cm^{-1}.[3] The four side chains that are observed prove invaluable in understanding protein dynamics during thermal stress.[17,25,27,29,31,35,37,46,47] These side-chain modes can provide information regarding weak interactions such as hydrophobic, ionic, and hydrogen bonding interactions. The peaks associated with key side chains serve as internal probes to provide unequivocal proof of a molecular event within the protein.

The use of a QCLM and custom slide cells has allowed for the direct 2D-COS evaluation of an array of proteins in aqueous (H_2O) solution.[48,49] This increases the number of residues that can be monitored to describe the behavior of a protein during a perturbation. Now, 10 of the 20 amino acids can be monitored, allowing a more detailed analysis and description of the behavior of the protein. These side-chain vibrational modes are: (i) aspartate $\nu(COO^-)$ at 1572 and 1567 cm^{-1} and ν(COOH) 1725 and 1714 cm^{-1}; (ii) glutamate $\nu(COO^-)$ at 1554, 1545, and 1534 cm^{-1} and ν(COOH) at 1712 and 1704 cm^{-1}; (iii) arginine $\nu_{as}(CN_3H_5^+)$ at 1675 cm^{-1} and $\nu_s(CN_3H_5^+)$ at 1630 cm^{-1}; (iv) lysine $\delta_{as}(NH_3^+)$ at 1625 cm^{-1} and $\delta_s(NH_3^+)$ at 1525 cm^{-1}; (v) asparagine ν(C═O) at 1678 cm^{-1} and $\delta(NH_2)$ at 1622 cm^{-1} and 1612 cm^{-1}; (vi) glutamine ν(C═O) at 1670 cm^{-1} and $\delta(NH_2)$ at 1590–1580 cm^{-1}; (vii) histidine ν(C═C) ring at 1600–1609 cm^{-1}; (viii) tyrosine ν(C═C) ring at 1518 cm^{-1}; (ix) phenylalanine ν(C═C) ring at 1504 cm^{-1}; and (x) tryptophan ν(C═C) ring at 1496 cm^{-1}. All peak positions are within ±3 cm^{-1}.[48,49] The ability to study proteins in a site-specific manner is now possible through the enhanced temporal and spatial resolution provided by 2D-COS.

5.2 Side Chains as Internal Probes in Proteins

The evaluation of fully hydrogen/deuterium exchanged model helical peptides during unfolding led to the understanding of the role of salt-bridge interactions in the stability of the helical motif. The peak shifts of the arginine $\nu_s(CN_3H_5^+)$ from 1582 to 1586 cm^{-1} and glutamate $\nu(COO^-)$ from 1565 to 1572 cm^{-1} indicated the disruption

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

of the salt-bridge interaction that led to the unfolding of the peptide.[29]

Side-chain modes are also sensitive to the presence of metal cofactors. As an example, calcium-binding protein belongs to a highly conserved protein class known as the EF-hand superfamily. This class of protein shares a motif that is comprised of a helix-loop-helix structure with multiple aspartate $\nu(COO^-)$ modes and at least one glutamate $\nu(COO^-)$ mode, which shift upon binding to calcium.[17,31] The binding of calcium causes structural stability within the motif, and multiple studies have investigated the extent to which calcium binding impacts the stability of the protein.[17,31,35,36,46]

6 THERMAL STABILITY OF PROTEINS (RANK ORDER)

Three main methods were employed that validated the 2D-COS analysis as an effective tool for the study of protein stability include the use of (i) comparability assessment for multiple protein samples that established similarities and differences amongst the samples, (ii) isotope editing to probe conformational changes during thermal unfolding, and (iii) step analysis, which separates the pretransition, thermal transition, and post-transition events that occurred within the full temperature range studied. The method is a valid approach that can be customized for each protein studied. Examples of all three are discussed below.

6.1 Unfolding of Proteins

The chain length dependence of helical stability during thermal perturbation led to an understanding of the mechanism of unfolding of helical segments.[27,29] The onset and thermal transition temperatures can be defined by monitoring the peptide bond ν(C═O) stretching modes within the amide I band, which are sensitive to conformational changes as a function of temperature, resulting in protein unfolding and the ability to rank order the stability of a series of proteins.[27,29,35,49] The chain length dependence of a series of synthetic model helical peptides comprised of the sequence Ac-W$(EAAAR)_n$A-NH_2 (where $n = 1$, 3, 5, and 7) were studied to rank order the stability of these peptides.[27] These peptides were observed to have greater helical content as the repeating segment increased and the transition from helical to unfolded state was monitored. These studies were extended to include several ^{13}C isotope-edited alanine residues for the $n = 5$ model helical peptide, allowing for the site-specific monitoring of $\nu(^{13}C{=}O)$ within the helical peptide. The site-specific isotope editing and 2D-COS enhanced peak-resolving power revealed that the unfolding of a peptide begins from the C-terminal end due to the differences in stability observed for each terminus during the unfolding of the peptide.[29] In the case of calcium-binding proteins, glutamates, and aspartates are sensitive to the coordination of calcium and serve to evaluate the stability of the EF-hand motif.[17,31] Detailed extensions of 2D-COS methods established differences in the stability of highly conserved proteins, such as human centrins (centrin 1, centrin 2, and centrin 3).[35] Side-chain probes have led to the possibility of rank ordering highly related proteins based on their conformational stability during thermal perturbation.[29,36,49] Side-chain modes have also served to probe weak interactions during thermal perturbation that, when disrupted, led to the unfolding of the protein,[29,30,33] hydrogen bonding,[32] and they have been an indicator of the pH environment for the protein by virtue of their pK_R.[23–25]

The list of proteins studied during thermal perturbation is extensive and has been reviewed elsewhere.[12,13] Among other things, these studies examined the differences between the unfolded CMP kinases from *Escherichia coli* and *Bacillus subtilis,*[28] established the different sequential molecular events that led to unfolding of cytochrome c from different mammalian sources,[30] and correlated the unfolding of methionine adenosyltransferase with its enzyme activity.[32] The latter protein is a dimer with an extensive hydrophobic surface in which the dimers associate to form the tetramer through polar interactions. Another protein kinase known as Cα was subjected to thermal unfolding in the presence and absence of calcium, ATP, and phospholipids, whereby less stability was conferred to the helical segments of the protein when compared to the β-sheets.[33]

Protein domain fragments have also been studied.[37] Peptides can be used for fragment-based drug design, as a pharmacologically active molecule to identify potential binding interfaces for the protein target of interest. In one example, the FF domain of human Prp40 Homolog A, a nuclear protein involved in pre-mRNA splicing, the sequence of the FF_3 domain peptide is comprised of $_{524}$KQLRKRNWEALKNILDNM$_{541}$. The FF_3 domain peptide was determined to bind with high affinity to calcium-binding proteins centrin[47]. The molecular description of the unfolding of the peptide included the disruption of a salt-bridge interaction between glutamate and arginine as a key event in the unfolding of the FF_3 domain.[37] As a second example, a peptide based on a GXXG loop motif comprised of the sequence $_{165}$KRRQRLIGPKGSTLKALELLTNCY$_{189}$ for human Krr1, which contains a KH domain, is present in a wide variety of nucleic acid-binding ribosomal proteins.[25] The KH domain comprised of the GXXG loop motif is highly conserved amongst mammalian organisms.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

The unfolding process was monitored under varying conditions.[25]

6.2 Weak Interactions and Their Effect on Stability of Proteins

Human serum albumin consists of three homologous domains comprised of 67% helical structure, 10% β-turns, and 23% random coil. The helical segments are thermally stable under varying pH conditions. One study found that pH-dependent conformational changes and hydrogen bonding interactions involving glutamates and aspartates with the backbone of the protein were responsible for the stability of the protein.[23] β-lactoglobulin serial dilutions were evaluated during thermal perturbation using ATR IR and 2D-COS, which resulted in the adsorption of the protein at high concentrations.[23]

6.3 The Effect of Posttranslational Modification of Proteins on Stability

The effect of posttranslational modifications (PTMs) on protein stability during thermal perturbation has provided insight into changes in local stability due to the chemical modification.[17,31,34] The effect phosphorylation had on the domain stability of centrin, a calcium-binding protein with two independent structural domains, was explored. Although a 1.1 °C difference in the thermal transition temperature was observed for the phosphorylated ($T_m = 111.0$ °C) and unphosphorylated ($T_m = 112.1$ °C) protein, as ascertained from differential scanning calorimetry, the instability of the EF-hand motif due to the phosphorylation modification was confirmed by 2D-COS, along with a complete description of the molecular events that led to the unfolding of the protein.[34] The effect of calcium binding on the local and overall stability of this conserved protein was also established through calcium titration of the domain fragments and the full-length protein. Differences in calcium binding affinity were correlated to differences in the molecular behavior of these domains when compared to each other and the full-length protein.[17,31] Taken together, these studies show that 2D-COS is not limited by protein type (aqueous or membrane-bound), size, structure, or PTM and is therefore an important method for the evaluation of protein unfolding.

7 CHARACTERIZATION OF PROTEIN–PROTEIN INTERACTIONS

7.1 Enabling Crystallization of Protein Complexes

The QCLM (*further discussion is provided in Section 11*) has proven useful in the evaluation of protein crystallization screens. The advantages of hyperspectral (HS) image acquisition, are that it is 200× faster than an FT-IR microscope, it provides enhanced signal-to-noise ratio spectral data, that can be used to investigate both the processes of crystallization and aggregation, and can help determine whether the crystallized protein is representative of the protein in solution.

The study of protein complex crystallization (human centrin (E32A)-Sfi1p$_{32}$ complex) in its crystallization solution without isotope editing was recently achieved. This study highlights 2D-COS's capability to distinguish the spectral peaks for each component solely based on differences in stability and dynamics during the thermal perturbation.[48] 2D-COS allowed the development of an understanding of the roles of weak interactions of the solvent-exposed side chains to enable crystallization. The approach has the potential to contribute in the evaluation of crystallization screens of protein complexes, thereby increasing the success rate of high-resolution X-ray structure determinations.

7.2 Validating Potential Drug Targets

Spectroscopically, the evaluation of a complex is challenging because both proteins will have peptide bonds. It is necessary to distinguish the carbonyl stretching vibrations within the amide I band ν(C═O) for each protein, independent of size and structural complexity. The approach has been to isotope label one of the protein components with ^{13}C using ^{13}C-enriched minimal media during the protein expression in cell cultures. The isotope causes the carbonyl stretching mode $\nu(^{13}C{=}O)$ to shift to lower wavenumbers, for example, from 1645 to 1596 cm^{-1} ($\nu(^{12}C{=}O)$ to $\nu(^{13}C{=}O)$, respectively) for a helical peptide. This shift is due to the reduced mass and can be calculated for a stretching vibration. More importantly, isotope editing effectively separates the main conformationally sensitive vibrational modes in a two-component protein mixture. The *Cr* centrin–melittin complex at a 1:1 molar ratio was used as a proof-of-concept model.[46] Centrin is a highly conserved calcium-binding protein, while melittin is a well-characterized 26-residue amphipathic peptide. Melittin is comprised of the sequence GIGAILKVLATGLPTLISWIKNKRKQ, of which the first 20 residues are hydrophobic and the C-terminal end is positively charged, providing surface area for interaction with a calcium-binding protein with EF-hand motifs. The two types of carbonyl stretching modes, $\nu(^{13}C{=}O)$ assigned to centrin and $\nu(^{12}C{=}O)$ assigned to melittin, provided the opportunity to monitor both protein components simultaneously. The requirements for peptide binding were defined to be the recognition of hydrophobic contacts for specificity, followed by ionic interactions leading to complex formation within the temperature

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

range of 5–60 °C; the centrin–melittin complex then proceeded to dissociate at temperatures of 60–95 °C. Centrin maintained an extended conformation while interacting with melittin. The centrin–melittin complex was also validated by isothermal titration calorimetry, where the interaction was determined to be exothermic with a $K_d = 7.4 \pm 0.07\,\text{nmol}\,\text{L}^{-1}$ at 25 °C, and by X-ray crystallography (PDB ID: 3QRX).[46]

The second case is the centrin–Prp40pA complex.[47] Prp40A was found to be a centrin target by yeast-two-hybrid screening assays. This pre-mRNA processing protein is a component of the spliceosome found in eukaryotic cells. Prp40 has an essential role in the initiation step of the pre-mRNA splicing process. There are two homologs, A and B, and homolog A was the subject of investigation. This protein is comprised of two WW domains followed by 6 tandem FF domains (FF_1–FF_6), which mediate Pro40-target interactions. Furthermore, human Prp40A has been implicated in genetic disorders such as Huntington's disease. Human centrin 2 is involved in mRNA and protein export from the nucleus. Also, human centrin 2 shares 70% amino acid sequence identity with *Cr* centrin. For this study, the Prp40 FF_3 domain is comprised of the sequence $_{524}$KQLRKRNWEALKNILDNM$_{541}$, which was found to contain a hydrophobic triad ($W_1L_4L_8$) as a potential binding site for centrin. Comparative analysis of the binding affinity of *Cr* centrin and human centrin 2 with the FF_3 domain of Prp40 using isothermal titration calorimetry resulted in binding isotherms that were exothermic, with $K_d = 278 \pm 31\,\text{nmol}\,\text{L}^{-1}$ and $588 \pm 69\,\text{nmol}\,\text{L}^{-1}$ for human centrin 2 and *Cr* centrin, respectively. 2D-COS was used to define the molecular mechanism of dissociation of the centrin–Prp40Ap complex (1:1 molar ratio) and unfolding during thermal stress.

The sequential order of molecular events revealed the key to complex formation was the intermolecular salt-bridge interaction between Prp40Ap's negatively charged residues, glutamate (E532) at $1541\,\text{cm}^{-1}$ and aspartate (D539) at $1567\,\text{cm}^{-1}$, and centrin's arginine ($\nu_s(^{13}\text{C}{-}\text{N})$ at $1557\,\text{cm}^{-1}$ and $\nu_{as}(^{13}\text{C}{-}\text{N})$ at $1577\,\text{cm}^{-1}$) residues, as well as the requirement of hydrophobic interaction within the binding interface. The backbone and side-chain vibrational modes were crucial to understanding complex formation and dissociation. Furthermore, the side-chain modes were used to probe the interactions and validate the thermodynamic studies performed for the different protein complexes. 2D-COS analysis defined the interactions between the proteins through the correlation of peaks from each protein component during thermal perturbation by taking advantage of isotope editing of one of the protein components.

8 MECHANISM OF AGGREGATION

There are multiple mechanisms of protein aggregation, which can be differentiated by 2D-COS due to its ability to characterize detailed molecular events during a perturbation.[25,43–47] These aggregation mechanisms are diverse and can lead to different manifestations of the protein's physical and biochemical state. The list includes, but is not limited to self-association[25] dominated by weak interactions (discussed below for the National Institute of Standards and Technologies (NIST) mAb in Section 12), the precipitation of the protein leading to insoluble aggregates,[48] the generation of amyloid fibrils,[44] and soluble aggregates that can be due to partial or complete unfolding of the protein that impacts its function. Soluble aggregates have been induced by deamidation at high pH and under thermal stress for clinical mAbs (*data not shown*). Therapeutic protein aggregation can lead to unwanted immune responses and has been reviewed elsewhere.[50–52]

8.1 Determining the Type of Aggregate

2D-COS studies have described the mechanisms of aggregation for a series of proteins, including myoglobin,[43] cytochrome c,[42] human Krr1 KH domain,[25] the activation domain of human procarboxypeptidase A2,[44] methionine adenosyltransferase,[32] insulin,[26] and melittin within the centrin–melittin complex.[25,26,32,42–45]

For myoglobin, the thermal unfolding of the protein was studied using step analysis whereby the entire thermal perturbation range was separated into smaller temperature segments (step analysis) to investigate molecular events resulting in a clear definition of the conformational state of myoglobin protein during each phase of the transition. This analysis provided unequivocal evidence as to the temperature where the aggregation occurred.[43] Cytochrome c is a membrane-associated protein and was studied in the presence of dimyristoylphosphatidylglycerol (DMPG) in a (50:1) lipid:protein, molar ratio. The reconstituted sample was subject to thermal perturbations from 10 to 65 °C, followed by the monitoring of the reconstituted sample during a 120-min incubation period at 65 °C. A two-step aggregation process was defined using 2D-COS, which began with self-association (reversible) leading to thermal unfolding and aggregation (irreversible).[42]

The KH domain of human Krr1, referred to as the GXXGlp synthetic peptide ($_{165}$KRRQRLIGPKGSTLK ALELLTHCW$_{189}$), was studied in the presence and absence of trifluoroacetic acid (TFA), a by-product of the of solid-phase synthesis. For the purpose of the aggregation evaluation, GXXGlp samples were at two different concentrations, 50 and $16\,\text{mg}\,\text{mL}^{-1}$. At the

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

higher concentration, self-association, distinct from aggregation, was observed under thermal perturbation (5–90 °C). Self-association was determined to be due to an intermolecular salt-bridge interaction between arginines (R165 and/or R166) and glutamate (E182) located near the C-terminal end of the peptide.[25]

The formation of amyloid fibrils was observed for procarboxypeptidase A2 (ADA_2h). The loss of structure with increasing temperature followed by the dramatic appearance of the aggregation peak at 1619 cm^{-1} was observed for the wild-type as well as three variants of the protein. The only differentiator in the amyloid fibril formation for ADA_2h samples was the kinetics of the process.[44] The evaluation of the methionine adenosyltransferase enzyme under thermal perturbation led to the determination of protein unfolding leading to aggregation.[32] The monitoring of insulin for the formation of fibrils at 74 °C required a period of 18 hours.[26] Finally, the selective aggregation of melittin in the centrin–melittin complex in the presence of TFA demonstrated the role of an excipient,[45] a by-product of the solid phase synthesis of melittin. In this case, melittin was observed to have increased helicity upon binding to centrin, and as the thermal perturbation progressed, the selective aggregation of melittin was observed within the ternary complex (further discussion in Section 8.3 below).

A QCLM was used for the evaluation of the developability of a mAb fragment, the IgG CH_2 domain. The IgG1 CH_2 domain has the potential to be the basis of a new class of protein therapeutics, because of its bifunctional role as an effector and biomarker protein. Three similar mAb fragments (C_H2, C_H2s, and m01s) comprised of residues Pro238 to Lys340 of the IgG1 heavy chain (HC) sequence were studied during thermal perturbation within the temperature range of 28–60 °C. Regions within the protein that were prone to self-association were identified for C_H2 and C_H2s. The sequential molecular order of events allowed for a detailed description of the mechanism of self-association and unfolding for C_H2, C_H2s, and m01s. Furthermore, the engineered disulfide bond in m01s created a stable protein free of aggregation and consequently the most stable mAb fragment candidate.[49]

8.2 Determining the Mechanism and Region Prone to Aggregation

2D-COS provides an in-depth analysis of the spectral data of proteins during a perturbation, offering an opportunity to understand the mechanism by which the protein of interest aggregated, but more importantly, to identify the region within the protein prone to aggregation. This information enhances our understanding of the risk factors associated with aggregation and provides valuable information for the mitigation of the risk of aggregation. Several case studies will be presented to substantiate this approach.

The first example is heat-induced aggregation of glutamyl-tRNA synthetase, both wild-type and a variant that differs by only one amino acid. The FT-IR study for both proteins was limited to the amide I band within the spectral region of 1700–1600 cm^{-1}.[38] The mechanism of aggregation determined by 2D-COS involved the helical segments of the protein, thus effectively identifying the region prone to aggregation within glutamyl-tRNA synthetase.

Cytochrome c, a peripheral protein involved in electron transport in the mitochondrial inner membrane, was investigated to understand its aggregation mechanism in the presence of DMPG during thermal perturbation.[42] The group identified for the first time that the cytochrome c protein underwent partial unfolding that led to aggregation. The helical regions were linked to aggregation via the formation of fibrils upon incubation at 40 °C. The aggregation process was monitored to track the conformational changes of the protein. The monitoring of conformational changes within myoglobin, a well-known heme protein, using the amide I band led to a description of steps defined as pretransition (30–58 °C), during transition (60–68 °C), and posttransition (70–80 °C). This study provided a complete molecular description of the phases of the aggregation process during thermal stress of the protein and gave an unequivocal assessment of the mechanism of aggregation involving multiple destabilized helical segments and β-sheets.[43]

The study of amyloid formation has remained a key interest for protein spectroscopists, resulting in the study of a well-characterized model for protein folding and misfolding, known as procarboxypeptidase A_2 (ADA_2h). The wild-type and 3 variants were studied to understand the kinetics of fibril formation. The nucleation of the protein followed by conformational reorganization rendered the amyloid fibril. The different kinetics of the processes were dependent on the initial stabilities of each variant protein when compared to the wild-type.[44]

8.3 The Extent of Aggregation

An understanding of the mechanism of aggregation is a prerequisite to assessing the extent of aggregation. Knowledge of the region prone to aggregation within the protein is also necessary. The secondary structural segment that becomes unstable leading to the protein's structural collapse is the fundamental basis for the determination of the extent of aggregation. One example is the case of the centrin–melittin–TFA complex, where

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

melittin underwent selective aggregation. The extent of aggregation was determined to be 12.0%.[45] The assessment of the molar absorptivity of the aggregate is difficult, but the cumulative experience describing the phenomenon of protein aggregation has led to a method for the determination of the extent of aggregation that is applicable to any protein and again requires a 2D-COS analysis.[50,51]

Given the wide variety of 2D-COS studies describing processes related to protein drugs, this work could focus on many areas. However, we chose to focus on protein deamidation because, like oxidation and glycosylation, it is one of the critical quality attributes (CQAs) evaluated during therapeutic protein development. By focusing on deamidation, we will illustrate and validate the application of 2D-COS as a site-specific method to study therapeutic proteins and formulations.

9 PROTEIN DEAMIDATION

9.1 The Mechanism of Deamidation

Deamidation is a nonenzymatic two-step chemical reaction involving an asparagine or glutamine residue that is subject to a nucleophilic attack by the amide of the $N+1$ peptide bond on the carbonyl of the side chain to form a cyclic intermediate. The cyclic intermediate undergoes hydrolysis, causing ring opening and the generation of the corresponding carboxylate (aspartate or glutamate, respectively), effectively introducing a negative charge in the amino acid sequence of the protein. The deamidation event has been observed to occur when the protein is subject to acidic or basic pH and/or thermal stress. Furthermore, deamidation has been found to occur during production, processing, and storage of proteins.[52] The introduction of a negative charge can affect the protein's biological activity, half-life, and immunogenicity.[53,54]

These significant modifications of a protein's properties must be discovered, tracked, and evaluated to improve drug quality, safety, and efficacy. Protein developers have categorized deamidation as a product of a degradation pathway that leads to drug substance heterogeneity.[55,56] The US Food and Drug Administration requires a comprehensive characterization of PTMs, such as deamidation, and the structural and compositional heterogeneities they cause in order for a protein drug substance or biosimilar to become a new biotherapeutic protein.[52]

The mechanisms of these deamidation processes are described in Figure 1 for asparagine and Figure 2 for glutamine.[57,58] Yellow highlights define the peptide

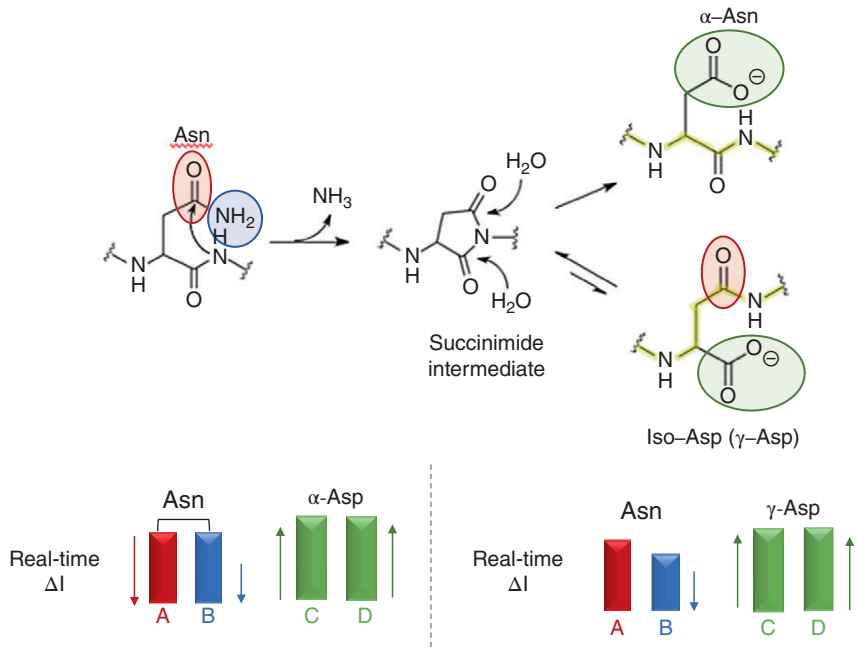

Figure 1 Schematic representation of the mechanism for asparagine deamidation. Top panel: The backbone is highlighted in yellow. Bottom panel: Bar graphs correspond to intensity changes for different functional groups (vibrational modes) that are monitored in real-time during the reaction: (A) Asn and iso-Asn ν(C═O) at 1678 cm^{-1} (red), (B) Asn $\delta(NH_2)$ at 1612 cm^{-1} (blue), (C) Asp and (D) Iso-Asp $\nu(COO^-)$ at 1576 or 1567 cm^{-1} (green). The arrows define the direction and the proportionate change in intensity. Source: Based on Riggs et al.[58]

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

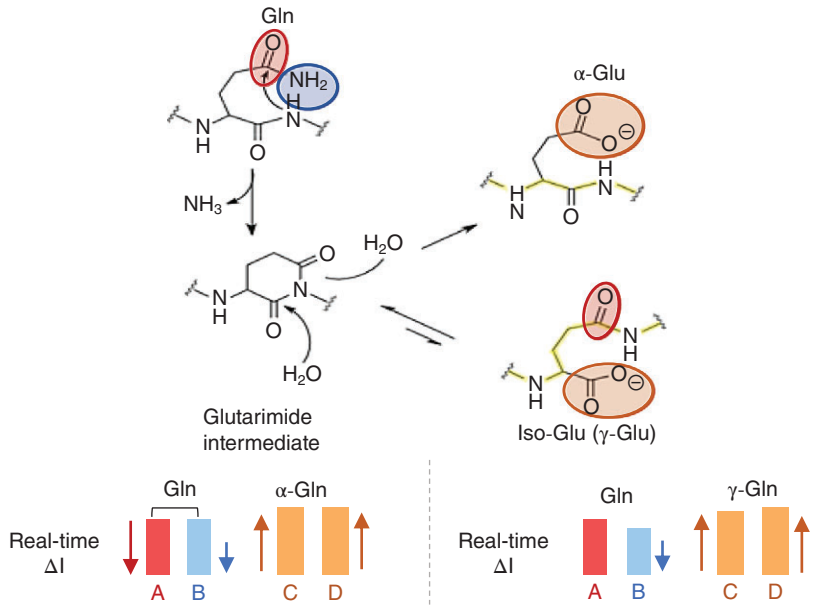

Figure 2 Schematic representation of the mechanism for glutamine deamidation. Top panel: The backbone is highlighted yellow. Bottom panel: Bar graphs correspond to intensity changes monitored in real-time: (A) Gln and iso-Gln ν(C═O) at 1670 cm^{-1} (red), (B) Gln $\delta(NH_2)$ at 1581 cm^{-1} (blue), (C) Glu and (D) Iso-Glu $\nu(COO^-)$ at 1554 or 1545 cm^{-1} (orange). Source: Based on Pace et al.[57]

backbones. Temperature, pH, the identities of the residues near the reacting amino acid, and the higher-order structure near the reaction site influence the deamidation reaction rate. In addition to the rate, the pH of the solution influences which specific reaction mechanism takes place. At higher pHs, deamidation of asparagine is base-catalyzed. The reaction involves the nucleophilic attack of the $N+1$ nitrogen of the protein backbone on the carbonyl carbon of the asparagine side chain (Figure 1). The bond formation produces a cyclic five-membered succinimide intermediate and releases ammonia. Spontaneous hydrolysis occurs at either the α or β carbonyl groups to form isoaspartic acid or aspartic acid in a ratio of approximately 3:1.[59] In contrast, at lower pH near 5.5 or 6.0, an Asu-mediated (2-aminosuberic acid) isomerization pathway (*not shown*) is accelerated.[60]

Glutamine deamidation occurs by direct hydrolysis to glutamate under acidic conditions.[60,61] At neutral or alkaline conditions, glutamine deamidates via a six-membered glutarimide ring to glutamic acid or α-glutamic acid (Figure 2), similar to the five-membered ring mechanism of the asparagine deamidation.[57,60,61] While deamidation of glutamine is often slower than asparagine, rates are dependent on many factors beyond residue identity.[58,62,63]

The rates of both types of deamidation are dependent on buffer conditions, pH, the amino acids before and after the glutamine/asparagine residue, and additional higher-order structural factors. Small charged amino acids on either side of the asparagine or glutamine residue can accelerate deamidation.[62] In addition, despite different rates, both types of deamidation are often observed together. For example, one previous report described 10 asparagine and 7 glutamine deamidations. Further, deamidation rates are difficult to predict.[64] For example, despite reduced and denatured conditions, only 4 of 10 asparagines and 1 of 7 glutamines were significantly deamidated in Liu.[62]

The deamidation of asparagine or glutamine residues increases the protein molecular weight by ≈1 Da and adds one negative charge to the protein by generating an acidic species.[60] The formation of isoaspartate or aspartate isomerization introduces a carbon into the protein backbone (Figure 1).[58] These changes introduce heterogeneity into the amino acid composition and structure of a mAb.[64,65] These changes can occur at any time between molecular synthesis until the antibody leaves the body. While a complete understanding of all types of PTMs and the heterogeneities they cause in a protein drug to allow complete drug formulation, this article focuses on deamidation as one of the most prevalent types of PTMs.[66]

9.2 Current Approach to the Evaluation of Deamidation

The NISTmAb Reference Material (RM) 8671 is intended to be an industry-standard representative mAb for precompetitive harmonization of best practices and designing next-generation characterization

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

technologies for identity, quality, and stability testing.[67] RM 8671 represents a typical therapeutic protein and is available in a stable format. A stratified sampling and analysis plan using a series of qualified analytical and biophysical methods is described that assures RM 8671 meets industry standard criteria.[67] RM 8671 was verified to be homogeneous both within and between vialing lots, demonstrating the robustness of the life cycle management plan.[67] RM 8671 can be compared with the in-house primary sample 8670 (PS 8670) to produce lot-to-lot variability studies. The precompetitive nature of NISTmAb RM 8671 makes it a practical, open-innovation platform to demonstrate novel analytical and biophysical technologies.[67] RM 8671 offers an unprecedented opportunity for developing tools to answer foundational questions such as determining the presence and site of protein deamidation. RM 8671 provides a common platform for integrating, modeling, and data visualization of complex, multi-approach data to guide the next-generation biopharmaceutical analytical tools to determine the developability of a protein candidate. The presence of a widely available industry-quality material, analyzed by countless permutations of every new and available analytical tool, will produce the most comprehensive and dynamic single protein characterization to date.[67]

While recombinant DNA technology can allow the production of well-defined proteins with a high level of consistency, the testing of the product is required to ensure the proper starting material is selected.[66] Unlike traditional chemical synthesis, recombinant DNA technology occurs in living cells, making modifications of the desired characteristics (modalities) of the therapeutic protein possible. Newly developed proteins are subjected to stress conditions during the manufacturing process that can lead to alterations in the desired protein product. The structural and compositional complexity requires new analytical methods to identify and characterize these alterations. Generally, the first analytical methods employed are chromatography and electrophoresis, which allow the separation and quantitation of the major and minor isoforms.[66]

While necessary, knowledge of the fraction sizes and biochemical characteristics described by chromatography and electrophoresis of the intact molecule or its subunits are insufficient. Information that describes the specific site of deamidation is desirable. Knowledge of the specific location of deamidation allows the importance and role of each deamidation to be determined. For example, deamidations that occur in the complementarity-determining regions (CDRs) of a mAb are possibly more important than those that occur in the Fc region because they are more likely to influence antigen binding.[66]

LC-MS-based techniques, described in Guideline Q6B by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, may be used to establish the identity of the protein drug substance through the determination of the amino acid sequence.[52] Once the amino acid sequence is confirmed, deamidation sites can be localized to specific residues, but they are still challenging to characterize.

While there are limits to finding deamidation sites using top-down MS on the intact protein, it is usually possible to do so with peptide-level analyses.[68] These sophisticated MS techniques have many strengths that make these analyses possible and some weaknesses that make them difficult. For example, absolute differentiation of isomeric amino acids like isoleucine and leucine is impossible. In addition, even in complete protein analysis, near-isomeric compositions cannot be unequivocally differentiated (e.g. lysine vs. glutamine, leucine/isoleucine vs. asparagine, or aspartic acid). Further, unequivocal amino acid assignments get more complicated with additional fragment size; for example, amino acid pairs EG, DA, VS, and single residue W all have a nominal mass of 186 Da.[52] Finally, deamidations at low levels may not be observable regardless of instrumental resolving power.[52]

Bottom-up MS approaches are commonly used to localize PTMs to a particular amino acid site.[69] Multiple sample preparative steps are required that result in the generation of a plethora of peptides that are expected to be representative of the full-length protein. Such methods involve the reduction and alkylation of the protein, then enzymatic digestion prior to analysis. Trypsin is most often used for digestion because it, cleaves proteins on the carboxyl side of arginine and lysine (except when followed by proline).[67] Since most proteins contain numerous arginine and lysine residues, the use of trypsin typically generates an abundance of peptides ideally sized for MS analysis. In addition, the basic amino acid at the C-terminus of each peptide allows them to ionize readily in the MS source. After tryptic peptides are analyzed by LC-MS, software algorithms are used to identify each peptide based on the observed mass to charge ratios (m/z) and their correlation to the theoretical m/z values calculated from the known protein sequence. Tandem MS (MS/MS) methods are typically used to fragment, and essentially sequence, the peptides, which provides not only identification of the peptide but also localization of PTMs to a specific residue.[69]

MS signal detected over time as tryptic peptides elute from the chromatography column and pass through the detector can be visualized as a chromatographic trace, or 'peptide map' can be used for monitoring deamidation and other PTMs.[52,67,70] The method requires the comparison to a reference peptide map to survey for changes in peak retention time or peak height, or the appearance of a

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

new peak, that would be indicative of sample degradation (i.e. new or increased levels of PTMs).

Peptide mapping does not always result in full sequence coverage of the protein, for example typical sequence coverage for the HC of the NISTmAb has been reported at 96.89%, thus a modification may be missed if it resides on an undetected site.[67,52] It can be challenging to detect a deamidated peptide and localize its deamidation site if it is not chromatographically resolved from its unmodified counterpart because their overlapping isotope envelopes make *m/z* resolution difficult.[52] Finally, incomplete fragmentation of a deamidated peptide during MS/MS may not provide confident localization of the deamidation site if multiple candidates are present in the same peptide.

9.3 Implications of Deamidation for Therapeutic Protein Development

Deamidation can occur in the cells where the protein is produced, during purification, during storage, and under different stress conditions. Buffer composition, ionic strength, pH, and higher-order structure can all affect the deamidation rate.[60] Structurally, the formation of isoaspartate or aspartate from asparagine and subsequent isomerization introduces a carbon atom into the protein backbone, which is more significant than the addition of the single negative charge.[52] Deamidation is also involved in creating charge heterogeneity (gain of negative charge) of proteins and mAbs.[68,52] Global mapping of charge isomers can determine the effects of modifications such as deamidation on charge isomers.[71,72] At the protein level, deamidation is second only to N-glycoforms containing sialic acid in terms of the effect on charge isomers. At the peptide level, deamidation is more critical.[73]

Asparagine deamidation is an almost universal modification of mAbs.[52] Asparagine residues in the CDR are inherently susceptible to deamidation because of their relatively higher flexibility and greater exposure to solvents than at other locations. In addition to deamidation in the CDR, deamidation also occurs in susceptible asparagine residues in the constant region.[74,76] For example, the most widely observed deamidation sites are in the amino acid sequence SNGQPENNY, located in the Fc region. Variants containing deamidation products are less hydrophobic than those with the original asparagine residues. When measured by differential scanning calorimetry, the antigen-binding (Fab) fragment with the deamidation product, isoaspartate, is less stable than Fab with the original asparagine residue.[74]

Deamidation does not always impact in vivo clearance.[75,78] However, mAbs have long half-lives in the body, so there is more time for heterogeneity to both happen and to have an effect. Natural human IgG has 23% deamidation at the conserved site in the Fc region, which is consistent with the molecule's in vivo half-life.[77,78] The presence of such high levels of deamidation in natural human IgG suggests that deamidation, at least in the conserved fragment, is not foreign to the immune system and, therefore, would not present an increased risk of immunogenicity.[78] On the other hand, a case is known where the in vivo deamidation of a HC CDR-based asparagine significantly reduced the drug's efficacy.[76–78]

Shelf life can be affected by many other factors, deamidation is the primary degradation pathway in mAbs, and deamidation is mainly affected by hydroxide ion concentration.[57] Knowledge of deamidation under different forced degradation conditions is commonly employed to accelerate testing and provide predictive results for industry formulation screens. It is crucial to optimize tryptic digestion procedures and distinguish procedure-induced artifacts from actual alterations for the accurate determination of deamidation levels.[72] For example, acidic and basic buffers exhibited different trends in mAb deamidation rates at 40 and 5 °C, suggesting that a formulation that renders the protein more stable at 40 °C may result in reduced stability at 5 °C.[57] The presence of two trends suggests a second deamidation site tentatively assigned to Asn323 in the VSNK motif. Combining the two trends suggests that the deamidation rate is lowest at pH 6.3.[57] A change in the mechanism of deamidation related to changes in higher-order structural rigidity rather than a new site is unlikely because acid-catalyzed mechanisms usually require close to pH 3.

10 WHERE AND WHEN DOES DEAMIDATION HAPPEN?

Deamidation modifications occur during production, processing, and storage.[52] Most modifications are low abundance minor species. Therefore, modifications must be discovered, tracked, and evaluated for quality, safety, and efficacy. Historically, the analytical methods of chromatography and electrophoresis allow the separation and quantitation of the major and minor isoforms. While necessary, knowledge of the fraction, sizes, and the sequence related to the deamidation site is insufficient. Therefore, many MS techniques are used to localize the modifications' positions in the protein backbone. For example, LC-MS peptide mapping produces a unique chromatographic peptide trace for a particular protein.[67,79] Then, the mass spectra of each peptide are used to determine the specific residue undergoing deamidation.

There are many reports of asparagine deamidation but fewer reports of glutamine deamidation. There are several

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

Table 1 (a) Comparison of relative abundance (%) deamidation for different sites identified using different digestion methods for NISTmAb (RM 8671) may belong to the LC or HC. (adapted). (b) Comparison of relative abundance (%) deamidation for different sites identified using different digestion methods for NISTmAb (RM 8671) heavy chain.[51]

Deamidation site	Alternative rapid digestion ($n = 6$)	SMART digest 45 min ($n = 6$)	Magnetic SMART digest 45 min ($n = 6$)	In-solution, heating ($n = 6$)
(a)				
Q1	0.041	—	—	0.041
~Q36	**0.690**	**1.553**	**1.440**	**—**
N78	0.116	0.584	0.690	—
Q79	0.380	0.845	1.059	0.477
Q146	0.284	0.724	0.603	0.310
~Q154	**0.517**	**0.244**	**0.664**	**0.464**
~N157	**—**	**1.478**	**0.977**	**0.840**
Q198	0.367	0.786	0.784	0.544
N209	—	0.195	0.241	0.555
(b)				
N289	0.180	1.056	0.869	0.591
N300	—	0.230	0.106	—
N318	**—**	**5.502**	**3.206**	**11.942**
N328	0.650	0.294	—	2.253
N364	0.040	1.849	0.440	0.288
Q365	0.304	0.853	1.694	0.324
~Q422	**—**	**0.625**	**0.353**	**—**

~The modification was found on the tryptic peptide but could not be localized to a specific amino acid using MS/MS.
Bold font is used to highlight the deamidation sites that have been independently validated by the use of the combination of a Quantum Cascade Laser Microscope and the Correlation Dynamics software.[51]

of both types of deamidation-prone sites reported in the NISTmab RM 8671, as shown in Table 1. These sites come from deamidations of both spontaneous and forced degradation origins. These analyses of NISTmAb RM 8671 produced excellent quality data with high confidence in the results. Sequence coverage was high, and sample preparation-induced PTMs, while observed, were low, except for deamidation of N318 and N387, which are potentially more susceptible to deamidation.[79–81]

10.1 Prediction of Deamidation Sites

Over the last ten years, substantial efforts have determined many factors responsible for deamidation.[60,62–64,85] Generally, these studies concentrated on the structural properties of CDRs, which correlated with deamidation observed by MS. The deamidation hot spots generally have conformational flexibility and certain C-terminal flanking residues within the sequence. These efforts also strived to determine the roles of primary sequence, solvent dielectric constant, temperature and pH.

One of these studies focused on a total of about 20 structural properties.[64] The authors built an *in silico* prediction method for CDR degradation by asparagine deamidation. Data from 37 different mAbs and 60 residue changes identified 21 hot spots (i.e. residues found to have >3% modification after stress treatment), 14 weak spots (1–3% modification), and 24 reactive spots (<1% modification). All of the hot spots were in CDR loops. Asparagine residues flanked by glycine, serine, threonine, asparagine, or histidine were the most likely to deamidate, with 67% of the NG pairs, followed by 21% of NS pairs, 14% of NN pairs, and 11% at NT pairs being identified as hotspots. The most degradation-prone site, NG, was found to be comparatively infrequent in the ImMunoGeneTics database of human mAb CDRs.[85]

Another study applied machine learning to an LC-MS/MS dataset comprised of 776 peptides from mAbs using random decision forest models to predict the likelihood and rates of deamidation.[86,87] The resulting model predicted the overall observed deamidation rate in the training set an R^2 of 0.96 and is intended to be used for improving the developability of drug candidates.

LC-MS was employed to aid in the development of these models to optimize the chemical stability of mAbs at an early stage of development. In this study, a combination of low pH, high pH, and heat stresses were applied to 131 mAbs designed with variable regions representative of the clinical-stage mAbs. The study was limited to the tryptic peptide mapping of the CDRs exclusively. The authors observed high frequency of asparagine deamidations in CDRs H2 and L1. Multiple asparagine sequence motifs were observed to deamidate. NS, NN and NT had > 1% modification as compared to 5% modification for NG based on the total motif occurrence within the CDRs.[86]

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

The extensive experimental study of the deamidation of the asparagine and glutamine residues in these 131 mAbs spawned a modeling study to understand and predict a protein sequence's propensity for such PTMs and ultimately help expedite therapeutic protein discovery and development.[59] To that end, proton-affinity calculations with semi-empirical quantum mechanics and μs-long equilibrium molecular dynamics simulations were done to investigate mechanistic roles of the structural and chemical environment in dictating spontaneous degradation of asparagine residues in 131 clinical-stage-like therapeutic antibodies. Backbone secondary structure, side-chain rotamer conformation, and solvent accessibility were shown to be three key molecular predictors of asparagine deamidation.[59] In addition, comparative analysis of backbone dihedral angles and N–H proton affinity calculations provided a mechanistic explanation for the strong influence of the identity of the $N+1$ residue on the rate of asparagine deamidation. The authors leveraged their calculations into a minimalistic physics-based classification model to predict the deamidation propensity of therapeutic proteins.

10.2 Developability and Deamidation Hot Spots

The importance of developability is increasing. The concept of developability relies on the knowledge gained from the successful development of approximately 80 marketed antibodies and Fc-fusion protein drug products and the lessons learned from many failed development programs over the last three decades. The goal of developability is to utilize thorough evaluations of mAb lead candidates at early stages to bring this body of knowledge into play to avoid delays during late-stage development.[82–88]

Developability relies on the measurement of attributes critical to developing protein drug products such as the sites and extents of protein deamidation.[87,88] Traditional and state-of-the-art analytical methods to monitor those attributes are not enough. A practical early developability assessment will require new analytical methods to characterize protein deamidation. Here we describe our efforts to apply IR spectroscopy-based methods to improve the early-stage detection of protein deamidation.

11 METHODS

11.1 HS Image Acquisition

A QCLM allows for the rapid collection of a series of HS images at different sets of temperatures for a series of protein samples in a slide array. The advantage is the real-time comparability assessment of the protein samples within the spectral range of 1700–1500 cm^{-1}, which includes both the amide I and II bands, evaluated under the exact same temperature conditions in H_2O. The QCLM is 200× faster than current FT-IR microscopes. The QCLs were developed by the Capasso[89] group using a series of thin semiconductor layers of varying material composition that, upon the application of an electric field, inject electrons into the upper state of a quantum well system. The relaxation of the radiated electron generates mid-IR photons. A single electron can undergo multiple radiative transitions, generating the spectral region of interest. The thickness of the semiconductor layers is used to define the emission wavelength, allowing for a broad spectral range of about 100 wavenumbers. This type of laser is known as an external cavity QCL and has been designed for the QCLM. The tuning of these lasers using a diffraction grating relative to the QCL chip, along with the combination of up to four external cavity QCLs, provides an increased spectral range of 1000 wavenumbers within the mid-IR range (MIRcat-QT from Daylight Solutions).

To evaluate stability, NISTmAb Primary Sample (PS 8670) was subject to a dilution using 25 mM L-histidine at pH 6.0. The resulting concentrations were as high as 100–40 $\mu g\,\mu L^{-1}$ and as low as 10–2.5 $\mu g\,\mu L^{-1}$. One μL of each sample was applied to a predefined well in a custom milled CaF_2 slide array with fixed pathlength from Protein Dynamic Solutions, Inc. (Wakefield, MA). The slide cell was then placed in a heated accessory with (±0.03 °C) thermal control. HS images were acquired within the temperature range of 28–60 °C with 8 °C intervals and 4 min equilibration periods using the ProteinMentor™, a QCL IR transmission microscope from Protein Dynamic Solutions, Inc. (Wakefield, MA). The QCLM is comprised of four broadly tunable external cavity QCLs, a low magnification objective (4×) with a numerical aperture of 0.3, and a $2\times2\,mm^2$ field of view providing a pixel size of $4.25\times4.25\,\mu m^2$ and a linear response microbolometer focal plane array (480×480 pixels) detector. The HS images are composed of 223 000 QCLM spectra collected at 4 cm^{-1} spectral resolution within the spectral region of 1775–1500 cm^{-1}. The background was collected at each temperature once thermal equilibrium (4 min) was achieved. The multiple HS image acquisitions for an array of mAbs in solution at different set temperatures were possible due to the real-time HS image acquisition, the design of the slide cell array, and the accurate thermally controlled accessory. Furthermore, a complete HS image is comprised of 223 000 QCL IRM spectra, requiring just 40 sec per sample.

11.2 HS Image and QCLM Spectral Analysis

The HS images were analyzed using DataPDS™ software from Protein Dynamic Solutions, Inc. (Wakefield, MA).

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

QCLM spectral baseline correction, difference spectra calculations, overlays, 2D-CDS, and 2D-COS plots were generated using Correlation Dynamics™ software from Protein Dynamic Solutions, Inc. (Wakefield, MA). Finally, the thermal dependence difference spectra and slice plots were generated using Excel in Microsoft Office Professional 2016 (Microsoft; Redmond, WA).

12 RESULTS

12.1 HS Images and QCLM Spectra

A series of HS images were acquired at every set temperature for every NISTmAb sample concentration in the slide cell array. Typical HS images are shown in Figure 3. These HS images are featureless and aggregate-free at the spatial resolution of 4.3 μm. Each HS image is comprised of 223 000 QCLM IR spectra. The spectral overlay for the different NISTmAb concentrations within the spectral region of 1775–1500 cm^{-1} is shown in Figure 4. The relationship between protein concentration and intensity is evident. In addition, the reproducibility and the high signal-to-noise ratio of the spectra allow the generation of high-quality difference spectra shown in Figure 5 which allows for the subsequent 2D-COS and 2D-CDS analysis.

The NISTmAb samples examined were in an aqueous buffered solution. The HOH bending vibration of water must be considered due to its high absorptivity of 55.5 M^{-1} cm^{-1} at 1643 cm^{-1}, which overlaps with the protein amide I band within the spectral region of 1700–1600 cm^{-1}.[14] Like any aqueous protein sample, the protein effectively dilutes the contribution of the HOH bending mode within the spectral region of 1700–1600 cm^{-1}.[9,12,13] By adequately matching the cell pathlength to the brightness of the QCL, samples with high absorptivity's can be measured using the spectral difference approach. Difference spectra were generated by subtracting the initial spectra at low temperature from all subsequent spectra to effectively subtract the bulk H_2O bending vibrational band that overlaps the amide I band. This treatment of the spectral data also subtracts all protein vibrational modes that did not change during the thermal perturbation. We overcome the common deterrents of using IR spectroscopy for aqueous samples in these ways.

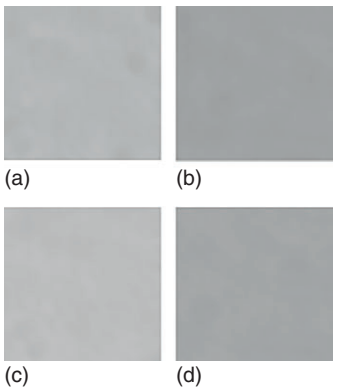

Figure 3 Representative HS images of NISTmAb in 25 mM L-histidine at pH 6.0. Images were acquired at a spatial resolution of 4.3 μm at the initial and final set temperature: (a, c) 28 °C and (b, d) 60 °C for the high and low concentrations of (a, b) 100 and (c, d) 2.5 μg $μL^{-1}$ mAb.

QCLM IR peak assignments were based on the 2D-COS asynchronous correlations, which provide the highest peak-resolving power. Comparisons to the synchronous and baseline-corrected spectral overlay for each data set verified those assignments. We split the peaks into two categories of vibrational modes: (i) peptide bonds ν(C═O) sensitive to secondary structure changes and located within the amide I band (1700–1600 cm^{-1}); and (ii) peaks associated with weak interactions. The weak interactions include 10 of the 20 amino acid side-chain modes located within the entire spectral region studied (1780–1500 cm^{-1}). Table 2 summarizes these peak assignments.

All peak assignments have been validated by comparison to the available high-resolution structure (PDB ID:

Table 2 Summary of peak assignments for the amide I and II bands in the spectral region of 1780–1500 cm^{-1}

Peak assignment	Vibrational mode	Peak position (cm^{-1})
Asp/Glu	ν(COOH)	1725, 1717, 1712, 1704
β-turns	**Peptide bond ν(C═O)**	**1696, 1684, 1680**
Asparagine (Asn)	Side chain ν(C═O)	1678
Arginine (Arg)	$\nu_{as}(CN_3H_5^+)$	1675
Glutamine (Gln)	Side chain ν(C═O)	1670
Hinge Loop	**Peptide bond ν(C═O)**	**1663**
Helical segments (π-, α-)	**Peptide bond ν(C═O)**	**1653, 1646**
β-sheets	**Peptide bond ν(C═O)**	**1636**
Arginine (Arg)	$\nu_s(CN_3H_5^+)$	1630
Lysine (Lys)	$\delta_{as}(NH_3^+)$	1625
Aggregation/ self-association	**Peptide bond ν(C═O)**	**1616***
Asparagine (Asn)	Side chain $\delta(NH_2)$	1612
Histidine (His)	ν(C═C) ring	1606, 1600
C-terminal end	$\nu(COO^-)$	1595
Glutamine (Gln)	Side chain $\delta(NH_2)$	1581, 1590
Aspartate (Asp^-)	$\nu(COO^-)$	1572, 1567
Glutamate (Glu^-)	$\nu(COO^-)$	1554, 1545
Lysine (Lys)	$\delta_s(NH_3^+)$	1525
Tyrosine (Tyr)	ν(C═C) ring	1518, 1512
Tryptophan (Trp)	ν(C═C) ring	1500, 1496
Phenylalanine (Phe)	ν(C═C) ring	1504

Bold font is used to highlight backbone peak assignments.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

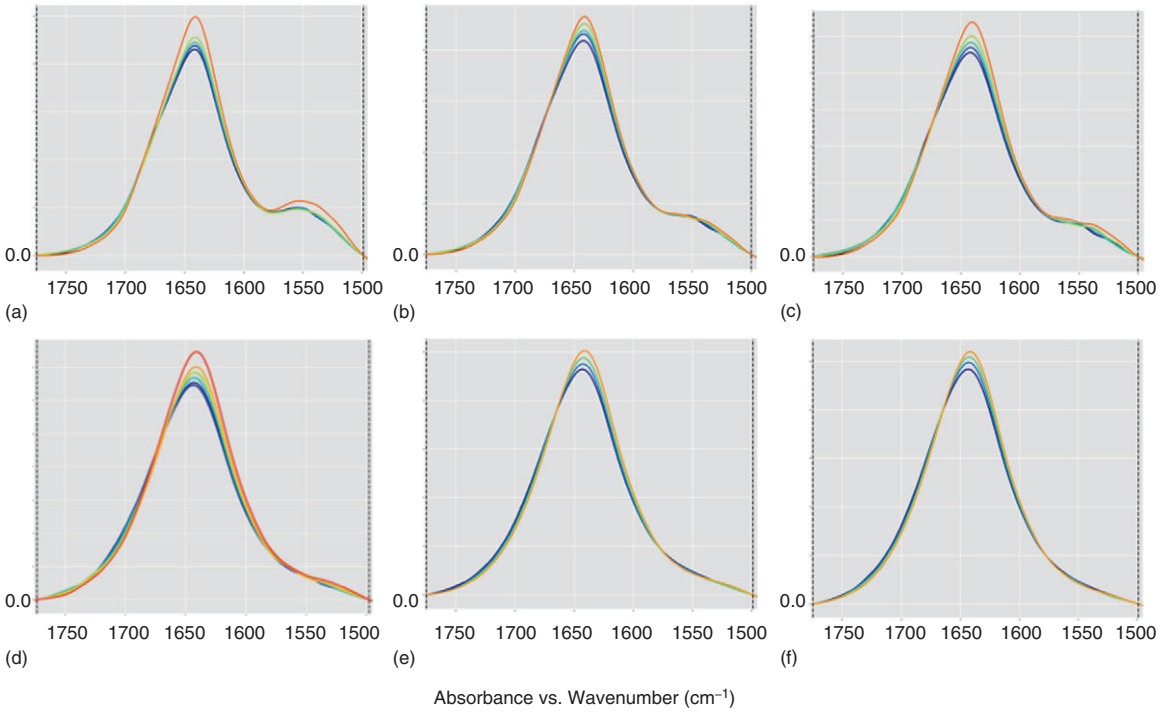

Figure 4 QCLM overlayed spectra of NISTmAb in 25 mM L-histidine at pH 6.0. QCL IRM spectra illustrate the region of 1775–1500 cm^{-1} and a temperature perturbation range of 28–60 °C for concentrations of: (a) 100.0, (b) 70.0, (c) 40.0, (d) 10.0, (e) 5.0, and (f) 2.5 $\mu g\,\mu L^{-1}$. Vertical dashed lines represent the selected baseline points (1775 and 1500 cm^{-1}) for linear baseline correction. Furthermore, the intensity differences observed for each sample correspond to the QCLM spectrum obtained for each set temperature at 8 °C intervals.

5K8A and 5VGP)[90,91] for the NISTmAb, its amino acid sequence, and prior IR assessment of the NISTmAb standard (RM 8671) and primary sample (PS 8670).[52] The side-chain mode assignment validation is based on previous studies.[5,11]

12.2 2D-CDS

Our first approach was to evaluate the population distributions of NISTmAb in solution under different concentrations to evaluate colloidal stability and the weak interactions (Figure 6) using 2D-CDS. For the 100 $\mu g\,\mu L^{-1}$ sample, the intense cross peaks associated with salt-bridge interactions were assigned to lysine residues located within β-sheets and helical regions of the mAb (1636, 1525 and 1653, 1525 cm^{-1}, respectively, Table 2) and glutamates and/or aspartates located within the same secondary structure regions. Hydrogen bonding interactions involving tyrosines were observed as intense cross peaks located within β-sheets and helical regions (1636, 1518 and 1653, 1518 cm^{-1}) and to a lesser extent, lysines and tyrosines located within the β-turns (1696, 1525 and 1696, 1518 cm^{-1}) that were disrupted during the thermal perturbation. For the *70 $\mu g\,\mu L^{-1}$ sample*, the cross peaks associated with the salt-bridge interactions are limited and the hydrogen-bonding interactions involving the tyrosines and lysines have decreased. These factors are indicated by the attenuated intensity changes observed for the corresponding cross peaks (1653, 1518 and 1636, 1525 and 1653, 1525 cm^{-1}, respectively). For the *40 $\mu g\,\mu L^{-1}$ sample,* the cross peaks (1636, 1518; 1653, 1518 and 1636, 1525 and 1653, 1525 cm^{-1}, respectively) associated with the lysine and tyrosine residues located within the β-sheets and helical regions are now involved in hydrogen bonding interactions as opposed to salt-bridging interactions between side chains when compared to the higher concentration samples. Presumably, these hydrogen bonds are with their aqueous environment. For the *10 $\mu g\,\mu L^{-1}$ sample*, slight intensity changes are observed for the cross peaks (1636, 1518; and 1636, 1525 cm^{-1}) assigned to lysine and tyrosine residues located within β-sheets (1636, 1518 and 1636, 1525 cm^{-1}, respectively). These residues appear to be exclusively involved in hydrogen bonding interactions. Meanwhile, at low NISTmAb concentrations (5 and 2.5 $\mu g\,\mu L^{-1}$), the intensity changes for the cross peaks associated with the backbone vibrational modes are observed, suggesting that at these low concentrations the mAbs are undergoing conformational changes due to the thermal stress. In summary, two different types of governing interactions were observed: antibody-antibody

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

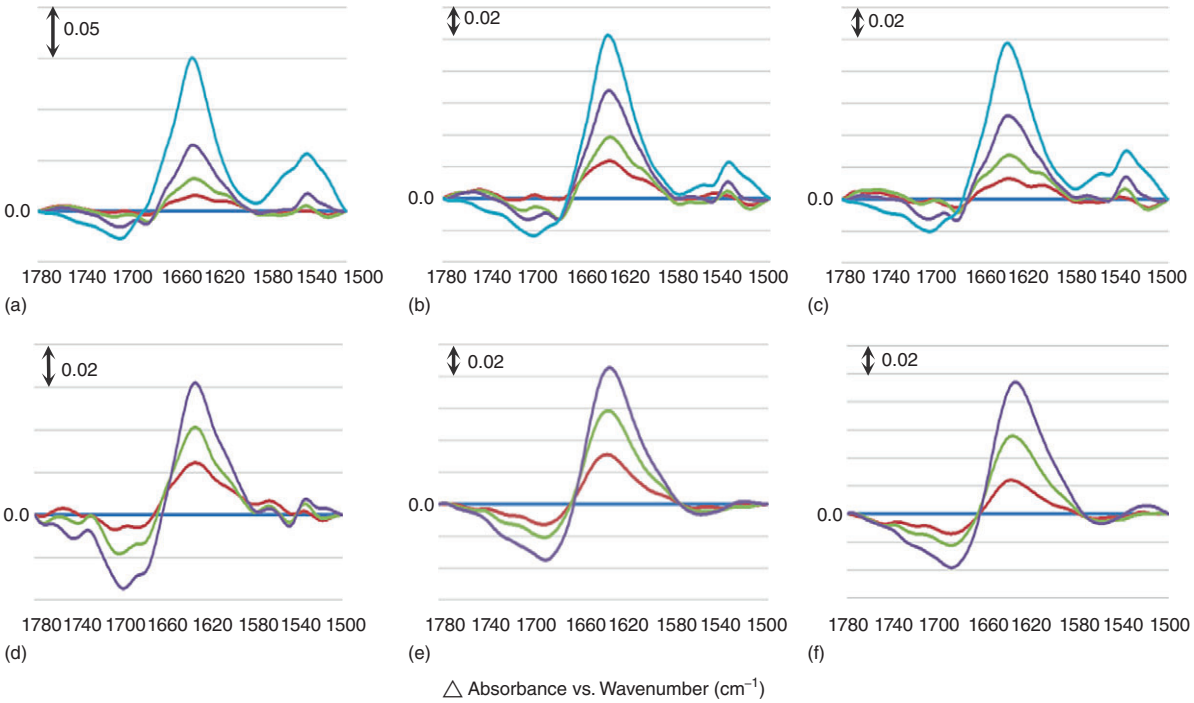

Figure 5 Overlayed QCLM Difference spectra of NISTmAb in 25 mM L-histidine at pH 6.0 for the spectral region of 1775–1500 cm^{-1} and within a temperature perturbation range of 28–60 °C and 8 °C temperature intervals, for concentrations of: (a) 100.0, (b) 70.0, (c) 40.0, (d) 10.0, (e) 5.0, and (f) 2.5 $\mu g\,\mu L^{-1}$. The intensity differences observed for each sample correspond to the QCL IRM spectrum obtained for each set temperature minus the spectrum of the same sample at 28 °C.

interactions (100 $\mu g\,\mu L^{-1}$) and weak interactions between the antibodies and their aqueous environment (40 and 10 $\mu g\,\mu L^{-1}$). In contrast, only conformational changes are observed at low concentrations (5–2.5 $\mu g\,\mu L^{-1}$).

12.3 2D-COS

A detailed dynamic description of the protein is obtained by applying 2D-COS analysis to the difference QCLM IR spectra within the spectral region of 1775–1500 cm^{-1} (Figure 7).

The synchronous plots contain similarities in the overall peak pattern at higher concentration 100–40 $\mu g\,\mu L^{-1}$ (Figure 8), while for the 5 and 2.5 $\mu g\,\mu L^{-1}$ samples two isosbestic points could be discerned (Figure 9). Upon closer examination, the 100 $\mu g\,\mu L^{-1}$ sample has a greater degree of perturbation involving certain side-chain vibrational modes (lysines (1525 cm^{-1}), tyrosines (1518 cm^{-1}), tryptophans (1504 cm^{-1}), glutamates (1534, 1554 cm^{-1}) and aspartates (1572, 1567 cm^{-1})), suggesting perturbations of weak interactions. For the lower concentrations (10–2.5 $\mu g\,\mu L^{-1}$ (Figure 9)) the β-turns were perturbed to a greater extent when compared to the higher concentrations (Figure 8).

For the asynchronous plots, differences in dynamics were observed for the 100 $\mu g\,\mu L^{-1}$ NISTmAb sample, including but not limited to glutamine and asparagine deamidation within the β-turns, that led to self-association when compared to the samples with lower concentrations. More in-depth discussion regarding the analysis of the synchronous and asynchronous plots is presented in Section 12.4 describing the sequential order of molecular events.

Slice plots shown in Figures 10 and 11 were generated specifically from the 2D-COS asynchronous analysis provide additional evidence of the robustness of the analysis providing confidence in the measurements performed. These slice plots serve to evaluate the changes in intensity with respect to each secondary structure evaluated. These plots serve as a tool to highlight the key signature peak intensity changes associated with both deamidation and if present, self-association/aggregation at high (Figure 10) and low (Figure 11) NISTmab concentrations.

12.4 Sequential Order of Molecular Events

The sequential molecular order of events allows for the description of the behavior of the protein during thermal stress, which includes the deamidation event and self-association (Figure 12). For NISTmAb at 100 $\mu g\,\mu L^{-1}$, initially, the perturbation of the β-turns

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

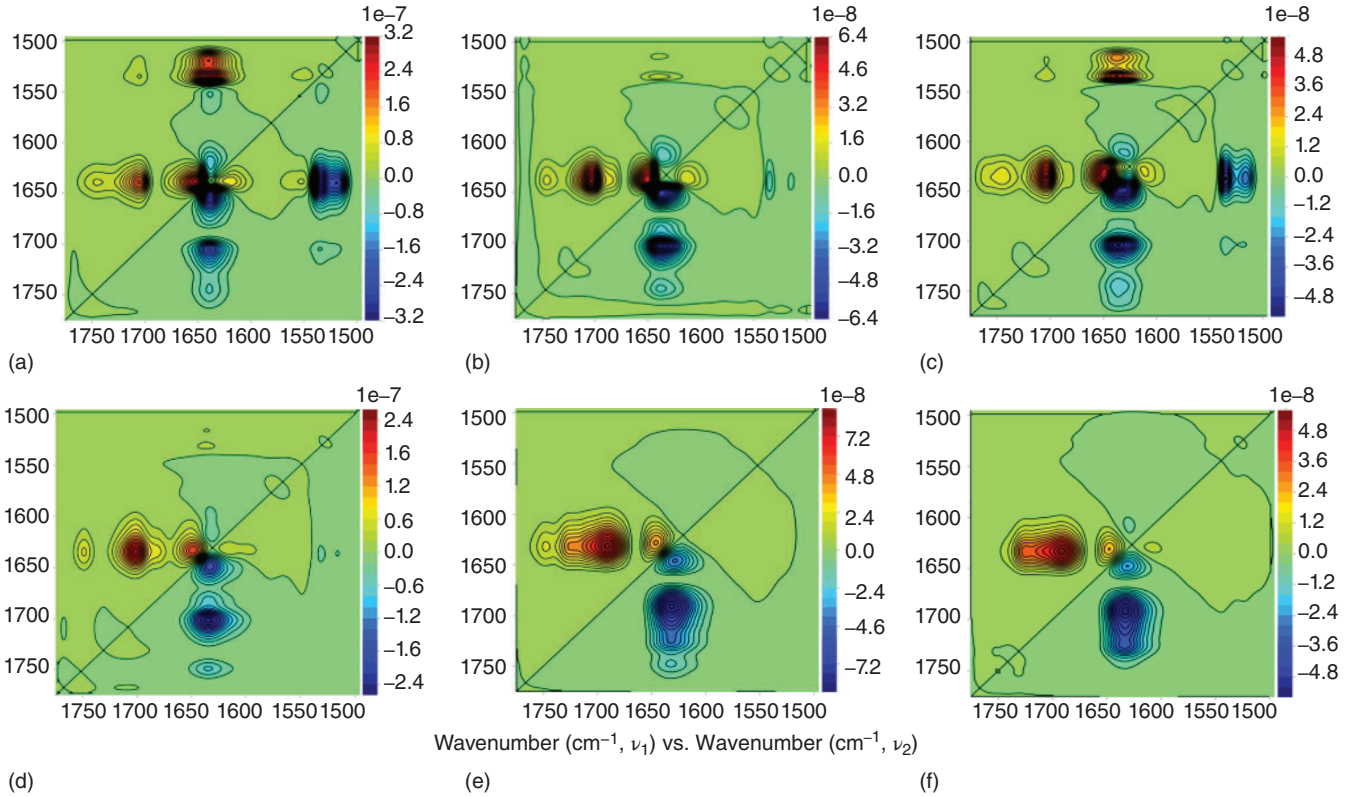

Figure 6 Co-distribution analysis for the distribution population of NISTmAb in solution in the spectral region of 1775–1500 cm^{-1} and within the temperature range of 28–60 °C at: (a) 100.0, (b) 70.0, (c) 40.0, (d) 10.0, (e) 5.0, and (f) 2.5 μg μL^{-1}.

(1684, 1680 cm^{-1}) is observed, followed by hydrogen bonding disruption involving arginine (1676 cm^{-1}) and glutamine (1581 cm^{-1}), then by glutamine (1581 cm^{-1}) deamidation within the β-turns, followed by perturbation of negatively charged aspartates (1572, 1567 cm^{-1}) and glutamates (1554 cm^{-1}). Then the disruption of hydrophobic (π-π stacking) interactions involving tyrosines (1518 cm^{-1}) and tryptophans (1504 cm^{-1}) is observed, followed by the disruption of hydrogen bonding interactions involving glutamates (1545 cm^{-1}) and tyrosines (1512 cm^{-1}) and a water molecule, then by the disruption of salt-bridge interactions involving lysines (1525 cm^{-1}) and glutamates (1534 cm^{-1}), followed by the perturbation of the backbone α- and π-helical segments (1645 and 1653 cm^{-1}, respectively), β-sheets (1636 cm^{-1}) and the hinge loop region (1663 cm^{-1}). Next, the disruption of hydrogen bonding interactions occurs, involving lysines (1625 cm^{-1}), arginines (1630 cm^{-1}), and glutamines (1590 cm^{-1}), followed by self-association (1616 cm^{-1}), then by disruption of hydrogen bonding interactions involving glutamines (1590 cm^{-1}), asparagines (1612 cm^{-1}), and histidines (1606, 1600 cm^{-1}), and the C-terminal end (1595 cm^{-1}). Finally, the disruption of hydrogen bonding interactions of glutamates/aspartates (1725, 1717, 1712, and 1704 cm^{-1}) with backbone and the surrounding aqueous environment is observed, which led to the perturbations of the β-turns (1696, 1684 cm^{-1}).

For the NISTmAb 70 μg μL^{-1} sample (Figure 12b), initially the disruption of hydrophobic (π–π stacking) interactions involving tryptophans (1504 cm^{-1}) and tyrosines (1518, 1512 cm^{-1}) was observed, followed by the disruption of weak interactions involving lysines (1525, 1625 cm^{-1}) and negatively charged residues such as aspartates (1572, 1567 cm^{-1}) and glutamates (1534 cm^{-1}), then by the disruption of hydrogen bonding interactions involving positively charged lysines (1625 cm^{-1}) and arginines (1630 cm^{-1}) within the β-sheets (1636 cm^{-1}) and α-helical segments (1645 cm^{-1}). This was followed by self-association (1616 cm^{-1}), then by disruption of hydrogen bonding interactions involving asparagines (1612 cm^{-1}), glutamines (1590 cm^{-1}), and π-helical segments (1653 cm^{-1}), followed by the disruption of hydrogen bonding interactions involving π-helical segments (1653 cm^{-1}) and histidines (1606, 1600 cm^{-1}), then by the perturbation of the C-terminal end (1595 cm^{-1}), followed by the perturbation of the hinge loop (1663 cm^{-1}), then by the perturbation of glutamines

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

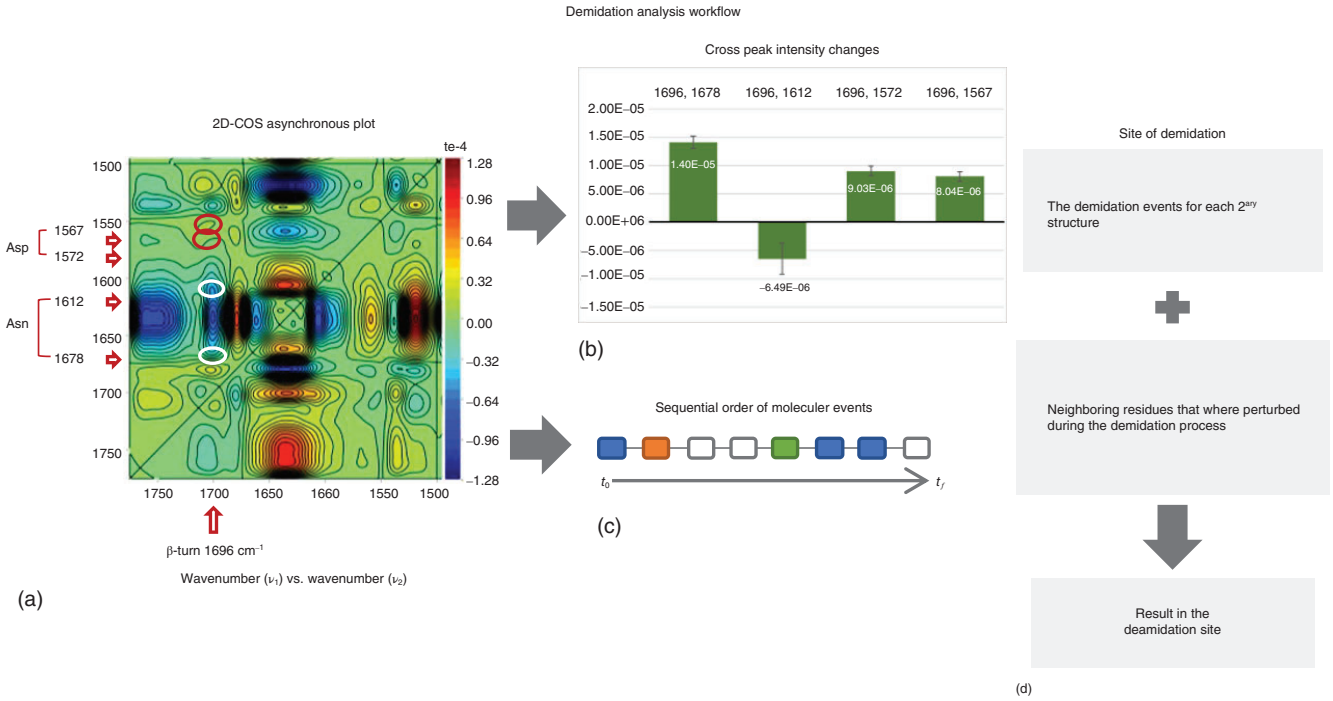

Figure 7 Schematic representation of the deamidation analysis workflow. From the actual 2D-COS asynchronous plot that is used to determine the extent of asparagine and glutamine deamidation, we show a representative NISTmAb 2D-COS asynchronous plot and interpretation of the same for a deamidation event (a). We focus first on the ν_2 axis on the β-turn ν(C═O) at 1696 cm^{-1}, then identify the cross peaks associated with asparagine and aspartate on the ν_1 axis within the β-turn at 1696 cm^{-1} (along this slice) of the ν_2 axis. The cross peaks for asparagine deamidation: Cross peaks (1696, 1678 and 1696, 1612 defined as (ν_2, ν_1)) assigned to the asparagine ν(C═O) at 1678 cm^{-1} and $\delta(NH_2)$ at 1612 cm^{-1} within the β-turn 1696 cm^{-1}, and are highlighted with white circles. For deamidation to be self-evident, the aspartate cross peaks must have a peak intensity change that is increasing in intensity (see color bar). The cross peaks associated with the β-turn 1696 cm^{-1} on the ν_2 axis aspartate $\nu(COO^-)$ at 1572 and 1567 cm^{-1} on the ν_1 axis (1696, 1572 and 1696, 1567 defined as (ν_2, ν_1)) have been highlighted with red circles. (b) To confirm deamidation, the aspartate cross peaks should be increasing in intensity while the asparagine peaks decrease. A bar graph summarizing the cross peak intensity changes demonstrates that the deamidation event has occurred for an asparagine residue located within the β-turn at 1696 cm^{-1}. The error propagation is defined in the error bar for each peak intensity change. (c) Both the synchronous (*not shown*) and asynchronous 2D-COS plots are analyzed using Noda's rules to define the sequential order of molecular events for the NISTmAb, (d) The cross peak intensity changes from panel (b) for a deamidation event within a secondary structure are investigated for neighboring residues that are perturbed during the deamidation process were identified using (c) the sequential order of molecular events. The schematic representation of the threaded squares are: (white) secondary structures, (green) asparagine residue that is undergoing deamidation, (blue) positively charged, and (orange) negatively charged residues. This information is used to define the site of deamidation within the amino acid sequence of the NISTmAb. We also used the available high-resolution structure to generate representative models of the site of deamidation to confirm the aqueous solvent accessibility required for the second step of the deamidation process.

(1581 cm^{-1}) by deamidation, followed by the perturbation of glutamates (1545, 1554 cm^{-1}) and of β-turns (1696 cm^{-1}). Finally, we see the disruption of hydrogen bonding interactions involving glutamates/aspartates (1725, 1717, 1712, and 1704 cm^{-1}) with the protein backbone and the surrounding aqueous environment, which led to perturbations of the β-turns (1696, 1684, 1680 cm^{-1}) and arginine (1676 cm^{-1}).

For the RM 8670 40 μg μL^{-1} sample (Figure 12c), initially the disruption of hydrogen bonding interactions involving glutamate (1554 cm^{-1}) and tyrosine (1518 cm^{-1}) through a water molecule was observed, followed by the disruption of hydrophobic (π–π stacking) interactions involving tryptophans (1504 cm^{-1}) and tyrosines (1518, 1512 cm^{-1}), then by the disruption of weak interactions involving lysines (1525, 1625 cm^{-1}) and negatively charged residues such as aspartates (1572, 1567 cm^{-1}) and glutamates (1534 cm^{-1}), followed by the disruption of hydrogen bonding interactions involving positively charged lysines (1625 cm^{-1}) and arginines (1630 cm^{-1}) with the backbone of the β-sheets (1636 cm^{-1}) and α-helical segments (1645 cm^{-1}). Then, self-association (1616 cm^{-1}) was disrupted, followed by disruption of hydrogen bonding interactions involving asparagines (1612 cm^{-1}), glutamines (1590 cm^{-1}), and π-helical segments (1653 cm^{-1}), then by the disruption of hydrogen

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

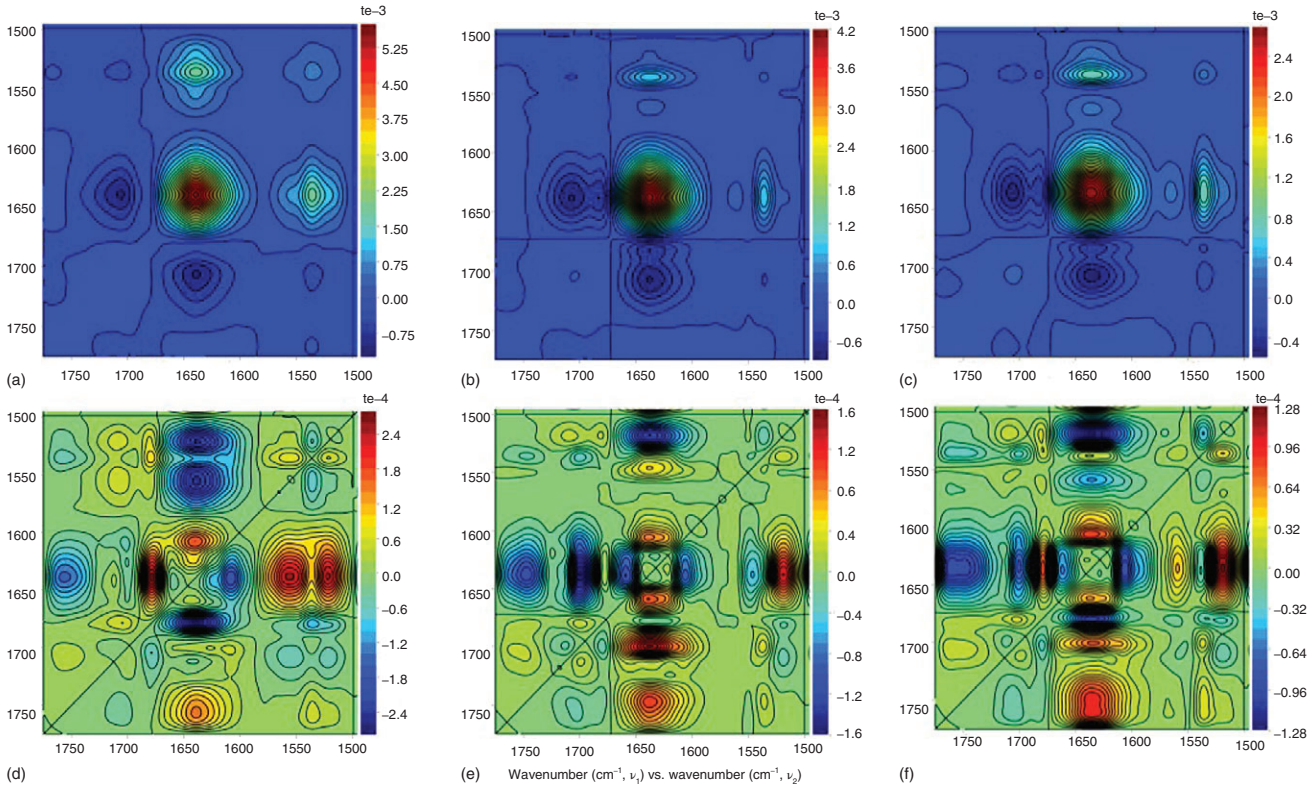

Figure 8 2D-COS for NISTmAb at high concentrations in 25 mM L-histidine at pH 6.0 within the spectral region of 1775–1500 cm^{-1} for the (a–c) synchronous and (d–f) asynchronous plots of (a, d) 100.0 μg $μL^{-1}$, (b, e) 70.0 μg $μL^{-1}$, and (c, f) 40.0 μg $μL^{-1}$, respectively.

bonding interactions involving π-helical segments (1653 cm^{-1}) and histidines (1606, 1600 cm^{-1}) and by the perturbation of the C-terminal end (1595 cm^{-1}). Next, the perturbation of the hinge loop occurred (1663 cm^{-1}), then the perturbation of glutamines (1581 cm^{-1}) by deamidation, followed by the perturbation of glutamates (1545, 1554 cm^{-1}) and of β-turns (1696 cm^{-1}). Finally, the disruption of hydrogen bonding interactions involving glutamates/aspartates (1725, 1717, 1712, and 1704 cm^{-1}) with the backbone and the surrounding aqueous environment was observed, which led to perturbations of the β-turns (1696, 1684, 1680 cm^{-1}) and arginine (1676 cm^{-1}).

For the 10 μg $μL^{-1}$ sample (control) shown in Figure 12(d), initial perturbations of hydrophobic (π-π stacking) interactions involving tyrosines (1518, 1512 cm^{-1}) and tryptophans (1504 cm^{-1}) were observed, followed by the disruption of hydrogen bonding interactions involving tryptophans (1504 cm^{-1}) and lysines (1525 cm^{-1}) and by the disruption of salt-bridge interactions involving lysines (1525 cm^{-1}), glutamates (1534, 1545, 1554 cm^{-1}) and aspartates (1567, 1572 cm^{-1}). Perturbation of the hinge loops (1663 cm^{-1}) was observed next, and then the disruption of hydrogen bonding interactions between glutamines (1581 cm^{-1}) and the backbone of the hinge loop (1663 cm^{-1}) and helical structures (1653 and 1645 cm^{-1}), followed by self-association (1616* cm^{-1}) due to the disruption of these weak interactions, followed by asparagine (1612 cm^{-1}) perturbation due to the disruption of hydrogen bonding interactions within the β-sheets (1636 cm^{-1}) or deamidation events. Next, we observe perturbation of the β-sheets (1636 cm^{-1}), followed by the disruption of the hydrogen bonding between glutamines (1590 cm^{-1}) and the β-sheets (1636 cm^{-1}) and histidines (1606, 1600 cm^{-1}), then by the disruption of hydrogen bonding interactions involving arginines (1630 cm^{-1}) and lysines (1625 cm^{-1}) with the C-terminal end (1595 cm^{-1}) and histidines (1600 cm^{-1}), then by the perturbation of aspartates (1567, 1572 cm^{-1}), then by the disruption of hydrogen bonding interactions involving glutamates/aspartates (1726, 1717, 1712 and 1704 cm^{-1}), arginine (1676 cm^{-1}) and glutamine (1670 cm^{-1}) with the backbone and the surrounding aqueous environment. This led to the perturbations of the β-turns (1696, 1684 cm^{-1}), and finally to the perturbation of glutamates (1554, 1545 cm^{-1}) located within the β-turns (1696, 1684 cm^{-1}).

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

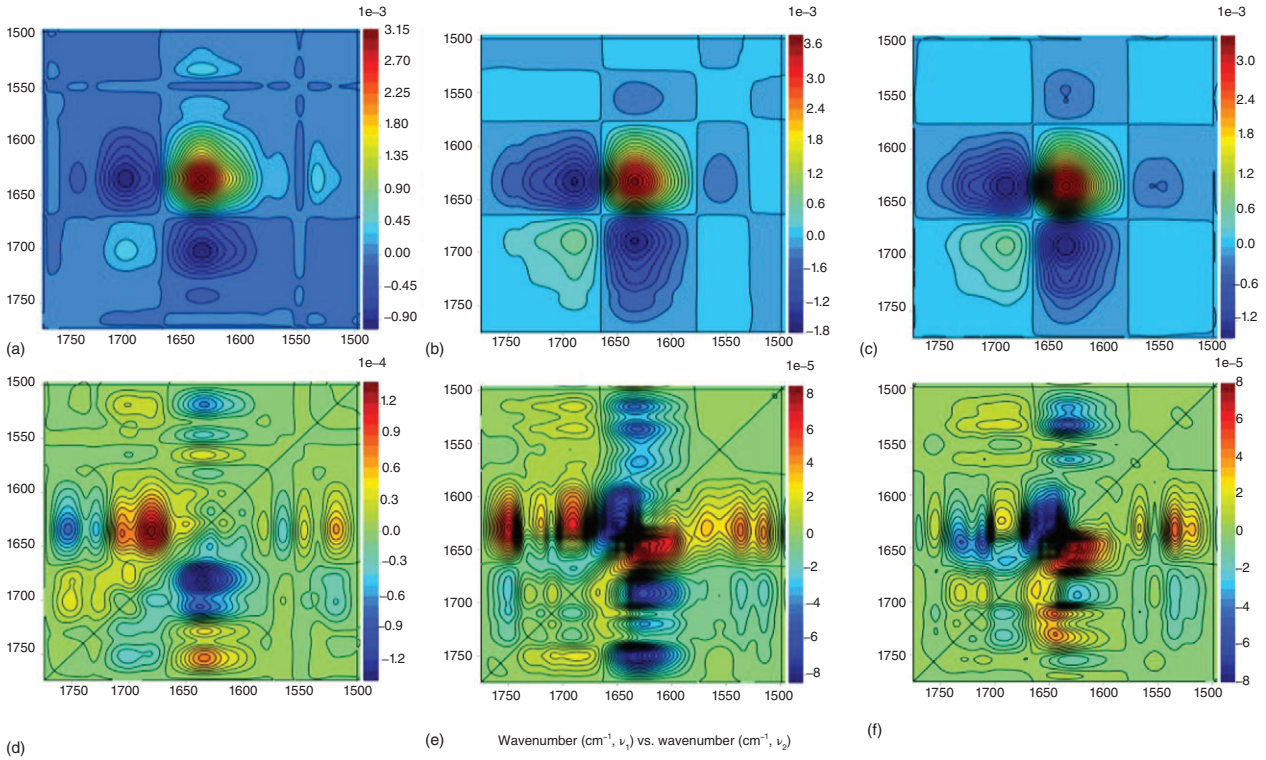

Figure 9 2D-COS for NISTmAb at low concentrations in 25 mM L-histidine at pH 6.0 within the spectral region of 1775–1500 cm^{-1} for the (a–c) synchronous and (d–f) asynchronous plots with concentrations of (a, d) 10.0 μg μL^{-1}, (b, e) 5.0 μg μL^{-1}, and (c, f) 2.5 μg μL^{-1}, respectively. The β-turns undergo a greater extent of perturbation when compared to samples at a higher concentration range.

For NISTmAb at 5 μg μL^{-1} shown in Figure 12(e), the initial perturbation involves the disruption of a salt-bridge interaction involving glutamate (1534 cm^{-1}) and lysine (1525 cm^{-1}), followed by perturbation of hydrophobic (π-π stacking) interactions involving tyrosines (1518, 1512 cm^{-1}) and tryptophans (1504 cm^{-1}), then by perturbation of negatively charged residues such as glutamates (1534, 1545, 1554 cm^{-1}) and aspartates (1567, 1572 cm^{-1}) and by disruption of salt-bridge interactions involving lysines (1525 cm^{-1}). This is followed by disruption of hydrogen bonding interactions involving the glutamines (1581 cm^{-1}) and C-terminal end (1596 cm^{-1}), other glutamines (1590 cm^{-1}) and histidines (1600, 1606 cm^{-1}), then by disruption of hydrogen bonding interactions involving histidines (1606 cm^{-1}) and asparagine (1612 cm^{-1}), and then by self-association (1616 cm^{-1}) involving lysine (1625 cm^{-1}) and arginine (1630 cm^{-1}). Next, we see the perturbation of the backbone β-sheets (1636 cm^{-1}), α-helical segments (1645 cm^{-1}), π-helices (1653 cm^{-1}), and the hinge loop (1661 cm^{-1}), followed by disruption of hydrogen bonding interactions of glutamates/aspartates (1726, 1717, 1712, 1708, and 1704 cm^{-1}), arginine (1676 cm^{-1}), and glutamine (1670 cm^{-1}) with the backbone and the surrounding aqueous environment, which led to the perturbations of the β-turns (1696, 1684 cm^{-1}), and finally to the perturbation of glutamates (1554, 1545, 1534 cm^{-1}) and aspartates (1572, 1567 cm^{-1}) located within the β-turns (1696, 1684, and 1680 cm^{-1}). The greater level of hydrogen bonding disruption involving charged residues within the β-turns (1696, 1684, and 1680 cm^{-1}) suggests there may be greater aqueous solvent accessibility of the diluted mAb. This relationship may also be observed through the presence of two isosbestic points at 1665 and 1577 cm^{-1}, which are not observed at higher mAb concentrations.

For NISTmAb at 2.5 μg μL^{-1} shown in Figure 12(f), initially, the negatively charged residues are perturbed, i.e. glutamates (1554, 1545 cm^{-1}) and aspartates (1567, 1572 cm^{-1}), followed by the disruption of a salt-bridge interaction involving lysine (1525 cm^{-1}) and glutamate (1534 cm^{-1}), then by the disruption of hydrophobic (π-π stacking) interactions involving tyrosines (1518, 1512 cm^{-1}) and tryptophans (1504 cm^{-1}), followed by disruption of hydrogen bonding interactions involving the glutamine (1581 cm^{-1}) and C-terminal end (1596 cm^{-1}), glutamines (1590 cm^{-1}) and histidines (1600, 1606 cm^{-1}).

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

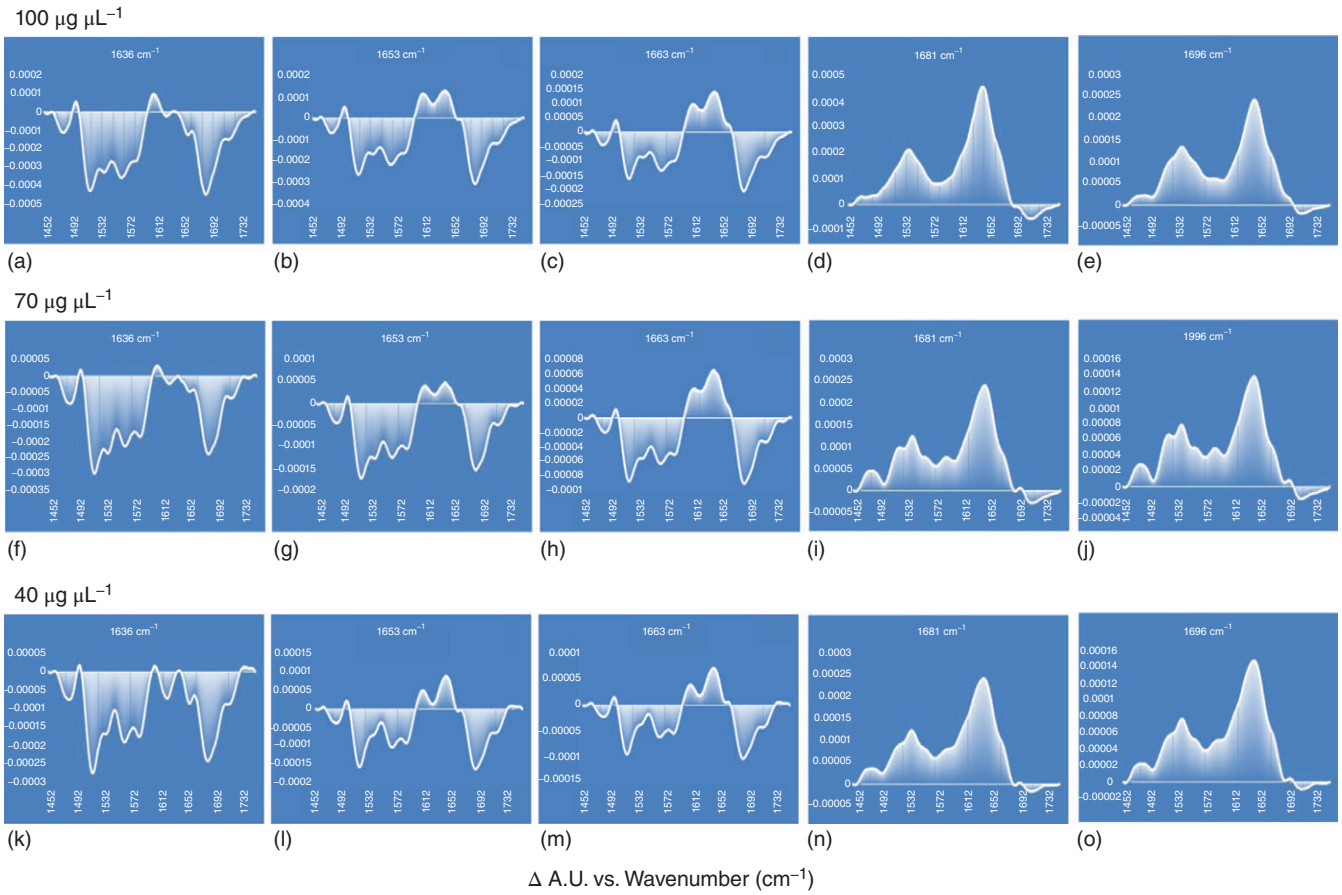

Figure 10 2D-COS asynchronous slice spectra NISTmAb at high concentration within the thermal temperature range 28–60 °C: (top row) 100.0, (middle row) 70.0, (bottom row) 40.0 μg μL^{-1}, respectively. Each spectrum shown is comprised of the intensity changes with respect to each secondary structure component analyzed: (a,f,k) β-sheet at 1636 cm^{-1} (b,g,l) helical segment 1653 cm^{-1}, (c,h,m) hinge loops at 1663 cm^{-1}, (d,i,n) β-turn at 1681 cm^{-1} and (e,j,o) β-turn at 1696 cm^{-1}. Also, observed within these slice plots are the key signature peaks associated with the deamidation process and self-association/aggregation (1616 cm^{-1}).

Next, we observe disruption of hydrogen bonding interactions involving histidines (1606 cm^{-1}) and asparagine (1612 cm^{-1}), followed by self-association (1616 cm^{-1}) involving lysine (1625 cm^{-1}) and arginine (1630 cm^{-1}), then by disruption of hydrogen bonding interactions involving positively charged lysines (1625 cm^{-1}) and arginines (1630 cm^{-1}) with the β-sheets (1636 cm^{-1}) and helical segments (1645, 1653 cm^{-1}), followed by backbone perturbations in β-sheets (1636 cm^{-1}), α-helical segments (1645 cm^{-1}), π-helices (1653 cm^{-1}), and the hinge loop (1661 cm^{-1}). Disruption of hydrogen bonding interactions involving glutamines (1670 cm^{-1}) occurs next, followed by the disruption of hydrogen bonding interactions of glutamates/aspartates (1725, 1717, 1712, and 1704 cm^{-1}) with the backbone and the surrounding aqueous environment, which led to perturbations of the β-turns (1696, 1684, 1680 cm^{-1}), then by disruption of weak interactions involving glutamates (1554, 1545, 1534 cm^{-1}), arginine (1676 cm^{-1}), aspartates (1572, 1567 cm^{-1}), lysines (1525 cm^{-1}), tyrosines (1518, 1512 cm^{-1}) and tryptophan (1504 cm^{-1}), all located within the β-turns.

12.5 The Extent of Deamidation

The real-time monitoring of the loss of intensity of side-chain vibrational modes for asparagine and/or glutamine (ν(C═O), $\gamma(NH_2)$: 1670, 1612 or 1678, 1581 cm^{-1}, respectively) and the gains in the intensity of aspartate and/or glutamate ($\nu(COO^-)$: 1572, 1567 or 1554, 1545, cm^{-1}, respectively) within the thermal perturbation range of 28–60 °C allowed for assessment of the risk of deamidation. The amount of each type of deamidation for each NISTmAb sample at high (100–40 μg μL^{-1}) and low (10–2.5 μg μL^{-1}) concentration range is summarized in Tables 3–6. The amount of deamidation of glutamine and

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

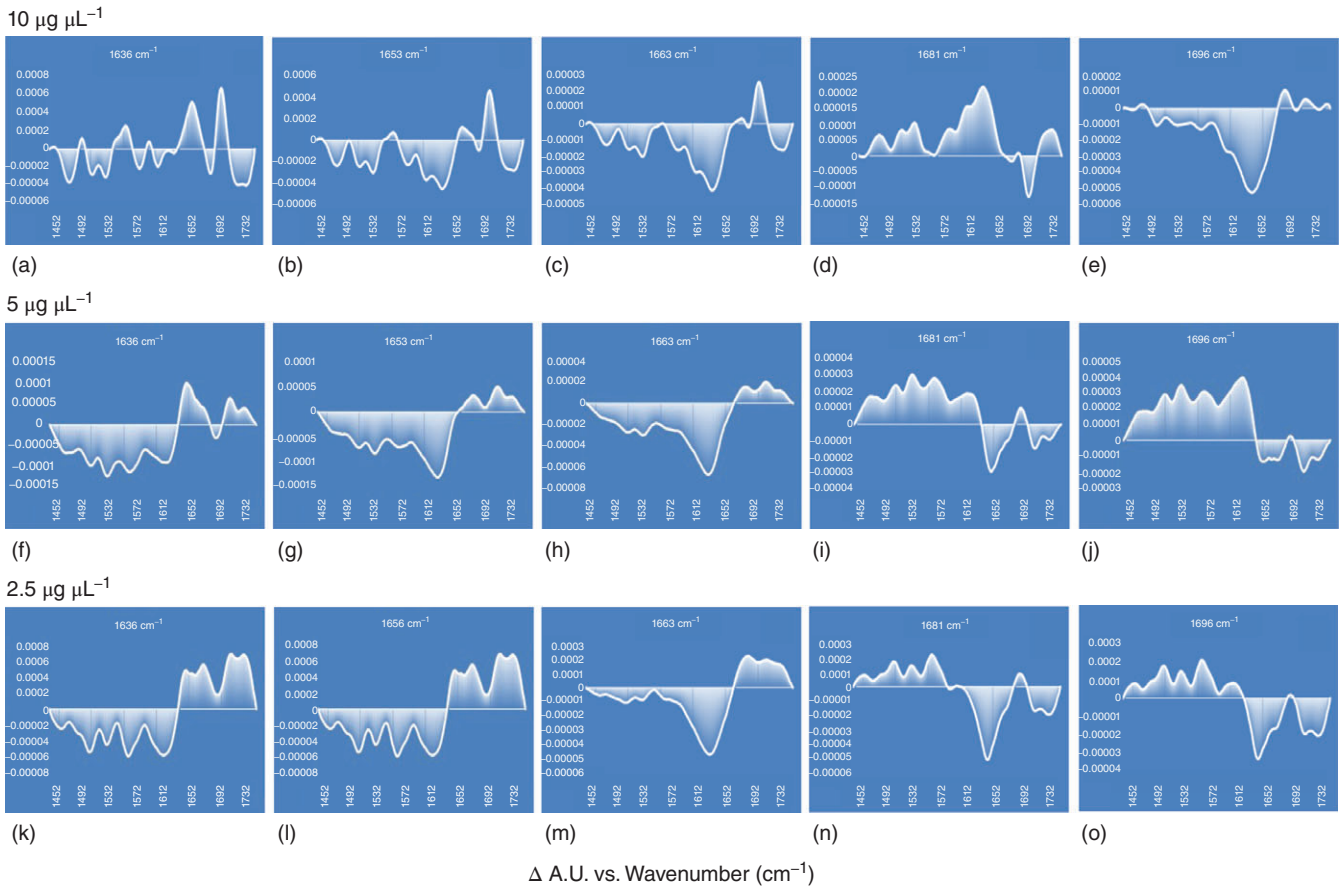

Figure 11 2D-COS asynchronous slice spectra NISTmAb at low concentration within the thermal temperature range 28–60 °C: (top row) 10.0, (middle row) 5.0, (bottom row) 2.5 μg μL^{-1}, respectively. Each spectrum is comprised of the intensity changes with respect to each secondary structure component analyzed: (a,f,k) β-sheet at 1636 cm^{-1} (b,g,l) helical segment 1653 cm^{-1}, (c,h,m) hinge loops at 1663 cm^{-1}, (d,i,n) β-turn at 1681 cm^{-1}, and (e,j,o) β-turn at 1696 cm^{-1}. Also, the aggregation peak at 1616 cm^{-1} can readily be observed in these slice plots.

asparagine was determined for each secondary structure component within the sample.

12.6 Sites of Deamidation

Theoretically, the potential deamidation sites can be evaluated by amino acid sequence analysis as defined by steric hindrance based on the work done by Robinson and others.[79,61] However, here, the sites of deamidation were empirically determined by taking into consideration both the extent of deamidation associated with its secondary structure as summarized in Tables 3–6 and the sequential order of molecular events to define the neighboring residues. The perturbations to the surrounding amino acids are significant due to the deamidation process introducing a negative charge. This cumulative information is used to determine the deamidation sites, summarized in Figure 13. In addition, we generated structural models for the Fab and Fc regions using the available high-resolution structures PDB ID 5K8A[90] and PDB ID: 5VGP[91], respectively to highlight the spatial location of the deamidation sites and their solvent accessibility Figures 14 and 15.

We show (Figure 13) the different deamidation sites identified for all of the NISTmAb concentrations studied. The asparagine deamidation sites are located within the constant Fc region of the HC N437 near the C-terminal end and HC N318. For glutamine deamidation, which resulted in higher deamidation risk, its site(s) within the HC region were also located within the constant Fc region at Q421 and/or Q422. A greater risk of deamidation was observed for glutamine located within the β-turns exposed to their aqueous environment at 100 μg μL^{-1} for the NISTmAb within the conditions evaluated. Furthermore, deamidation risk was also observed for the LC, including the two glutamines, Q36 and Q154, and a single

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

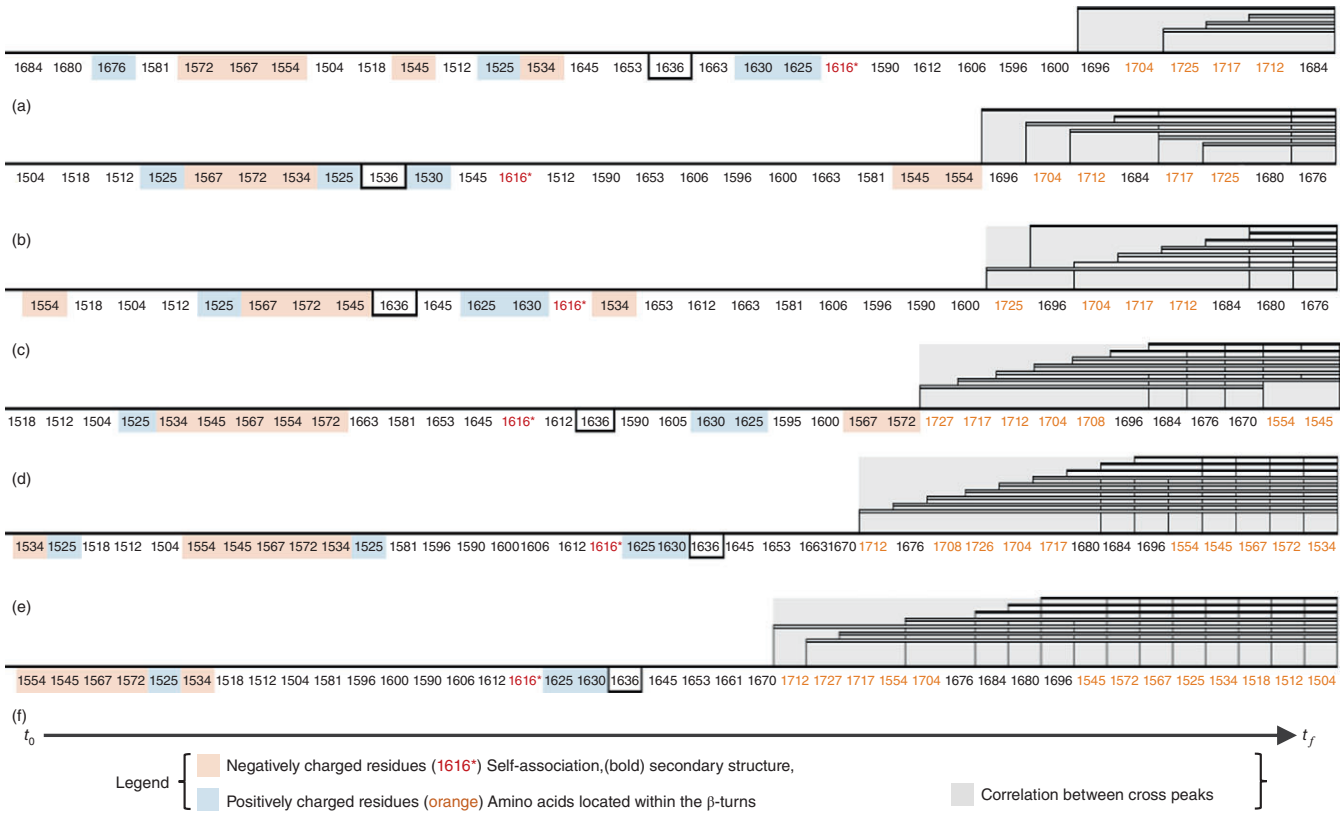

Figure 12 Sequential order of molecular events for the NISTmAb samples within a concentration range of 100–2.5 μg μL^{-1}. Samples were perturbed within the temperature range of 32–60 °C with 8 °C intervals for: (a) 100.0, (b) 70.0, (c) 40.0, (d) 10.0, (e) 5.0, and (f) 2.5 μg μL^{-1}. All charge variant samples exhibit self-association. For the 100 μg μL^{-1} sample, the β-turns (1684 and 1680 cm^{-1}) were presumably perturbed due to glutamine deamidation events observed within these secondary structures at this high concentration. For the lower concentrations, weak interactions are perturbed first. Also, a disruption of hydrogen bonding interactions involving tyrosines (1518, 1512 cm^{-1}) and tryptophans (1504 cm^{-1}) are observed within the β-turns for the 2.5 μg μL^{-1} sample, presumably due to greater solvent accessibility.

Table 3 Summary of the asparagine/glutamine deamidation per secondary structure for NISTmAb at high concentration range (100–40 μg μL^{-1}) in 25 mM L-histidine at pH 6.0 as a function of concentration

	β-turn		β-turn		Hinge loop		α-helix		β-sheet	
aa	1696 cm^{-1}		1681 cm^{-1}		1663 cm^{-1}		1653 cm^{-1}		1636 cm^{-1}	
	Value[a]	Error	Value	Error	Value	Error	Value	Error	Value	Error
100 μg μL^{-1}										
N	0.33	±0.02	0.09	±0.00	0.00	±0.00	0.00	±0.00	0.00	±0.00
Q	**1.35**	**±0.10**	**2.01**	**±0.24**	0.00	±0.00	0.00	±0.00	0.00	±0.00
70 μg μL^{-1}										
N	0.54	±0.17	0.14	±0.01	0.00	±0.00	0.00	±0.00	0.00	±0.00
Q	0.82	±0.17	0.26	±0.05	0.00	±0.00	0.00	±0.00	0.00	±0.00
40 μg μL^{-1}										
N	0.26	±0.02	0.12	±0.00	0.00	±0.00	0.00	±0.00	0.00	±0.00
Q	0.78	±0.08	0.54	±0.04	0.00	±0.00	0.00	±0.00	0.00	±0.00

[a]The number of deamidated residues for the NISTmAb during the thermal perturbation.
Bold font is used to highlights a significant deamidation event.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

Table 4 Summary of the overall extent of deamidation for NISTmAb at high concentration range (100–40 $\mu g\,\mu L^{-1}$) in 25 mM L-histidine at pH 6.0 as a function of concentration

Deamidation Type	Overall extent of deamidation[a]	Error propagation	comment[b]
100 $\mu g\,\mu L^{-1}$			
Asparagine	0.42	±0.02	No deamidation risk
Glutamine	**3.36**	**±0.26**	Deamidation risk
70 $\mu g\,\mu L^{-1}$			
Asparagine	0.68	±0.17	No deamidation risk
Glutamine	**1.08**	**±0.18**	Deamidation risk
40 $\mu g\,\mu L^{-1}$			
Asparagine	0.38	±0.02	No deamidation risk
Glutamine	**1.32**	**±0.09**	Deamidation risk

[a]The sum of the deamidation per secondary structure values.
[b]The threshold value for the deamidation risk has been set to ≥1.0.
Bold font is used to highlights a significant deamidation event.

Table 5 Summary of the asparagine/glutamine deamidation per secondary structure for NISTmAb at low concentration range (10–2.5 $\mu g\,\mu L^{-1}$) in 25 mM L-histidine at pH 6.0 as a function of concentration

	β-turn		β-turn		Hinge loop		α-helix		β-sheet	
aa	**1696 cm^{-1}**		**1681 cm^{-1}**		**1663 cm^{-1}**		**1653 cm^{-1}**		**1636 cm^{-1}**	
	Value	Error	Value	Error	Value	Error	Value	Error	Value	Error
10 $\mu g\,\mu L^{-1}$										
N	0.10	±0.00	0.03	±0.00	0.08	±0.01	0.12	±0.01	0.10	±0.01
Q	0.10	±0.02	0.02	±0.01	0.00	±0.00	0.00	±0.00	0.00	±0.00
5 $\mu g\,\mu L^{-1}$										
N	0.16	±0.02	0.11	±0.01	0.00	±0.00	0.00	±0.00	0.00	±0.00
Q	0.84	±0.08	0.60	±0.05	0.00	±0.00	0.00	±0.00	0.00	±0.00
2.5 $\mu g\,\mu L^{-1}$										
N	0.04	±0.01	0.03	±0.00	0.01	±0.00	0.02	±0.00	0.06	±0.01
Q	0.17	±0.01	0.04	±0.01	0.42	±0.02	0.00	±0.00	0.00	±0.00

Table 6 Summary of the overall extent of deamidation for NISTmAb at low concentration range (10–2.5 $\mu g\,\mu L^{-1}$) in 25 mM L-histidine at pH 6.0 as a function of concentration

Deamidation type	Overall extent of deamidation	Error propagation	Comment
10 $\mu g\,\mu L^{-1}$			
Asparagine	0.43	±0.02	No deamidation risk
Glutamine	0.12	±0.02	No deamidation risk
5 $\mu g\,\mu L^{-1}$			
Asparagine	0.27	±0.02	No deamidation risk
Glutamine	**1.44**	**±0.09**	Deamidation risk
2.5 $\mu g\,\mu L^{-1}$			
Asparagine	0.16	±0.01	No deamidation risk
Glutamine	0.63	±0.02	No deamidation risk

Bold font is used to highlights a significant deamidation event.

asparagine, N157, located within the constant and variable Fab regions of the NISTmAb. Also, no deamidation risk was observed for the CDRs at the lower concentrations; therefore, glutamine deamidation does not impact target binding under the conditions evaluated.

12.7 Does the Deamidation Event Induce Self-association?

All the 2D-COS plots for NISTmAb, a charge variant protein, exhibited self-association (1616 cm^{-1}).[25] Upon closer examination, the self-association was preceded by deamidation of the glutamine residues (1581 cm^{-1}) located within the β-turns (1696 and 1681 cm^{-1}). Those steps led to a perturbation of the secondary structure and disruption of hydrogen bonding interactions between positively charged residues, lysines (1625 cm^{-1}) and/or arginines (1630 cm^{-1}), and glutamine (1590 cm^{-1}) for the 100 $\mu g\,\mu L^{-1}$ sample. For the 70 and 40 $\mu g\,\mu L^{-1}$ samples, the extent of deamidation was smaller; therefore, a different event, the disruption of salt-bridge interactions that led to perturbations within the β-sheets (1636 cm^{-1}) and helical regions (1645 cm^{-1}), led to the self-association (1616 cm^{-1}). Meanwhile, for the low concentration range (10–2.5 $\mu g\,\mu L^{-1}$) samples, the instability caused by the disruption of weak interactions (salt-bridge and hydrogen bonding) also led to self-association.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

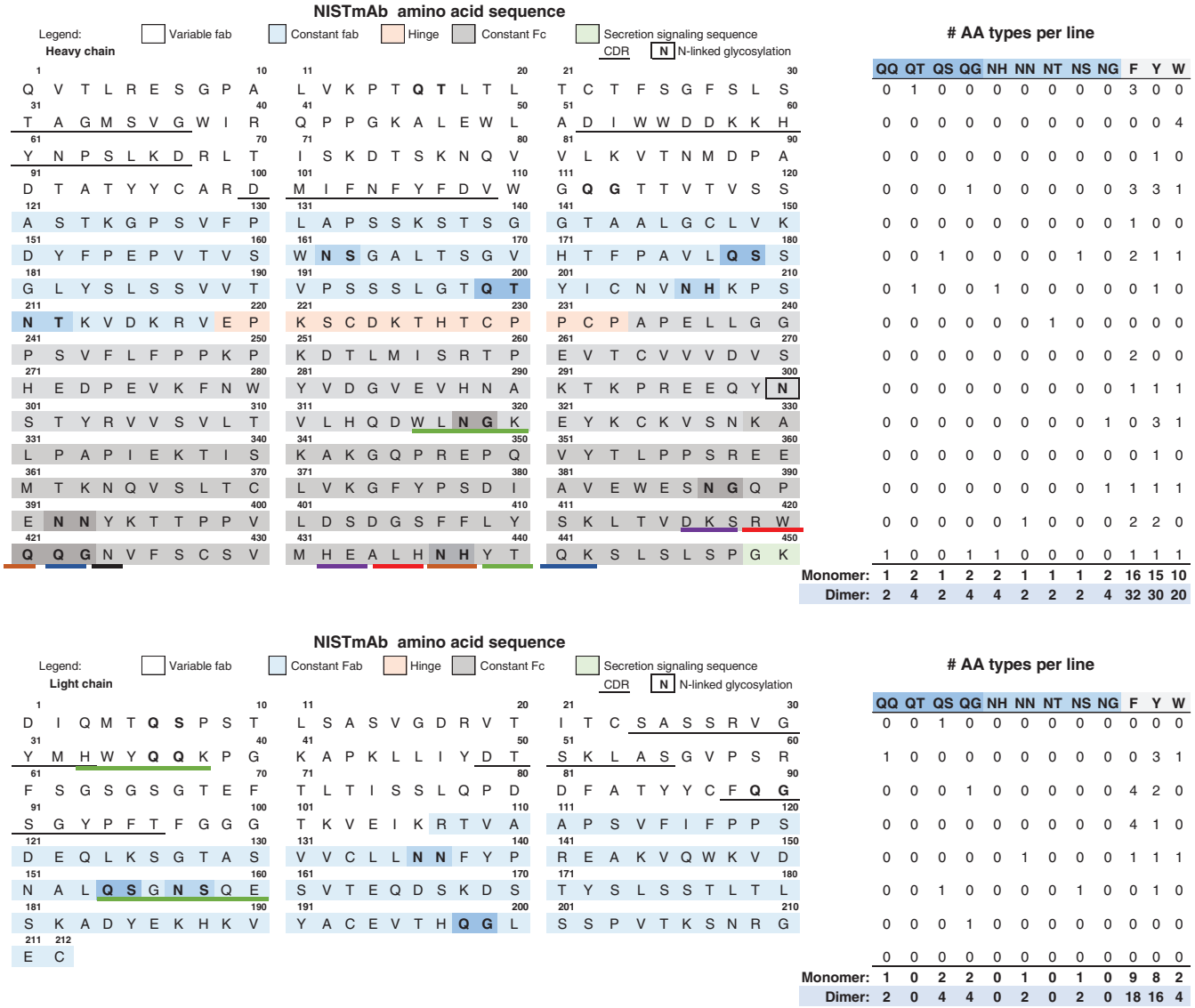

Figure 13 Schematic representation of the amino acid (AA) sequence along with asparagine and glutamine deamidation sites for the NISTmAb. Specifically: (purple underline) 100.0, (red underline) 70.0, (orange underline) 40.0, (green underline) 10.0, (blue underline) 5.0, and (thick black underline) 2.5 μg μL^{-1}, respectively. The 10 μg μL^{-1} sample also contains the deamidation sites identified for the sample under varying storage conditions (*data not shown*). Also shown are tables for the Heavy and Light chain summarizing the deamidation site sequence selection criteria and aromatic residue composition for the (top right) heavy and (bottom right) light chain components, respectively.

13 STABILITY

Protein stability involves 'weak interactions', including salt-bridge, hydrogen bonding, and hydrophobic interactions, that can be destabilized due to a deamidation which introduces a negative charge into a region of the protein where there was none. Furthermore, disruption of weak interactions or the loss of secondary structure can lead to self-association or aggregation, respectively. All of these events must be considered to evaluate the stability of a protein. The most significant contribution to secondary structure for the NISTmAb arises from the β-sheets.[92] We compared the stability of the remaining structural components with respect to the β-sheets. This assessment was performed using a 2D-COS asynchronous plot, specifically the slice analysis[23] for the β-sheet (1636 cm^{-1}) within the temperature range of 28–60 °C (Figure 16).

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

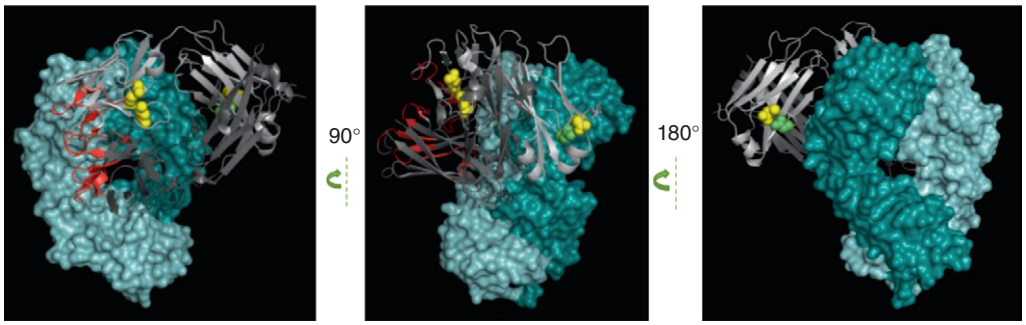

Figure 14 NISTmAb Fab high-resolution structure representation (PDB ID: 5K8A) is a combined (grey) ribbon and (cyan) solid surface model generated to highlight the symmetry of the homodimer. In addition, two different shades were used to distinguish between the (lighter shade) Light chain (LC) and (darker shade) heavy chain (HC). Only one of the dimer components (grey) was used to highlight: (red) the complementarity determining regions (CDR's) where antigen binding occurs for the light and heavy chains and the deamidation sites (green sphere) asparagine (N) and (yellow sphere) glutamine (Q) where found to be limited to the LC region. Specifically, Fab LC: Q36, Q37 and Q154, and N157 were all aqueous solvent-accessible deamidation sites. The model representation was generated using the PyMOL Molecular Graphics System, Version 2.3, Schrödinger, LLC. Source: Based on Gallagher et al.[90]

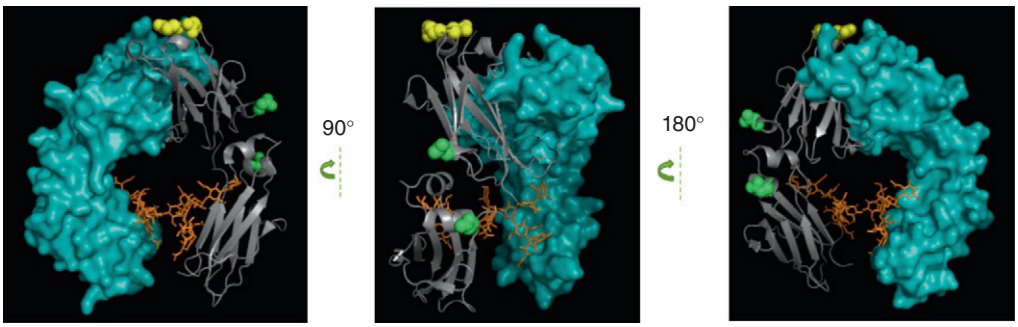

Figure 15 NISTmAb Fc high-resolution structure representation (PDB ID: 5VGP). The Fc region is comprised exclusively of the heavy chain (HC) component. The model was generated as a combined (dark grey) ribbon and (dark cyan) solid surface representation to highlight the symmetry of the homodimer. Also, represented as orange sticks are the glycosylation components of the NISTmAb. Only one of the dimer components (dark grey) chain A, will be used to highlight the deamidation sites within the model. The deamidation sites for the Fc region are (green sphere) asparagine (N) and (yellow sphere) glutamine (Q). Specifically, the Fc HC Q421, Q422, N318, and N437 are all aqueous solvent-accessible deamidation sites. In addition, these asparagine and glutamine deamidation sites do not impact the CDR regions, yet induce self-association of the NISTmAb. The model representation was generated using the PyMOL Molecular Graphics System, Version 2.3, Schrödinger, LLC. Source: Based on Gallagher et al.[91]

The bar graphs can be used to represent the intensity changes that occurred for the cross peaks associated with the secondary structure of the mAb protein. Negative peaks represent unfolding events, while positive peaks represent the stable structure or evidence of aggregation. Evidence of the impact of self-association during the thermal perturbation was observed for the β-turns of the NISTmAb within the 100–40 $\mu g\,\mu L^{-1}$ samples. A schematic diagram summarizing the key molecular perturbations for the 100 $\mu g\,\mu L^{-1}$ sample compared to the 2.5 $\mu g\,\mu L^{-1}$ sample is shown in Figure 17. The schematic diagram highlights the molecular events that lead to self-association, which results from glutamine deamidation.

14 DISCUSSION

The HS images of the NISTmAb in solution were analyzed using 2D-CDS and 2D-COS to assess further asparagine and/or glutamine deamidation, aggregation risk, and its impact on stability. These are key CQAs for the developability assessment of therapeutic proteins. 2D-COS synchronous and asynchronous plots were used to determine the concentration effects on the stability of the NISTmAb, if any, due to its charge variance after deamidation. We provide a full description of the sequential molecular order of events, the extent of deamidation, the identification of the deamidation sites, and their role in inducing self-association which would lead to altered

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

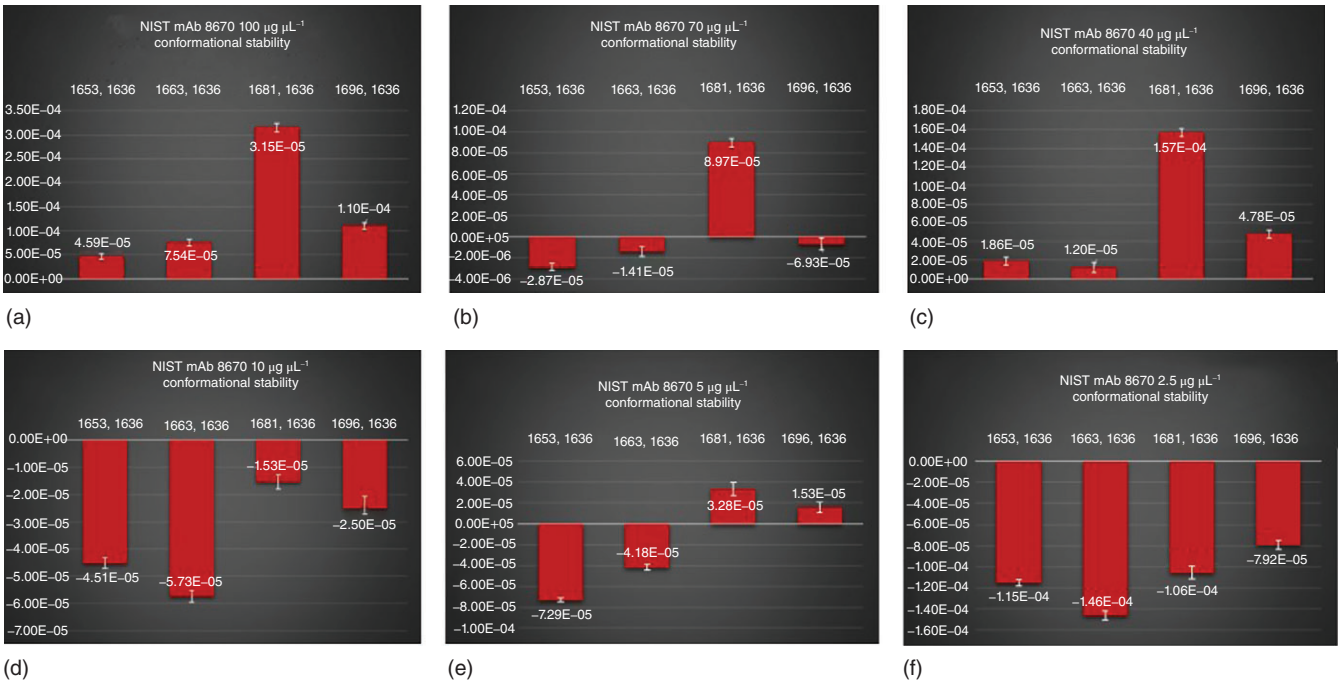

Figure 16 Bar graph summarizing the intensity changes of key asynchronous cross peaks used to evaluate conformational stability with respect to the β-sheet (1636 cm^{-1}) of NISTmAb: (a) 100.0, (b) 70.0, (c) 40.0, (d) 10.0, (e) 5.0, and (f) 2.5 $\mu g\,\mu L^{-1}$, respectively; within the temperature range (32–60 °C). Observed differences in stability suggest the extent of glutamine deamidation within the β-turns.

stability. The spatial and temporal resolution of IR spectral peaks associated with both conformational changes and the amino acid side chains of proteins obtained by 2D-COS and 2D-CDS has provided an unprecedented understanding of protein stability. Also, slice representations of the asynchronous 2D-COS provided confirming evidence of the spectral changes that occurred with respect to each secondary structure. The general stability of the distribution of proteins in solutions was assessed by 2D-CDS analysis for the NISTmAb under varying concentrations that led to identifying differences in the weak interactions for these samples. Two different types of governing interactions were observed: the first was intermolecular interactions between NISTmAb molecules at 100 $\mu g\,\mu L^{-1}$ and the second was weak interactions between the NISTmAb and its aqueous environment at concentrations ≤10 $\mu g\,\mu L^{-1}$.

The real-time monitoring of both asparagine and glutamine deamidation during the thermal perturbation of 32–60 °C allowed for assessment of the risk of deamidation for the NISTmAb at different concentrations. The extent of deamidation and sites of glutamine and asparagine deamidation were determined for each sample. In addition, we generated models for the Fab and Fc region using the available high-resolution structure[90,91] (PDB ID: 5K8A and 5VGP) to highlight the spatial location of the deamidation sites, evaluate the impact the deamidation event would have on the stability, and assess solvent accessibility. In summary, no deamidation risk was observed for the CDRs, therefore both glutamine and asparagine deamidation does not impact target binding under the conditions evaluated. However, *deamidation sites were identified within the HC including, N318 and N437 near the C-terminal end, and Q421/422 within the FC region. For the LC, three additional sites were determined: N157, Q36, and Q154 were observed within the variable and constant Fab regions. All sites were validated using LC-MS by independent groups using multiple instruments and different sample preparative methods*.[52,56,67,78,88–91] The significance of this result is that our approach did not involve the sample digestion and separation techniques involving peptide mapping, but rather the real-time monitoring of eight key signature peaks involved in the deamidation process during the thermal perturbation. Through the 2D-COS analysis, we correlated: (i) the key signature peaks with (ii) their corresponding secondary structure and (iii) the identification of neighboring residues that were perturbed during the deamidation process. These three results were then used to identify the sites of deamidation for the full-length protein in its standard formulation. Deamidation may also lead to self-association due to the perturbation of the weak interactions, as observed from

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

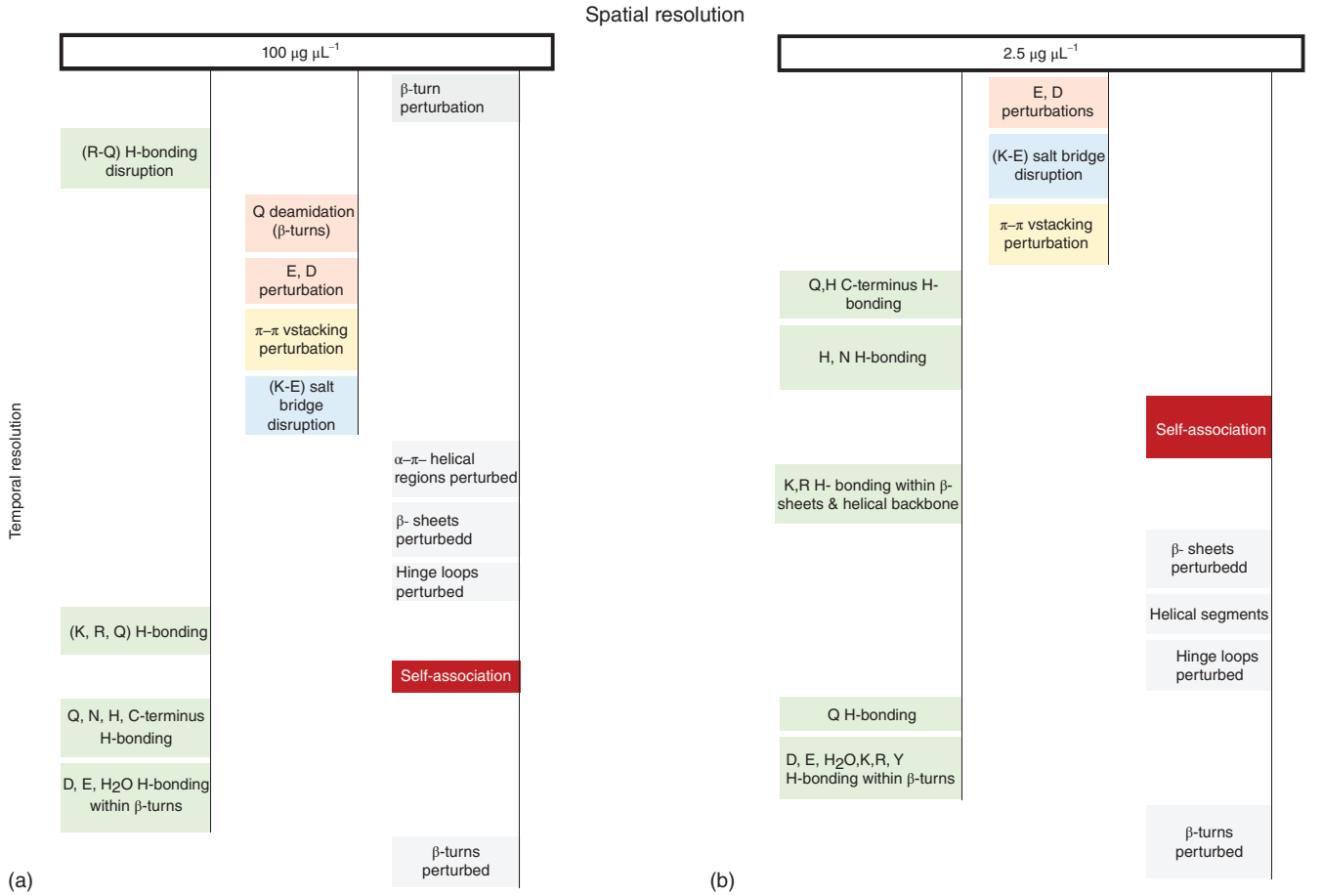

Figure 17 Schematic diagram summarizing the molecular dynamics observed for the NISTmAb during the thermal perturbation at (a) 100.0 and (b) 2.5 μg μL^{-1}.

the analysis of the sequential order of molecular events obtained from the analysis of the 2D-COS plots.

15 CONCLUSION

A review of the proven capabilities of both 2D-COS and 2D-CDS to study proteins was presented. For far too long, only protein secondary structures and conformational changes have been studied by IR spectroscopy. Thus, leaving obscured the information in the IR spectral data that describes all of the changes governed by amino acid side chains obscured. An analysis of the side-chain information has been proven possible by taking advantage of the enhanced resolution through 2D-COS. More importantly, original work toward the understanding of both glutamine and asparagine deamidation was presented. Deamidation of asparagine and glutamine residues in proteins is a concern for both protein developers and regulatory agencies. Here the direct monitoring of the deamidation process of the full-length protein in its formulation without any sample preparative steps, providing unprecedented understanding of this degradation event. Furthermore, the use of a QCLM and a slide array allow for minimal sample requirements and comparability assessment of up to 23 samples in real-time under controlled thermal conditions. The analysis extends to the identification of the deamidation sites within the sequence of the protein and the direct evaluation of the protein's stability. We have presented a quantitative and comprehensive 2D-COS and 2D-CDS evaluation of the deamidation processes of the NISTmAb validated by LC-MS. Our results suggest this QCLM-based technology will be an essential step in establishing protein developability and comparability, while enabling proteomics and drug design today and in the future.

ACKNOWLEDGMENTS

The authors would like to thank Sherly Nieves (Protein Dynamic Solutions, Inc) for performing the actual

DOI: 10.1002/9780470027318.a9792

experiments and for her contribution to the analysis of the spectral data. We would also like to acknowledge Dr. Melissa Stauffer (Scientific Editing Solutions, Walworth, WI) for editing the manuscript. The authors would like to acknowledge Dr. Young Jong Lee and Dr. Trina Mouchahoir for their efforts as internal reviewers at NIST.

Commercial instruments and materials are identified in this paper to specify the experimental procedure adequately. Such identification does not imply an endorsement by the National Institute of Standards and Technology.

FUNDING

The work presented herein was made possible by support from the National Science Foundation (SBIR PII Award 1632420 (BP)).

ABBREVIATIONS AND ACRONYMS

2D-CDS	Two-dimensional Co-distribution Spectroscopy
2D-COS	Two-dimensional Infrared Correlation Spectroscopy
ATR	Attenuated Total Reflectance
CDR	Complementarity-determining Region
DMPG	Dimyristoylphosphatidylglycerol
HC	Heavy Chain
HS	Hyperspectral
IR	Infrared
IRM	Infrared Microscope
LC	Light Chain
LC–MS	Liquid Chromatography–Mass Spectrometry
mAb	Monoclonal Antibody
MS/MS	Tandem Mass Spectrometry
NIST	National Institute of Standards and Technologies
QCLM	Quantum Cascade Laser Microscope
QCL IRM	Quantum Cascade Laser Microscopy
TFA	Trifluoroacetic Acid

RELATED ARTICLES

Biomolecules Analysis
Infrared Spectroscopy of Biological Applications

Peptides and Proteins
Fourier Transform Infrared Spectroscopy in Peptide and Protein Analysis • Proteolytic Mapping • Matrix-assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis• High-performance Liquid Chromatography/Mass Spectrometry in Peptide and Protein Analysis • High-performance Liquid Chromatography-Mass Spectrometry in Peptide and Protein Analysis

General Articles
Multivariate and Hyperspectral Image Analysis

Biomedical Spectroscopy
Two-dimensional Vibrational Correlation Spectroscopy in Biomedical Sciences

Application of Two-dimensional Correlation Spectroscopy in Protein Research • Instrumentation and Applications of Cascade Laser Spectroscopy • Quantitative and Qualitative Analysis of Posttranslationally Modified Proteins

REFERENCES

1. J. Bandekar, 'Amide Modes and Protein Conformation', *Biochim. Biophys. Acta*, **1120**, 123–143 (1992). DOI: 10.1016/0167-4838(92)90261-B.
2. Y.N. Chirgadze, O.V. Fedorov, N.P. Trushina, 'Estimation of Amino Acid Residue Side-Chain Absorption in the Infrared Spectra of Protein Solutions in Heavy Water', *Biopolymers*, **14**, 679–694 (1975). DOI: 10.1002/bip.1975.360140402.
3. A. Barth, 'Infrared Spectroscopy of Proteins', *Biochim. Biophys. Acta*, **1767**, 1073–1101 (2007). DOI: 10.1016/j.bbabio.2007.06.004.
4. J.L. Arrondo, A. Muga, J. Castresana, F.M. Goni, 'Quantitative Studies of the Structure of Proteins in Solution by Fourier-Transform Infrared Spectroscopy', *Prog. Biophys. Mol. Biol.*, **59**, 23–56 (1993). DOI: 10.1016/0079-6107(93)90006-6.
5. P.I. Haris, D. Chapman, 'The Conformational Analysis of Peptides Using Fourier Transform IR Spectroscopy', *Biopolymers*, **4**, 251–263 (1995). DOI: 10.1002/bip.360370404.
6. S.Y. Venyaminov, F.G. Prendergast, 'Water (H_2O, and D_2O) Molar Absorptivity in the 1000 – 4000 cm^{-1} Range and Quantitative Infrared Spectroscopy of Aqueous Solutions', *Anal. Biochem.*, **248**, 234–245 (1997). DOI: 10.1006/abio.1997.2136.
7. P.I. Haris, G.T. Robillard, A.A. van Dijk, D. Chapman, 'Potential of 13C and 15N Labeling for Studying Protein-Protein Interactions Using Fourier Transform-Infrared Spectroscopy', *Biochemistry*, **31**, 6279–6284 (1992). DOI: 10.1021/bi00142a016.

DOI: 10.1002/9780470027318.a9792

8. I. Noda, A.E. Dowrey, C. Marcott, G.M. Story, Y. Ozaki, 'Generalized Two-Dimensional Correlation Spectroscopy', *Appl. Spectrosc.*, **54**, 236A–248A (2000). DOI: 10.1366/0003702001950454.
9. I. Noda, Y. Ozaki, 'Principle of Two-dimensional Correlation Spectroscopy', in *Two-Dimensional Correlation Spectroscopy – Applications in Vibrational and Optical Spectroscopy*, Chapter 2, pp 15–38, John Wiley & Sons, Chichester, 2004. DOI: 10.1002/0470012404.
10. I. Noda, 'Frontiers of Two-Dimensional Correlation Spectroscopy. Part 1. New Concepts and Noteworthy Developments', *J. Mol. Struct.*, **1069**, 3–22 (2014). DOI: 10.1016/j.molstruc.2014.01.025.
11. I. Noda, 'Frontiers of Two-Dimensional Correlation Spectroscopy. Part 2. Perturbation Methods, Fields of Applications, and Types of Analytical Probes', *J. Mol. Struct.*, **1069**, 23–49 (2014). DOI: 10.1016/j.molstruc.2014.01.016.
12. Y. Park, I. Noda, Y.M. Jung, 'Novel Developments and Applications of Two-Dimensional Correlation Spectroscopy', *J. Mol. Struct.*, **1124**, 11–28 (2016). DOI: 10.1016/j.molstruc.2016.01.028.
13. Y. Park, S. Jin, I. Noda, Y.M. Jung, 'Recent Progress in Two-Dimensional Correlation Spectroscopy (2D-COS)', *J. Mol. Struct.*, **1168**, 1–21 (2018). DOI: 10.1016/j.molstruc.2018.04.099.
14. I. Noda, 'Two-Dimensional Co-distribution Spectroscopy to Determine the Sequential Order of Distributed Presence of Species', *J. Mol. Struct.*, **1069**, 50–59 (2014). DOI: 10.1016/j.molstruc.2014.01.024.
15. I. Noda, 'Techniques Useful in Two-Dimensional Correlation and Co-distribution Spectroscopy (2DCOS and 2DCDS) Analyses', *J. Mol. Struct.*, **1124**, 29–41 (2016). DOI: 10.1016/j.molstruc.2016.01.089.
16. L. Smeller, K. Heremans, '2D FT-IR Spectroscopy Analysis of the Pressure-Induced Changes in Proteins', *Vib. Spectrosc.*, **19**, 375–378 (1999). DOI: 10.1016/S0924-2031(98)00075-7.
17. W. Ocaña, B. Pastrana-Ríos, 'Calcium Titration of *Chlamydomonas reinhardtii* Centrin and its Structural Changes', *J. Molec. Struct.*, **1069**, 73–78 (2014). DOI: 10.1016/j.molstruc.2014.01.029.
18. A. Nabet, M. Pézolet, 'Two-Dimensional FT-IR Spectroscopy: A Powerful Method to Study the Secondary Structure of Proteins Using H-D Exchange', *Appl. Spectrosc.*, **51**, 466–469 (1997). DOI: 10.1366/00037029 71940701.
19. S. Meskers, J.M. Ruysschaert, E. Goormaghtigh, 'Hydrogen-Deuterium Exchange of Streptavidin and its Complex with Biotin Studied by 2D-Attenuated Total Reflection Fourier Transform Infrared Spectroscopy', *J. Am. Chem. Soc.*, **121**, 5115–5122 (1999). DOI: 10.1021/ja984208k.
20. Y. Wu, K. Murayama, Y. Ozaki, 'Two-Dimensional Infrared Spectroscopy and Principal Component Analysis Studies of the Secondary Structure and Kinetics of Hydrogen-Deuterium Exchange of Human Serum Albumin', *J. Phys. Chem. B*, **105**, 6251–6259 (2001). DOI: 10.1021/jp010225b.
21. M. Ortiz, Z. Sanoguet, C.T. McMurray, J. Salisbury, H. Haitao, W.J. Chazin, B. Pastrana-Ríos, 'Dynamics of Hydrogen/Deuterium exchange in *Chlamydomonas* Centrin', *Biochemistry*, **44**, 2409–2418 (2005). DOI: 10.1021/bi0484419.
22. I. Iloro, D. Narváez, N. Guillén, C. Camacho, L. Guillén, E. Cora, B. Pastrana-Ríos, 'Kinetics of Hydrogen/Deuterium Exchange of EGFR Ligands', *Biophys. J.*, **94**, 4041–4055 (2008). DOI: 10.1529/biophysj.107.125856.
23. B. Czarnik-Matusewicz & Y.M. Jung, 'Two-Dimensional Mid-Infrared Correlation Spectroscopy in Protein Research', in *Optical Spectroscopy and Computational Methods in Biology and Medicine* (pp.213-250). Springer, Dordrecht (2014). DOI: 10.1007/978-94-007-7832-0_8.
24. Y. Ozaki, K. Murayama, Y. Wu, B. Czarnik-Matusewicz, 'Two-Dimensional Infrared Correlation Spectroscopy Studies on Secondary Structures and Hydrogen Bondings of Side Chains of Proteins', *Spectrosc. Int. J.*, **17**, 79–100 (2003). DOI: 10.1155/2003/191875.
25. A. Rodríguez-Nassif, I. Arada, J.L. Arrondo, B. Pastrana-Ríos, 'Krr1 KH Domain Peptide Aggregation in the Presence and Absence of an Excipient', *Anal. Chem.*, **89**, 5765–5775 (2017). DOI: 10.1021/acs.analchem.6b04800.
26. B. Chon, S. Xu, Y. Jong Lee, 'Compensation of Strong Water Absorption in Infrared Spectroscopy Reveals the Secondary Structure of Proteins in dilute Solutions', *Anal. Chem.*, **93**, 2215–2225 (2021). DOI: 10.1021/acs.analchem.0c04091.
27. D.K. Graff, B. Pastrana-Ríos, S. Yu Venyaminov, F.G. Prendergast, 'Fourier Transform Infrared Spectroscopy and Two-Dimensional Infrared Analysis of Model Helical Peptides', *J. Am. Chem. Soc.*, **119**, 11282–11294 (1997). DOI: 10.1021/ja970512m.
28. C.P. Schultz, O. Bârzu, H.H. Mantsch, 'Two-Dimensional Infrared Correlation Analysis of Protein Unfolding: Use of Spectral Simulations to Validate Structural Changes During Thermal Denaturation of Bacterial CMP Kinases', *Appl. Spectrosc.*, **54**, 931–938 (2000). DOI: 10.1366/0003702001950562.
29. B. Pastrana-Ríos, 'Understanding the Mechanism of Peptide Unfolding', *Biochemistry*, **40**, 9074–9081 (2001). DOI: 10.1021/bi0155145.
30. A. Filosa, Y. Wang, A.A. Ismail, A.M. English, 'Two-Dimensional Infrared Correlation Spectroscopy as a Probe of Sequential Events in the Thermal Unfolding of Cytochromes c', *Biochemistry*, **40**, 8256–8263 (2001). DOI: 10.1021/bi002710n.
31. B. Pastrana-Ríos, W. Ocaña, M. Ríos, G. Lorenzo Vargas, G. Ysa, G. Poynter, J. Tapia, J.L. Salisbury, 'Centrin:

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

Its Secondary Structure in the Presence and Absence of Cations', *Biochemistry*, **41**, 6911–6919 (2002). DOI: 10.1021/bi0157971.

32. I. Iloro, R. Chehín, F.M. Goñi, M.A. Pajares, J.L.R. Arrondo, 'Methionine Adenosyltransferase α-helix Structure Unfolds at Lower Temperatures than β-Sheets: A 2D-IR Study', *Biophys. J.*, **86**, 3951–3958 (2004). DOI: 10.1529/biophysj.103.028373.
33. A. Torrecillas, S. Corbalán-García, J.C. Gómez-Fernández, 'An Infrared Spectroscopic Study of the Secondary Structure of Protein Kinase Cα and its Thermal Denaturation', *Biochemistry*, **43**, 2332–2344 (2004). DOI: 10.1021/bi035128i.
34. Z. Sanoguet-Carrero, M. Campbell, S. Ramos, C. Seda, B. Pastrana-Ríos, 'Effects of Phosphorylation in *Chlamydomonas* Centrin Ser 167', *Calcium Bind. Prot.*, **1**, 108–114 (2006).
35. B. Pastrana-Ríos, M. Reyes, V. Meza, J. De Orbeta, A. Rodríguez Nassif, D. Narváez, A.M. Gómez, R. Almodovar, A. Díaz Casas, J. Robles, L. Irizarry, M. Campbell, M. Colón, 'Relative Stability of Centrins and its Relationship with Calcium Binding', *Biochemistry*, **52**, 1236–1248 (2013). DOI: 10.1021/bi301417z.
36. B. Pastrana-Ríos, 'Thermal Perturbation Correlation of Calcium Binding Human Centrin 3 and its Structural Changes', *J. Molec. Struct.*, **1069**, 85–88 (2014). DOI: 10.1016/j.molstruc.2013.12.085.
37. A. Díaz Casas, G. Casanova Sepúlveda, O. Sánchez Negrón, A.P. Caro Muñiz, S.R. Malavé Ramos, A.R. Cebollero López, B. Pastrana-Ríos, 'Molecular Biophysical Characterization of the Third FF Domain of *Homo sapiens* Prp40 Homolog A', *J. Molec. Struct.*, **1167**, 174–179 (2018). DOI: 10.1016/j.molstruc.2018.04.059.
38. J.L.R. Arrondo, I. Iloro, J. Aguirre, J. Castresana, F.M. Goñi, 'A Two-Dimensional IR Spectroscopic (2D-IR) Simulation of Protein Conformational Changes', *J. Spectrosc.*, **18**, 49–58 (2004). DOI: 10.1155/2004/406126.
39. T. Lefèvre, K. Arseneault, M. Pézolet, 'Study of Protein Aggregation using 2-Dimensional Correlation Infrared Spectroscopy and Spectral Simulations', *Biopolymers*, **73**, 705–715 (2004). DOI: 10.1002/bip.20010.
40. I. Iloro, B. Pastrana-Ríos, 'Simulation of FT-IR Spectra and 2D-COS Analysis for the H/D Exchange of Two Related Ligands', *J. Molec. Struct.: Special issue of 2D-COS*, **799**, 153–157 (2006). DOI: 10.1016/j.molstruc.2006.03.032.
41. B. Pastrana-Ríos, 'Simulation of FT-IR Spectra and 2D-COS Analysis for the Thermal Perturbation of Apo-Centrin', *J. Molec. Struct.: Special issue of 2D-COS*, **799**, 163–167 (2006). DOI: 10.1016/j.molstruc.2006.03.007.
42. M.J. Paquet, M. Laviolette, M. Pézolet, M. Auger, 'Two-Dimensional Infrared Correlation Spectroscopy Study of Aggregation of Cytochrome c in the Presence of Dimyristoylphosphatidylglycerol', *Biophys. J.*, **81**, 305–312 (2001). DOI: 10.1016/S0006-3495(01)75700-1.
43. Y.-B. Yan, Q. Wang, H.-W. He, X.-Y. Hu, R.-Q. Zhang, H.-M. Zhou, 'Two-Dimensional Infrared Correlation Spectroscopy Study of Sequential Events in the Heat-Induced Unfolding and Aggregation Process of Myoglobin', *Biophys. J.*, **85**, 1959–1967 (2003). DOI: 10.1016/S0006-3495(03)74623-2.
44. N. Cerdà-Costa, I. De la Arada, F.X. Avilés, J.L.R. Arrondo, S. Villegas, 'Influence of Aggregation Propensity and Stability on Amyloid Fibril Formation as Studied by FT-IR Spectroscopy and 2D-COS Analysis', *Biochemistry*, **48**, 10582–10590 (2009). DOI: 10.1021/bi900960s.
45. B. Pastrana-Ríos, L.D.V. Sosa, J. Santiago, 'TFA as Excipient Destabilizes Melittin Causing the Selective Aggregation of Melittin within the Centrin-Melittin-TFA Complex', *Struct. Dyn.*, **2**, 041711 (2015). DOI: 10.1063/1.4921219.
46. L.D.V. Sosa, E. Alfaro, J. Santiago, D. Narváez, M.C. Rosado, A. Rodríguez, A.M. Gómez, E. Schreiter, B. Pastrana-Ríos, 'The Structure, Molecular Dynamics, and Energetics of Centrin-Melittin Complex', *Proteins: Struct. Funct. Bioinf.*, **79**, 3132–3143 (2011). DOI: 10.1002/prot.23142.
47. A. Díaz Casas, W.J. Chazin, B. Pastrana-Ríos, 'Prp40 Homolog A is a Novel Centrin Target', *Biophys. J.*, **12**, 2529–2539 (2017). DOI: 10.1016/j.bpj.2017.03.042.
48. S. Nieves, B. Pastrana, 'Quantum Cascade Laser Infrared Microscopy and 2D IR Correlation Spectroscopy Improves Crystallization Screening of a Protein Complex', *Spectroscopy*, **34**, 34–39 (2019).
49. B. Pastrana, S. Nieves, W. Li, X. Liu, D. Dimitrov, 'Developability Assessment of an Isolated C_H2 Immunoglobulin Domain', *Anal. Chem.*, **93**, 1342–1351 (2021). DOI: 10.1021/acs.analchem.0c02663.
50. B. Pastrana-Ríos, J. Rodriguez-Toro, 'Method and System for Spectral Data Analysis'. (2020). US Patent No. 10,714,209 B2. U.S. Patent and Trademark Office. https://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=10714209.PN.&OS=PN/10714209&RS=PN/10714209.
51. B. Pastrana-Ríos, I. Noda, 'System and Method for Determining Deamidation and Immunogenicity of Polypeptides'. 1–89 (2019). USPTO: provisional application 62/750,022. U.S. Patent and Trademark Office.
52. T. Formolo, M. Ly, M. Levy, L. Kilpatrick, S. Lute, K. Phinney, L. Marzilli, K. Brorson, M. Boyne, D. Davis, J. Schiel, 'Determination of the NISTmAb Primary Structure', in *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 2. Biopharmaceutical Characterization: The NISTmAb Case Study*, Chapter 1, pp 1–62, ACS Symposium Series, American Chemical Society, Washington, DC, 2015. DOI: 10.1021/bk-2015-1201.ch001.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792

53. M.K. Joubert, M. Hokom, C. Eakin, L. Zhou, M. Deshpande, M. Baker, T. Goletz, B.A. Kerwin, N. Chirmule, L.O. Narhi, J. Jawa, 'Highly Aggregated Antibody Therapeutics Can Enhance the In Vitro Innate and Late-Stage T-cell Immune Response', *J. Biol. Chem.*, **287**, 25266–25279 (2012). DOI: 10.1074/jbc.M111.330902.

54. K.D. Ratanji, J.P. Derrick, R.J. Dearman, I. Kimber, 'Immunogenicity of Therapeutic Proteins: Influence of Aggregation', *J. Immunotoxicol*, **11**, 99–109 (2014). DOI: 10.3109/1547691X.2013.821564.

55. M.L. Chiu, D.R. Goulet, A. Teplyakov, G.L. Gilliland, 'Antibody Structure and Function: The Basis for Engineering Therapeutics', *Antibodies*, **8** (2019). DOI: 10.3390/antib8040055.

56. C. Nowak, A. Tiwari, H.-C. Liu, 'Asparagine Deamidation in a Complementarity Determining Region of a Recombinant Monoclonal Antibody in Complex with Antigen', *Anal. Chem.*, **90**, 6998–7003 (2018). DOI: 10.1021/acs.analchem.8b01322.

57. A.L. Pace, R.L. Wong, Y.T. Zhang, Y.H. Kao, Y.J. Wang, 'Asparagine Deamidation Dependence on Buffer Type, pH, and Temperature', *Pharm. Sci.*, **102**, 1712–1723 (2013). DOI: 10.1002/jps.23529.

58. D.L. Riggs, J.W. Silzel, Y.A. Lyon, A.S. Kang, R.R. Julian, 'Analysis of Glutamine Deamidation: Products, Pathways, and Kinetics', *Anal. Chem.*, **91**, 13032–13038 (2019). DOI: 10.1021/acs.analchem.9b03127.

59. S. Sinha, L. Zhang, S. Duan, T.D. Williams, J. Vlasak, R. Ionescu, E.M. Topp, 'Effect of Protein Structure on Deamidation Rate in the Fc Fragment of an IgG1 Monoclonal Antibody', *Protein Sci*, **18**, 1573–1584 (2009). DOI: 10.1002/pro.173.

60. X. Lu, R.P. Nobrega, H. Lynaugh, T. Jain, K. Barlow, T. Boland, A. Sivasubramanian, M. Vásquez, Y. Xu, 'Deamidation and Isomerization Liability Analysis of 131 Clinical-Stage Antibodies', *MABS*, **11**, 45–57 (2019). DOI: 10.1080/19420862.2018.1548233.

61. H.T. Wright and D.W. Urry, 'Nonenzymatic Deamidation of Asparaginyl and Glutaminyl Residues in Proteins', *Crit. Rev. Biochem. Mol. Biol.*, **26**, 1–52 (1991). DOI: 10.3109/10409239109081719.

62. H.-C. Liu, G. Gaza-Bulseco, C. Chumsae, 'Glutamine Deamidation of a Recombinant Monoclonal Antibody', *Rapid Commun Mass Spectrom.*, **22**, 4081–4088 (2008). DOI: 10.1002/rcm.3831.

63. V. Zabrouskov, X. Han, E. Welker, H. Zhai, C. Lin, K.J. van Wijk, H.A. Scheraga, F.W. McLafferty, 'Stepwise Deamidation of Ribonuclease A at Five Sites Determined by Top-Down Mass Spectrometry', *Biochemistry*, **45**, 987–992 (2006). DOI: 10.1021/bi0517584s.

64. J.F. Sydow, F. Lipsmeier, V. Larraillet, M. Hilger, B. Mautz, J.K. M. Mølhøj, J. Kuentzer, S. Klostermann, J. Schoch, H.R. Voelger, J.T. Regula, P. Cramer, A. Papadimitriou, H. Kettenberger, 'Structure-Based Prediction of Asparagine and Aspartate Degradation Sites in Antibody Variable Regions', *PLoS ONE*, **9**, e100736 (2014). DOI: 10.1371/journal.pone.0100736.

65. J.E. Schiel, A. Turner, T. Mouchahoir, K. Yandrofski, S. Telikepalli, J. King, P. DeRose, D. Ripple, K. Phinney, 'The NISTmAb Reference Material 8671 Value Assignment, Homogeneity, and Stability', *Anal. Bioanal. Chem.*, **410**, 2127–2139 (2018). DOI: 10.1007/s00216-017-0800-1.

66. Z. Hao, Q. Hong, F. Zhang, S.-L. Wu, P. Bennett, 'Current Methods for the Characterization of Posttranslational Modifications in Therapeutic Proteins Using Orbitrap Mass Spectrometry', in *Analysis Using Mass Spectrometry: Accelerating Protein Biotherapeutics from Lab to Patient*, 1st edition, eds. M.S. Lee, Q.C. Ji, John Wiley & Sons Inc, (2017). DOI: 10.1002/9781119371779.CH3.

67. T. Mouchahoir, J.E. Schiel, 'Development of an LC-MS/MS Peptide Mapping Protocol for the NISTmAb', *Anal. Bioanal. Chem.*, **410**, 2111–2126 (2018). DOI: 10.1007/s00216-018-0848-6.

68. Y.-T. Jin, Y.-Y. Yi, B. Yeung, 'Mass Spectrometric Analysis of Protein Deamidation – A Focus on Top-Down and Middle-Down Mass Spectrometry', *Methods*, **200**, 58–66 (2022). DOI: 10.1016/j.ymeth.2020.08.002.

69. D. Rathore, A. Faustino, J. Schiel, E. Pang, M. Boyne, S. Rogstad, 'The Role of Mass Spectrometry in the Characterization of Biologic Protein Products', *Exp. Rev. Proteomics*, **15**, 431–449 (2018). DOI: 10.1080/14789450.2018.1469982.

70. T. Shantha Raju, 'Deamidation of Proteins', in *Co and Post-Translational Modifications of Therapeutic Antibodies and Proteins*, Chapter 5. John Wiley & Sons, pp 39–49 (2019). DOI: 10.1002/9781119053354.ch5.

71. S.K. Sze, G. JebaMercy, S.C. Ngan, 'Profiling the 'Deamidome' of Complex Biosamples using Mixed-Mode Chromatography-Coupled Tandem Mass Spectrometry', *Methods*, **200**, 31–41 (2022). DOI: 10.1016/j.ymeth.2020.05.005.

72. W. Li, J.L. Kerwin, J. Schiel, T. Formolo, D. Davis, A. Mahan, S.A. Benchaar, 'Biopharmaceutical Characterization: The NISTmAb Case Study', in *Structural Elucidation of Posttranslational Modifications in Monoclonal Antibodies State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization*, Volume 2, ACS Symposium Series, American Chemical Society, Washington, DC, pp 119–183 (2019). DOI: 10.1021/bk-2015-1201.ch003.

73. X. Li, C. Lin, P.B. O'Connor, 'Glutamine Deamidation: Differentiation of Glutamic Acid and γ-Glutamic Acid in Peptides by Electron Capture Dissociation', *Anal. Chem.*, **82**, 3606–3615 (2010). DOI: 10.1021/ac9028467.

74. X. Cui, W. Mi, Z. Hu, X. Li, B. Meng, X. Zhao, X. Qian, T. Zhu, W. Ying, 'Global Characterization of

DOI: 10.1002/9780470027318.a9792

Modifications to the Charge Isomers of IgG Antibody', *J. Pharmaceut. Anal.*, **12**, 156–163 (2022). DOI: 10.1016/j.jpha.2020.11.006.

75. E.B. Dunkelberger, L.E. Buchanan, P. Marek, P. Cao, D.P. Raleigh, M.T. Zanni, 'Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure', *J. Am. Chem. Soc.*, **134**, 12658–12667 (2012). DOI: 10.1021/ja3039486.
76. M. Haberger, K. Bomans, K. Diepold, M. Hook, J. Gassner, T. Schlothauer, A. Zwick, C. Spick, J.F. Kepert, B. Hienz, M. Wiedmann, H. Beck, P. Metzger, M. Mølhøj, C. Knoblich, U. Grauschopf, D. Reusch, P. Bulau, 'Assessment of Chemical Modifications of Sites in the CDRs of Recombinant Antibodies Susceptibility vs. Functionality of Critical Quality Attributes', *MABS*, **6**, 327–339 (2014). DOI: 10.4161/mabs.27876.
77. P. Bults, R. Bischoff, H. Bakker, J.A. Gietema, N.C. van de Merbel, 'LC-MS/MS-Based Monitoring of In Vivo Protein Biotransformation: Quantitative Determination of Trastuzumab and Its Deamidation Products in Human Plasma', *Anal. Chem.*, **88**, 1871–1877 (2016). DOI: 10.1021/acs.analchem.5b04276.
78. L. Huang, J. Lu, V.J. Wroblewski, J.M. Beals, R.M. Riggin, 'In Vivo Deamidation Characterization of Monoclonal Antibody by LC/MS/MS', *Anal. Chem.*, **77**, 1432–1439 (2005). DOI: 10.1021/ac0494174.
79. Q. Dong, Y. Liang, X. Yan, S.P. Markey, Y.A. Mirokhin, D.V. Tchekhovskoi, T.H. Bukhari, S.E. Stein, 'The NISTmAb Tryptic Peptide Spectral Library for Monoclonal Antibody Characterization', *MABS*, **10**, 354–369 (2018). DOI: 10.1080/19420862.2018.1436921.
80. S. Millán-Martín, C. Jakes, N. Dorival-García, N. McGillicuddy, S. Carillo, A. Farrell, J. Bones. 'Comparison of Alternative Approaches to Trypsin Protein Digestion for Reproducible and Efficient Peptide Mapping Analysis of Monoclonal Antibodies', AN21782-EN 0718S (2018). https://www.separatedbyexperience.com/documents/an-21782-lc-ms-peptide-map-digest-comparison-an21782-en.pdf.
81. N.E. Robinson, Z.W. Robinson, B.R. Robinson, A.L. Robinson, J.A. Robinson, M.L. Robinson, A.B. Robinson, 'Structure-dependent Nonenzymatic Deamidation of Glutaminyl and Asparaginyl Pentapeptides', *J. Pept. Res.*, **63**, 426–436 (2004). DOI: 10.1111/j.1399-3011.2004.00151.x.
82. H. Liu, G. Gaza-Bulseco, D. Faldu, C. Chumsae, J. Sun, 'Heterogeneity of Monoclonal Antibodies', *J Pharm Sci*, **97**, 2426–2447 (2008). DOI: 10.1002/jps.21180.
83. Y. Xu, D. Wang, B. Mason, T. Rossomando, N. Li, D.-J. Liu, J.K. Cheung, W. Xu, S. Raghava, A. Katiyar, C. Nowak, T. Xiang, D.D. Dong, J. Sun, A. Beck, H. Liu, 'Structure, Heterogeneity and Developability Assessment of Therapeutic Antibodies', *MABS*, **11**, 239–264 (2019). DOI: 10.1080/19420862.2018.1553476.
84. F.J. Irudayanathan, J. Zarzar, J. Lin, S. Izadi. 'Divining Deamidation and Isomerization in Therapeutic Proteins: Effect of Neighboring Residue', (2021) bioRxiv 2021.07.26.453885. DOI: 10.1101/2021.07.26.453885.
85. M.-P. Lefranc, V. Giudicelli, P. Duroux, J. Jabado-Michaloud, G. Folch, Aouinti, E. Carillon, H. Duvergey, A. Houles, T. Paysan-Lafosse, S. Hadi-Saljoqi, G. Sasorith, G. Lefranc, S. Kossida, 'IMGT, the International ImMunoGeneTics Information System® 25 years on', *Nucleic Acids Res.*, **43**, D413–D422 (2015). DOI: 10.1093/nar/gku1056.
86. J.A. Delmar, J. Wang, S.W. Choi, J.A. Martins, J.P. Mikhail, 'Machine Learning Enables Accurate Prediction of Asparagine Deamidation Probability and Rate', *Mol. Ther. Methods Clin. Dev.*, **15**, 264–274 (2019). DOI: 10.1016/j.omtm.2019.09.008.
87. H. Qiu, R. Wei, J. Jaworski, E. Boudanova, H. Hughes, S. VanPatten, A. Lund, J. Day, Y. Zhou, T. McSherry, C.Q. Pan, R. Sendak, 'Engineering Anti-CD52 Antibody for Enhanced Deamidation Stability', *MABS*, **11**, 1266–1275 (2019). DOI: 10.1080/19420862.2019.1631117.
88. A. Beck, H. Liu, 'Macro- and micro-heterogeneity of natural and recombinant IgG antibodies', *Antibodies*, **8**, 18 (2019). DOI: 10.3390/antib8010018.
89. J. Faist, F. Capasso, D.L. Sivco, C. Sirtori, A.L. Hutchinson, A.Y. Cho, 'Quantum Cascade Laser', *Science*, **264**, 553–556 (1994). DOI: 10.1126/science.264.5158.553.
90. D.T. Gallagher, I. Karageorgos, J.W. Hudgens, C.V. Galvin, 'Data on Crystal Organization in the Structure of the Fab Fragment from the NIST Reference Antibody RM8671', *Data Brief.*, **16**, 29–36 (2017). DOI: 10.1016/J.DIB.2017.11.013.
91. D.T. Gallagher, V.G. Connor, I. Karageorgos, 'Structure of the Fc Fragment of the NIST Reference Antibody RM8671', *Acta Crystallogr. F Struct. Biol. Commun.*, **74**, 524–529 (2018). DOI: 10.1107/S2053230X18009834.
92. J. Kill, M. Bern, K. Crowell, D. Kletter, N. Bern, W. Tang, E. Carlson, and C. Becker. 'Towards a Comprehensive Bioinformatics Analysis of the NIST Reference mAb', in *State-of-the-art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*, Chapter 14 eds. J.E. Schiel, D.L. Davis, O.V. Borizov, Volume 14, American Chemical Society, 395–414, (2015). DOI: 10.1021/bk-2015-1202.ch014.

Encyclopedia of Analytical Chemistry, Online © 2006–2022 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2022 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a9792