Chemical and Physical Characterization of Therapeutic Proteins in Solution and Amorphous Solids

By

Sandipan Sinha

B.S. Pharmaceutical Technology, Jadavpur University, 2001
M.S. Industrial Pharmacy, the University of Toledo, 2003
M.S. Pharmaceutical Chemistry, the University of Kansas, 2006

Submitted to the Department of Pharmaceutical Chemistry and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dissertation Committee:

_____

Chairperson: Elizabeth M. Topp

_____

Teruna J. Siahaan

_____

Eric J. Munson

_____

Susan M. Lunte

_____

Kyle Camarda

Dissertation Defended on May 22nd, 2008

The Dissertation Committee for Sandipan Sinha certifies
that this is the approved version of the following dissertation:

Chemical and Physical Characterization of Therapeutic Proteins in Solution and Amorphous Solids

_____

Chairperson: Elizabeth M. Topp

_____

Teruna J. Siahaan

_____

Eric J. Munson

_____

Susan M. Lunte

_____

Kyle Camarda

Date approved: May 22nd , 2008

Chemical and Physical Characterization of Therapeutic Proteins in Solution and Amorphous Solids

Sandipan Sinha

The University of Kansas, 2008

The chemical and physical stability of proteins in solution and solids was addressed in this dissertation. Protein-excipient interactions in lyophilized solids were studied by hydrogen/deuterium exchange with mass spectrometry (chapter 3) while glycosylation quanitification (chapter 4) and deamidation (chapter 5) was characterized in antibodies in solution. LC/ESI-MS was the method of choice for all studies. Hydrogen/deuterium exchange study showed that the method can be used to obtain region specific information about protein-excipient interactions in solids. It was demonstrated that exchange protection did not occur uniformly along the backbone of the protein and was dependant on excipient type and protein structure. The glycosylation quanitification study demonstrated that the Fc/2 (limited proteolysis followed by reduction) method was relatively quick and accurate and showed comparable values to the standard sugar release assay. Antibody deamidation study demonstrated that secondary structure played a pivotal role in determination of the deamidation products in antibodies.

Dedicated to:
My parents
Ashoka & Salil Kumar Sinha
My brothers
Suman & Sovan Sinha
My wife
Bhaswati Sinha

Acknowledgements

Life in graduate school has been a great learning experience. It would not have been possible without the contributions of a few gifted people. First and foremost, I would like to express my gratitude towards my advisor, Dr. Elizabeth M. Topp for her encouragement and excellent mentorship. Her immense contribution to my development in graduate school as a scientist, and the latitude provided to shape my dissertation, is sincerely appreciated. I would like to thank Dr. Todd Williams, Dr. Himanshu Gadgil, Dr. Roxana Ionescu and Dr. Josef Vlasak – their guidance and insights in the research projects has been invaluable. I am especially thankful to Dr. Teruna Siahaan (reader), Dr. Eric Munson (reader), Dr Susan M. Lunte and Dr. Kyle Camarda for accepting to be members of my dissertation committee. I would also like to thank Dr. Gerald Lushington, Dr. Victor Day, Eric Gorman and Dr. Jong Gu Park for assistance with various experiments.

I express my deepest gratitude to the faculty, staff and students of KU Pharmaceutical Chemistry department for my education. It really did live up to its reputation of being one of the best in the country and I thoroughly enjoyed my time in Lawrence-especially with the Jayhawks claiming the national title for 2008. I would like to thank Dr. Ajit D'Souza, Dr. Mary Houchin and Dr. Stephanie Krogmeier – all past members of the Topp group for their support in the first few years of graduate life. Also, it has been my pleasure to have Dr. Yunsong Frank Li as a co-worker and friend. I would also like to thank the present members – Dr. Hong Zhao and Dr. Lei Zhang for their help and encouragement. Friends play a big role in keeping your spirits high and egging you to keep going – it has been no different for me. I have made good friends and the list is a long one – Rashida Banerjee, Deb Banerjee, Ramu Gopalan, Sanjibani Banerjee, Dr. Mrinal Das, Barnali Das, Sumit Majumdar, Sandeep Dhareshwar, Anurupa Shrestha,

Nadim Asrar, Gagandeep K. Somal, Diptesh Sil, Sasi Maganti, Deepti Mandava, Aparna Kher, Vinya Sankaran, Naveen Raja – I thank you all for the good times that we spent together in Lawrence. Also, my old friends Indranil Rao, Arpana Acharya, Tarak Das, Sumit Goswami, Monojit Mondol, Manash Sinha, Bitansu Biswas – they have a continuous presence in my life and I am deeply indebted to them.

Finally, I owe everything to my family. My parents, Salil Kumar and Ashoka Sinha's contribution is something that words cannot do justice to. They still patiently continue to support me in my pursuits. My brothers, Suman Sinha and Sovan Sinha have equally contributed to my career and life. Their unflinching support and faith on me has helped me get whatever little I have achieved in life. My wife, Bhaswati, has been a constant source of sustenance. Her unconditional love and encouragement has made me go the extra mile. My sisters-in-law – Paromita Sinha and Bhaswati Sinha, its been great to have them as a part of my family along with my two sweet nephew and niece, Ayush and Asmita. I would also like to express my gratitude to my parents-in law, Prithwish Dattachowdhury and Kumkum Dattchowdhury for their faith on me. Last but not the least, my brother-in-law, Atish Dattachowdhury, sister-in-law Anindita Dattachowdhury and my very cute nephew Arnab – thanks for all your support and inspiration. In a nutshell, I could not have asked for a more caring and loving family.

Table of Contents

Chapter 1

Introduction to protein characterization in solution and solids

1.1. Introduction

The objective of this dissertation is to address some key issues associated with protein chemical stability (in solution) and protein-excipient interactions in the lyophilized state. The issues addressed are – i) most common pathway of protein degradation – deamidation in the Fc portion of a recombinant human antibody (IgG), ii) one of the common forms of post translational modification – glycosylation in a recombinant human antibody and iii) analysis of protein-excipient interaction in lyophilized solids using hydrogen/deuterium exchange with mass spectrometry.

In the past decade, there has been a significant rise in number of therapeutic peptides and proteins passing through rigorous clinical trials and finally reaching the market[1]. The trend shows no signs of letting up. A major advantage of proteins compared to small molecules is that they are highly specific in their therapeutic effects able to exert these effects at low concentrations (i.e., they are potent). This aspect of protein therapeutics has made them an important area of research in the pharmaceutical industry. Protein drug products include enzyme activators and inhibitors, poly and monoclonal antibodies, interferons, interleukins and vaccines.

Even though they are highly potent, protein drugs are subject to a variety of chemical and physical degradation processes during manufacturing and storage that reduce potency, limit shelf-life and increase the potential for immunogenic side effects. The development of a stable protein formulation thus is imperative for safe administration of protein drugs, but presents serious challenges for the pharmaceutical scientist.

Physical degradation changes protein conformation or phase behavior through processes such as aggregation, denaturation, precipitation and adsorption to

surfaces. Aggregation is the most common pathway of physical aggregation observed in protein formulations. For example, the development of high concentration protein formulations that are to be administered through the subcutaneous route faces considerable aggregation and viscosity issues[2-4]. Surface adsorption is particularly important in the development of pre-filled syringe formulations[5]. Proteins have been observed to denature at the interfaces leading to loss of efficacy.

Chemical degradation produces covalent changes in particular amino acids through reactions and are generally irreversible. Common pathways include deamidation[6, 7], oxidation[8], hydrolysis[9], isomerization[10, 11] and disulfide exchange[9]. Studies presented in Chapter 3 address the effects of secondary structure on deamidation in the Fc portion of a human IgG molecule which includes an introduction on deamidation. Among other degradation reactions oxidation is another common degradation reaction and typically occurs in the side-chains of His, Met, Cys, Trp and Tyr residues.

Among the various post-translational modifications (PTMs) involving additional of another functional group in their biosynthetic pathway, glycosylation has emerged to be one of the most important[12, 13]. Glycosylated proteins generally are more stable than other aglycan forms, particularly with respect to physical stability[14-16]. Chapter 4 includes an introduction on glycosylation and deals with LC-MS methods for quantitation of different glycoforms present in a recombinant human IgG molecule.

Due to their inherent instability in solution, proteins are often formulated in their lyophilized form. Though lyophilization is intended to stabilize the protein, it can

still undergo chemical and physical degradation during lyophilization or storage in the solid state[17]. Excipients are often added in these formulations as lyo- or a cryoprotectant in an attempt to preserve protein structure during the freeze-drying process. Chapter 5 describes the characterization of protein-excipient interactions in lyophilized solids using hydrogen/deuterium exchange with ESI-MS. Also, chapter 2 is a comprehensive review of techniques currently employed to analyze protein structure in lyophilized solids including hydrogen/deuterium exchange.

Overall, the chemical and physical stability of therapeutic proteins irrespective of it being in the form of solution or solid will determine the efficacy of a protein drug. Thus characterization and analysis of intra or intermolecular interaction between molecules assumes immense importance. Thus, in this dissertation, our emphasis has been on characterizing and analyzing certain important aspects of protein instability both in solution and lyophilized solids.

1.2. Specific aims

The work presented in this dissertation addresses three Specific Aims related to the chemical and physical characterization of protein drugs in solution and in the solid state:

1.2.1. Specific Aim 1. – To determine the effect of secondary structure on deamidation in the Fc portion of an IgG (Chapter 3) – The work reported in Chapter 3 is a detailed study of deamidation at N_{382} and N_{387} of the tryptic fragment G_{369} -K_{390}, (GFYPSDIAVEWESNGQPENNYK) located in the CH3 domain of the Fc portion of a humanized IgG1 antibody. The objective of the study was to detect the deamidation sites and to elucidate the effects of secondary structure on the reaction

products and kinetics. The results demonstrate that the product profile for the tryptic fragment and the intact protein are indeed different, supporting the hypothesis that higher order structure of the protein plays a significant role in determining both deamidation kinetics and product distribution.

1.2.2. Specific Aim 2. – To compare LC and LC/MS-based methods for quantifying glycosylation in a recombinant IgG (Chapter 4) - The studies reported here compare six LC and LC/MS-based methods for quantifying glycosylation in two production lots of a IgG. The studies test the hypothesis that LC/MS-based methods provide identification and quantitation of glycoforms that is equivalent to the convenational normal phase LC-based sugar release assay. The results demonstrate that LC/MS analysis of Fc/2 fragments yields results that are statistically comparable to the sugar release assay and can be used as a complimentary technique due to its ease of sample preparation and reduced analysis time.

1.2.3. Specific Aim 3. – To develop hydrogen/deuterium exchange with +ESI/MS analysis as a method for characterizing protein conformation and excipient interactions in lyophilized solids (Chapter 5) - In previous studies by our group[18-20], HDX with tandem liquid chromatography / electrospray ionization mass spectrometry (LC/+ESI-MS) and peptic mapping were used to provide site specific information on HDX in solid samples of calmodulin. The results demonstrated that low molecular weight sugars (i.e., trehalose, sucrose) provided significant protection from exchange relative to excipient-free controls, and that this protection was exerted preferentially in the α-helical fragments of calmodulin. The studies presented here extend this method to other model proteins having differing

secondary structure and in the presence of excipients with differing size and H-bond donor and/or acceptor capacities. The studies test the hypothesis that the ability of excipients to protect proteins from HDX in amorphous solids depends on both excipient type and protein structure, and that this effect is exerted non-uniformly along the protein sequence.

1.3. References

1. PhRMA, 418 biotechnology medicines in testing promise to bolster the arsenal against disease. *Pharmaceutical Research and Manufacturers of America: Washington, DC* **2006**, 52.
2. Alford, J. R.; Kendrick, B. S.; Carpenter, J. F.; Randolph, T. W., High concentration formulations of recombinant human interleukin-1 receptor antagonist: II. aggregation kinetics. *J Pharm Sci* **2007**.
3. Alford, J. R.; Kwok, S. C.; Roberts, J. N.; Wuttke, D. S.; Kendrick, B. S.; Carpenter, J. F.; Randolph, T. W., High concentration formulations of recombinant human interleukin-1 receptor antagonist: I. physical characterization. *J Pharm Sci* **2007**.
4. Harn, N.; Allan, C.; Oliver, C.; Middaugh, C. R., Highly concentrated monoclonal antibody solutions: direct analysis of physical structure and thermal stability. *J Pharm Sci* **2007,** 96, (3), 532-46.
5. Deechongkit, S.; Aoki, K. H.; Park, S. S.; Kerwin, B. A., Biophysical comparability of the same protein from different manufacturers: a case study using Epoetin alfa from Epogen and Eprex. *J Pharm Sci* **2006,** 95, (9), 1931-43.
6. Robinson, N. E.; Robinson, A. B., Molecular clocks. *Proc Natl Acad Sci U S A* **2001,** 98, (3), 944-9.
7. Robinson, N. E.; Robinson, A. B., Deamidation of human proteins. *Proc Natl Acad Sci U S A* **2001,** 98, (22), 12409-13.

8. Lam, X. M.; Yang, J. Y.; Cleland, J. L., Antioxidants for prevention of methionine oxidation in recombinant monoclonal antibody HER2. *J Pharm Sci* **1997,** 86, (11), 1250-5.

9. Shahrokh, Z.; Eberlein, G.; Buckley, D.; Paranandi, M. V.; Aswad, D. W.; Stratton, P.; Mischak, R.; Wang, Y. J., Major degradation products of basic fibroblast growth factor: detection of succinimide and iso-aspartate in place of aspartate. *Pharm Res* **1994,** 11, (7), 936-44.

10. Capasso, S.; Di Cerbo, P., Kinetic and thermodynamic control of the relative yield of the deamidation of asparagine and isomerization of aspartic acid residues. *J Pept Res* **2000,** 56, (6), 382-7.

11. Geiger, T.; Clarke, S., Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J Biol Chem* **1987,** 262, (2), 785-94.

12. Delorme, E.; Lorenzini, T.; Giffin, J.; Martin, F.; Jacobsen, F.; Boone, T.; Elliott, S., Role of glycosylation on the secretion and biological activity of erythropoietin. *Biochemistry* **1992,** 31, (41), 9871-6.

13. Keusch, J.; Lydyard, P. M.; Delves, P. J., The effect on IgG glycosylation of altering beta1, 4-galactosyltransferase-1 activity in B cells. *Glycobiology* **1998,** 8, (12), 1215-20.

14. Krapp, S.; Mimura, Y.; Jefferis, R.; Huber, R.; Sondermann, P., Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* **2003,** 325, (5), 979-89.

15. Raju, T. S.; Scallon, B. J., Glycosylation in the Fc domain of IgG increases resistance to proteolytic cleavage by papain. *Biochem Biophys Res Commun* **2006,** 341, (3), 797-803.

16. Wright, A.; Morrison, S. L., Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol* **1997,** 15, (1), 26-32.

17. Carpenter, J. F.; Crowe, J. H., An infrared spectroscopic study of the interactions of carbohydrates with dried proteins. *Biochemistry* **1989,** 28, (9), 3916-22.

18. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Characterizing protein structure in amorphous solids using hydrogen/deuterium exchange with mass spectrometry. *Anal Biochem* **2007,** 366, (1), 18-28.

19. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Trehalose and calcium exert site-specific effects on calmodulin conformation in amorphous solids. *Biotechnol Bioeng* **2007,** 97, (6), 1650-3.

20. Li, Y.; Williams, T. D.; Topp, E. M., Effects of Excipients on Protein Conformation in Lyophilized Solids by Hydrogen/Deuterium Exchange Mass Spectrometry. *Pharm Res* **2007**.

Chapter 2

Methods for assessing protein structure in lyophilized solids

2.1. Introduction

Protein drugs have seen unprecedented growth over the past decade. These complex and labile molecules are subject to a variety of physical and chemical degradation processes during manufacturing and storage which can lead to loss of activity, severe side effects and immunogenicity. Development of a stable protein formulation thus is imperative for safe administration of protein drugs, but presents serious challenges for the pharmaceutical scientist. To stabilize the protein to meet the long term shelf-life requirement, many protein drug products are formulated in lyophilized (i.e., freeze-dried) forms. Other techniques of drying are also available (e.g., spray drying) but lyophilization is still the method of choice.

Though lyophilization is intended to stabilize the protein, the protein can still undergo chemical and physical degradation during the lyophilization process or during storage in the solid state. The physical degradation changes can be reversible on reconstitution, but often are not. Chemical degradation reactions such as deamidation, oxidation, and the Maillard reaction occur in the solid state, leading to net changes in the chemical composition of the protein. It is generally accepted that the native structure of the protein has to be maintained in the solid state for the protein to maintain potency and minimize any adverse reactions (e.g., immune response). Excipients are generally added to the formulations to keep the protein in its native form by providing protection during freeze drying and storage. To assess the stability of solid protein formulations, it is imperative to have analytical techniques capable of determining protein structure in the solid state. Although there are many techniques available for studying protein structure in solutions, far fewer are available for solids. The techniques available do not provide useful/interpretable

information in many cases. This chapter presents a review of techniques available with their advantages and limitations, together with examples of their applications. The techniques described here are Fourier Transform Infrared Spectrocopy (FTIR, Section 2.2), solid-state Nuclear Magnetic Resonance (ssNMR, Section 2.3), Fourier Transform Raman Spectroscopy (FT-Raman, Section 2.4) and hydrogen/deuterium exchange with mass spectrometry (HDX, Section 2.5).

2.2. Fourier Transform Infrared Spectroscopy (FTIR)

FTIR has emerged over the last decade as one of the most important technique for analyzing protein structure in solid formulation. Many research groups have employed FTIR to study protein structure in amorphous solid formulations[1-8]. The technique's greatest advantage is that it is not limited by the physical form of the sample, and can be effectively used for to dispersions[9, 10] and hydrated solids[11] as well as solutions.

FTIR measures the light energy absorbed by a protein molecule in the infrared region between transitions of vibrational modes and therefore is characteristic of bonds or bond types. Only three of a possible nine interpretable absorption bands[12] are used for present day analysis. The Amide I (1600 - 1700 cm^{-1}), Amide II (1500 - 1600 cm^{-1}) and Amide III (1220 - 1330 cm^{-1}) bands are highly sensitive to changes in the protein backbone. The Amide I band is directly correlated to the protein backbone conformation. It arises due to C=O stretching vibrations with minor contributions from C-N stretching in the region of 1600 to 1700 cm^{-1}. Secondary structure measurements of proteins in both solutions and solids are generally performed by deconvolution of the Amide I band. The Amide II band arises

due to N-H bending coupled with C-N stretching vibrations and usually appears between 1500 and 1600 cm^{-1}. Since it involves N-H bending, the Amide II band has been used to characterize protein secondary structure by employing hydrogen/deuterium exchange experiments. On exchange of the imide hydrogen with deuterium, the band shifts to 1450 cm^{-1}, thus enabling quantitation of H / D exchange kinetics. The Amide III band is comprised of a number of coordinated vibrations and is reflected in the 1220 to 1330 cm^{-1} region. Though it has been used to quantitate secondary structure, it has failed to gain importance due to the low signal quality[13-16].

There are three distinct FTIR sampling methods that are generally used for amorphous solid samples, each with its own advantages: KBr pellet, Diffused Reflectance (DRIFTS), and Attenuated Total Reflectance (ATR). The KBr pellet technique involves the application of force to form thin pellets of the protein-KBr mixture which are then introduced into a sample holder for transmission mode measurement. Although this technique has a long history, there is some concern regarding possible changes in protein secondary structure by the pressure applied during pellet formation[17]. However, other reports claim that KBr pellet formation has little effect on protein structure based on comparisons with other techniques like IR microscopy[18-20]. ATR has been has been used extensively in solution studies but less widely for solids. ATR is purported to have minimal sample preparation and does not affect the secondary structure of proteins. The method involves spreading the sample over a crystal made of ZnSe or Ge or diamond and then applying enough pressure to ensure close contact of the sample with the crystal surface. The infrared beam enters one end of the crystal and is totally reflected multiple times because of

the high reflective index of the crystal. At the total reflection point, the IR beam is able to penetrate into the sample for ~0.5 mm; this penetrating light is called the evanescent wave. The sample absorbs the evanescent wave at this point. The beam exiting the crystal is recorded as the sample absorption and provides structural information. DRIFTS has emerged has an excellent alternative to KBr pellet and ATR techniques, especially for powders that are not easily spread. DRIFTS has been used by many groups to study the effect of excipients on secondary structure in the amorphous state. Van de Weert et al.[20] have demonstrated that sampling techniques had a minor effect on the spectrum and therefore all methods are acceptable.

2.3. Solid State Nuclear Magnetic Resonance Spectroscopy (ssNMR)

ssNMR is fast emerging as a technique for characterizing macromolecular formulations as more and more lyophilized protein formulations are being developed in the industry. Previously, ssNMR has been used almost exclusively for characterizing small molecules.

In solution NMR, small molecules are tumbling rapidly, which serves to average many of the interactions that occur. For example, only the isotropic value of the chemical shift is observed, rather than full chemical shift tensors. In ssNMR, the molecules are no longer tumbling rapidly, and therefore are characterized by broad peaks which can be attributed to the strong coupling interactions and chemical shift anisotropy (CSA) in the solid. These problems are being overcome for dilute nuclei such as ^{13}C by applying high-power ^{1}H decoupling and by spinning the sample at an angle of 54^{o} 44' between the spin axis and the direction of the magnetic field, the so-

called “magic angle”. Through cross polarization (CP) of magnetization from protons to carbons (^{13}C), the ^{13}C sensitivity can be enhanced by up to a factor of four.

For biological systems, ssNMR applications have been primarily used for the elucidation of the structure of proteins with low water solubility such as membrane proteins and beta amyloid fibrils[21-33] since these systems cannot be studied by conventional solution NMR or X-ray crystallography. In lyophilized powders, the molecular orientation is more random, making it more difficult to obtain protein structure information. Although no full structure elucidation has been reported for a lyophilized protein using ssNMR, the method can provide useful structural information in solid formulations[34]. Though limited in its applications for lyophilized formulations it has been used for secondary structure measurement, chemical reactivity and measurement of molecular mobility of both protein and excipient in amorphous solids[35-41].

2.4. Near Infrared Spectroscopy (NIR)

NIR has gained popularity as a simple, rapid, noninvasive and nondestructive technique that allows protein structure monitoring, and has found widespread acceptance in the food[42] and pharmaceutical industry[43-46]. Since NIR provides information on both physical and chemical properties of the sample, it has been used for analysis of residual moisture content[47-49], crystallinity[50, 51] and molecular interactions in solids[52, 53].

The greatest advantage of NIR is the need for little or no sample preparation: samples in any glass vial can be subjected to NIR to yield reproducible information on protein structure in less than a minute. This can be attributed to the low molar

absorptivity of NIR bands which facilitates operation in reflectance mode[46, 51, 54]. Wavenumbers of 13300-4000 cm^{-1} are generally referred to as the NIR region. NIR bands are overtone and combination bands of absorption bands involving C-H, N-H and O-H bonds and thus can be used for investigating protein conformation in the solid state. These bands are the outcome of forbidden transitions from the ground vibrational energy levels and are therefore several fold less intense than their corresponding mid-IR absorption bands[55]. Higher wavenumber regions (5000 - 13000 cm^{-1}) are assigned mainly to the overtone bands while the lower wavenumber regions (4000 – 5000 cm^{-1}) are assigned to combination bands. It has been proposed that protein structure is best elucidated from the 4000 – 5000 cm^{-1} region[56-59]. NIR studies of lyophilized proteins have shown certain bands at 4369 and 4604 cm^{-1} corresponding to α-helix and at 4323, 4417, 4525-4535 cm^{-1} corresponding to β-sheet domains [56, 60]. A recent report by Bai et. al[61] suggested the possibility of distinguishing proteins with respect to their secondary structure and also the protective effects of sucrose in amorphous solids using NIR. Izutsu et al.[62] demonstrated the effect of freeze drying on secondary structure using seven proteins having different secondary structures and comparing their NIR spectra in solution and in the solid state. The authors showed a slight reduction in α-helical structure with concomitant rise in β-sheet structure upon freeze drying and also at elevated temperatures. β -sheet proteins did not show any effect.

NIR may be most useful technique in a manufacturing or production setting that demands ready analysis of the secondary structure and other physico-chemical properties of amorphous solids[63]. Further development in data analysis procedures would improve the quality of information obtained from NIR spectra.

2.5. Fourier Transform Raman Spectroscopy (FT-Raman)

Raman spectroscopy has emerged as an important tool to monitor secondary structure of proteins. Like FTIR, FT-Raman has the flexibility to handle samples with different physical forms, and has been used to study protein conformation in aqueous solution[64-66], organic solvents[67] and solids[68-70]. The method has also been used to study polymers, ceramics, semiconductors and other biological molecules like DNA and carbohydrates.

Raman spectroscopy basically involves the inelastic scattering or Raman scattering of monochromatic light usually from a laser beam in the visible, near infrared or in the near UV region. It measures the frequency differences in scattered light between the ground and excited vibrational energy levels of a molecule on exposure to a beam of light. This in turn gives rise to a characteristic set of bands or Raman spectra for that particular chemical entity. It involves the excitation of a molecule to virtual energy states by an incident photon and then emitting a photon during relaxation to a vibrational excited state. Three types of transitions are possible. Raman scattering arises when the energy of the emitted photon is less than that of the incident photon indicating that the molecule has gained energy during this process to an excited state (Fig. 2.1). This is defined as the Stokes Raman scattering. If the energy of the emitted photon is greater than that of the incident photon, indicating gain of energy by the photon from the molecule in an excited vibrational state while it falls back to its ground state, then it is defined as the anti-Stokes Raman scattering. Stokes Raman scattering is generally measured in conventional Raman spectrometers. The third transition involves same energy for

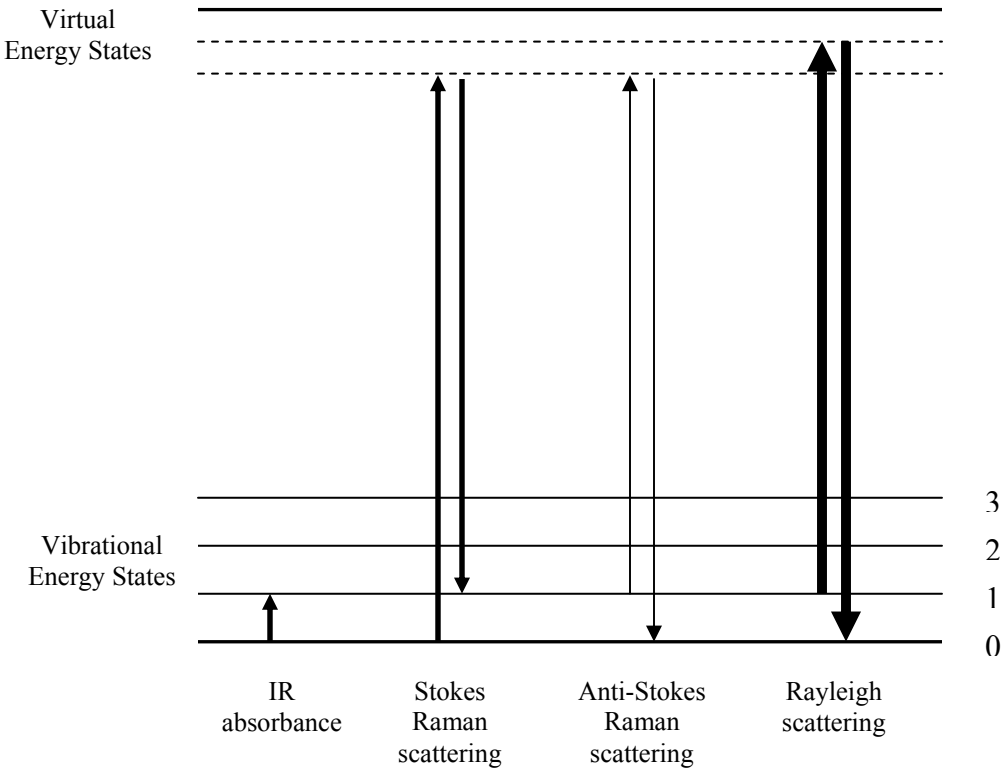


Figure 2.1. Energy diagram for different vibrational transitions possible. Thickness of lines is approximately proportional to the signal strength of each transition.

both the incident and emitted photon and is stronger than the previous two transitions. It is known as Rayleigh scattering and is a major obstacle to obtaining meaningful Raman data. For FT-Raman, a 1064 nm near infrared laser is generally used as the source, and has found widespread application as it can completely block Rayleigh scattering and selectively allow Raman scattering for analysis. As in FTIR, the secondary structure of proteins in FT-Raman is analyzed mainly by Amide I bands (C=O stretching) which are sensitive to protein conformation. Typically, α-helical structures have the Amide I band at 1655 cm^{-1} while β-sheet structures appear at 1670 cm^{-1} and random coil are generally observed at approximately 1640 cm^{-1}. Research has revealed that with UV resonance Raman, the Amide III band is also very sensitive to protein secondary structure and could be used for analysis[64].

Raman spectroscopy enjoys a few advantages over FTIR but is necessarily a complementary vibrational spectroscopy technique rather than a replacement for FTIR. An advantage of Raman spectroscopy is that the water signal in Raman is weak, particularly in comparison to its marked IR absorbance which can interfere with analysis, especially in solution. Raman spectroscopy also provides information on highly polarizable groups having symmetrical vibrational modes, which may or may not be IR active. This attribute of Raman spectroscopy provides information on the disposition of protein side-chains e.g., whether they are hydrogen bonded or if the S-H bonds are free. Raman also has the advantage of using light of any single wavelength from deep UV to near infra-red, thus offering great selectivity. A disadvantage of the technique is that Raman scattering is a weak physical phenomenon and therefore signal intensity in terms of quantum yield is low. As a result, a relatively concentrated solution (1-5 mg/ml) or solid is required to obtain a

good quality spectrum. An excellent review by Wen[71] lists recent advances of Raman spectroscopy along with their applications.

FT Raman has been used to analyze lyophilized proteins for the past three decades. The literature primarily addresses the change in secondary structure of proteins on lyophilization[69, 70, 72, 73]. Sane et. al.[68] examined the extent of protection provided by sugars for a therapeutic antibody and also compared spray drying to freeze drying by Raman Spectroscopy. The antibody was found to lose some β-sheet structure with a corresponding gain in the turn and unordered content in the absence of sugars. Trehalose and sucrose were found to be equally protective of the predominantly β-sheet structure of the antibody, with the degree of protection increasing with higher sugar:protein molar ratios. Histidine had a similar effect, but was not as protective as the sugars. Aggregation was also reported to be reduced in presence of sugars and occurred more rapidly in spray dried formulations than in lyophilized formulations. Tattini et. al.[73] studied the PEGylation of bovine serum albumin (BSA) and found out that a ratio of 1:0.25 (BSA:PEG) provided the best maintenance of the secondary structure. Miller et. al.[74], studied the effect of photolysis on reactive functional groups and/or reaction products of lyophilized recombinant bovine somatotropin with Raman spectroscopy. They observed a significant loss of intensity in the 400-550 cm^{-1} and 650-725 cm^{-1} regions, indicating modifications of the Cys disulfide groups. The data was supported by a photolysis-dependent increase of intensity between 1060-1125 cm^{-1}, indicating the presence of S=O. These results demonstrate that Raman spectroscopy can be used as a complementary technique to FTIR for protein secondary structure characterization.

2.6. Hydrogen / Deuterium Exchange (H/D Exchange)

Hydrogen / deuterium exchange has long been used to analyze protein structure[75], protein folding-unfolding mechanisms[76-79] and protein ligand interactions[80, 81] in solution. Traditionally H/D exchange was used in conjunction with NMR analysis,[82] but in recent years H/D exchange with mass spectrometry has emerged as a powerful technique capable of investigating protein structure[83-85] and dynamics under different conditions. H/D exchange with mass spectrometry has certain advantages over exchange studies with NMR, in that mass spectrometry: i) enables superior sensitivity (ESI, MALDI), ii) requires minute quantities of protein for analysis while NMR is performed in high protein concentration, and iii) is adaptable to large proteins and their complexes. NMR offers the advantage of providing information at the amino acid level i.e., determining exchange rates at specified amide linkages, and therefore affords greater spatial resolution than mass spectrometry.

H/D exchange is based on the rate at which the protein backbone amide hydrogens undergo exchange with deuterons when they are incubated in the presence of D_2O vapor or solution. The amide hydrogens are located at every amino acid along the backbone and are involved in numerous hydrogen bonds that stabilize the secondary and tertiary structure. It usually takes 1-10 s for an amide hydrogen to exchange with deuterium when exposed to D_2O solution. For a structured protein, the exchange rate will depend on the solvent accessibility of that particular amide hydrogen (i.e., whether that particular residue is buried or surface exposed) and also on its involvement in intramolecular hydrogen bonding [86]. Typically, an H/D exchange experiment involves exposing the protein sample to a D_2O environment (vapor or

solution) for predetermined time intervals followed by withdrawal of samples and quenching by a low pH buffer (~ 2.4) and low temperature (~0^{o}C) to arrest back exchange with non-deuterated solvent[87]. Intact protein analysis can be performed to monitor the total uptake of deuterium. Alternatively, the protein can be subjected to digestion with an acid-stable protease (usually pepsin) and deuterium uptake by different regions of the protein determined by measuring the increase in mass by each fragment with LC-ESI MS.

2.6.1. H/D exchange in lyophilized solids. - Lyophilization has been reported to have adverse effect on protein structure and can lead to aggregation and denaturation, causing increased chemical degradation[88] and immunogenicity[89]. H/D exchange with mass spectrometry offers considerable promise for providing detailed and region specific information on protein structure and interactions in amorphous solids. When applied to solids, the technique involves a procedure similar to the well-established methods used in solution, except that the lyophilized powder is exposed to D_2O vapor at controlled relative humidity followed by reconstitution of the sample before analysis by LC-ESI-MS.

In recent years, there have been reports of the use of H/D exchange methods for monitoring protein structure and protein-excipient interactions in amorphous solids with NMR[90], FTIR[91] and mass spectrometry being the analytical techniques of choice[92-94]. Desai et. al[90] employed H/D exchange technique with NMR analysis to study and confirm the unfolding of bovine pancreatic trypsin inhibitor (BPTI) on lyophilization which led to its loss of activity when reconstituted in organic solvents. Six aromatic residues (three buried and three on the surface of the protein) which were well resolved in the ^{1}H NMR spectrum were used for analysis. The authors

reasoned that if the protein unfolds, some of these buried protons would become exposed and exchange with the D_2O in the vapor phase. They observed faster exchange for certain NH protons in the lyophilized form compared to solution, and found that on addition of sorbitol (lyoprotectant), no exchange occurred. They argued that the sorbitol promotes the maintenance of native structure on lyophilization, while the protein unfolds when it is lyophilized without this excipient.

French et. al.,[91] used the isotopic shifts for the Amide II/II' band in the FTIR spectra to characterize the solvent exposure of proteins and trehalose in spray-dried powders. Human granulocyte colony stimulating factor (rhG-CSF) and recombinant consensus interferon-α (rConIFN) were studied in the presence of different weight ratios of trehalose at 33 % and 75% RH. At 33% RH, the H/D exchange was found to decrease with increasing percent of trehalose in the formulation, suggesting increased protection of the backbone amide bonds due to hydrogen bonding between the sugar and the protein. At 75% RH, H/D exchange increased for both proteins with trehalose having a mild protective effect. The exchange data correlated well with conformational changes that were observed at the two RH values by conventional FTIR. H/D exchange rates were also monitored for trehalose. It was observed that the trehalose exchange rate was not affected by RH indicating that the protons in the sugar are highly protected in the powder.

More recently, Li et al., [92, 93] used LC ESI-MS to analyze H/D exchange in amorphous solid samples. Calmodulin, a calcium binding protein with predominantly α-helical structure, was used as a model protein with trehalose and calcium chloride as the stabilizing excipients. Pepsin digestion was performed to dissect the effect of trehalose and calcium on different regions of the protein. At 33%RH the stabilizing

effects of the excipients were not exerted uniformly along the protein backbone. The protective effect of calcium chloride was primarily observed in the calcium binding loops, while that of trehalose was observed in the α-helices. At 75% RH, trehalose was not able to reduce the exchange rate in any of the fragments. FTIR data did not show any appreciable perturbation of the secondary structure in any of the formulations.

Li et al.[94] also studied calmodulin-excipient interactions in the solid state, using trehalose, raffinose, sucrose, dextran, mannitol and guanidine hydrochloride (negative control) as excipients. Trehalose, raffinose and sucrose showed the greatest protection from exchange, with the effect primarily exerted on the α-helical portions. Dextran also showed mild protection. Mannitol showed no protection, which was attributed to its recrystallization on storage.

2.6. Conclusions

This chapter has reviewed five common methods for assessing protein structure in lyophilized solids: FTIR, ssNMR, NIR, FT-Raman and H/D exchange. In the pharmaceutical and biotechnology industry, FTIR remains the most widely used method for characterizing protein secondary structure changes and in lyophilized solids. With further development, emerging techniques like ssNMR and H/D exchange have the potential to provide detailed information about protein-excipient interactions at the peptide or amino acid level. NIR offers the advantage of minimal sample preparation and rapid analysis time, and is finding increasing application in manufacturing and quality control. Though not discussed here, recent studies have employed solid state fluorescence of tryptophan residues[95, 96] to analyze protein

tertiary structure, an approach that may add an important dimension to solid state analysis. Further evolution of these techniques may ultimately make it possible to determine exact interactions between protein and excipients in amorphous solids, facilitating the development of stable well-characterized formulations.

2.7 References

1. Griebenow, K.; Klibanov, A. M., Lyophilization-induced reversible changes in the secondary structure of proteins. *Proc Natl Acad Sci U S A* **1995,** 92, (24), 10969-76.
2. Costantino, H. R.; Carrasquillo, K. G.; Cordero, R. A.; Mumenthaler, M.; Hsu, C. C.; Griebenow, K., Effect of excipients on the stability and structure of lyophilized recombinant human growth hormone. *J Pharm Sci* **1998,** 87, (11), 1412-20.
3. Garzon-Rodriguez, W.; Koval, R. L.; Chongprasert, S.; Krishnan, S.; Randolph, T. W.; Warne, N. W.; Carpenter, J. F., Optimizing storage stability of lyophilized recombinant human interleukin-11 with disaccharide/hydroxyethyl starch mixtures. *J Pharm Sci* **2004,** 93, (3), 684-96.
4. Imamura, K.; Ogawa, T.; Sakiyama, T.; Nakanishi, K., Effects of types of sugar on the stabilization of protein in the dried state. *J Pharm Sci* **2003,** 92, (2), 266-74.
5. Andya, J. D.; Hsu, C. C.; Shire, S. J., Mechanisms of aggregate formation and carbohydrate excipient stabilization of lyophilized humanized monoclonal antibody formulations. *AAPS PharmSci* **2003,** 5, (2), E10.
6. Souillac, P. O.; Middaugh, C. R.; Rytting, J. H., Investigation of protein/carbohydrate interactions in the dried state. 2. Diffuse reflectance FTIR studies. *Int J Pharm* **2002,** 235, (1-2), 207-18.

7. Lu, J.; Wang, Y.-X.; Ching, C.-B., Thermal and FTIR investigation of freeze-dried protein-excipient mixtures. *Journal of Thermal Analysis and Calorimetry* **2007,** 89, 913-919.

8. Izutsu, K.; Aoyagi, N.; Kojima, S., Protection of protein secondary structure by saccharides of different molecular weights during freeze-drying. *Chem Pharm Bull (Tokyo)* **2004,** 52, (2), 199-203.

9. Meng, G. T.; Ma, C. Y., Fourier-transform infrared spectroscopic study of globulin from Phaseolus angularis (red bean). *Int J Biol Macromol* **2001,** 29, (4-5), 287-94.

10. Ma, C. Y.; Rout, M. K.; Mock, W. Y., Study of oat globulin conformation by Fourier transform infrared spectroscopy. *J Agric Food Chem* **2001,** 49, (7), 3328-34.

11. Wellner, N.; Belton, P. S.; Tatham, A. S., Fourier transform IR spectroscopic study of hydration-induced structure changes in the solid state of omega-gliadins. *Biochem J* **1996,** 319 (Pt 3), 741-7.

12. Cooper, E. A.; Knutson, K., Fourier transform infrared spectroscopy investigations of protein structure. *Pharm Biotechnol* **1995,** 7, 101-43.

13. Cai, S.; Singh, B. R., A distinct utility of the amide III infrared band for secondary structure estimation of aqueous protein solutions using partial least squares methods. *Biochemistry* **2004,** 43, (9), 2541-9.

14. Carrasquillo, K. G.; Sanchez, C.; Griebenow, K., Relationship between conformational stability and lyophilization-induced structural changes in chymotrypsin. *Biotechnol Appl Biochem* **2000,** 31 (Pt 1), 41-53.

15. Vecchio, G.; Zambianchi, F.; Zacchetti, P.; Secundo, F.; Carrea, G., Fourier-transform infrared spectroscopy study of dehydrated lipases from candida antarctica B and pseudomonas cepacia. *Biotechnol Bioeng* **1999,** 64, (5), 545-51.

16. Costantino, H. R.; Nguyen, T. H.; Hsu, C. C., Fourier-transform infrared spectroscopy demonstrates that lyophilization alters the secondary structure of recombinant human growth hormone. *Pharmaceutical Sciences* **1996,** 2, (5), 229-232.

17. Chan, H. K.; Ongpipattanakul, B.; Au-Yeung, J., Aggregation of rhDNase occurred during the compression of KBr pellets used for FTIR spectroscopy. *Pharm Res* **1996,** 13, (2), 238-42.

18. Kendrick, B. S.; Dong, A.; Allison, S. D.; Manning, M. C.; Carpenter, J. F., Quantitation of the area of overlap between second-derivative amide I infrared spectra to determine the structural similarity of a protein in different states. *J Pharm Sci* **1996,** 85, (2), 155-8.

19. Costantino, H. R.; Andya, J. D.; Shire, S. J.; Hsu, C. C., Fourier-transform infrared spectroscopic analysis of the secondary structure of recombinant humanized immunoglobulin G. *Pharmaceutical Sciences* **1997,** 3, (3), 121-128.

20. van de Weert, M.; Haris, P. I.; Hennink, W. E.; Crommelin, D. J., Fourier transform infrared spectrometric analysis of protein conformation: effect of sampling method and stress factors. *Anal Biochem* **2001,** 297, (2), 160-9.

21. Makin, O. S.; Serpell, L. C., Structures for amyloid fibrils. *Febs J* **2005,** 272, (23), 5950-61.

22. Lorigan, G. A., Solid-state nuclear magnetic resonance spectroscopic studies of magnetically aligned phospholipid bilayers. *NMR Spectroscopy of Biological Solids* **2006**, 237-259.

23. Hong, M.; Wi, S., Torsion angle determination in biological solids by solid-state nuclear magnetic resonance. *NMR Spectroscopy of Biological Solids* **2006**, 87-122.

24. Yamada, K.; Shimizu, T.; Yoshida, M.; Asanuma, M.; Tansho, M.; Nemoto, T.; Yamazaki, T.; Hirota, H., Solid-state 17O NMR study of small biological compounds. *Zeitschrift fuer Naturforschung, B: Chemical Sciences* **2007,** 62, (11).

25. Gong, X. M.; Franzin, C. M.; Thai, K.; Yu, J.; Marassi, F. M., Nuclear magnetic resonance structural studies of membrane proteins in micelles and bilayers. *Methods Mol Biol* **2007,** 400, 515-29.

26. Baxa, U.; Wickner, R. B.; Steven, A. C.; Anderson, D. E.; Marekov, L. N.; Yau, W. M.; Tycko, R., Characterization of beta-sheet structure in Ure2p1-89 yeast prion fibrils by solid-state nuclear magnetic resonance. *Biochemistry* **2007,** 46, (45), 13149-62.

27. De Angelis, A. A.; Opella, S. J., Bicelle samples for solid-state NMR of membrane proteins. *Nat Protoc* **2007,** 2, (10), 2332-8.

28. Curtis-Fisk, J.; Preston, C.; Zheng, Z.; Worden, R. M.; Weliky, D. P., Solid-state NMR structural measurements on the membrane-associated influenza fusion protein ectodomain. *J Am Chem Soc* **2007,** 129, (37), 11320-1.

29. Luo, W.; Mani, R.; Hong, M., Side-chain conformation of the M2 transmembrane peptide proton channel of influenza a virus from 19F solid-state NMR. *J Phys Chem B* **2007,** 111, (36), 10825-32.

30. Naito, A.; Kawamura, I., Solid-state NMR as a method to reveal structure and membrane-interaction of amyloidogenic proteins and peptides. *Biochim Biophys Acta* **2007,** 1768, (8), 1900-12.

31. Hong, M., Structure, topology, and dynamics of membrane peptides and proteins from solid-state NMR spectroscopy. *J Phys Chem B* **2007,** 111, (35), 10340-51.

32. Tycko, R., Characterization of amyloid structures at the molecular level by solid state nuclear magnetic resonance spectroscopy. *Methods Enzymol* **2006,** 413, 103-22.

33. Aisenbrey, C.; Bertani, P.; Henklein, P.; Bechinger, B., Structure, dynamics and topology of membrane polypeptides by oriented 2H solid-state NMR spectroscopy. *Eur Biophys J* **2007,** 36, (4-5), 451-60.

34. McDermott, A. E., Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. *Curr Opin Struct Biol* **2004,** 14, (5), 554-61.

35. Yoshioka, S.; Miyazaki, T.; Aso, Y.; Kawanishi, T., Significance of local mobility in aggregation of beta-galactosidase lyophilized with trehalose, sucrose or stachyose. *Pharm Res* **2007,** 24, (9), 1660-7.

36. Yoshioka, S.; Miyazaki, T.; Aso, Y., Beta-relaxation of insulin molecule in lyophilized formulations containing trehalose or dextran as a determinant of chemical reactivity. *Pharm Res* **2006,** 23, (5), 961-6.

37. Lam, Y. H.; Bustami, R.; Phan, T.; Chan, H. K.; Separovic, F., A solid-state NMR study of protein mobility in lyophilized protein-sugar powders. *J Pharm Sci* **2002,** 91, (4), 943-51.

38. Yoshioka, S.; Aso, Y.; Kojima, S.; Sakurai, S.; Fujiwara, T.; Akutsu, H., Molecular mobility of protein in lyophilized formulations linked to the molecular mobility of polymer excipients, as determined by high resolution 13C solid-state NMR. *Pharm Res* **1999,** 16, (10), 1621-5.

39. Yoshioka, S.; Aso, Y.; Kojima, S., Determination of molecular mobility of lyophilized bovine serum albumin and gamma-globulin by solid-state 1H NMR and relation to aggregation-susceptibility. *Pharm Res* **1996,** 13, (6), 926-30.

40. Tishmack, P. A.; Bugay, D. E.; Byrn, S. R., Solid-state nuclear magnetic resonance spectroscopy--pharmaceutical applications. *J Pharm Sci* **2003,** 92, (3), 441-74.

41. Yoshioka, S., Molecular mobility of freeze-dried formulations as determined by NMR relaxation, and its effect on storage stability. *Drugs and the Pharmaceutical Sciences* **2004,** 137 187-212.

42. Alomar, D.; Gallo, C.; Castaneda, M.; Fuchslocher, R., Chemical and discriminant analysis of bovine meat by near infrared reflectance spectroscopy (NIRS). *Meat Science* **2003,** 63, 441-450.

43. Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; de la Pezuela, C., Near-infrared spectroscopy in the pharmaceutical industry. *Analyst* **1998,** 123, (8), 135R-150R.

44. Guidance for industry: PAT, A framework for innovative pharmaceutical development, manufacturing and quality assurance. **2004,** Food & Drug Administration.

45. Ciurczak, E. W., Uses of Near-infrared spectroscopy in pharmaceutical analysis. **1987,** 23, 147-163.

46. Morisseau, K. M.; Rhodes, C. T., Near-infrared spectroscopy as a nondestructive alternative to conventional tablet hardness testing. *Pharm Res* **1997,** 14, (1), 108-11.

47. Savage, M.; Torres, J.; Franks, L.; Masecar, B.; Hotta, J., Determination of adequate moisture content for efficient dry-heat viral inactivation in lyophilized factor VIII by loss on drying and by near infrared spectroscopy. *Biologicals* **1998,** 26, (2), 119-24.

48. Lin, T. P.; Hsu, C. C., Determination of residual moisture in lyophilized protein pharmaceuticals using a rapid and non-invasive method: near infrared spectroscopy. *PDA J Pharm Sci Technol* **2002,** 56, (4), 196-205.

49. Kamat, M. S.; Lodder, R. A.; DeLuca, P. P., Near-infrared spectroscopic determination of residual moisture in lyophilized sucrose through intact glass vials. *Pharm Res* **1989,** 6, (11), 961-5.

50. Bai, S. J.; Rani, M.; Suryanarayanan, R.; Carpenter, J. F.; Nayar, R.; Manning, M. C., Quantification of glycine crystallinity by near-infrared (NIR) spectroscopy. *J Pharm Sci* **2004,** 93, (10), 2439-47.

51. Luner, P. E.; Majuru, S.; Seyer, J. J.; Kemper, M. S., Quantifying crystalline form composition in binary powder mixtures using near-infrared reflectance spectroscopy. *Pharm Dev Technol* **2000,** 5, (2), 231-46.

52. Bruun, S. W.; Holm, J.; Hansen, S. I.; Jacobsen, S., Application of near-infrared and Fourier transform infrared spectroscopy in the characterization of ligand-induced conformation changes in folate binding protein purified from bovine milk: influence of buffer type and pH. *Appl Spectrosc* **2006,** 60, (7), 737-46.

53. Izutsu, K.; Fujimaki, Y.; Kuwabara, A.; Aoyagi, N., Effect of counterions on the physical properties of l-arginine in frozen solutions and freeze-dried solids. *Int J Pharm* **2005,** 301, (1-2), 161-9.

54. Miller, C. E.; Honigs, D. E., Discrimination of different crystalline phases using near-infrared diffuse reflectance spectroscopy. *Appl. Spectrosc* **1989,** 43, 44-47.

55. Brittain, H. G., *Physical characterization of pharmaceutical solids*. Mercel Dekker.

56. Miyazawa, M.; Sonoyama, M., Second derivative near infra red studies on the structural characterization of proteins. *J. Near infrared spectroscopy* **1998,** 6, A253 – A257.

57. Wold, S.; Berglund, A.; Kettaneh, N., New and old trends in chemometrics. How to deal with the increasing data volumes in R & D & P (research, development and production) with examples from pharmaceutical research and process modeling. *J Chemometrics* **2002,** 16, 377-386.

58. Workman, J. J., Interpretive Spectroscopy for Near Infrared. *Appl Spectrosc Rev* **1996,** 31, 251-320.

59. Robert, P.; Devaux, M. F.; Mouhous, N.; Dufour, E., Monitoring the secondary structure of proteins by near-infrared spectroscopy. *Appl Spectroscopy* **1999,** 53, 226-232.

60. Liu, Y.; Cho, R. K.; K., S.; Miura, T.; Ozaki, Y., Studies on spectra/structure correlations in near-infrared spectra of proteins and polypeptides. Part I: A marker band for hydrogen bonds. *Applied Spectroscopy* **1994,** 48, 1249-1254.

61. Bai, S.; Nayar, R.; Carpenter, J. F.; Manning, M. C., Noninvasive determination of protein conformation in the solid state using near infrared (NIR) spectroscopy. *J Pharm Sci* **2005,** 94, (9), 2030-8.

62. Izutsu, K.-I.; Fujimaki, Y.; Kuwabara, A.; Hiyama, Y.; Yomota, C.; Aoyagi, N., Near-infrared analysis of protein secondary structure in aqueous solutions and freeze-dried solids. *Journal of Pharmaceutical Sciences* **2006,** 95, (4), 781-789.

63. Brulls, M.; Folestad, S.; Sparen, A.; Rasmuson, A., In-situ near-infrared spectroscopy monitoring of the lyophilization process. *Pharm Res* **2003,** 20, (3), 494-9.

64. William, R. W., Protein secondary structure analysis using Raman amide I and amide III spectra. . *Method enzymology* **1986,** 130, 311-331.

65. Thamann, T. J., Raman spectroscopic studies of a dimeric form of recombinant bovine growth hormone. *Anal Biochem* **1998,** 265, (2), 202-7.

66. Krimm, S.; Bandekar, J., Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv Protein Chem* **1986,** 38, 181-364.

67. Guo, W.; Mabrouk, P. A., Raman evidence that the lyoprotectant poly(ethylene glycol) does not restore nativity to the heme active site of horseradish peroxidase suspended in organic solvents. *Biomacromolecules* **2002,** 3, (4), 846-9.

68. Sane, S. U.; Wong, R.; Hsu, C. C., Raman spectroscopic characterization of drying-induced structural changes in a therapeutic antibody: correlating structural changes with long-term stability. *J Pharm Sci* **2004,** 93, (4), 1005-18.

69. Prescott, B.; Renugopalakrishnan, V.; Glimcher, M. J.; Bhushan, A.; Thomas, G. J., Jr., A Raman spectroscopic study of hen egg yolk phosvitin: structures in solution and in the solid state. *Biochemistry* **1986,** 25, (10), 2792-8.

70. Yu, N. T., Comparison of protein structure in crystals, in lyophilized state, and in solution by laser Raman scattering. 3. Alpha-Lactalbumin. *J Am Chem Soc* **1974,** 96, (14), 4664-8.

71. Wen, Z. Q., Raman spectroscopy of protein pharmaceuticals. *J Pharm Sci* **2007,** 96, (11), 2861-78.

72. Al-Azzam, W.; Pastrana, E. A.; Ferrer, Y.; Huang, Q.; Schweitzer-Stenner, R.; Griebenow, K., Structure of poly(ethylene glycol)-modified horseradish peroxidase in organic solvents: infrared amide I spectral changes upon protein dehydration are largely caused by protein structural changes and not by water removal per se. *Biophys J* **2002,** 83, (6), 3637-51.

73. Tattini, V., Jr.; Parra, D. F.; Polakiewicz, B.; Pitombo, R. N., Effect of lyophilization on the structure and phase changes of PEGylated-bovine serum albumin. *Int J Pharm* **2005,** 304, (1-2), 124-34.

74. Miller, B. L.; Hageman, M. J.; Thamann, T. J.; Barron, L. B.; Schoneich, C., Solid-state photodegradation of bovine somatotropin (bovine growth hormone): Evidence for tryptophan-mediated photooxidation of disulfide bonds. *Journal of Pharmaceutical Sciences* **2003,** 92, (8), 1698-1709.

75. Yan, X.; Zhang, H.; Watson, J.; Schimerlik, M. I.; Deinzer, M. L., Hydrogen/deuterium exchange and mass spectrometric analysis of a protein containing multiple disulfide bonds: Solution structure of recombinant macrophage colony stimulating factor-beta (rhM-CSFbeta). *Protein Sci* **2002,** 11, (9), 2113-24.

76. Wang, F.; Li, W.; Emmett, M. R.; Marshall, A. G.; Corson, D.; Sykes, B. D., Fourier transform ion cyclotron resonance mass spectrometric detection of small Ca(2+)-induced conformational changes in the regulatory domain of human cardiac troponin C. *J Am Soc Mass Spectrom* **1999,** 10, (8), 703-10.

77. Zhu, M. M.; Rempel, D. L.; Zhao, J.; Giblin, D. E.; Gross, M. L., Probing Ca2+-induced conformational changes in porcine calmodulin by H/D exchange and ESI-MS: effect of cations and ionic strength. *Biochemistry* **2003,** 42, (51), 15388-97.

78. Lanman, J.; Lam, T. T.; Barnes, S.; Sakalian, M.; Emmett, M. R.; Marshall, A. G.; Prevelige, P. E., Jr., Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J Mol Biol* **2003,** 325, (4), 759-72.

79. Rist, W.; Jorgensen, T. J.; Roepstorff, P.; Bukau, B.; Mayer, M. P., Mapping temperature-induced conformational changes in the Escherichia coli heat

shock transcription factor sigma 32 by amide hydrogen exchange. *J Biol Chem* **2003,** 278, (51), 51415-21.

80. Yan, X.; Deinzer, M. L.; Schimerlik, M. I.; Broderick, D.; Leid, M. E.; Dawson, M. I., Investigation of ligand interactions with human RXRalpha by hydrogen/deuterium exchange and mass spectrometry. *J Am Soc Mass Spectrom* **2006,** 17, (11), 1510-7.

81. Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; He, Y.; Hendrickson, C. L.; Marshall, A. G.; Griffin, P. R., Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **2006,** 78, (4), 1005-14.

82. Dyson, H. J.; Wright, P. E., Unfolded proteins and protein folding studied by NMR. *Chem Rev* **2004,** 104, (8), 3607-22.

83. Abzalimov, R. R.; Kaltashov, I. A., Extraction of local hydrogen exchange data from HDX CAD MS measurements by deconvolution of isotopic distributions of fragment ions. *J Am Soc Mass Spectrom* **2006,** 17, (11), 1543-51.

84. Kaltashov, I. A.; Eyles, S. J., Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom Rev* **2002,** 21, (1), 37-71.

85. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* **2006,** 25, (1), 158-70.

86. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q Rev Biophys* **1983,** 16, (4), 521-655.

87. Englander, S. W.; Poulsen, A., Hydrogen-tritium exchange of the random chain polypeptide. *Biopolymers* 7, 379-393.

88. Costantino, H. R., *Excipients for use in lyophilized pharmaceutical peptide , protein, and other bioproducts. Biotechnology: Pharmaceutical Aspects , 2(Lyophilization of Biopharmaceuticals)*. 2004; p 139-228.

89. Stotz, C. E.; Winslow, S. L.; Houchin, M. L.; D'Souza, A. J. M.; Ji, J.; Topp, E. M., Degradation pathways for lyophilized peptides and proteins. Biotechnology: Pharmaceutical Aspects, 2(Lyophilization of Biopharmaceuticals). **2004**, 443-479.

90. Desai UR, O. J., Klibanov AM, Protein structure in the lyophilized state: A hydrogen isotope exchange/NMR study with bovine pancreatic trypsin inhibitor. *Journal of American Chemical Society* **1994,** 116, 9420-9422.

91. French, D. L.; Arakawa, T.; Li, T., Fourier transform infared spectroscopy investigation of protein conformation in spray-dried protein/trehalose powders. *Biopolymers* **2004,** 73, (4), 524-31.

92. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Characterizing protein structure in amorphous solids using hydrogen/deuterium exchange with mass spectrometry. *Anal Biochem* **2007,** 366, (1), 18-28.

93. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Trehalose and calcium exert site-specific effects on calmodulin conformation in amorphous solids. *Biotechnol Bioeng* **2007,** 97, (6), 1650-3.

94. Li, Y.; Williams, T. D.; Topp, E. M., Effects of Excipients on Protein Conformation in Lyophilized Solids by Hydrogen/Deuterium Exchange Mass Spectrometry. *Pharm Res* **2007**.

95. Sharma, V. K.; Kalonia, D. S., Steady-state tryptophan fluorescence spectroscopy study to probe tertiary structure of proteins in solid powders. *J Pharm Sci* **2003,** 92, (4), 890-9.

96. Ramachander, R.; Dimitrova, M.; Young, M.; Narhi, L.; Jiang, Y., Solid - state fluorescence to analyze lyophilized protein formulations. In *232nd ACS National Meeting*, San Francisco, United States, 2006.

Chapter 3

Protein-Excipient Interactions in Amorphous Solids by Hydrogen/Deuterium Exchange with Mass Spectrometry

3.1. Introduction

Proteins and other biotech drugs are among the fastest growing sectors of the pharmaceutical industry[1]. To protect these labile molecules from chemical and physical degradation, protein drugs are often formulated and marketed as solids. The properties of proteins and formulation additives ("excipients") together with the processing methods used (e.g., lyophilization) typically produce solids that are amorphous rather than crystalline. Though amorphous solids are lower in energy than solutions, there nevertheless is ample evidence that proteins undergo a variety of degradation processes in the amorphous solid state.[2-10] Understanding and controlling these processes is central to the effective development of solid protein drug products.

Though the mechanisms of protein degradation in amorphous solids are far from clear, maintaining native conformation is generally considered critical to preventing degradation[12-16]. Various methods have been used to assess protein structure in amorphous solids, though far fewer methods are available than for proteins in solution. By far the most commonly used technique is Fourier transform infrared spectroscopy (FTIR)[12-14, 16-19]. FTIR offers the advantages of applicability to both solid and solution samples, ease of sample preparation and rapid analysis. However, though FTIR can detect gross changes in protein secondary structure, the method lacks sufficient resolution to detect more subtle structural changes[17-19]. Other spectroscopic methods such as near infra-red[20], circular dichroism[21] and Raman spectroscopy[22], and thermal methods such as differential scanning calorimetry (DSC)[23] also have been used to acquire information on protein structure in amorphous solids but share the limited resolution of FTIR. Solid-state nuclear

magnetic resonance spectroscopy (ssNMR) has made it possible to solve the structures of membrane proteins at atomic resolution[24]. Current ssNMR methods allow the complete assignment of backbone and side chain signals for solid proteins in the 5-10 kD range[25, 26], but generally require that the sample possess a degree of microscopic order (e.g., crystallinity) and/or isotopic labeling[25, 27]. Since protein drugs are often far larger than 10 kD, lack microscopic order in the amorphous solid state, and are not routinely expressed in isotopically enriched forms, the routine application of current ssNMR methods to determine protein drug conformation in amorphous solids is impractical.

Hydrogen/deuterium exchange (HDX) has emerged as a new method for studying protein conformation and excipient interactions in the solid state. In solution, HDX has been used for more than 50 years to study protein conformation, folding and ligand binding[28-33]. Recently, efforts have been made to extend hydrogen/deuterium exchange to proteins in the solid state. Generally, lyophilized formulations are exposed to D_2O vapor for variable lengths of time at a predetermined RH value before being analyzed by a suitable technique. French et. al.[34] used HDX with FTIR analysis to characterize the solvent exposure of human granulocyte colony stimulating factor (rhG-CSF) and recombinant consensus interferon-α (rConIFN) in spray-dried powders containing trehalose, using isotopic shifts in the Amide II/II' bands. Desai et. al[35] employed HDX with ^{1}H NMR analysis to study the unfolding of bovine pancreatic trypsin inhibitor (BPTI) on lyophilization which led to its loss of activity when reconstituted in organic solvents. In previous studies by our group[11, 36, 37], HDX with tandem liquid chromatography / electrospray ionization mass spectrometry (LC/+ESI-MS) and peptic mapping were used to

provide site specific information on HDX in solid samples of calmodulin. The results demonstrated that low molecular weight sugars (i.e., trehalose, sucrose) provided significant protection from exchange relative to excipient-free controls, and that this protection was exerted preferentially in the α-helical fragments of calmodulin. The studies presented here extend this method to other model proteins having differing secondary structure and in the presence of excipients with differing size and H-bond donor and/or acceptor capacities. The studies test the hypothesis that the ability of excipients to protect proteins from HDX in amorphous solids depends on both excipient type and protein structure, and that this effect is exerted non-uniformly along the protein sequence.

3.2. Materials and Methods

3.2.1. Materials. **–** Myoglobin, lysozyme, ribonuclease A, β-lactoglobulin, concanavalin A, raffinose, trehalose, polyvinyl alcohol (PVA; avg. mol. wt 30,000), polyvinyl pyrrolidone (PVP; avg. mol. wt. 10,000), tris[2-carboxyethyl] phosphine (TCEP), urea-d_4 and guanidine hydrochloride (Gdn•HCl) were obtained from Sigma-Aldrich Co. (St. Louis, MO). Dextran 5000 was obtained from Fluka (Milwaukee, WI). Isopropyl-β-D-thiogalactopyranoside (IPTG) for protein expression was purchased from Amresco, Inc. (Solon, OH). Pepsin was obtained from Worthington Biochemical Corp. (Lakewood, NJ) and formic acid from Acros Organics (Morris Plains, NJ). All materials were of reagent grade or higher and used without further purification.

3.2.2. Expression and purification of E-cadherin 5 (EC-5). - E-cadherin 5 (EC-5) was expressed in *E. coli* by transforming the recombinant plasmid into BL21 (DE3)-$T1^R$ competent cells according to a previously reported protocol[38]. Briefly, cells were

incubated in self-made LB agar plates with kanamycin as the inhibiting antibiotic. A colony was selected with a sterile loop and then incubated further in LB media containing kanamycin (100mg/L) at 37 ^{o}C until OD_{600} reached 0.5-0.8. IPTG (100 mg/ml) was added to induce over-expression to achieve a final concentration of 1 mM. After 4 h of further incubation, the cells were transferred to 4 ^{o}C to stop growth followed by centrifugation to obtain the pellets which were stored at -80 ^{o}C overnight. The pellet was then reconstituted and French-pressed to lyse the cells followed by centrifugation at 20,000 rpm for 60 minutes. The supernatant was transferred to 2 ml microcentrifuge tubes, incubated in 80 ^{o}C water bath for 10 min and then kept in ice for 5 minutes. The suspension formed was then centrifuged at 20,000 rpm for 30 minutes. The supernatant was loaded onto a Q-Sepharose column (Buffer A: 50 mM Tris-HCl, pH 7.5, Buffer B: 50 mM Tris-HCl, 1 M NaCl, pH 7.5) connected to a fast protein liquid chromatography (FPLC) system (Amersham Biosciences, NJ). EC-5 fractions collected after separation were concentrated using Amicon Ultra-15 tubes with a 5000 MW cut-off (Millipore Corp., Billerica, MA). The concentrated protein was then loaded onto a Superdex™ 200 column (GE Healthcare, Piscataway, NJ) for final purification. The EC-5 concentration was determined by UV absorption at 280 nm.

3.2.3. Sample preparation. – Solid samples were prepared by lyophilization from aqueous solution. 100 µl samples of protein in solution (4 mg/ml) were directly lyophilized or co-lyophilized with one of the excipients (i.e., trehalose, raffinose, dextran, PVA, or PVP) in a ratio of 1:1 (w/w). For samples containing Gdn•HCl, 100 µl of a 3 M Gdn•HCl solution was added to the 100 µl protein solution before lyophilization. All samples were lyophilized by first freezing at -35 ^{o}C for 2 h. Drying

was then performed under a vacuum of 15 mT at a shelf temperature of -35 ^{o}C for 2 h, -5 ^{o}C for 8h, 5 ^{o}C for 6 h and 25 ^{o}C for 10 h.

3.2.4. Solid-state hydrogen/deuterium exchange (ssHDX). - ssHDX experiments were performed according to the protocol previously reported by our group[11, 36, 37]. Briefly, lyophilized samples were placed in sealed desiccators at room temperature at 33% relative humidity (RH) over D_2O, achieved by storing the samples over a saturated solution of $MgCl_2$ in D_2O. Samples were collected in triplicate at designated times for immediate analysis or stored at -80^{o} C for later analysis. ssHDX experiments were performed for intact protein and for peptic digests, as described below.

3.2.4.1. ssHDX for intact protein. – ssHDX studies of intact protein were performed to determine the total deuterium uptake by the protein upon exposure to D_2O vapor at 33% RH for 72 h following our previously reported protocol[11, 36, 37]. Preliminary data (not shown) and our previous studies[11, 36, 37] showed that proteins generally reach a plateau in deuterium uptake following 72 h at 33%RH; this standard storage time was applied to all proteins. Lyophilized formulations were reconstituted with Solvent A (94.5% H_2O, 5% acetonitrile and 0.5% formic acid, pH 2.3) to a protein concentration of 4 mg/ml. A 2 μl aliquot was then removed and diluted with an additional 48 μl of Solvent A. The sample was then injected into a short C18 trap column (Upchurch Scientific, Oak Harbor, WA) and washed with the aqueous phase for 1.3 minutes before being eluted with Solvent B (19.5% H_2O, 80% acetonitrile and 0.5% formic acid) into the mass spectrometer. A Micromass Q-Tof II mass spectrometer (Waters Corp., Milford, MA) was used in +ESI mode for analysis. The intact protein was detected within 3 minutes following reconstitution. MassLynx

software (Version 4.0, Waters Corp.) was used to deconvolute the mass spectrum for the intact molecules, and the mass was taken as the cetroid of each deconvoluted peak. Back exchange was corrected using the method of Zhang and Smith[39] and total deuterium incorporation was calculated using the equation:

$$D = (m - m_o) / (m_{100} - m_o) * N \quad (3.1)$$

where m is the mass of the sample protein at any time, m_o is the mass of the native protio form of the protein, m_{100} is the mass of the fully deuterated protein and N is the total number of exchangeable amides on the protein backbone.

3.2.4.2. ssHDX of peptic digests. – In addition to analysis of the intact protein, solid samples of selected proteins (i.e., EC-5, β-lactoglobulin, myoglobin, concanavalin A) were subjected to proteolytic digestion with pepsin to determine the distribution of deuterium incorporation along the protein sequence. Lyophilized pepsin (Worthington Biochemical Corp., Lakewood, NJ) was dissolved in 10 mM sodium acetate to a final concentration of 15 mg/ml. For each digestion, lyophilized proteins without disulfide bonds (i.e., myoglobin, concanavalin A) were reconstituted with 40 µl solvent A followed by the addition of 12 µl pepsin (1:3, protein:pepsin w/w). Proteins with disulfide bonds (i.e., EC-5, β-lactoglobulin) were reconstituted with 35 µl 1M TCEP/4M urea-d4 solution[40] and placed in an ice bath for 4 minutes to denature the protein and reduce the disulfide bonds prior to the addition of 12 µl of pepsin solution. In either case (i.e., proteins with or without disulfide bonds), the mixture was introduced into the injection loop (on ice) and digestion allowed to occur on-line for 3 minutes before LC $^{+}$ESI-MS. The peptides were separated on a C4 reverse phase column (Grace Vydac, Hesperia, CA) with a gradient from 5 to 50 % B in 7 minutes followed by a wash and re-equilibration step. Data analysis was performed using MS

scan with the highest ion count; the cluster was then smoothed and centralized to calculate the deuterium uptake after correcting for back exchange by the method of Zhang and Smith[39]. The injection port, column and tubing were kept on ice and low pH solvents were used to minimize back exchange.

3.2.5. Solids characterization. – The lyophilized solids were analyzed to determine moisture content, glass transition temperature (Tg), crystallinity and protein secondary structure using the methods described below.

3.2.5.1. Thermogravimetric analysis (TGA). - TGA was used to measure the water content in the lyophilized samples after exposure to 33 %RH conditions for 72 h. Samples were analyzed by a Q50 TGA (TA Instruments, New Castle, DE) with a thermal scan from ambient to 200 ^{o}C at a scan rate of 10 ^{o}C per minute in an open platinum pan with nitrogen purge. Universal Analysis software (Version 4.1, TA instruments) was used to determine water content from measured mass loss.

3.2.5.2. Differential scanning calorimetry (DSC). - DSC was used to determine the Tg of the lyophilized formulations after exposure to 33% RH for 72h. A modulated DSC (MTDSC) was used to distinguish glass transition events from other kinetic thermal events such as dehydration and degradation. Samples were analyzed by a Q100 DSC (TA Instruments, New Castle, DE). Samples were held isothermally at 25 ^{o}C for 5 minutes before increasing the temperature at a ramp rate of 1 ^{o}C per minute with modulation amplitudes of ± 0.32 ^{o}C and a modulation period of 60 s. Universal Analysis software (Version 4.1, TA Instruments) software was used for analysis.

3.2.5.3. Powder X-ray diffraction (PXRD). - PXRD was performed to ascertain whether the lyophilized samples were in the crystalline or the amorphous state after exposure to 33% RH for 72 h. The 25 mm diameter shallow well of a 50 mm.

diameter circular Plexiglas sample holder was filled with sample and then mounted on the diffractometer (Bruker AXS, Madison, WI). A locked-coupled $\theta/2\theta$ scan was performed for $10° \leq 2\theta(\text{ CuK}\alpha) \leq 50°$ using graphite-monochromated CuKα radiation (λ = 1.54184 Å) in 0.02° increments with a 4°/min. scan rate. X-rays were provided by a normal-focus sealed X-ray tube operated at 40 kV and 40 mA. No slits were inserted in the X-ray path

3.2.5.4. Fourier transform infrared spectroscopy (FTIR). - FTIR was performed to detect significant changes in protein secondary structure after lyophilization. A PerkinElmer FTIR One spectrometer (PerkinElmer Life and Analytical Sciences, Inc., Waltham, MA) with universal attenuated total reflectance accessory (UATR) was used to acquire the spectra. The solid sample was placed on a diamond crystal surface and covered with a stainless steel slide. Pressure greater than 100 Torr was applied on the steel slide to ensure good contact between the protein and the crystal. GRAM AI (Thermo Electron Corp., Waltham, MA) software was used to analyze the spectra. The raw absorption spectra (Amide I band) were derivatized followed by area normalization using the GRAMS software to enable direct qualitative and quantitative comparison between spectra. Quantitation was performed using the method of Kendrick et al. [41] Previous FTIR studies[34, 42-49] on lyophilized solids have identified bands that reflect a specific secondary structure of the protein. In keeping with these reports, bands at ~1650 cm^{-1} have been assigned to α-helical structures, those at ~1630 cm^{-1} to β-sheet structures and those at~ 1640 cm^{-1} to unstructured regions.

3.3. Results

3.3.1. Solid-state hydrogen/deuterium exchange (ssHDX). -

3.3.1.1. ssHDX for intact protein. - Deuterium incorporation for seven proteins following 72 h of exposure to D_2O vapor at 33% RH and room temperature is shown in Figure 1. The proteins were lyophilized without excipients (Fig. 3.1, "None"), in the presence of polymeric excipients at a protein:excipient ratio of 1:1 (w/w) (Fig. 3.1, "+Dex", "+PVA", "+PVP"), in the presence of di- or trisaccharide excipients at a protein:excipient ratio of 1:1 (w/w) (Fig. 3.1, "+Tre", "+Raf"), or following exposure to 1.5 M guanidine hydrochloride prior to lyophilization (Fig. 3.1, "+Gdn•HCl"). To normalize the data and facilitate comparisons among proteins with different mass, deuterium incorporation is reported as a percentage of the maximum (i.e., as 100 × (D/N), Eqn. 3.1). In Figure 3.1, proteins with the greatest α-helix content in the native form are presented to the left, while those with the greatest β-sheet content in the native form are presented to the right. Detailed structural information for all the model proteins is given in Table 3.1.

For each of the proteins studied, the greatest deuterium incorporation was observed following exposure to Gdn•HCl (70-95%; Fig. 3.1). Particularly high deuterium incorporation was observed for α-helical proteins without disulfide bonds (i.e., myoglobin, calmodulin; Fig. 3.1, Table 3.1). Proteins containing one or more disulfide bonds showed lower percentage deuterium uptake following Gdn•HCl exposure (i.e., lysozyme, ribonuclease A, β-lactoglobulin; Fig. 3.1, Table 3.1). E-cadherin 5, a β-sheet protein, with disulfide bonds shows comparable exchange to the α-helical proteins with no disulfide bonds probably because 60% of it is unstructured and thus allowing 70% of its amide hydrogens to exchange (Fig. 3.1,

Table 3.1. Molecular weight and secondary structure of proteins used for HDX analysis

Protein	Molecular Weight (kD)	Number of S-S bonds	α-Helical Percent	β-Sheet Percent	Unstructured
Myoglobin	16.95	0	83	0	17
Calmodulin	17.0	0	64	2	34
Lysozyme	14.3	4	40	6	54
Ribonuclease A	13.7	4	27	38	35
B-lactoglobulin	18.3	2	12	35	53
E Cadherin	12.6	2	7	30	60
Concanavalin A	25.6	0	7	79	14

Information source PDB: myoglobin (1WLA), Lysozyme (1CXV), RNase A (1RBX), β-lactoglobulin (1CJ5), Con A (1GKB). EC5 information from reports in literature[38]

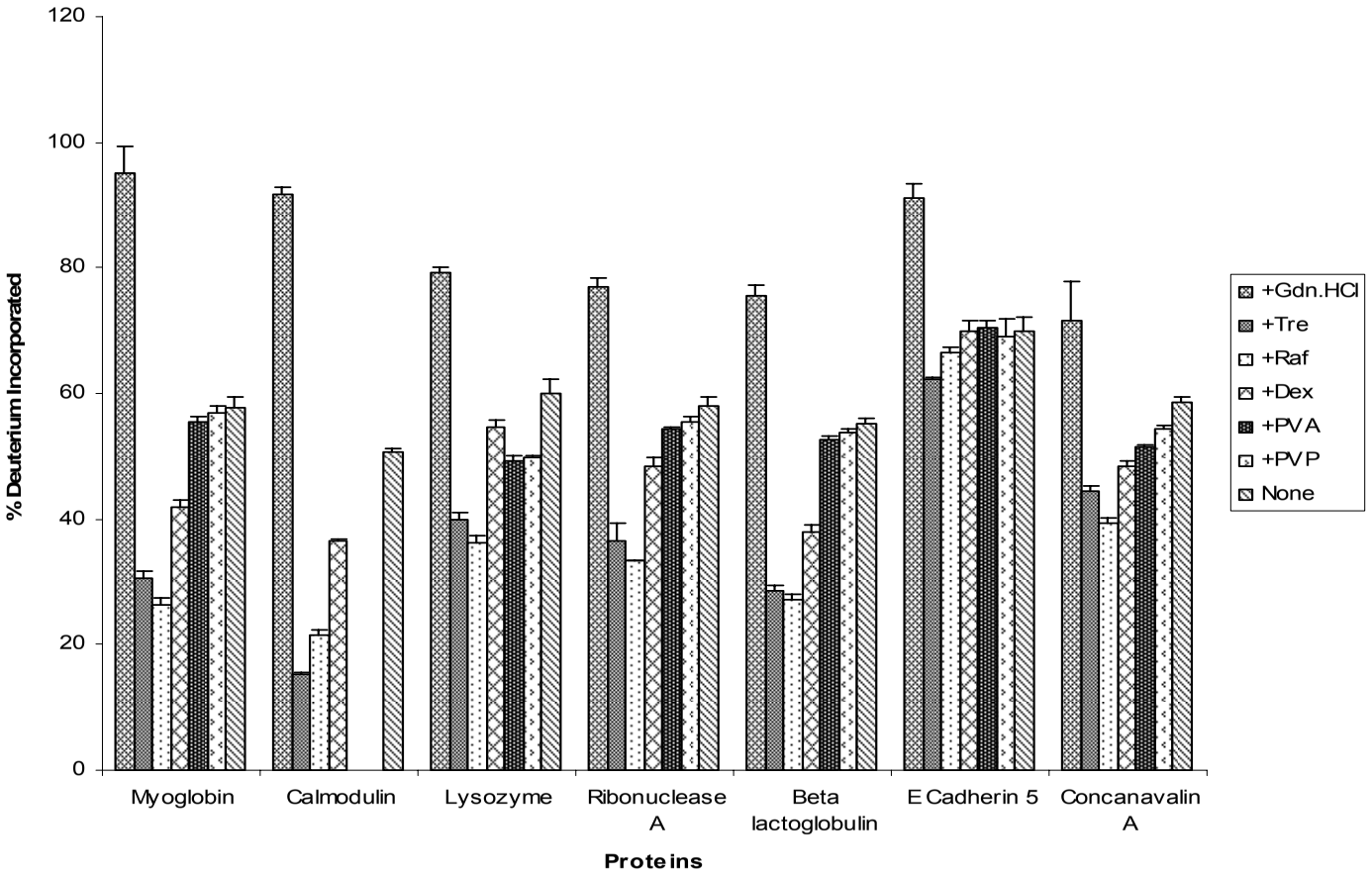


Figure 3.1. Deuterium incorporation following exposure of lyophilized protein samples to D_2O vapor at 33%RH and room temperature for 72 h. Proteins with high α-helix content are presented to the left, proteins with high β-sheet content to the right (see Table 1). Samples were lyophilized without excipients (None), with various excipients at 1:1 (w/w) (+Tre = trehalose, +Raf = raffinose, +Dex = dextran, +PVA = polyvinyl alcohol, +PVP = polyvinyl pyrrolidone), or were exposed to 1.5 M guanidine hydrochloride (+Gdn.HCl) prior to lyophilization. Data for calmodulin taken from[11]. n = 3+/- S.D.

Table 3.1). In solution, Gdn•HCl is known to promote protein unfolding and is often used as a denaturing agent. The high deuterium incorporation for proteins lyophilized from Gdn•HCl solutions suggests that they remain unfolded in the solid state with a high percentage of the backbone amide nitrogens accessible to D_2O from the vapor phase. That proteins with disulfide bonds showed somewhat lower deuterium incorporation suggests that the disulfide bonds, which were not reduced here, provided some protection from exchange. Interestingly, concanavalin A (con A) showed relatively low deuterium incorporation on ssHDX following exposure to Gdn•HCl in solution, though the protein contains no disulfide bonds. Con A forms dimers, tetramers and aggregates in aqueous solution[50] which may be retained during lyophilization and may protect the Gdn•HCl-treated protein from deuterium exchange in the solid state. Unlike the other proteins, solids containing con A were hazy upon reconstitution, further supporting the presence of aggregation and/or oligomerization. LC/+ESI-MS spectra showed peaks corresponding to dimeric and trimeric forms of con A following exposure to 33% RH for 72 h. The areas of these deconvulated mass spectrometric peaks were approximately 8% and 4%, respectively, of the area of intact con A. LC/+ESI-MS spectra of myoglobin also showed evidence of dimer formation (~4%) following storage at 33% RH for 72 h. Spectra for RNase A, lysozyme, calmodulin and β-lactoglobulin did not show evidence of aggregation.

Protein samples lyophilized in the absence of excipients (Fig. 3.1, “None”) or with polymeric excipients in a 1:1 w/w ratio (Fig. 3.1, “+Dex”, “+PVP”, “+PVA”) generally showed 50 to 70% deuterium incorporation following exposure to D_2O vapor. For each protein, deuterium incorporation was greater in excipient-free

samples than in samples containing polymeric excipients. This suggests that the polymeric excipients protect the proteins from exposure to D_2O vapor by promoting a more compact structure and/or through protein/excipient interactions. Among the polymeric excipients, the greatest protection from exchange was afforded by dextran for all proteins except lysozyme and EC-5, perhaps due to dextran's relatively low molecular weight (5 kD vs. 30 kD for PVA, 10 kD for PVP) and propensity to form H-bonds[51]. Though PVA and PVP were selected as an H-bond donor and an H-bond acceptor, respectively, differences in protection from exchange by the two materials were minimal.

For all seven proteins studied, the greatest protection from exchange was provided by trehalose (a disaccharide) and raffinose (a trisaccharide) (Fig. 3.1). Trehalose and raffinose reduced exchange to less than 50% of the value in the absence of excipients in some cases, with the strongest inhibition generally observed in proteins having moderate to high α-helix content (Fig. 3.1). Proteins having high β-sheet structure (i.e., EC-5 and concanavalin A, Fig. 3.1) also showed inhibition of exchange in the presence of trehalose and raffinose, but to a lesser extent. Low molecular weight sugars such as trehalose and raffinose have long been used to stabilize proteins during lyophilization and storage, an effect that has been attributed to intermolecular hydrogen bonds between the sugar and the protein in the solid [51-53]. It is reasonable to expect that such intermolecular protein-excipient hydrogen bonds and/or intramolecular hydrogen bonds within the protein influence hydrogen/deuterium exchange on exposure to D_2O vapor.

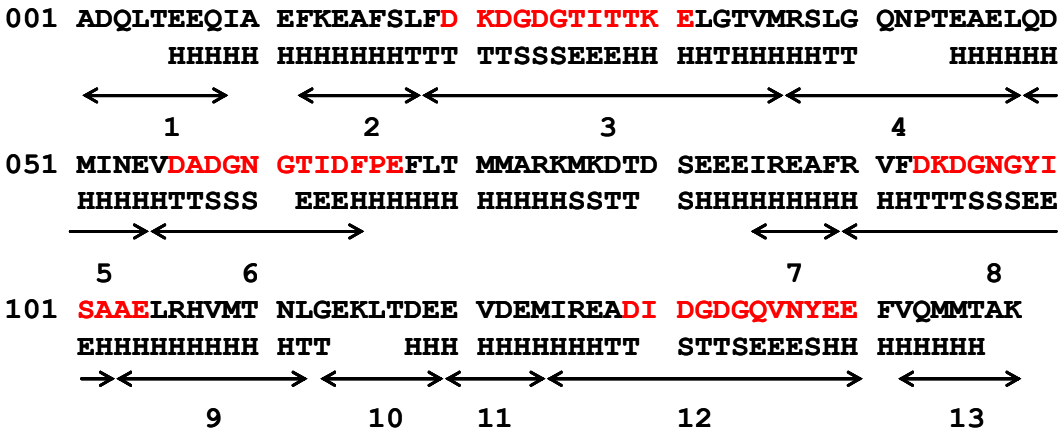
ssHDX studies of intact proteins provide global information on deuterium uptake in the presence of various excipients, but do not address whether this uptake

occurs uniformly along the protein sequence. Analysis of peptic digests of selected proteins was performed to provide information on the uniformity of exchange.

3.3.1.2. ssHDX of peptic digests. – Solid samples of selected proteins exposed to D_2O vapor were subjected to digestion with pepsin to assess the distribution of deuterium incorporation. Our previous studies of calmodulin[11, 36, 37] showed that the protective effects of the excipients were exerted primarily in its α-helices and calcium binding loops. Here, we evaluate whether this site specific protection from exchange occurs in other proteins. To that end, myoglobin (α-helical), β-lactoglobulin (mixed) and EC5 (mostly β-sheet) were subjected to peptic digestion and analysis following D_2O exposure in the solid phase. Lysozyme and Con A were resistant to pepsin digestion and could not be digested in a time frame short enough to allow reasonable retention of deuterium labeling. The remainder of this section summarizes previous digest results on calmodulin followed by current results on myoglobin, β-lactoglobulin and EC-5.

3.3.1.2.1. Calmodulin. - Calmodulin is an EF hand protein with high α-helix content (Table 3.1). In our group's previous work[11, 36, 37], calmodulin work was lyophilized with trehalose, raffinose, dextran 5000 and Gdn•HCl. Following exposure to D_2O(g) in the solid state and peptic digestion, thirteen peptic fragments were used to map the protein and the results grouped according to secondary structure (Fig. 3.2A). We define the "protective effect" of a particular excipient as the difference between the percent deuterium incorporation for the excipient-free control ("None", Fig. 3.2B) and the percent deuterium incorporation in samples containing that excipient. Low molecular weight carbohydrates (trehalose and raffinose) thus had a distinct protective effect in the α-helical regions of calmodulin (Fragments 2, 7, 9 and 11; Fig.

A



B

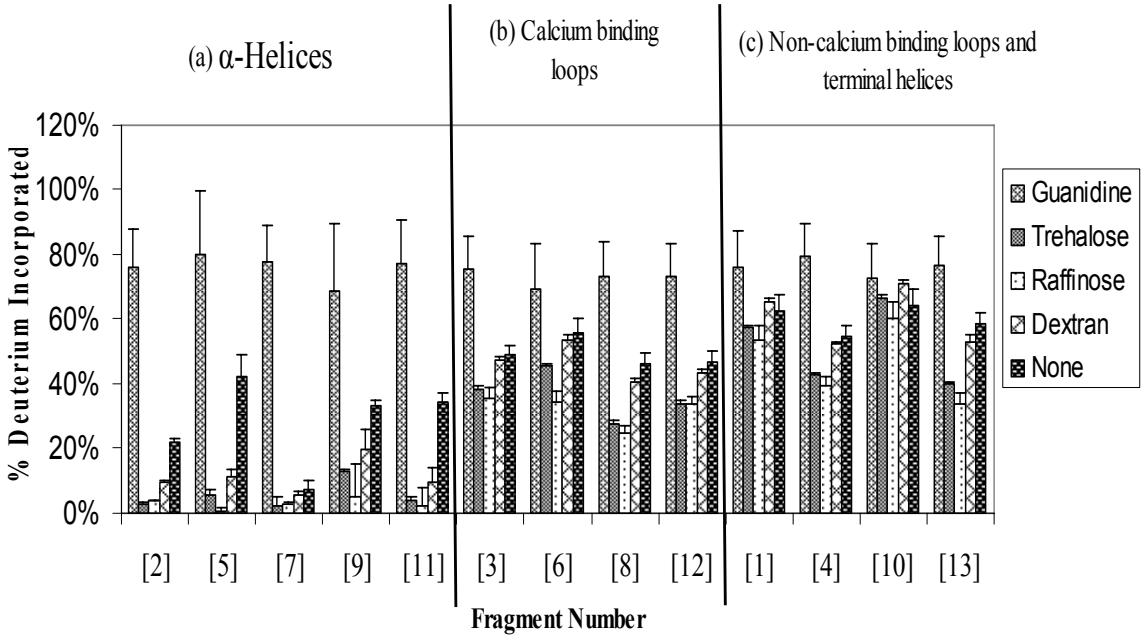


Figure 3.2. A) Sequence and predicted secondary structure of calmodulin obtained from PDB, entry CLL. B) Peptide fragments of calmodulin showing the percent of deuterium uptake in presence of various excipients. Data for calmodulin taken from[11]

3.2B) and to some extent in the calcium binding loops (Fragments 3, 6, 8 and 12; Fig. 3.2B) while the unstructured fragments were largely unaffected. Dextran had a moderate effect, mainly in the α-helical regions.

3.3.1.2.2. Myoglobin. - Myoglobin is a globular protein with high (83%) α-helix content (Table 3.1). Peptic digestion produced twenty-eight fragments detectable by LC/+ESI-MS, of which thirteen were selected to provide maximum sequence coverage (74%, Fig. 3.3A). As for calmodulin, the data were divided into two groups according to secondary structure, i.e., into fragments constituting the α-helical regions and those constituting the loops and the unstructured portions of the protein (Fig. 3.3B). The pattern of protection from exchange (Fig. 3.3B) is similar to that observed for calmodulin. Raffinose and trehalose showed the greatest protection from exchange for all the fragments, but the greatest protective effect of these low molecular weight sugars occurred for the α-helical regions (Fig. 3.3B, Fragments 2-9,12). Most of the unstructured fragments also showed some protective effect of the sugars (Fig. 3.3B, Fragments 1, 10, 11, 13); though these fragments were assigned to the "unstructured" group, all but Fragment 13 has some α-helical content (Fig. 3.3A). Dextran exhibited a moderate protective effect primarily in the α-helical domains, while PVA and PVP showed no protective effects in any portion of the protein. As expected, the greatest deuterium incorporation occurred in samples exposed to Gdn•HCl. For some fragments (Fig. 3.3B, Fragments 4-8,11,12) deuterium incorporation was greater than 100% in solid samples treated with Gdn•HCl prior to lyophilization. Since the percent deuterium incorporation is calculated relative to the fully folded and maximally deuterated protein in solution

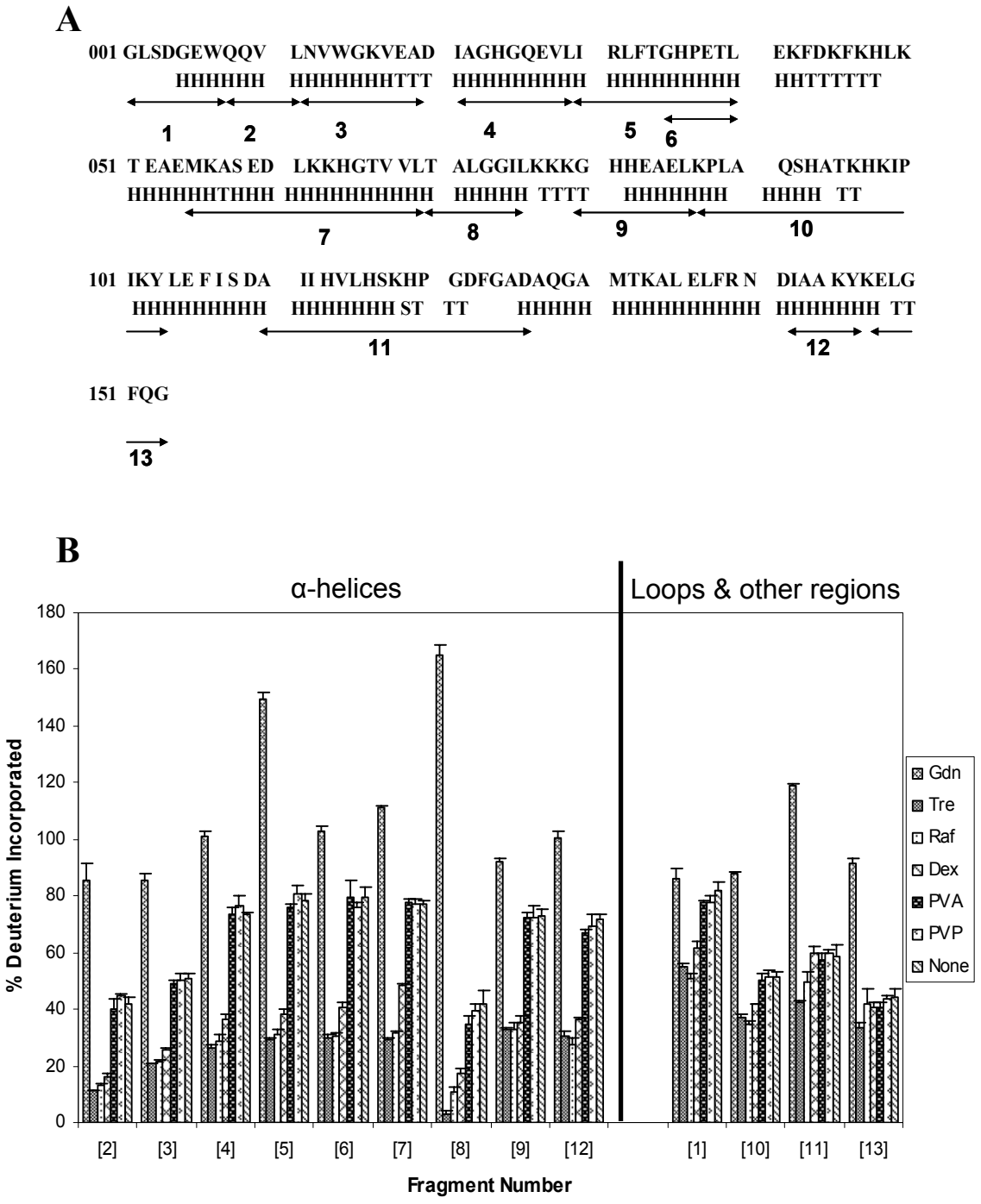


Figure 3.3. A) Sequence and predicted secondary structure of myoglobin obtained from PDB, entry 1wla. B) Peptide fragments of myoglobin showing the percent of deuterium uptake in presence of various excipients.

(see Eqn. 3.1), values greater than 100% suggest that the protein was denatured in Gdn•HCl solutions, remained unfolded upon lyophilization, and in that form showed greater deuterium uptake from D_2O(g) in the solid than the native protein in D_2O solution. Of the four proteins subjected to digestion, this effect was only observed for myoglobin.

3.3.1.2.3. β-lactoglobulin. - β-lactoglobulin is a lipocalin with mixed α-helix (14%) and β-sheet (35%) structure (Table 3.1). Sixteen of the thirty-two peptic fragments detected by LC/+ESI-MS were used to map the protein (Fig. 3.4A), providing 76% sequence coverage. Since this is a mixed protein, an attempt was made to select fragments with exclusively β-sheet or α-helix content (Fig. 3.4A). The β-sheet regions of β-lactoglobulin generally showed lower percent deuterium incorporation than other types of secondary structure, and values were less sensitive to excipient selection (Fig. 3.4B). Protection from exchange in the α-helical fragments followed the pattern observed for calmodulin and myoglobin, with the greatest protection from exchange provided by trehalose and raffinose. Exchange in fragments assigned to the "mixed" or "loops and other" categories showed variable exchange and protection from exchange. In some fragments, inclusion of PVA or PVP resulted in deuterium incorporation that was greater than the excipient-free control (e.g., Fragments 1, 12, 13, 16, Fig. 3.4B), suggesting that PVA and PVP promote exposure to D_2O (g) in these regions. Including Gdn•HCl produced the greatest exposure to exchange in all fragments.

3.3.1.2.4. E-Cadherin-5 (EC-5). – EC-5 has high β-sheet content (~50%). Eleven peptic fragments were selected for analysis, giving a total sequence coverage of

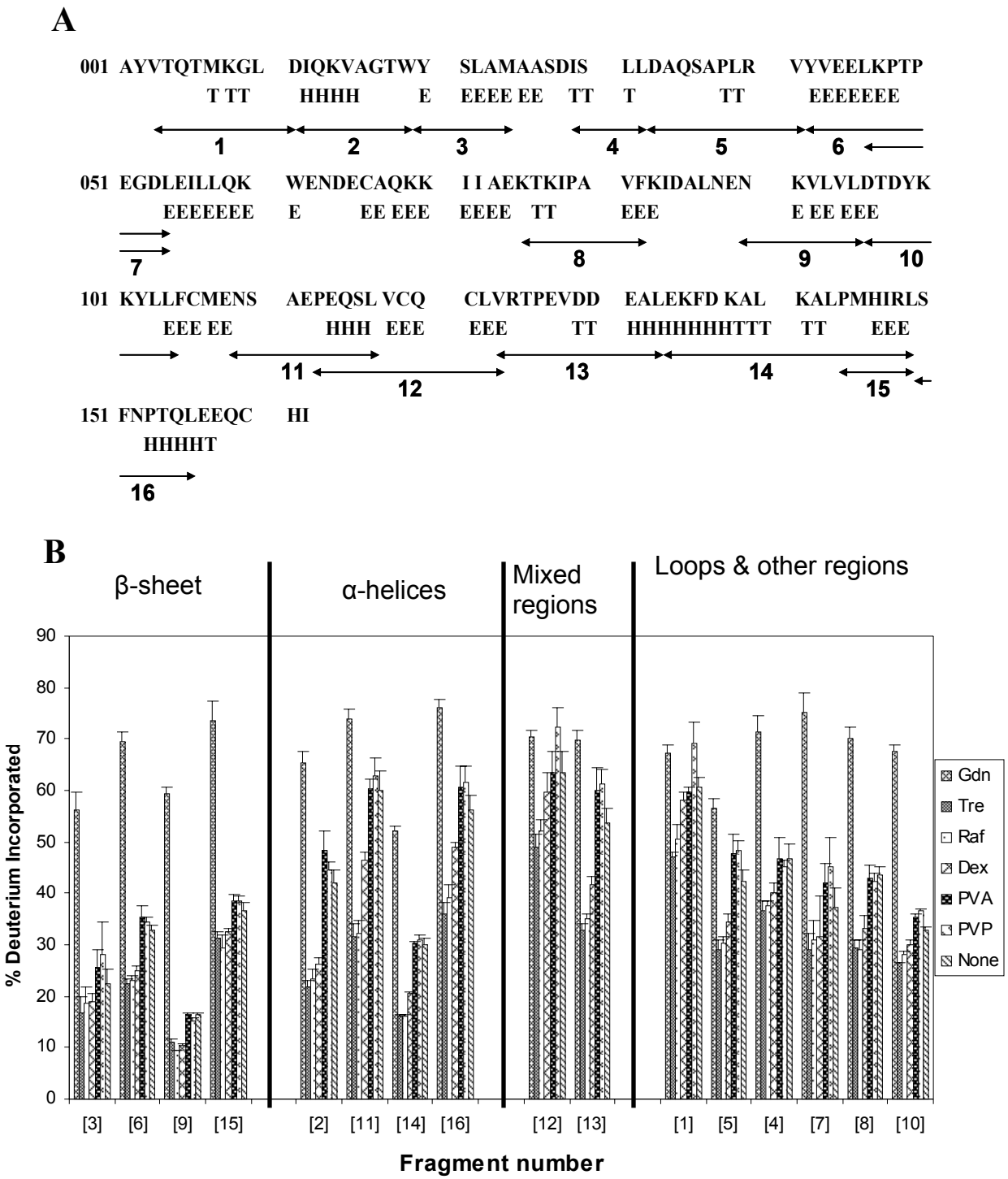


Figure 3.4. A) Sequence and predicted secondary structure of β-lactoglobulin obtained from PDB, entry 1cj5. B) Peptide fragments of β-lactoglobulin showing the percent of deuterium uptake in presence of various excipients.

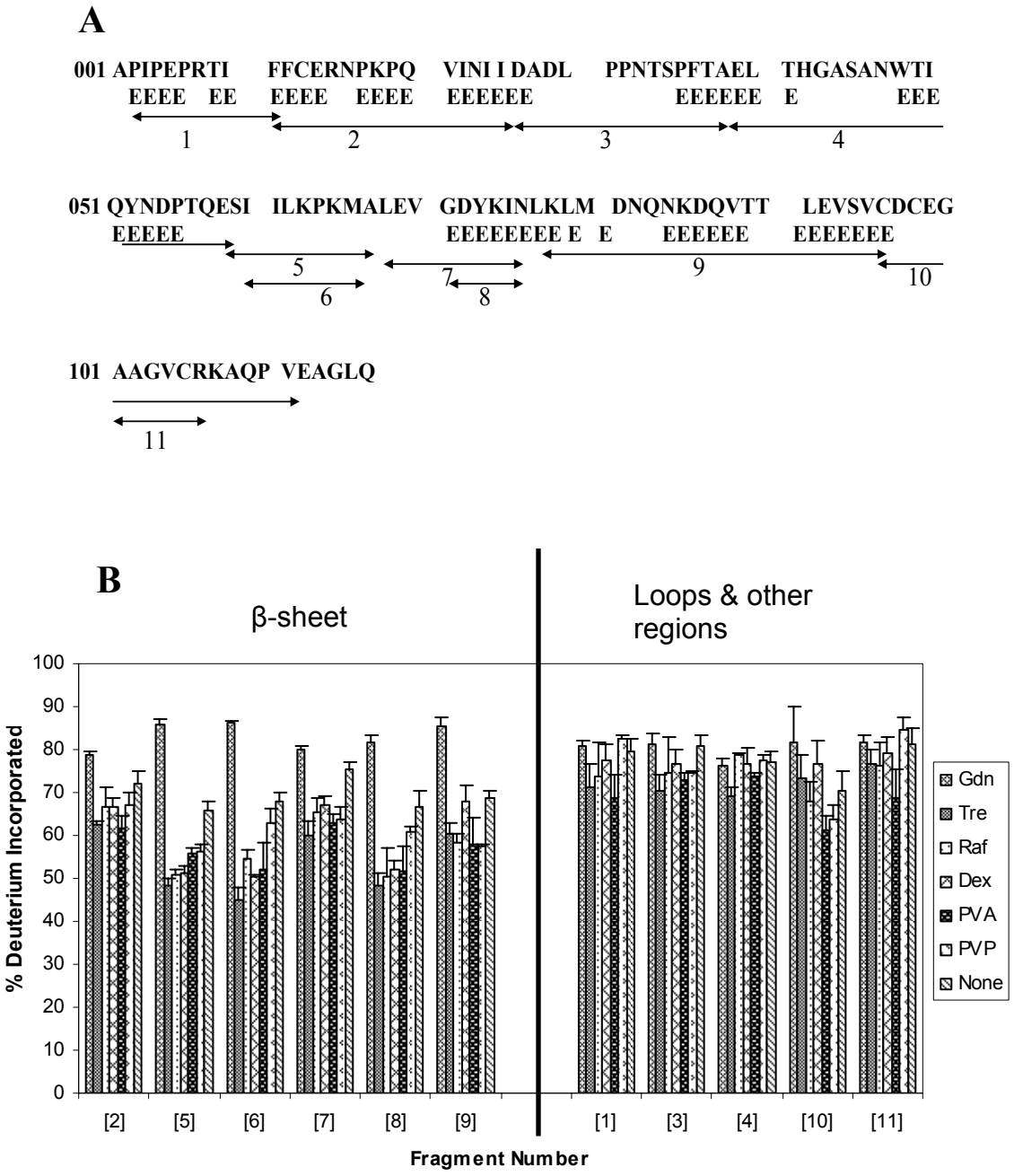


Figure 3.5. A) Sequence and predicted secondary structure of EC5. B) Peptide fragments of EC5 showing the percent of deuterium uptake in presence of various excipients.

97%, and assigned to either β-sheet or unstructured regions (Fig. 3.5A). Fragments with partial β-sheet character (i.e., Fragments 1, 3, 4, Fig. 3.5A) were included in the unstructured group on the basis of their exchange behavior. The peptic fragments of EC-5 generally showed greater percent deuterium incorporation than observed for other proteins, suggesting a less compact structure in the solid state. The β-sheet fragments (Fig. 3.5B, Fragments 2, 5-9) showed some protection from exchange by the various excipients, to a degree roughly comparable to the β-sheet regions of β-lactoglobulin (Fig. 3.5B). The unstructured fragments of EC-5 were insensitive to the excipients used, and exposure to Gdn•HCl did not increase deuterium incorporation relative to the other excipients (Fig. 3.5B, Fragments 1,3,4,10,11), again suggesting a loose structure.

3.3.2. Solid state characterization

3.3.2.1. Powder -ray diffraction (PXRD) – PXRD was performed on lyophilized samples exposed to 33% RH for 72 h to determine the physical state of the solids (i.e., amorphous vs. crystalline) and to detect partial crystalline character. - PXRD patterns for all samples were consistent with amorphous material (Figs. 3. 6, 3.7, 3.8, 3.9, 3.10), with the exception of samples treated with Gdn•HCl, which showed partial crystallinity (Figs. 3.11). The salt crystallizes out on lyophilization with the protein remaining in the amorphous form[11].

3.3.2.2. FTIR. - FTIR spectroscopy was performed to assess protein secondary structure in the solid state following exposure to 33% RH for 72 h, and to provide qualitative and quantitative information on the effects of lyophilization and excipient selection on protein secondary structure. Since the pyrrolidone ring of PVP absorbs in the amide I region, masking the protein signal, FTIR measurements were not

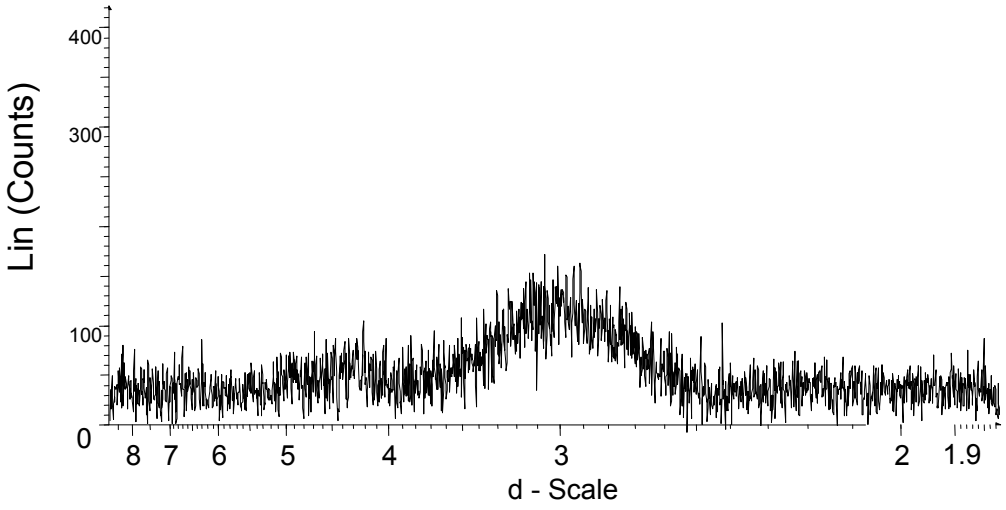


Figure. 3.6. PXRD pattern for a formulation containing β-lactoglobulin and trehalose in 1:1 ratio (w/w).

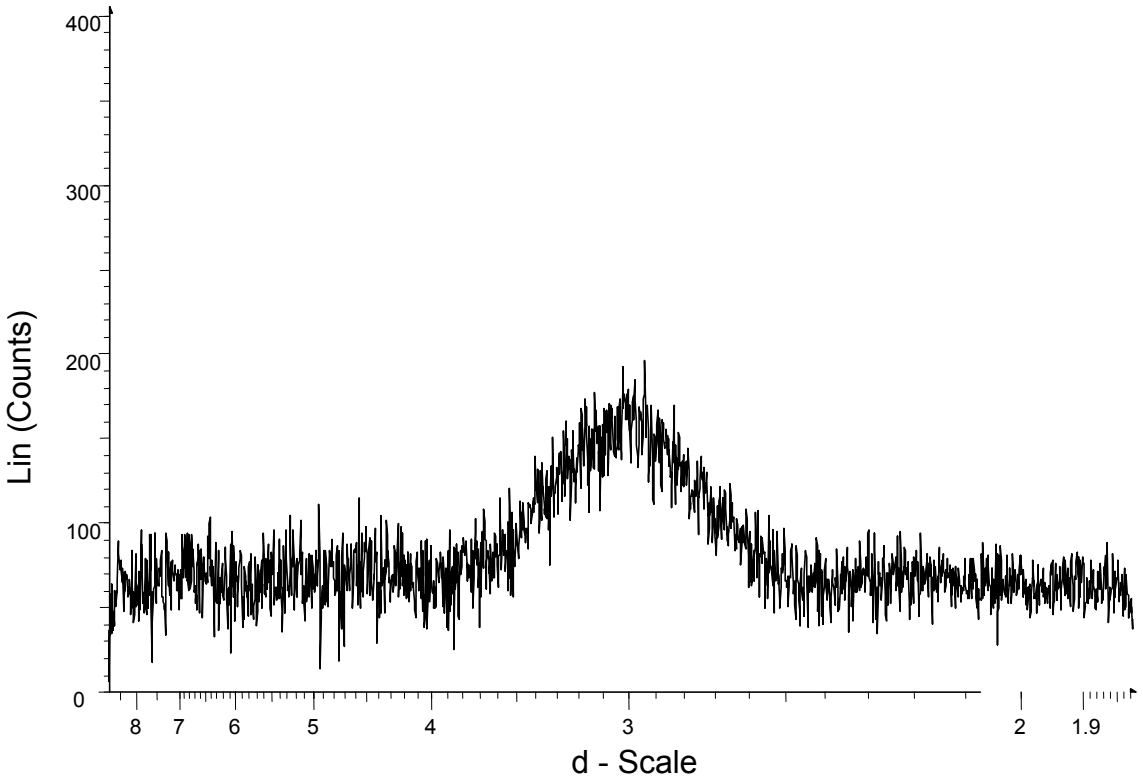


Figure. 3.7. PXRD pattern for a formulation containing β-lactoglobulin and raffinose in 1:1 ratio (w/w).

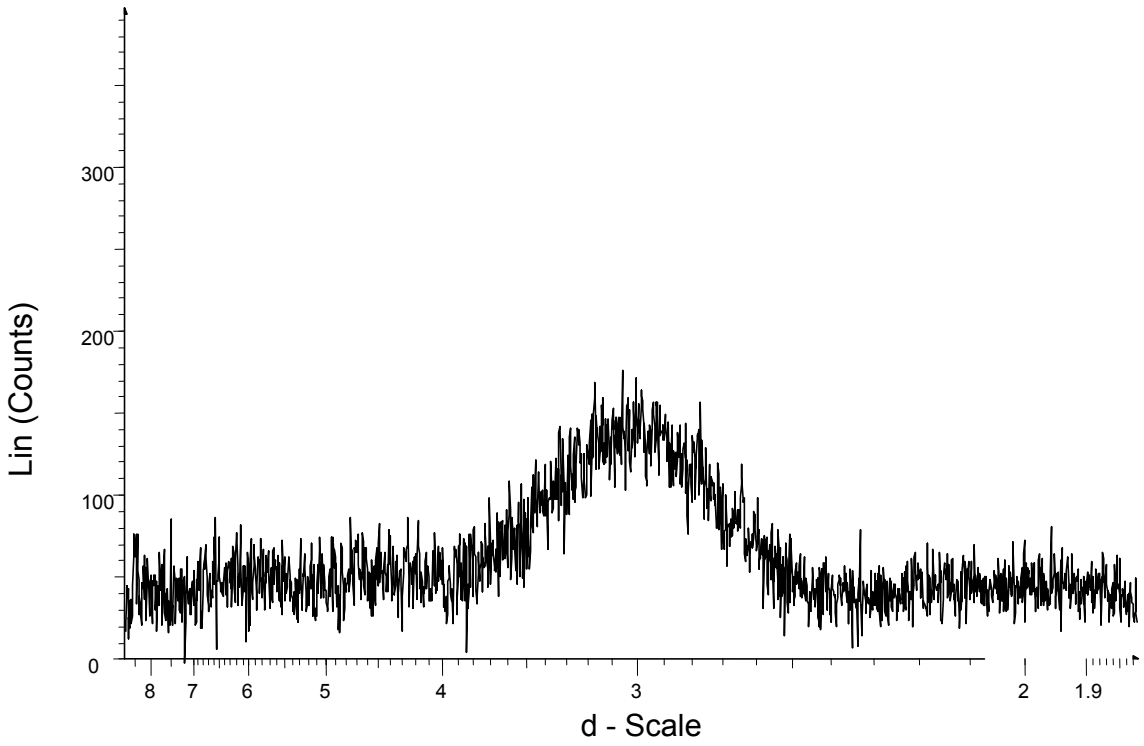


Figure. 3.8. PXRD pattern for a formulation containing β-lactoglobulin and dextran in 1:1 ratio (w/w).

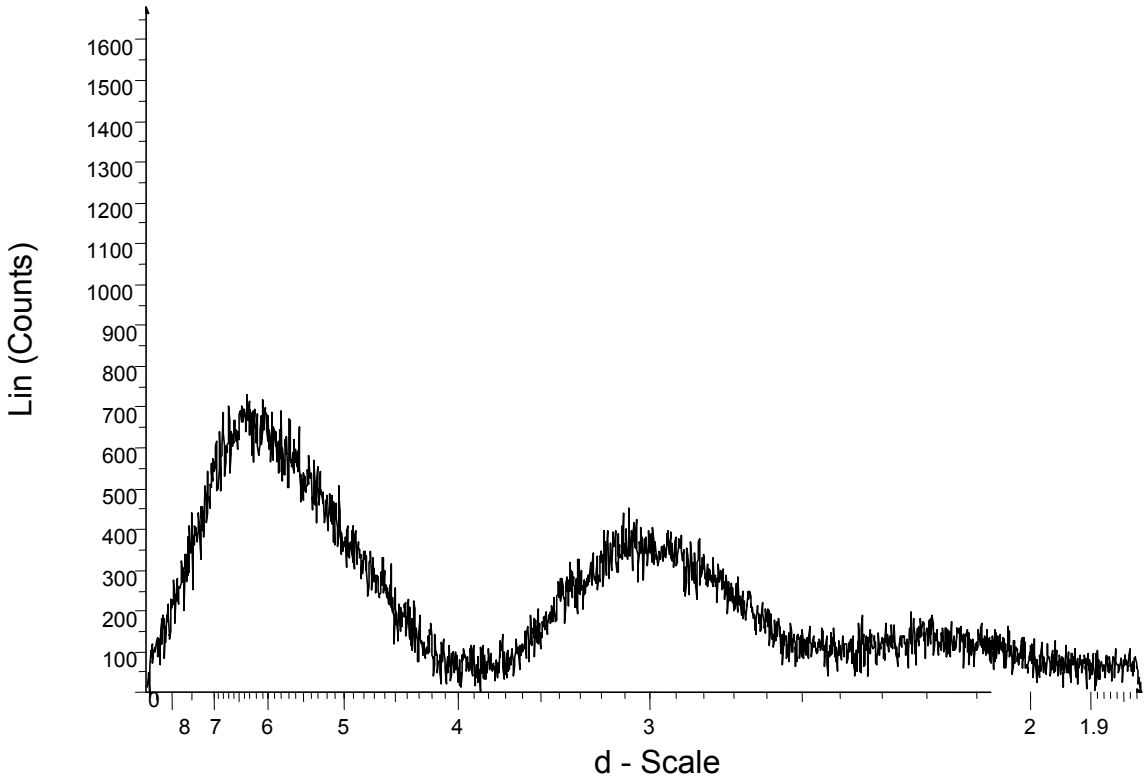


Figure. 3.9. PXRD pattern for a formulation containing β-lactoglobulin and PVA in 1:1 ratio (w/w).

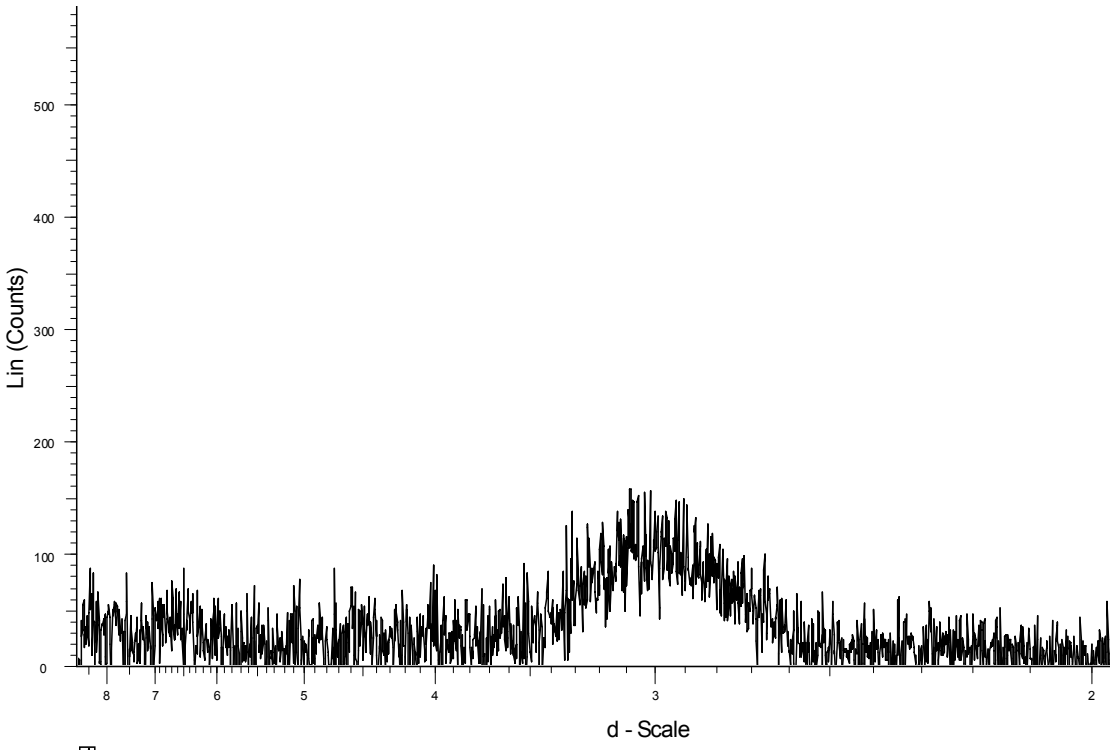


Figure. 3.10. PXRD pattern for a formulation containing β-lactoglobulin and PVP in 1:1 ratio (w/w).

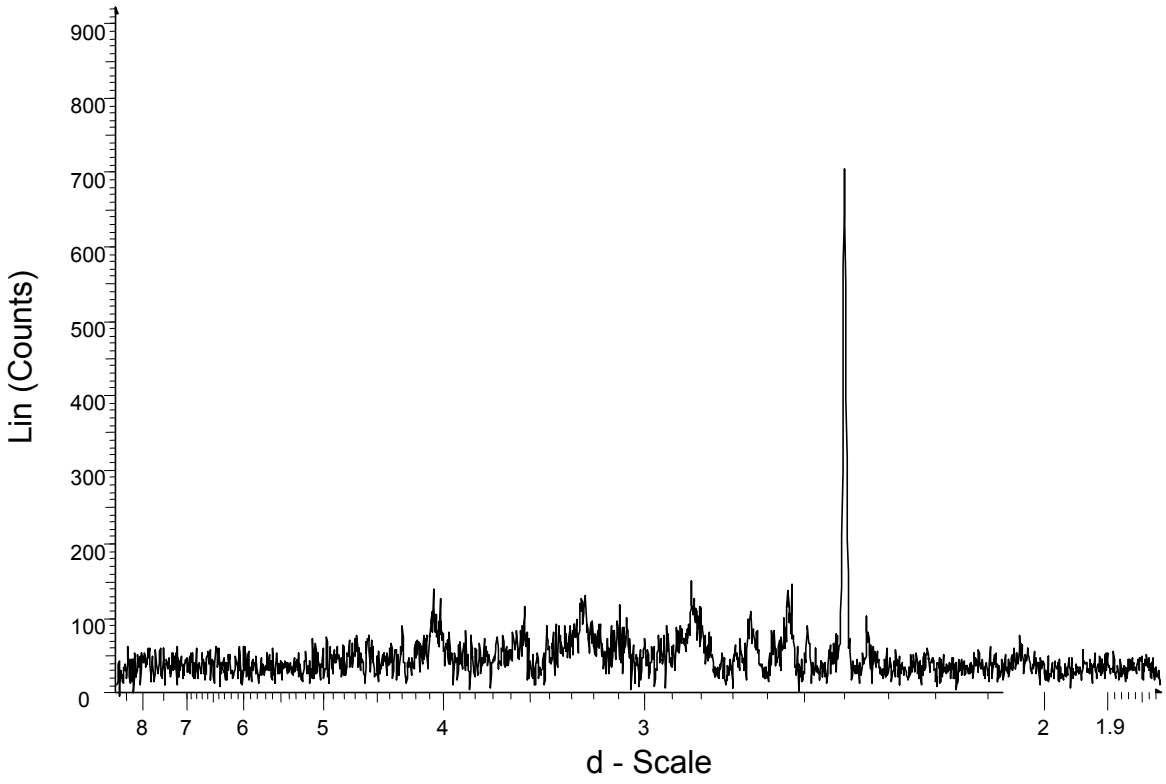


Figure. 3.11. PXRD pattern for a formulation containing β-lactoglobulin and Guanidine. HCl in 1:1 ratio (w/w).

performed on samples containing PVP. FTIR results are summarized and discussed below by each protein.

3.3.2.2.1 Myoglobin- Myoglobin's predominantly α-helical structure is retained in solid samples lyophilized in the absence of excipients, as reflected in the strong FTIR band at 1660 and1650 cm^{-1} (Fig. 3.12). The band intensity is greater for samples containing trehalose or raffinose, consistent with an increase in helix content and greater retention of structure. FTIR spectra for myoglobin samples containing dextran were similar to the excipient-free samples, while those containing PVA showed a loss in band intensity relative to controls, consistent with a loss in structure. The absorption band and the second derivative peak assignments are in agreement with previous FTIR reports for myoglobin[54-56].

3.3.2.2.2. Lysozyme – The FTIR spectra for solid samples of lysozyme without excipients showed moderately strong α-helical absorption bands at 1658 and 1648 cm^{-1} together with strong β-sheet bands at 1638 cm^{-1} (Fig. 3.13), consistent with previous FTIR studies of this protein.[51, 55] In samples containing excipients, the α-helical bands show increases in intensity that are generally considered consistent with increased secondary structure[57]. Samples containing PVA showed increased intensity for bands associated with both α-helical (1658, 1648 cm^{-1}) and β-sheet structure (1638 cm^{-1}), a seeming contradiction that may be associated with increased water content in these formulations. Lysozyme has been reported to undergo structural perturbation on drying leading to loss of secondary structure[58].

3.3.2.2.3. RNase A – Though RNase A has mixed α-helix and β-sheet structure in solution (Table 3.1), the dominant feature of the FTIR spectra for solid samples is a strong β-sheet absorption band (~ 1638 cm^{-1}) that increases in intensity when

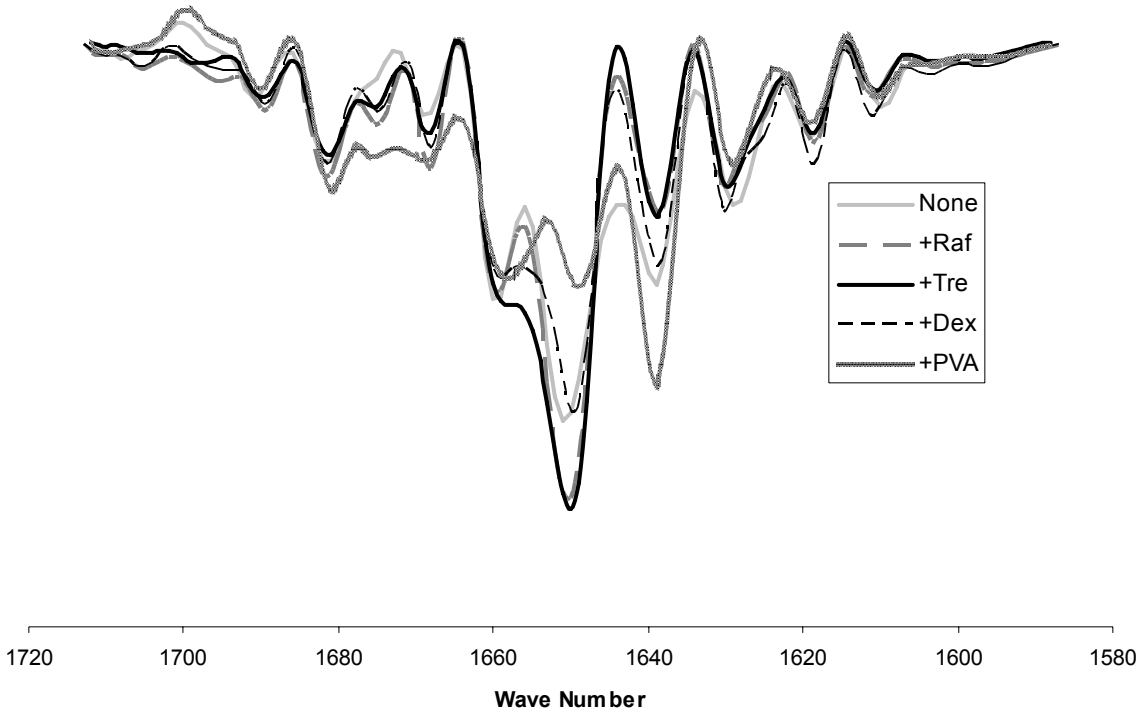


Figure 3.12. Overlaid FTIR spectra for myoglobin lyophilized formulations with various excipients (1:1, w/w).

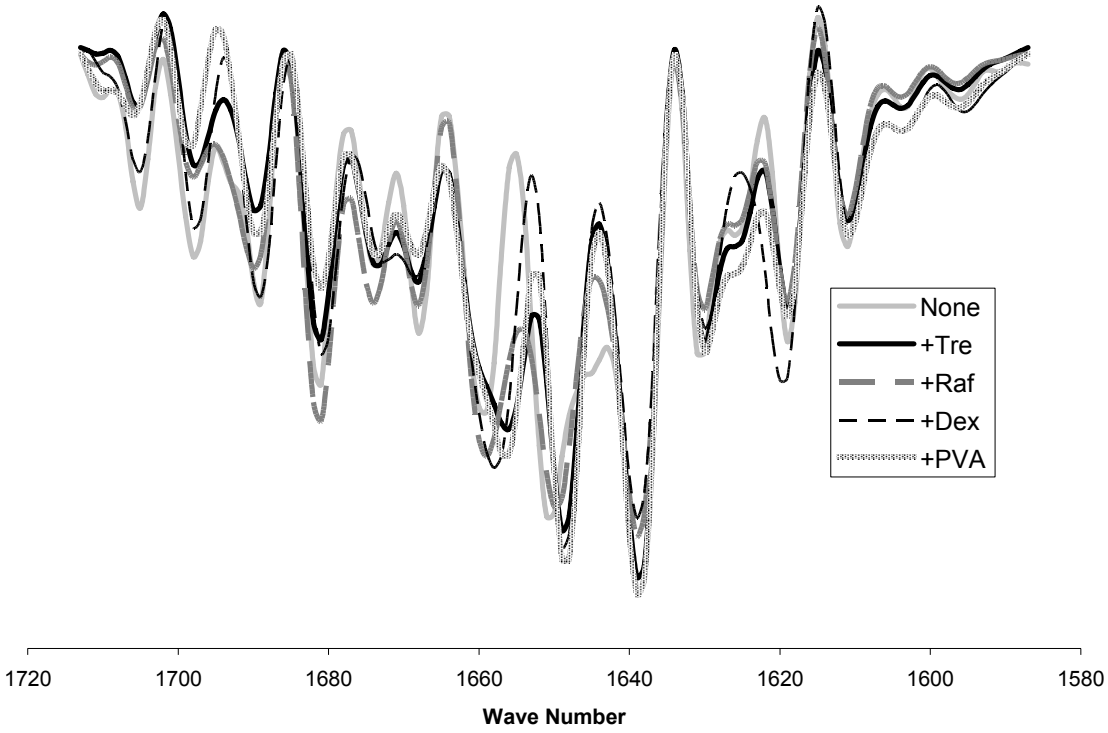


Figure 3.13. Overlaid FTIR spectra for lysozyme lyophilized formulations with various excipients (1:1, w/w).

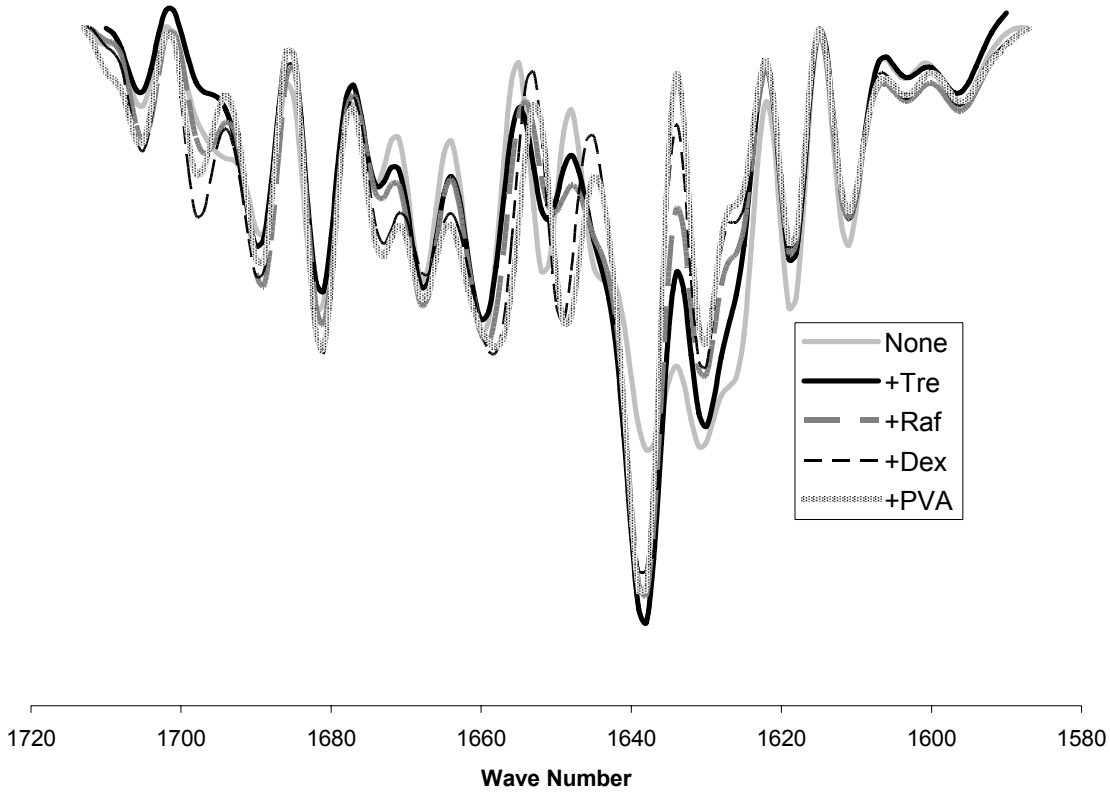


Figure 3.14. Overlaid FTIR spectra for RNase A lyophilized formulations with various excipients (1:1, w/w).

excipients are included in the sample (Fig. 3.14). Weak α-helical bands are also observed at 1658 and 1650 cm^{-1}, the latter of which shows minor increases in excipient-free control. The absorption band and the second derivative peak assignments agree with previous FTIR studies of myoglobin[54, 55, 59, 60]

3.3.2.2.4. β-lactoglobulin - β-lactoglobulin, a protein with mixed α-helix and β-sheet structure, showed a strong absorption band at 1626 cm^{-1} consistent with its β-sheet structure (Fig. 3.15). The position and intensity of the β-sheet band are relatively unaffected by the carbohydrate excipients, while inclusion of PVA is associated with a decrease in intensity of this band. Though β-lactoglobulin has some α-helix content (Table 3.1), these bands are not prominent in the spectra. The raw absorption spectra are in good agreement with those previously reported for β-lactoglobulin[59]

3.3.2.2.5 EC5 – FTIR spectra for EC5 show a strong β-sheet band at 1644 cm^{-1}, suggesting that this β-sheet containing protein retains structure in the solid samples (Fig. 5.16). The spectrum is relatively unaffected by the inclusion of excipients.

3.3.2.2.6 Con A - FTIR spectra for solid samples containing Con A show an intense β-sheet absorbtion band at 1633 cm^{-1}, consistent with retention of secondary structure for this lectin (Fig. 3.17). Relative to the excipient-free control, the intensity of the band is increased in solids containing trehalose or raffinose with a corresponding decrease in the shoulder at ~1625 cm^{-1}, suggesting increased secondary structure. The spectrum is relatively insensitive to other excipients. The absorption band and the second derivative peak assignments agreed with previous FTIR reports for con A[54].

For each protein, FTIR spectra for samples containing excipients were compared quantitatively to the excipient-free control using a correlation matrix

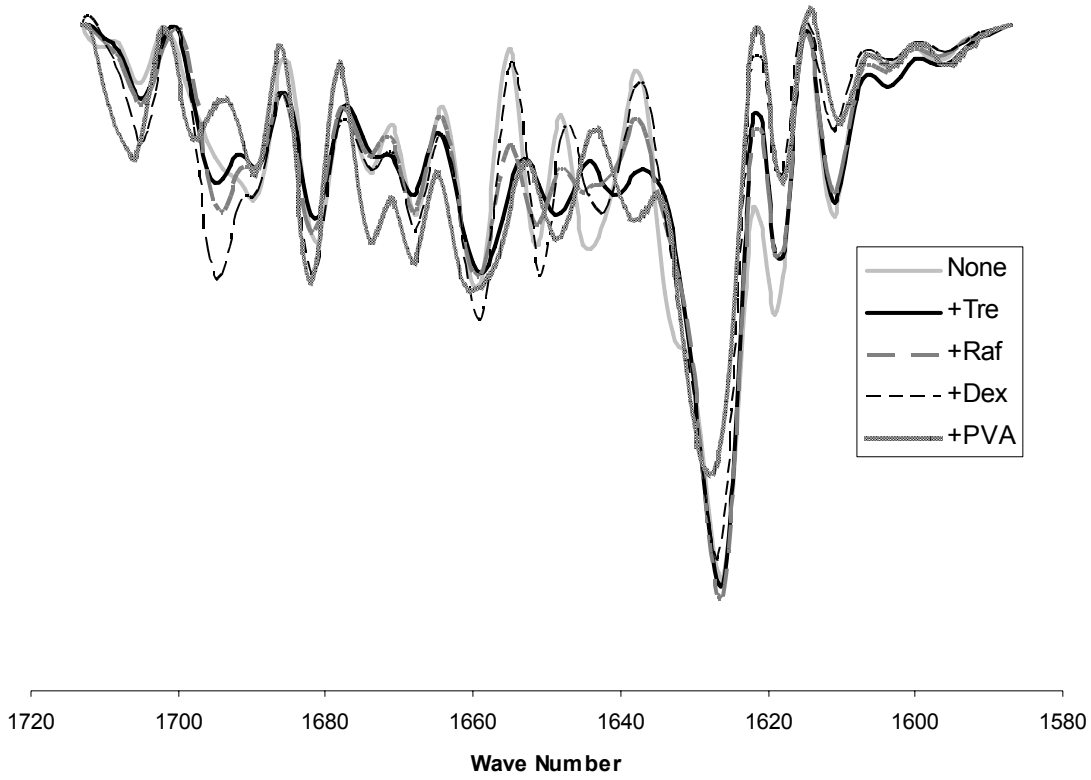


Figure 3.15. Overlaid FTIR spectra for β-lactoglobulin lyophilized formulations with various excipients (1:1, w/w).

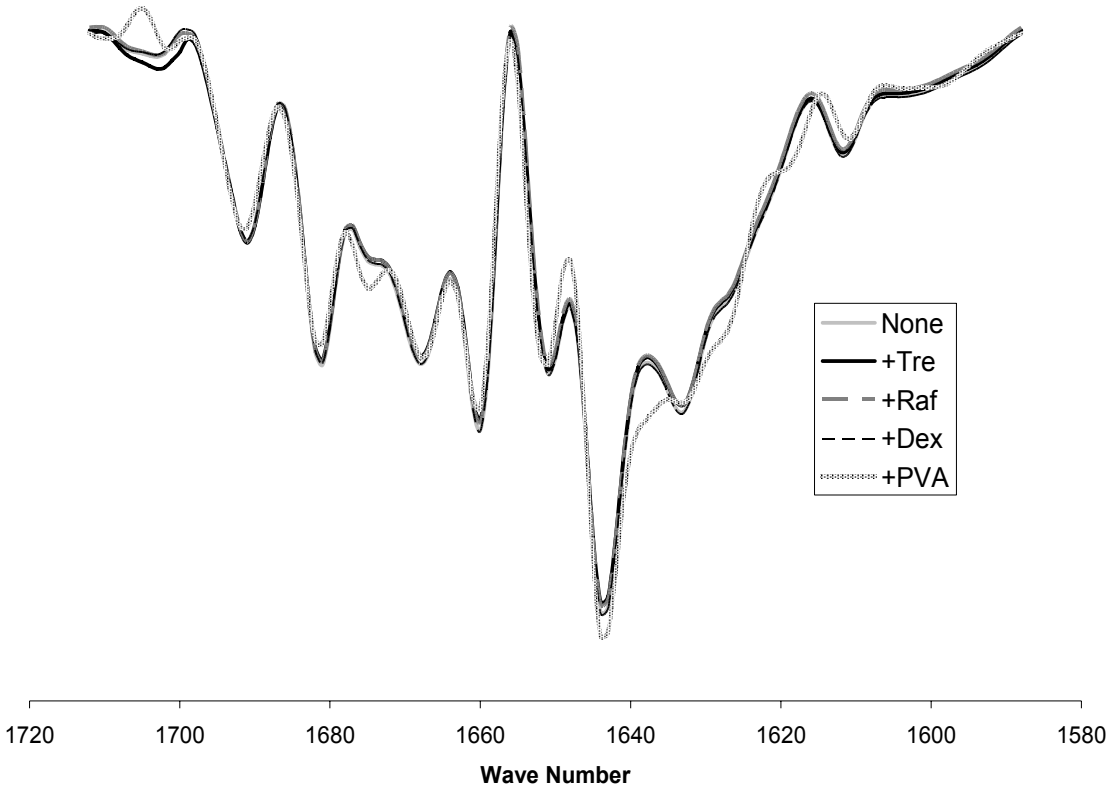


Figure 3.16. Overlaid FTIR spectra for EC5 lyophilized formulations with various excipients (1:1, w/w).

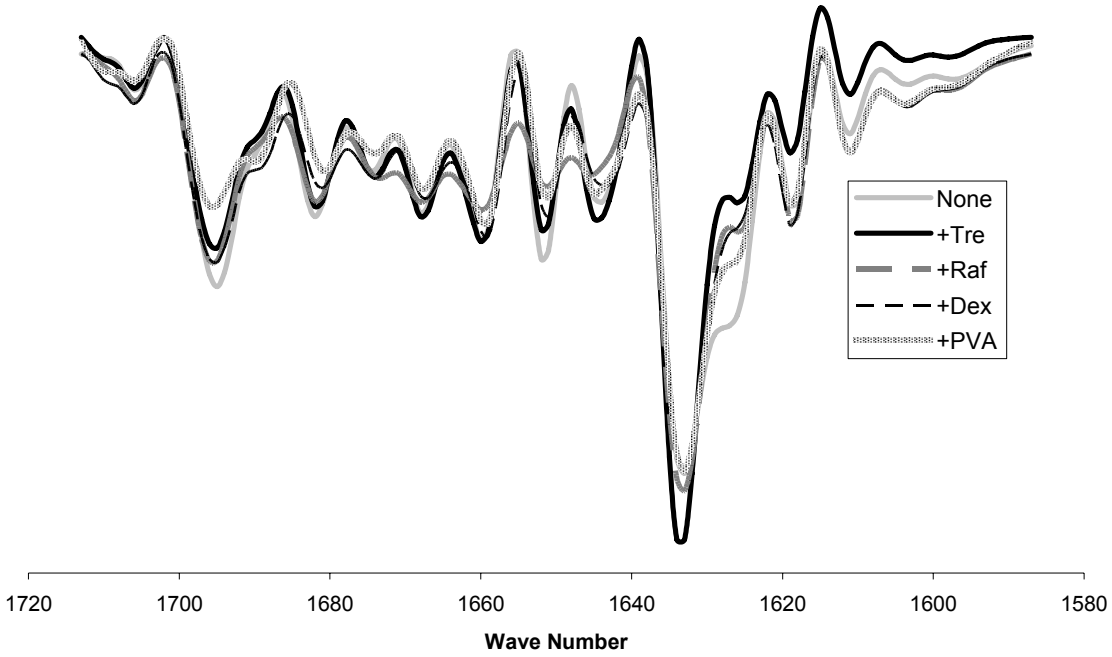


Figure 3.17. Overlaid FTIR spectra for con A lyophilized formulations with various excipients (1:1, w/w).

Table 3.2. FTIR correlation coefficients (R values) between protein with no excipients and with excipient.

Protein	**+Tre**	**+Raf**	**+Dex**	**+PVA**
Myoglobin	0.94	0.95	0.97	0.92
Lysozyme	0.95	0.97	0.94	0.94
RNase A	0.97	0.96	0.92	0.90
β-Lactoglobulin	0.97	0.98	0.98	0.91
EC5	0.99	0.99	0.99	0.99
Con A	0.97	0.98	0.98	0.98

(Table 3.2). All the table entries are greater than or equal to 0.90, indicating that the spectra for samples containing excipients are similar to the excipient-free controls for all seven of the proteins studied. Correlation coefficients were greatest for the EC-5 spectra, will all values ~ 0.99 indicating no effect of excipients. In general, correlation coefficients were greater for the β-sheet proteins than for the α-helical proteins (Table 3.2), indicating a greater effect of excipients on solid structure for the latter group.

3.3.3 Thermal analysis - Thermogravimetric analysis was performed to measure the water content in myoglobin and β-lactoglobulin samples after exposure to 33% RH for 72 . Regardless of excipient, the water content did not exceed 7.5% (Table 3.3), and was highest in dextran and PVP formulations.

DSC studies were performed on amorphous formulations to determine the glass transition temperature (T_g). At a given temperature, increasing T_g values are associated with a more glassy solid matrix with lower molecular mobility[61, 62]. Since formulations containing Gdn•HCl were partially crystalline, their T_g values were not measured. For myoglobin, the T_g values increased in the order PVA < trehalose < raffinose < dextran < PVP while for β-lactoglobulin the order was trehalose < PVA < raffinose < dextran ~ PVP.

3.4. Discussion

Hydrogen/deuterium exchange (HDX) has long been used to analyze protein structure[40], protein structural stability[63], protein folding-unfolding mechanisms[30, 32, 33, 64] and protein ligand interactions[65, 66] in solution. Traditionally HDX has been used in conjunction with

Table 3.3. Moisture content and glass transition temperatures for myoglobin and β-lactoglobulin formulations after exposure to 33% RH for 72 h.

		Trehalose	**Raffinose**	**Dextran**	**PVA**	**PVP**	**None**
Myoglobin	Moisture Content (%)	4.53 ± 0.25	4.42 ± 0.09	6.11 ± 0.79	5.12 ± 0.2	7.00 ± 0.85	3.94 ± 0.64
	Tg (^{O}C)	52.41 ± 0.98	56.18 ± 0.52	95.06 ± 0.95	41.72 ± 0.49	96.3 ± 2.09	NA
β-Lactoglobulin	Moisture Content (%)	6.18 ± 0.38	5.67 ± 0.24	7.23 ± 0.23	4.68 ± 0.59	6.32 ± 0.78	6.85 ± 0.44
	Tg (^{O}C)	43.53 ± 0.58	47.34 ± 2.17	94.93 ± 0.58	44.41 ± 0.06	94.69 ± 0.50	NA

NMR analysis[67], but in recent years HDX with mass spectrometry has emerged as a powerful technique capable of investigating protein structure[68, 69] and dynamics under a variety of conditions. This application of this technique to protein dynamics in solution has been covered in great detail and has been reviewed extensively[31, 70-73].

The study of protein-excipient interactions in amorphous solids by HDX with LC/+ESI-MS analysis was initiated in our labs using a predominantly α-helical protein, calmodulin, with no disulfide bonds. We observed that co-lyophilization with a variety of excipients protected calmodulin from HDX in the solid state[11, 36, 37]. Here, we have extended this method to solid samples of other proteins with a range of secondary structures, both with and without disulfide bonds. The studies test the hypothesis that the ability of excipients to protect proteins from HDX in amorphous solids depends on both excipient type and protein structure, and that this effect is exerted non-uniformly along the protein sequence. The results for intact proteins (Fig. 3.1) and for peptic digests of (Fig. 3.2, 3.3, 3.4, 3.5, 3.6) support this hypothesis and demonstrate that the technique provides detailed information on protein-excipient interactions in amorphous solids with peptide-level resolution.

Several mechanisms have been proposed to explain protein-excipient interactions during lyophilization and storage in amorphous solids, and the means by which these interactions influence protein conformation and stability. (i) The "vitrification hypothesis" asserts that effective excipients produce a solid with high Tg and limited mobility that "locks" the protein in its native conformation[61, 62]. The "water replacement hypothesis" proposes that effective excipients form hydrogen bonds to the protein, replacing the hydrogen bonds to water that stabilize protein conformation in solution[74, 75]. The "preferential hydration hypothesis" suggests that effective

excipients are excluded from the protein surface, preserving a hydration shell that promotes structure[74, 76-79]. The results presented here provide information on several of these hypotheses.

Comparison of the HDX data (Figs. 3.1) and Tg data (Table 3.3) for myoglobin and β-lactoglobulin provide a test of the vitrification hypothesis. The hypothesis is supported if the proteins show less exposure to exchange, consistent with a more compact conformation, in glassy matrices with high Tg. The data do not provide this support. Trehalose and raffinose showed significant protection against exchange for both proteins (Figs. 3.1) but have the lowest T_g indicating greater mobility of the matrix. In contrast, the polymers dextran and PVP showed higher T_g values (Table 3.3), but limited protection from exchange (Figs. 3.1). FTIR data provide additional evidence for retention of secondary structure for both proteins co-lyophilized with trehalose and raffinose, and a loss of structure with dextran (particularly for myoglobin).

Previous tests of the "water replacement hypothesis" have monitored the carboxylate absorption band in FTIR spectra, showing hydrogen bonding between the protein's carboxylate groups and the excipient in lyophilized solids. These hydrogen bonds may involve the side-chains or the backbone of the protein, and cannot be distinguished readily by FTIR. In solid-state HDX, only the protein backbone behavior can be monitored as any deuterium incorporation in the side-chains is immediately lost to back exchange as the solid is reconstituted for analysis. Thus, the two methods provide complementary information on hydrogen bonding interactions for proteins in lyophilized solids.

The peptic digest data for calmodulin (Fig. 3.2), myoglobin (Fig. 3.3) and β-lactoglobulin (Fig. 3.4) show that the α-helical domains of these proteins are preferentially protected from HDX exchange by trehalose and raffinose, and to some extent, by dextran. In an intact α-helix, all backbone '*i*' amide protons participate in hydrogen bonds to the *i* + *4* carbonyl residues. This suggests that excipients that protect α-helical regions from exchange promote these i → i + 4 hydrogen bonds rather than competing for them. Previous reports[51, 58, 80] have shown that protein-excipient hydrogen bonds involve side-chains groups and not the main chain. Hence, the low exchange for α-helical domains of the proteins studied here is consistent with hydrogen bonding of the excipients to protein side-chains, promoting the helix structure and inhibiting D_2O access to the backbone.

Some protection from exchange is also provided by the carbohydrate excipients in the β-sheet regions, albeit to a lesser extent than in the α-helical regions. β-sheets and strands are stabilized by hydrogen bonds involving alternate imino hydrogens and carbonyl groups. β-lactoglobulin has a β-barrel structure consisting of seven anti-parallel β-strands which form a closed structure, with the first strand H-bonded to the last. Low exchange is observed in the β-sheet fragments of β-lactoglobulin that is relatively insensitive to excipient type (Fig. 3.4, with the exception of Gdn.HCl), suggesting that the β-barrel is relatively robust. Exchange in the β-sheet portions of EC-5, is much higher than in the β-barrel of β-lactoglobulin (Fig. 3.5), and may reflect the looseness of the (unpublished) structure of this fragment of E-cadherin.

Overall, the HDX method has been shown to provide detailed, region specific information about the interactions of proteins and excipients in the lyophilized state. The use of peptic digests in addition to the intact protein analysis allows excipient effects to be assigned to secondary structural domains of the proteins. The intact HDX exchange studies are sensitive to the nature of the excipient, with the extent of exchange varying by 2-5 fold for the proteins studied here. In contrast, FTIR studies of identical formulations showed correlation coefficients greater than 0.90 in all cases (Table 3.2), indicating largely similar spectra that shows little variation with excipient type.

Though solid-state HDX with LC/+ESI-MS analysis is a promising technique, the technique has several limitations. High molecular weight proteins (>50,000) will definitely pose difficulties in +ESI/MS analysis, particularly with respect to peptic digestion followed by elution. FTIR is not subject to this size constraint. In the present studies, pepsin digestion was performed on proteins < 20 kD in three minutes followed by elution of all fragments within 7 minutes, keeping back-exchange to a minimum. For larger proteins, longer run times will be needed for full elution of the digest mixture so that back-exchange will be more prominent. Rapid digestion by pepsin is also necessary to minimize back exchange; lysozyme, con A and RNase A were not digestible in the present study. The use of an immobilized pepsin column which affords better digestibility could be an option[71]. Despite these limitations solid-state HDX analysis of intact proteins should be possible for proteins of all sizes, even for IgGs, studies may ultimately prove useful for excipient screening.

3.5. References

1. PhRMA, 418 biotechnology medicines in testing promise to bolster the arsenal against disease. *Pharmaceutical Research and Manufacturers of America: Washington, DC* **2006**, 52.
2. Byrn, S. R.; Xu, W.; Newman, A. W., Chemical reactivity in solid-state pharmaceuticals: formulation implications. *Adv Drug Deliv Rev* **2001,** 48, (1), 115-36.
3. Carpenter, J. F.; Chang, B. S.; Garzon-Rodriguez, W.; Randolph, T. W., Rational design of stable lyophilized protein formulations: theory and practice. In *Rational Design of Stable Protein Formulations*, Carpenter, J. F.; Manning, M. C., Eds. Kluwer Academic / Plenum Publishers: New York, 2002; pp 109-133.
4. Costantino, H. R.; Schwendeman, S. P.; Langer, R.; Klibanov, A. M., Deterioration of lyophilized pharmaceutical proteins. *Biochemistry (Mosc)* **1998,** 63, (3), 357-63.
5. Lai, M. C.; Topp, E. M., Solid-state chemical stability of proteins and peptides. *J Pharm Sci* **1999,** 88, (5), 489-500.
6. Roberts, C. J.; Debenedetti, P. G. E. p. s. w. a. s. A. J., 48, (6), 1140-1144., Engineering pharmaceutical stability with amorphous solids. *AIChE Journal* **2002,** 48, (6), 1140-1144.
7. Song, Y.; Wilson, A. D.; Li, R.; Hageman, M. J.; Schowen, R. L.; Topp, E. M., Solid-state chemical stability of peptides and proteins: application to controlled release formulations. In *Handbook of Pharmaceutical Controlled*

Release Technology,, Wise, D. L., Ed. Marcel Dekker, Inc: New York, 2000; pp 693-724.

8. Stotz, C. E.; Winslow, S. L.; Houchin, M. L.; D'Souza, A. J.; Ji, J. T., Degradation pathways for lyophilized peptides and proteins. In *Lyophilization of Biomaterials*, Pikal, M. J.; Costantino, H. R., Eds. 2003.

9. Xie, M.; Schowen, R. L., Secondary structure and protein deamidation. *J Pharm Sci* **1999,** 88, (1), 8-13.

10. Yu, L., Amorphous pharmaceutical solids: preparation, characterization and stabilization. *Adv Drug Deliv Rev* **2001,** 48, (1), 27-42.

11. Li, Y.; Williams, T. D.; Topp, E. M., Effects of Excipients on Protein Conformation in Lyophilized Solids by Hydrogen/Deuterium Exchange Mass Spectrometry. *Pharm Res* **2007**.

12. Abdul-Fattah, A. M.; Truong-Le, V.; Yee, L.; Pan, E.; Ao, Y.; Kalonia, D. S.; Pikal, M. J., Drying-induced variations in physico-chemical properties of amorphous pharmaceuticals and their impact on Stability II: stability of a vaccine. *Pharm Res* **2007,** 24, (4), 715-27.

13. Andya, J. D.; Hsu, C. C.; Shire, S. J., Mechanisms of aggregate formation and carbohydrate excipient stabilization of lyophilized humanized monoclonal antibody formulations. *AAPS PharmSci* **2003,** 5, (2), E10.

14. Katayama, D. S.; Kirchhoff, C. F.; Elliott, C. M.; Johnson, R. E.; Borgmeyer, J.; Thiele, B. R.; Zeng, D. L.; Qi, H.; Ludwig, J. D.; Manning, M. C., Retrospective statistical analysis of lyophilized protein formulations of progenipoietin using PLS: determination of the critical parameters for long-term storage stability. *J Pharm Sci* **2004,** 93, (10), 2609-23.

15. Klibanov, A. M.; Schefiliti, J. A., On the relationship between conformation and stability in solid pharmaceutical protein formulations. *Biotechnol Lett* **2004,** 26, (14), 1103-6.

16. Webb, S. D.; Cleland, J. L.; Carpenter, J. F.; Randolph, T. W., Effects of annealing lyophilized and spray-lyophilized formulations of recombinant human interferon-gamma. *J Pharm Sci* **2003,** 92, (4), 715-29.

17. Chen, B.; Costantino, H. R.; Liu, J.; Hsu, C. C.; Shire, S. J., Influence of calcium ions on the structure and stability of recombinant human deoxyribonuclease I in the aqueous and lyophilized states. *J Pharm Sci* **1999,** 88, (4), 477-82.

18. Prestrelski, S. J.; Tedeschi, N.; Arakawa, T.; Carpenter, J. F., Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophys J* **1993,** 65, (2), 661-71.

19. Schule, S.; Friess, W.; Bechtold-Peters, K.; Garidel, P., Conformational analysis of protein secondary structure during spray-drying of antibody/mannitol formulations. *Eur J Pharm Biopharm* **2007,** 65, (1), 1-9.

20. Izutsu, K.-I.; Fujimaki, Y.; Kuwabara, A.; Hiyama, Y.; Yomota, C.; Aoyagi, N., Near-infrared analysis of protein secondary structure in aqueous solutions and freeze-dried solids. *Journal of Pharmaceutical Sciences* **2006,** 95, (4), 781-789.

21. Hu, H. Y.; Li, Q.; Cheng, H. C.; Du, H. N., Beta-sheet structure formation of proteins in solid state as revealed by circular dichroism spectroscopy. *Biopolymers* **2001,** 62, (1), 15-21.

22. Forbes, R. T.; Barry, B. W.; Elkordy, A. A., Preparation and characterisation of spray-dried and crystallised trypsin: FT-Raman study to detect protein denaturation after thermal stress. *Eur J Pharm Sci* **2007,** 30, (3-4), 315-23.

23. Hill, J. J.; Shalaev, E. Y.; Zografi, G., Thermodynamic and dynamic factors involved in the stability of native protein structure in amorphous solids in relation to levels of hydration. *J Pharm Sci* **2005,** 94, (8), 1636-67.

24. Hong, M., Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR. *Structure* **2006,** 14, (12), 1731-40.

25. Zhou, D. H.; Shah, G.; Cormos, M.; Mullen, C.; Sandoz, D.; Rienstra, C. M., Proton-detected solid-state NMR spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J Am Chem Soc* **2007,** 129, (38), 11791-801.

26. Franks, W. T.; Kloepper, K. D.; Wylie, B. J.; Rienstra, C. M., Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J Biomol NMR* **2007,** 39, (2), 107-31.

27. Kovacs, F. A.; Fowler, D. J.; Gallagher, G. J.; Thompson, L. K., A practical guide for solid state NMR distance measurements in proteins. *Concepts in Magnetic Resonance, Part A: Bridging Education and Research* **2007,** 30A, (1), 21-39.

28. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q Rev Biophys* **1983,** 16, (4), 521-655.

29. Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G., Protein analysis by hydrogen exchange mass spectrometry. *Annu Rev Biophys Biomol Struct* **2003,** 32, 1-25.

30. Rist, W.; Jorgensen, T. J.; Roepstorff, P.; Bukau, B.; Mayer, M. P., Mapping temperature-induced conformational changes in the Escherichia coli heat shock transcription factor sigma 32 by amide hydrogen exchange. *J Biol Chem* **2003,** 278, (51), 51415-21.

31. Smith, D. L.; Deng, Y.; Zhang, Z., Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom* **1997,** 32, (2), 135-46.

32. Wang, F.; Li, W.; Emmett, M. R.; Marshall, A. G.; Corson, D.; Sykes, B. D., Fourier transform ion cyclotron resonance mass spectrometric detection of small Ca(2+)-induced conformational changes in the regulatory domain of human cardiac troponin C. *J Am Soc Mass Spectrom* **1999,** 10, (8), 703-10.

33. Zhu, M. M.; Rempel, D. L.; Zhao, J.; Giblin, D. E.; Gross, M. L., Probing Ca2+-induced conformational changes in porcine calmodulin by H/D exchange and ESI-MS: effect of cations and ionic strength. *Biochemistry* **2003,** 42, (51), 15388-97.

34. French, D. L.; Arakawa, T.; Li, T., Fourier transform infared spectroscopy investigation of protein conformation in spray-dried protein/trehalose powders. *Biopolymers* **2004,** 73, (4), 524-31.

35. Desai UR, O. J., Klibanov AM, Protein structure in the lyophilized state: A hydrogen isotope exchange/NMR study with bovine pancreatic trypsin inhibitor. *Journal of American Chemical Society* **1994,** 116, 9420-9422.

36. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Characterizing protein structure in amorphous solids using hydrogen/deuterium exchange with mass spectrometry. *Anal Biochem* **2007,** 366, (1), 18-28.

37. Li, Y.; Williams, T. D.; Schowen, R. L.; Topp, E. M., Trehalose and calcium exert site-specific effects on calmodulin conformation in amorphous solids. *Biotechnol Bioeng* **2007,** 97, (6), 1650-3.

38. Zheng, K.; Middaugh, C. R.; Siahaan, T. J., Evaluation of the physical stability of the EC5 domain of E-cadherin: Effects of pH, temperature, ionic strength, and disulfide bonds. *J Pharm Sci* **2008**.

39. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* **1993,** 2, (4), 522-31.

40. Yan, X.; Zhang, H.; Watson, J.; Schimerlik, M. I.; Deinzer, M. L., Hydrogen/deuterium exchange and mass spectrometric analysis of a protein containing multiple disulfide bonds: Solution structure of recombinant macrophage colony stimulating factor-beta (rhM-CSFbeta). *Protein Sci* **2002,** 11, (9), 2113-24.

41. Kendrick, B. S.; Dong, A.; Allison, S. D.; Manning, M. C.; Carpenter, J. F., Quantitation of the area of overlap between second-derivative amide I infrared spectra to determine the structural similarity of a protein in different states. *J Pharm Sci* **1996,** 85, (2), 155-8.

42. Ciurczak, E. W., Uses of Near-infrared spectroscopy in pharmaceutical analysis. **1987,** 23, 147-163.

43. Costantino, H. R.; Andya, J. D.; Shire, S. J.; Hsu, C. C., Fourier-transform infrared spectroscopic analysis of the secondary structure of recombinant humanized immunoglobulin G. *Pharmaceutical Sciences* **1997,** 3, (3), 121-128.

44. Costantino, H. R.; Carrasquillo, K. G.; Cordero, R. A.; Mumenthaler, M.; Hsu, C. C.; Griebenow, K., Effect of excipients on the stability and structure of lyophilized recombinant human growth hormone. *J Pharm Sci* **1998,** 87, (11), 1412-20.

45. Costantino, H. R.; Nguyen, T. H.; Hsu, C. C., Fourier-transform infrared spectroscopy demonstrates that lyophilization alters the secondary structure of recombinant human growth hormone. *Pharmaceutical Sciences* **1996,** 2, (5), 229-232.

46. Imamura, K.; Ogawa, T.; Sakiyama, T.; Nakanishi, K., Effects of types of sugar on the stabilization of protein in the dried state. *J Pharm Sci* **2003,** 92, (2), 266-74.

47. Izutsu, K.; Aoyagi, N.; Kojima, S., Protection of protein secondary structure by saccharides of different molecular weights during freeze-drying. *Chem Pharm Bull (Tokyo)* **2004,** 52, (2), 199-203.

48. Lu, J.; Wang, Y.-X.; Ching, C.-B., Thermal and FTIR investigation of freeze-dried protein-excipient mixtures. *Journal of Thermal Analysis and Calorimetry* **2007,** 89, 913-919.

49. van de Weert, M.; Haris, P. I.; Hennink, W. E.; Crommelin, D. J., Fourier transform infrared spectrometric analysis of protein conformation: effect of sampling method and stress factors. *Anal Biochem* **2001,** 297, (2), 160-9.

50. Xu, Q.; Keiderling, T. A., Trifluoroethanol-induced unfolding of concanavalin A: equilibrium and time-resolved optical spectroscopic studies. *Biochemistry* **2005,** 44, (22), 7976-87.

51. Allison, S. D.; Chang, B.; Randolph, T. W.; Carpenter, J. F., Hydrogen bonding between sugar and protein is responsible for inhibition of dehydration-induced protein unfolding. *Arch Biochem Biophys* **1999,** 365, (2), 289-98.

52. Cleland, J. L.; Lam, X.; Kendrick, B.; Yang, J.; Yang, T. H.; Overcashier, D.; Brooks, D.; Hsu, C.; Carpenter, J. F., A specific molar ratio of stabilizer to protein is required for storage stability of a lyophilized monoclonal antibody. *J Pharm Sci* **2001,** 90, (3), 310-21.

53. Wang, W., Lyophilization and development of solid protein pharmaceuticals. *Int J Pharm* **2000,** 203, (1-2), 1-60.

54. Dong, A.; Huang, P.; Caughey, W. S., Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry* **1990,** 29, (13), 3303-8.

55. Smith, B. M.; Franzen, S., Single-Pass Attenuated Total Reflection Fourier Transform Infrared Spectroscopy for the Analysis of Proteins in H2O Solution. *Anal. Chem.* **2002,** 74, (16), 4076-4080.

56. Meersman, F.; Smeller, L.; Heremans, K., Comparative Fourier transform infrared spectroscopy study of cold-, pressure-, and heat-induced unfolding and aggregation of myoglobin. *Biophys J* **2002,** 82, (5), 2635-44.

57. Luthra, S.; Kalonia, D. S.; Pikal, M. J., Effect of hydration on the secondary structure of lyophilized proteins as measured by fourier transform infrared (FTIR) spectroscopy. *J Pharm Sci* **2007,** 96, (11), 2910-21.

58. Carpenter, J. F.; Crowe, J. H., An infrared spectroscopic study of the interactions of carbohydrates with dried proteins. *Biochemistry* **1989,** 28, (9), 3916-22.

59. Ishida, K. P.; Griffiths, P. R., Comparison of the amide I/II intensity ratio of solution and solid-state proteins sampled by transmission, attenuated total reflectance, and diffuse reflectance spectrometry. *Applied Spectroscopy* **1993,** 47, (5), 584-589.

60. Bentaleb, A.; Abele, A.; Haikel, P.; Schaaf, P.; Voegel, J. C., FTIR-ATR and radiolabeling study of the adsorption of ribonuclease A onto hydrophilic surfaces: Correlation between the exchange rate and the interfacial denaturation. *Langmuir* **1998,** 14, 6493-6500.

61. Franks, F., Freeze drying: From empiricism to predictability. *Cryo-Letters* **1990,** 11, 93-110.

62. Franks, F.; Hatley, R. H. M.; Mathias, S. F., Materials science and the production of shelf-stable biologicals. *Pharm. Technol. Int.* **1991,** 3, 24-34.

63. Chi, Y. H.; Kumar, T. K.; Kathir, K. M.; Lin, D. H.; Zhu, G.; Chiu, I. M.; Yu, C., Investigation of the structural stability of the human acidic fibroblast growth factor by hydrogen-deuterium exchange. *Biochemistry* **2002,** 41, (51), 15350-9.

64. Lanman, J.; Lam, T. T.; Barnes, S.; Sakalian, M.; Emmett, M. R.; Marshall, A. G.; Prevelige, P. E., Jr., Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J Mol Biol* **2003,** 325, (4), 759-72.

65. Yan, X.; Deinzer, M. L.; Schimerlik, M. I.; Broderick, D.; Leid, M. E.; Dawson, M. I., Investigation of ligand interactions with human RXRalpha by hydrogen/deuterium exchange and mass spectrometry. *J Am Soc Mass Spectrom* **2006,** 17, (11), 1510-7.

66. Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; He, Y.; Hendrickson, C. L.; Marshall, A. G.; Griffin, P. R., Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **2006,** 78, (4), 1005-14.

67. Dyson, H. J.; Wright, P. E., Unfolded proteins and protein folding studied by NMR. *Chem Rev* **2004,** 104, (8), 3607-22.

68. Abzalimov, R. R.; Kaltashov, I. A., Extraction of local hydrogen exchange data from HDX CAD MS measurements by deconvolution of isotopic distributions of fragment ions. *J Am Soc Mass Spectrom* **2006,** 17, (11), 1543-51.

69. Kaltashov, I. A.; Eyles, S. J., Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom Rev* **2002,** 21, (1), 37-71.

70. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* **2006,** 25, (1), 158-70.

71. Wang, L.; Pan, H.; Smith, D. L., Hydrogen exchange-mass spectrometry: optimization of digestion conditions. *Mol Cell Proteomics* **2002,** 1, (2), 132-8.

72. Englander, S. W., Hydrogen exchange and mass spectrometry: A historical perspective. *J Am Soc Mass Spectrom* **2006,** 17, (11), 1481-9.

73. Yan, X.; Watson, J.; Ho, P. S.; Deinzer, M. L., Mass spectrometric approaches using electrospray ionization charge states and hydrogen-deuterium exchange for determining protein structures and their conformational changes. *Mol Cell Proteomics* **2004,** 3, (1), 10-23.

74. Tzannis, S. T.; Prestrelski, S. J., Moisture effects on protein-excipient interactions in spray-dried powders. Nature of destabilizing effects of sucrose. *J Pharm Sci* **1999,** 88, (3), 360-70.

75. Arakawa, T.; Kita, Y.; Carpenter, J. F., Protein--solvent interactions in pharmaceutical formulations. *Pharm Res* **1991,** 8, (3), 285-91.

76. Lee, J. C.; Timasheff, S. N., The stabilization of proteins by sucrose. *J Biol Chem* **1981,** 256, (14), 7193-201.

77. Timasheff, S. N., Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* **1992,** 31, (41), 9857-64.

78. Timasheff, S. N., Solvent stabilization of protein structure. *Methods Mol Biol* **1995,** 40, 253-69.

79. Carpenter, J. F.; Crowe, J. H., The mechanism of cryoprotection of proteins by solutes. *Cryobiology* **1988,** 25, (3), 244-55.

80. Remmele, R. L., Jr.; Stushnoff, C.; Carpenter, J. F., Real-time in situ monitoring of lysozyme during lyophilization using infrared spectroscopy: dehydration stress in the presence of sucrose. *Pharm Res* **1997,** 14, (11), 1548-55.

Chapter 4

LC/MS alternatives to oligosaccharide mapping for quantitation of N-linked glycoforms in recombinant IgG1

4.1. Introduction

Glycosylation is an important determinant of the stability and biodisposition of protein drugs, including recombinant immunoglobulins (IgGs). The challenges involved in characterizing glycosylation in recombinant therapeutic proteins differ somewhat from those associated with analyzing endogenous proteins. Studies of glycosylation in human and animal tissues are often focused on the general variation in glycosylation patterns of a wide array of proteins, as in proteomic experiments. In contrast, the analysis of therapeutic proteins is focused on the complete characterization and quantitation of glycosylation in the protein drug product. Glycosylation in recombinant bio-therapeutic proteins varies widely with the production conditions used during manufacturing [1,2] and can influence efficacy, folding, target binding and pharmacokinetic properties [3-5]. Both the variability and physiological effects of glycosylation make it important to accurately quantify the carbohydrate structures found in bio-therapeutic proteins.

Immunoglobulin G molecules (IgGs), a subclass of immunoglobulins, have become attractive as therapeutic proteins due to their high specificity and long circulation life [6]. An IgG is a multi-chain, symmetric protein consisting of two identical Fab arms (Fab = **F**ragment, **a**ntigen **b**inding) and a conserved Fc stem (Fc = **F**ragment, **c**rystallizable) connected through a flexible hinge [7]. The Fab arms are composed of a light chain (LC) connected through disulfide bonds to a portion of the heavy chain (HC). The remaining portions of the two HCs are linked to form the homo-dimeric Fc stem. The Fc sequence is highly conserved in IgG molecules and contains a single N-glycosylation site, Asn297 [7]. Glycosylation in the Fc defines the structure of the CH2 domain and has been shown to be important for the effector

functions of the Fc [8]. Unlike most proteins in which the carbohydrates are exposed, the carbohydrate moiety in the Fc is buried between the two CH2 domains [9] where space constraints restrict the extent of carbohydrate branching. Hence, the typical glycoform found in the Fc is the biantennary carbohydrate structure [8]. The common variability in glycosylation of IgG molecules is introduced by incomplete processing of the galactose and fucose residues from the biantennary oligosaccharide. In some cases, additional heterogeneity is introduced by the presence of high mannose glycoforms which are highly branched precursors of the biantennary carbohydrates [6]. While IgGs are symmetric with regard to the amino acid sequences of the light and heavy chains, glycosylation may be either symmetric or asymmetric[10]. Since each glycoform has a specific mass determined by its composition, mass spectrometry (MS) can be used to identify glycoforms. Recent advances in reversed-phase chromatography (rp-LC) and electrospray ionization mass spectrometry (ESI-MS) have made it possible to analyze glycoforms in samples of intact protein, as well as in protein fragments and in peptides generated after complete proteolysis with specific enzymes[11,12]. Each of these protein sample preparation methods offers potential advantages and disadvantages for glycoform analysis by MS, a topic that has been addressed in several recent reviews [11]. Briefly, the analysis of glycosylation in intact proteins offers the advantage of minimal sample preparation and the ability to identify asymmetry in glycosylation, but the wide natural isotopic distribution of proteins may limit resolution[13]. In addition, since IgGs are highly hydrophobic, solvents such as isopropanol or n-propanol may be required for their reversed-phase separation. Techniques based on protein fragments or digests may offer higher resolution due to lower sample mass, but require more extensive sample preparation

(e.g., digestion, reduction, alkylation) that may introduce artifacts. Furthermore, though the quantitation of glycoforms using peak intensities from deconvoluted ESI-TOF MS spectra has been reported by Gadgil[14,15] and others[16], concerns remain regarding the accuracy and reproducibility of concentration determinations by this method, regardless of sample preparation. An alternative approach involves chromatographic analysis of glycans released from the protein through enzymatic[17] (e.g., peptide N-glycosidases) or chemical (e.g., β-elimination) procedures [11]. This "sugar release assay" is relatively straightforward and well-established, but information on the site of the protein-carbohydrate bond and potential asymmetry in glycosylation of the two heavy chains is lost with this approach. Sample preparation is also more time-consuming for the sugar release assay than for many of the LC/MS methods. Despite these limitations, the sugar release assay is generally considered the standard for glycoform analysis in the biopharmaceutical industry.

The studies reported here compare six methods for quantifying glycosylation in two production lots of a IgG: (i) LC/ESI-MS analysis of intact IgG ("intact IgG method"), (ii) LC/ESI-MS analysis of the Fc fragment produced by limited proteolysis with Lys-C ("IgG Fc method"), (iii) LC/ESI-MS analysis of the IgG heavy chain produced by reduction ("IgG HC method"), (iv) LC/ESI-MS analysis of Fc/2 fragment produced by limited proteolysis and reduction ("IgG Fc/2 method"), (v) LC/MS analysis of the glycosylated tryptic fragment (293EEQYNSTYR301) using extracted ion chromatograms ("XIC method") and (vi) normal phase HPLC analysis of sugars cleaved from the IgG using PNgase F ("sugar release assay"). The studies test the hypothesis that the LC/MS-based methods (i.e., Methods i-v) provide identification and quantitation of glycoforms that is equivalent to the sugar release assay (i.e.,

Method vi). The studies were conducted at Amgen, Inc., Thousand Oaks, CA, under the direction of Dr. Himanshu Gadgil.

4.2. Materials and Methods

4.2.1. Materials. - Trifluoroacetic acid (TFA), formic acid (FA) and guanidine hydrochloride (GdnHCl) were obtained from Pierce (Rockford, IL). Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) and iodoaceatamide (IAM) were obtained from Sigma (St. Louis, MO). HPLC grade water and acetonitrile (ACN) were obtained from VWR international (West Chester, PA). Pepsin and trypsin were obtained from Roche (Indianapolis, IN). The IgG lots were produced and purified using processes proprietary to Amgen and kept frozen at –80^{o}C until used.

4.2.2. Sample pretreatment. - IgG samples were subjected to limited proteolysis and/or reduction to produce the IgG Fc, IgG HC and IgG Fc/2 fragments. Limited proteolysis was achieved by incubating the IgG samples with endoproteinase Lys-C at a protein:enzyme weight ratio of 400:1 in incubation buffer (0.1M Tris–HCl,pH 8.0). The incubation was carried out at 37 °C for 30 min. The reaction was quenched by lowering the pH to 4.5 with the addition of acetic acid. Reduction was achieved by incubating 0.5 mL of IgG or an IgG Fc fragment with Lys-C at a concentration of 2 mg/mL in denaturing buffer (7.5 M GdnHCl, 120 mM sodium acetate, pH 5.0) containing 5 mM TCEP, at 37 ºC for 30 min.

4.2.3. Reversed-phase chromatography. - Reversed-phase separation of intact IgG and IgG fragments was carried out on an Agilent 1100 HPLC system equipped with a Varian diphenyl 2 X 150mm column. 20 µg protein sample was typically injected and elution was achieved with a linear A-B gradient for 40 min where eluent

A was 0.1% aqueous TFA and eluent B was 0.1% TFA in 90% acetonitrile. The flow-rate and column temperature were maintained at 200 µl/min and 75 °C, respectively, throughout the run.

4.2.4. Mass spectrometry. - Mass spectrometric analysis was carried out on a Waters LCT premier equipped with an ESI source operated in the W mode. The capillary and cone voltages were set at 2500 and 80 V, respectively. The desolvation gas and source temperatures were set at 350° C and 80° C, respectively. All the other voltages were optimized to provide maximal signal intensity in each of the modes. All raw data was processed using Waters MassLynx MaxEnt 1 software to obtain deconvoluted mass.

4.2.5. Peptide mapping. - Reduced and alkylated IgG was buffer exchanged into digestion buffer (1 M Tris, 1 M urea and 20 mM hydroxylamine at pH 7.0) at a protein concentration of approximately 1 mg/mL using a NAP-5 column (Amersham Bioscience, Uppsala, Sweden) following the procedure described by the manufacturer. Trypsin digestion was carried out by incubating 1 mg/mL of sample (in digestion buffer) with 20 µg of trypsin at 37 ºC for 4 hours, followed by a second addition of 20 µg of trypsin. The mixture was allowed to incubate at 37 ºC for 4 additional hours. The resulting tryptic peptides were separated using a Waters Atlantis column, 2.0 mm x 250 mm. Approximately 20 µg of the digested material was injected on the column. Elution was achieved using a linear gradient from 100% buffer A (0.1%FA) to 50% buffer A and 50% buffer B (90% acetonitrile 0.085% FA) in 170 minutes. The flow rate was maintained at 0.2 ml/min and the column temperature was held at 50 ºC.

4.2.6. Sugar release assay. - The antibody samples were first diluted to 1 mg/ml in digestion buffer (provided with the kit) and deglycosylated by addition of PNGase F (Sigma St Louis, MO, USA) at a weight ratio of 1:100 (PNGase F : antibody) followed by incubation at 37^{o} C for 24 hrs. The cleaved glycoforms were then purified with a GlycoClean™ R cartridge from Prozyme (San Leandro, CA, USA) using the procedure described by the manufacturer. The purified glycoforms were then labeled with 2-aminobenzamide following the protocol in the Prozyme (San Leandro, CA, USA) labeling kit. Normal phase chromatography was used to separate the labeled carbohydrates. The separation was carried on an Agilent 1100 system equipped with an Amide-80, 4.6 mm X 250 mm, 5 µm pore size column from Tosoh Biosciences (Grove City, OH) and a fluorescence detector with the excitation wavelength set at 330 nm and the emission wave length set at 420 nm. Buffer A was 50 mM ammonium formate (pH 4.4) and buffer B was acetonitrile. The gradient employed was 20% to 53% Buffer A over 132 min at 0.4 mL/min, then 53% to 100% Buffer A over 5 min at 0.4 mL/min followed by 100% Buffer A for 5 min at 1 mL/min and re-equilibration in starting conditions for 5 min at 0.4 mL/min.

4.2.7. Statistical analysis. - Results of the LC/MS based assays were compared quantitatively with the standard sugar release assay using a paired t-test, $\alpha = 0.05$. The intact IgG and HC assays were excluded from this comparison because these assays detect paired glycoforms on dimeric proteins, and so cannot be compared quantitatively with the results of the sugar release assay.

4.3. Results

4.3.1. LC/MS analysis of intact IgG molecules. - Recent advances in rp-LC and ESI-MS have made LC/MS analysis of intact IgGs routine [12,13,18]. The diphenyl column used in this study allows IgG separation with acetonitrile and has been shown previously to resolve site-specific modifications in IgGs [19,20]. ESI is the preferred mode of ionization for the analysis of large proteins as it produces a multiply charged envelope in the m/z range of 2000 to 4000 that can be deconvoluted to obtain the nominal mass. A major constraint in the MS analysis of large proteins is their wide natural isotopic distribution [13]. Since the full maximum at half width (FMHW) of the isotopic distribution of an IgG molecule is approximately 40 Da, small mass changes introduced by modifications such as oxidation (+16 Da) are difficult to resolve for intact IgGs even with high resolution MS analysis. However, glycosylation variation in IgGs is usually associated with larger mass changes which can be analyzed by standard time of flight instruments with resolution between 5,000 and 15,000 Da [18].

The deconvoluted mass spectra of two different lots of a recombinant monoclonal IgG1 analyzed with rp-LC/ESI-TOF are shown in Figure 4.1. These spectra were obtained by deconvoluting the raw m/z spectra (not shown). Both lots of IgG showed multiple peaks which can be attributed to the galactose and fucose heterogeneity typically found on the N-linked sugar present in the conserved region of all IgG molecules. This typical sugar profile described in earlier reports [21] is summarized in Table 4.1. The $(G0F)_2$ peak from Figure 4.1 contains two biantennary sugars, one on each heavy chain. The G0F sugar form has three mannose (hexose), four N-acetylglucosamine residues and a fucose residue. This structure is the basic

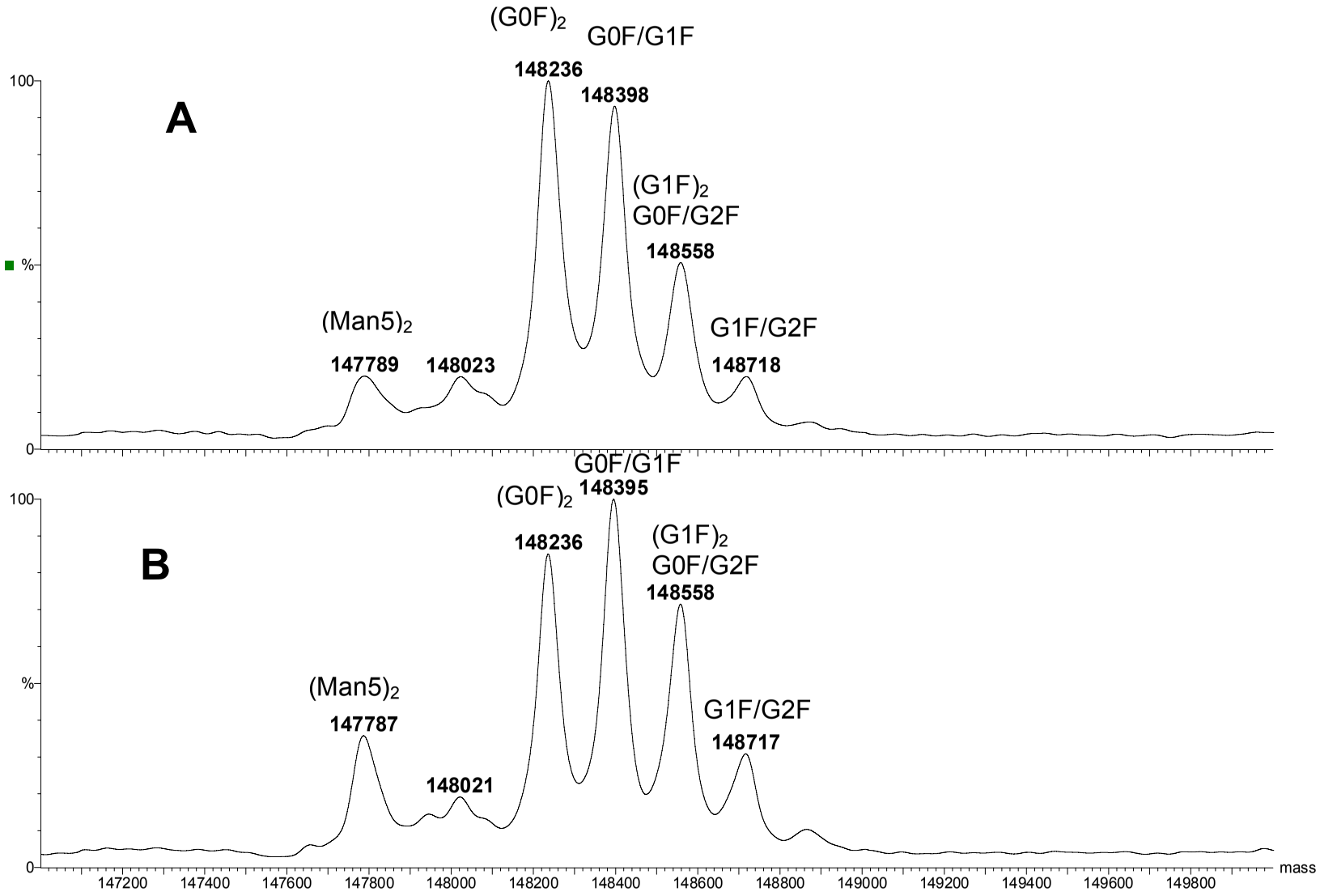
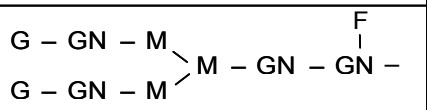
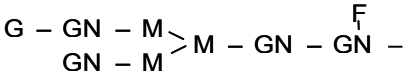
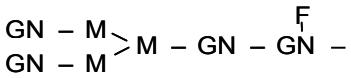
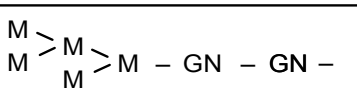
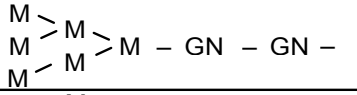
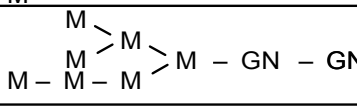
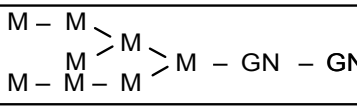
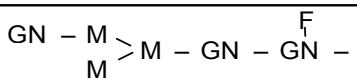
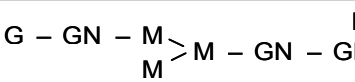
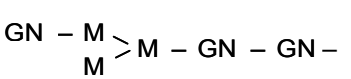
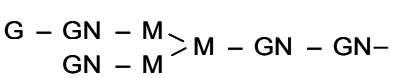
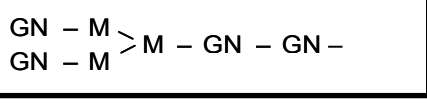


Figure 4.1: Deconvoluted spectra of intact A. IgGA and B. IgGB

Table 4.1: Structure, nomenclature and molecular weight of the carbohydrate moieties typically observed in recombinant monoclonal IgG molecules

Oligosaccharide Structure	Code	Mass (Da)
	G2F	1769.6
	G1F	1607.5
	G0F	1445.4
	Man5	1217.2
	Man6	1379.3
	Man7	1541.4
	Man8	1703.5
	G0F - GlcNac	1242.2
	G1F - GlcNac	1404.3
	G0 - GlcNac	1096.1
	G2	1623.5
	G1	1461.4
	G0	1299.3

core of the biantennary sugar. The mannose and N-acetylglucosamine residues in this structure are typically conserved during the production of IgG molecules and hence are unlikely to be the source of the identified heterogeneity. The terminal galactose residues however, show significant variability leading to the peaks G0F/G1F, $(G1F)_2$ etc, which are successively 162 Da apart. In addition to these galactose variants, Man5/Man5 and Man5/G0F structures were also observed. Man5, Man6, Man7, Man8 and Man9 are high mannose carbohydrate structures, which are precursors of the biantennary carbohydrates. The assignment of these peaks was based on their deconvoluted mass. The observed mass for the Man5/Man5 peak agrees well with its calculated mass of 147787. However, the Man5/G0F peak was nine Da less than its calculated mass. The mass for the Man5/G0F peak is within 100 Da of the mass of G0/G0, G0/G0F and Man5/Man6 glycoforms. As described earlier, the wide isotopic distribution for large proteins makes it difficult to fully resolve these forms. The incomplete resolution of the peaks could result in the larger mass error observed for the Man5/G0 form.

The deconvoluted spectra of the IgG from lots A and B show some differences in their glycosylation profiles. The IgG from lot B had a greater extent of the terminal galactose residues which was evident from the higher intensities of the G0F/G1F and $(G1F)_2$ peaks. We have shown previously that the intensities of the peaks from the deconvoluted spectra can be accurately used for quantitation of the hexose index (Hex_I) which is the molar ratio of galactose residues per molecule of IgG[18]. Additionally, the IgG from lot B also showed greater amounts of the high mannose oligosaccharide containing peaks than the IgG from Lot A.

These data indicate that LC/MS analysis of intact IgG can be an adequate method for the high level analysis of glycoforms on IgG molecules. Due to minimal sample preparation involved, this method is adaptable to high throughput analysis. A critical limitation that becomes apparent in comparison with the other methods is that low abundance glycoforms are not detected, particularly many of the mannose-containing glycoforms (see e.g., Fig. 4.2). In addition, individual glycoforms cannot be quantified, since the method only provides the total masses of paired glycoforms. Because of these limitations, the results of the intact IgG assay were excluded from further quantitative comparisons (see Table 4.2). Analysis of intact IgG could find application in lot release and product comparability assays, however.

4.3.2. LC/MS analysis of IgG-Fc. - The hinge region of an IgG is highly solvent exposed and susceptible to proteolytic cleavage. Limited proteolysis of IgG molecules has been widely used to generate IgG subunits which are generally more amenable to LC/MS analysis than the intact molecule[22,23]. Pepsin and papain have been classically used to clip below and above the hinge region to generate $F(ab)'_2$, Fab, and Fc fragments. Gadgil et al. have recently developed a method for the limited proteolysis of human IgG1 mAbs using endoproteinase Lys-C[12]. Limited proteolysis with Lys-C causes a single cleavage at the C-terminus of a lysine residue located in the hinge region to generate an Fc and two Fab fragments. Limited proteolysis with Lys-C maintains the disulfide structure of the Fc and Fab domains and hence allows the characterization of modifications in the disulfide architecture. In addition, limited proteolysis also conserves the oligosaccharide pairing in the CH2 domain.

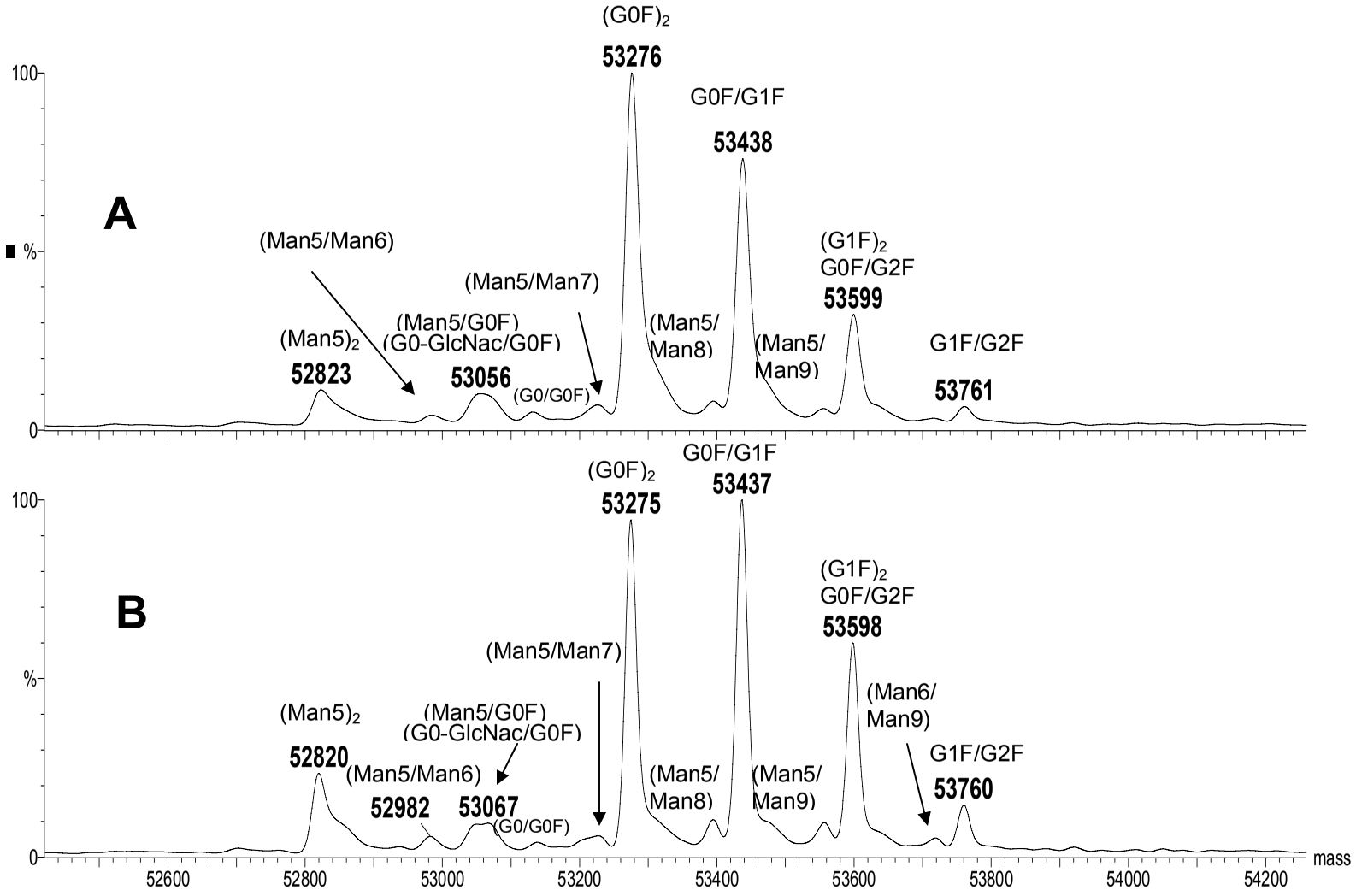


Figure 4.2: Deconvoluted spectra of intact IgG Fc

Table 4.2: Quantitation of the IgG glycoforms with various analytical methods: reduced (HC), limited and reduced, (Fc/2) extracted ion (XIC) and sugar release assay.

Sugars	Reduced (HC)		Limited & Reduced (Fc/2)		Extracted Ion (XIC)		Sugar Cleavage	
	IgG A (%)	IgG B (%)	IgG A (%)	IgG B (%)	IgG A (%)	IgG B (%)	IgG A (%)	IgG B (%)
Man5	9.6 ± 0.5	10.8 ± 0.5	9.0 ± 0.3	11.2 ± 0.05	17.2 ± 0.6	18.1 ± 1.4	11.0 ± 0.5	12.9 ± 0.5
G0F - GN	ND	ND	5.0 ± 0.8	5.4 ± 0.07	4.0 ± 0.2	6.2 ± 0.4	6.4 ± 0.2	6.0 ± 0.2
G1F - GN	ND	ND	ND	ND	ND	ND	1.8 ± 0.07	1.3 ± 0.2
G0	ND	ND	2.0 ± 0.02	ND	0.8 ± 0.1	ND	ND	ND
Man6	ND	ND	2.7 ± 0.2	2.6 ± 0.2	1.0 ± 0.2	1.3 ± 0.2	2.7 ± 0.2	3.0 ± 0.09
Man6*	ND	ND	ND	ND	ND	ND	0.6 ± 0.05	1.0 ± 0.04
G0F	55.8 ± 1.0	44.1 ± 0.3	39.7 ± 0.6	36.0 ± 0.3	68.3 ± 1.5	60.7 ± 1.2	46.9 ± 1.3	40.3 ± 0.7
G1	ND	ND	9.7 ± 0.3	6.67 ± 0.06	ND	ND	ND	ND
Man7	ND	ND	2.9 ± 0.09	1.86 ± 0.07	ND	ND	2.1 ± 0.3	2.8 ± 0.2
G1F	31.0 ± 1.4	38.0 ± 0.4	22.6 ± 0.1	26.4 ± 0.1	8.3 ± 0.6	12.9 ± 0.5	16.8 ± 0.3	19.3 ± 0.4
G1F*	ND	ND	ND	ND	ND	ND	6.4 ± 0.2	7.4 ± 0.4
G2	ND	ND	6.1 ± 0.3	4.9 ± 0.1	ND	ND	ND	ND
Man8	ND	ND	1.7 ± 0.05	0.73 ± 0.03	ND	ND	ND	ND
G2F	3.6 ± 0.5	7.0 ± 0.4	3.4 ± 0.1	4.21 ± 0.02	0.3 ± 0.06	0.7 ± 0.05	3.5 ± 0.2	4.7 ± 0.5

ND = not detected, * = isobaric form of that listed in the preceding row. N = 6 +/- standard deviation

The deconvoluted spectra of the IgG-Fc's from the two lots generated after limited proteolysis with Lys-C are shown in Figure 4.2. IgG-Fc, with a mass of approximately 50 kDa, is roughly one-third the molecular weight of the intact IgG molecule. Hence, IgG-Fc has a much smaller normal isotopic distribution with a FMHW of approximately 15 Da, which allows for improved resolution. This is evident from the spectra in Figure 4.2 which shows the isoform resolution of several additional peaks such as Man5/Man6, Man5/Man7, G0/G0F, etc. The pairing of different high mannose structures can lead to several isobaric peaks. For example, Man6/Man6 and Man5/Man7 both have the same mass, but for simplicity only one form (the most predominant) was used for labeling. Overall, this method of glycoform analysis of IgG-Fc yields at least 10 additional peaks as compared to the analysis of intact IgG. However, some forms of the carbohydrate could not be fully resolved in these analyses. The forms Man5/G0F and G-GlcNac/G0F vary in mass by only 25 Da and were not fully resolved. A partial resolution of these forms was obtained in the deconvoluted spectrum of lot B from Figure 4.2.

Analysis of intact IgG and IgG-Fc allows the determination of oligosaccharide pairing, an effect described previously by Masuda et al.[21]. Oligosaccharide pairing can lead to a symmetric or asymmetric Fc portion. Symmetric molecules have identical carbohydrates on each chain while asymmetric molecules have different carbohydrates on the two HCs. The study of pairing is important as each of the carbohydrates can independently control the effector functions of Fc. In both lots, the symmetric Man5/Man5 form was more abundant than the asymmetric Man5/G0F form. Since the G0F form was significantly greater than the Man5, a binomial distribution would lead to a greater amount of the asymmetric Man5/G0F form than

the Man5/Man5 form. However, in both lots, the Man5/Man5 form was greater in abundance than the Man5/G0F, indicating a preferential pairing of the Man/5Man5 form. This preferential pairing could be the result of structural limitations imposed on the asymmetric Man5/G0F form or could be caused by cell culture parameters such as antibody titer, production time or other factors inherent to the cell line itself. LC/MS analysis of the intact IgG and IgG-Fc is the only method that allows detection of oligosaccharide pairing in the IgG molecule.

While the detection of low abundance glycoforms is improved by the analysis of IgG-Fc rather than intact IgG, the method does not provide for the quantiation of individual glycoforms but only glycoform pairs (see below). The results of LC/MS analysis of IgG-Fc have thus been excluded from the quantitative comparison (see Table 4.2).

4.3.3. LC/MS analysis of IgG-HC. - Reduction of an IgG into individual HCs and LCs is another way to create IgG subunits and is often performed prior to analysis. The IgG HC which contains the carbohydrate is around 50 kDa, similar in size to the IgG-Fc. Analysis of the HC is more straightforward since reduction removes the complexity caused by the pairing of oligosaccharides when the two IgG HCs are linked, reducing the number of different analytes possible. For example, if five different glycoforms may be covalently attached to the heavy chains, the number of different masses expected for the HC fragment is five. In samples containing the dimeric heavy chain (i.e., IgG, IgG Fc), however, the number of possible masses is $2^5 = 32$, a consequence of the fact that different glycoforms may be linked to each of the heavy chains. Reduction of the IgG into monomeric HC fragments does result in fewer peaks, as shown in the deconvoluted spectra in Figure 4.3. The spectra of the

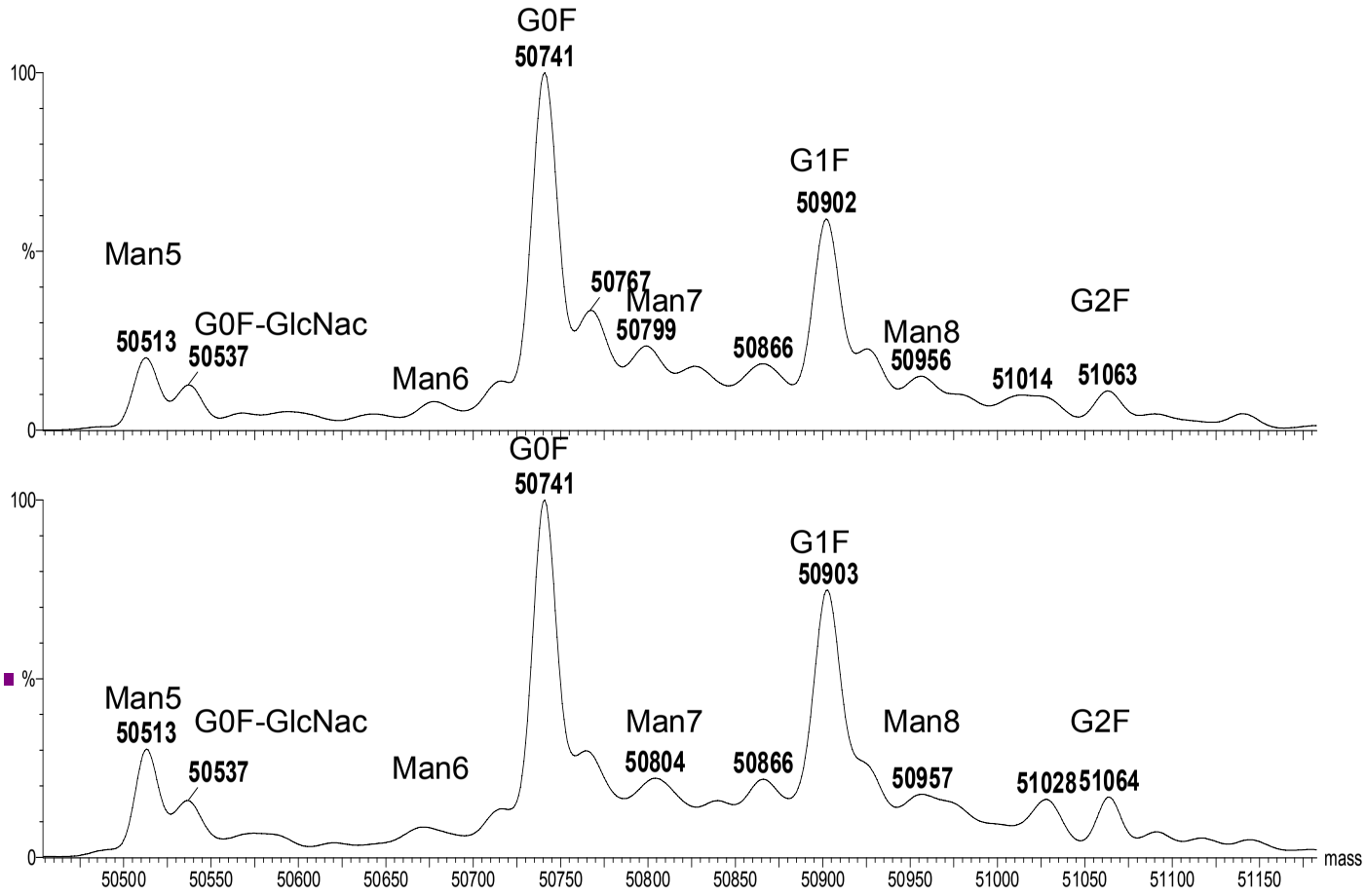


Figure 4.3: Deconvoluted spectra of A. IgGA HC and B. IgGB HC

HC show the biantennary sugars G0F, G1F and G2F along with smaller amounts of the high mannose forms. The paired glycoforms detected in intact IgG and IgG-Fc samples (e.g., Man5/Man5,G0F/G1F, Figs. 4.1, 4.2) are absent, however, as expected. Since in the analysis of IgG-HC the pairing effect is removed, the intensity of peaks for the various carbohydrate structures can be used to quantify the various glycoforms. The MaxEnt algorithm used for generating the deconvoluted spectra preserves the intensity information from the raw spectra, allowing accurate quantitation.

4.3.4. LC/MS analysis of IgG-Fc/2. - Fc/2 refers to the constant region of the single HC and is produced after reduction of the Fc. Fc/2 is approximately 25 kDa, is half the size of the HC and has a smaller normal isotopic distribution, which allows for greater resolution of modifications. The deconvoluted spectra of Fc/2 from the two different lots are shown in Figure 4.4. As compared to the deconvoluted spectra of HC (Fig. 4.3), the Fc/2 spectra showed improved resolution for the various glycoforms which was clearly observed in peaks such as Man5 and G0F-GlcNac. Additional low intensity peaks such as G0 were more clearly visible in the spectra for Fc/2. The improved detection of low intensity peaks could be the result of improved signal to noise of the more compact peaks in Fc/2. The higher sensitivity led to the identification of a greater number of glycoforms in the Fc/2 spectra. For example, while G0, G1 and G2 carbohydrates were not observed in the intact IgG, IgG-Fc or HC spectra, they were detected in the Fc/2 spectra. Similar to the previous analyses, the amount of the high mannose sugars was higher in lot B. Glycoforms with a mass difference as low as 25 Da were baseline resolved which subsequently allowed improved quantitation of these forms. All the peak assignments in the deconvoluted

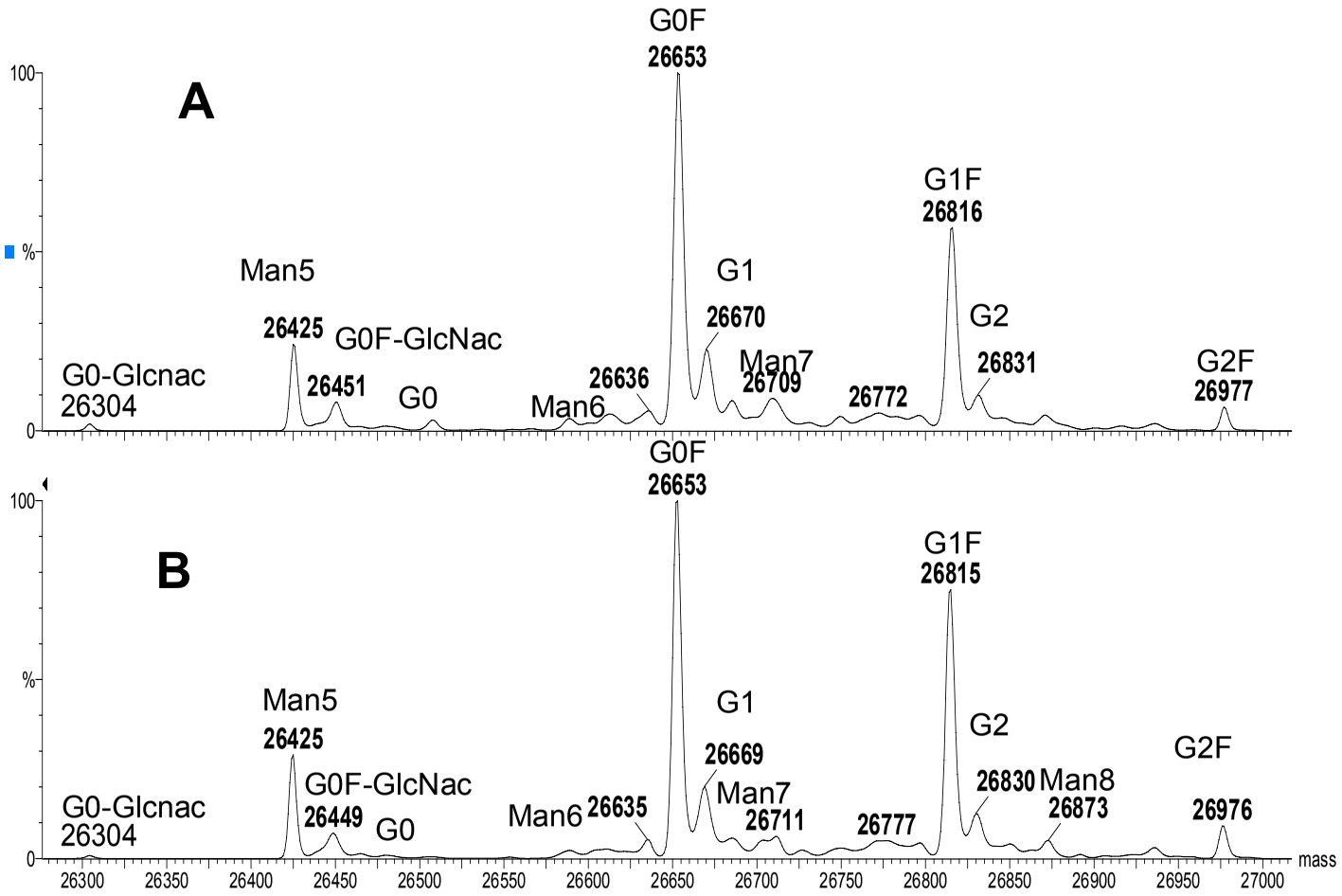


Figure 4.4: Deconvoluted spectra of A. IgGA Fc/2 and B. IgGB Fc/2

spectra were based on the calculated mass with errors less than 200 parts per million (ppm).

4.3.5. LC/MS analysis after trypsin digestion (XIC method). - Tryptic peptide mapping is commonly used to determine chemical modifications and sequence variants in proteins[24]. Peptide mapping relies on specific cleavage of the protein sequence with a proteolytic enzyme such as trypsin, giving rise to a known set of peptides. The LC/MS analysis allows determination of site specific modifications in proteins. Peptide mapping has been widely used for the characterization of IgG molecules. Complete trypsin cleavage of IgG1 molecules generates the peptide 293EEQYNSTYR301, which contains the N-linked carbohydrate moiety on N297. Standard reversed-phase separation methods may not be able to resolve the various glycoforms on the peptide. However, each glycoform (apart from isobaric structures) can be distinguished by its specific mass. The intensities specific to the glycoforms can be obtained from the total ion chromatogram (TIC) by generating extracted ion chromatograms (XIC). XICs are generated by extracting the ion signal for the mass of a particular peptide from the total ion chromatogram acquired on the mass spectrometer. This method allows the analysis of a specific compound in a mixture of analytes. Figure 4.5 shows the peptide maps (inlays) and XICs for the various glycoforms in the two lots. This method had a low sensitivity and only five glycoforms (Man5, Man6, G1F, G0F and G0F-GlcNAC) could be detected. XICs of other glycoforms, which were detected in the previous analyses, did not show measurable peaks (not shown). A difference in retention of the glycoforms was observed and highly branched structures (high mannose) had a shorter retention time than the

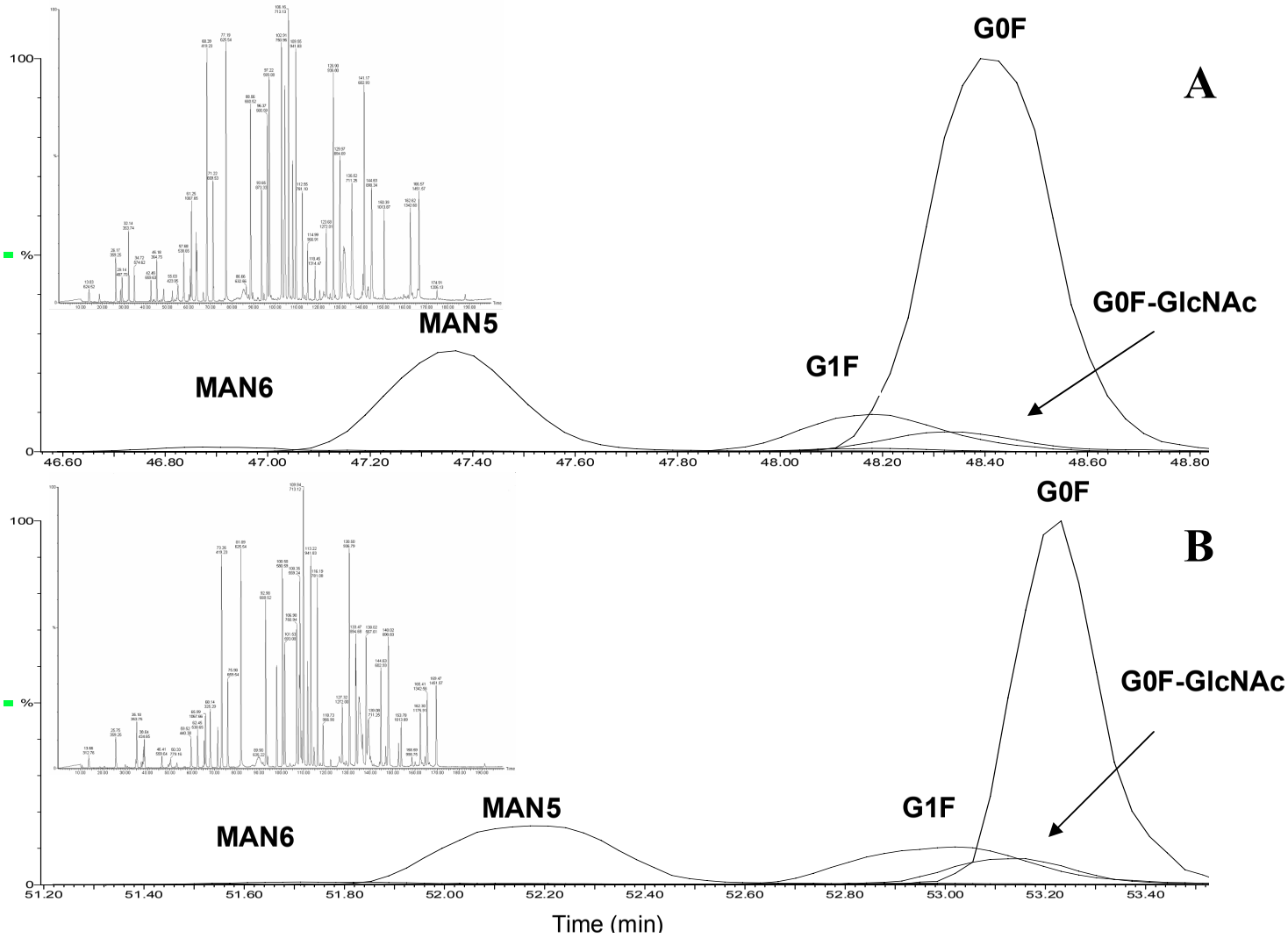


Figure 4.5: Extracted ion chromatograms for the various glycoforms. A. IgGA and B. IgGB

biantennary structures. Similarly, the size of the carbohydrate moiety also affected their retention.

4.3.6. Sugar release assay. - A sugar release assay is the most commonly used method for the quantitation of glycoforms in IgG molecules and other glycoproteins. For this assay, the N-linked carbohydrate is released from the protein with PNgase F or other glycanases specific for N-linked sugars. The released sugars are purified from the protein and analyzed with normal-phase chromatography[25], MALDI or other techniques[26,27]. In most cases, the released oligosaccharides are derivatized through their reactive reducing end prior to analysis. Derivatization is used to introduce fluorescent tags which improve the normal-phase separation as well as the sensitivity of detection[17]. The chromatograms of PNgase F released oligosaccharides from the two lots, derivatized with 2-aminobenzamide and separated with normal-phase chromatography, are shown in Figure 4.6A. The glycoform profile shown in Figure 4.6A agrees very well with that published by Hill et al. [17]. Additional online MS analysis was carried out to identify the peaks separated by normal-phase chromatography. For simplicity, only the mass spectra for G0F-GlcNAc, Man5 and G0F peaks are shown in Figure 4.6B. Similar mass spectra were obtained for other peaks as well. In a previous study by Hill et al.[17], retention of a standard dextran ladder and glucose unit values for each oligosaccharide were used for assignment of peaks from the normal-phase chromatogram. According to that assignment, Man5 was reported to elute just before the G0F peak while the peak eluting after G0F was assigned as Man6. However, our LC/MS data clearly shows Man5 to elute after the G0F peak while the peak eluting before G0F was assigned as a mixture of G0F-GlcNAc and G0 (Fig. 4.6A). The MS analysis allowed a more

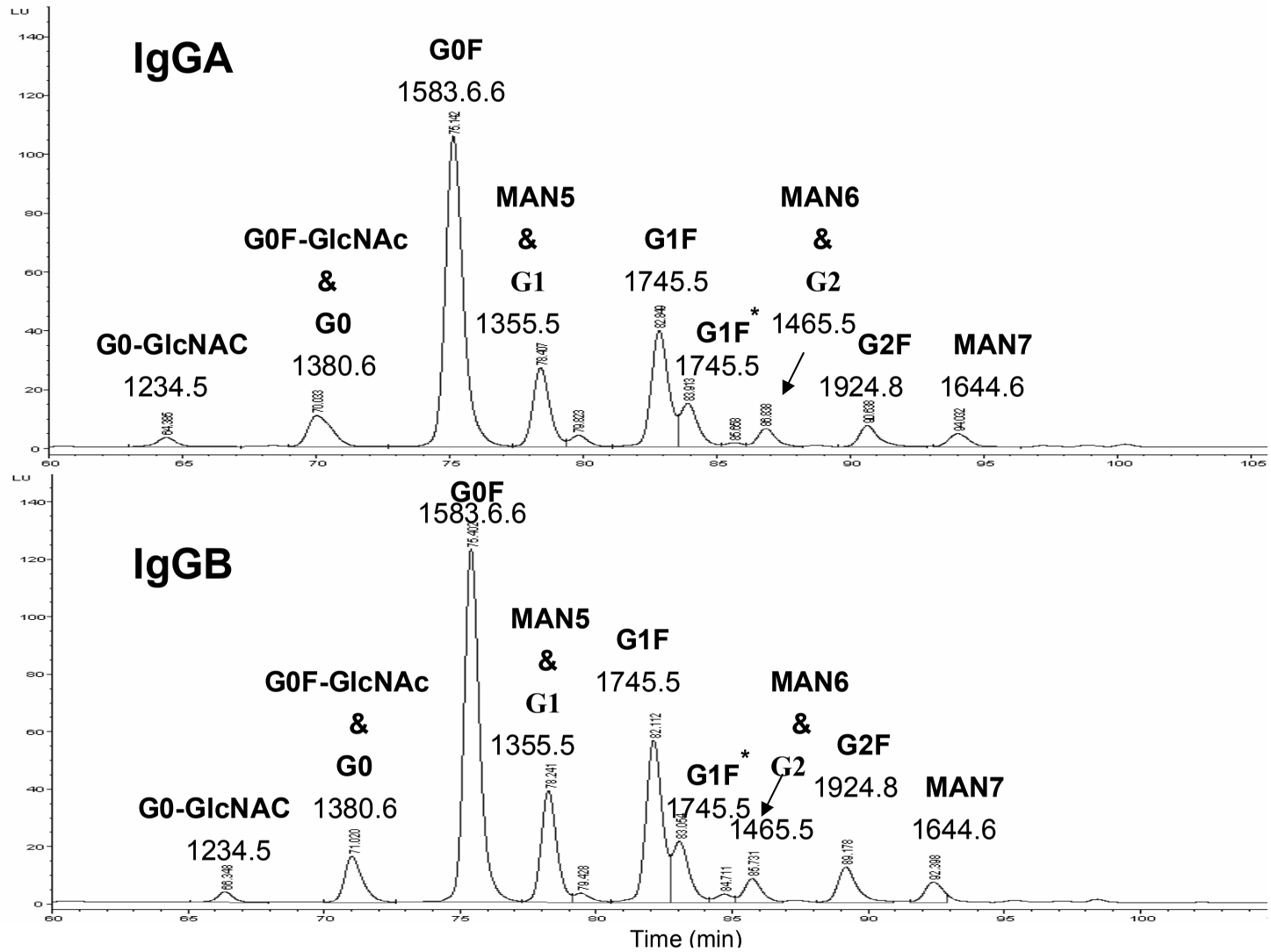


Figure 4.6A: Sugar release of the two IgG lots

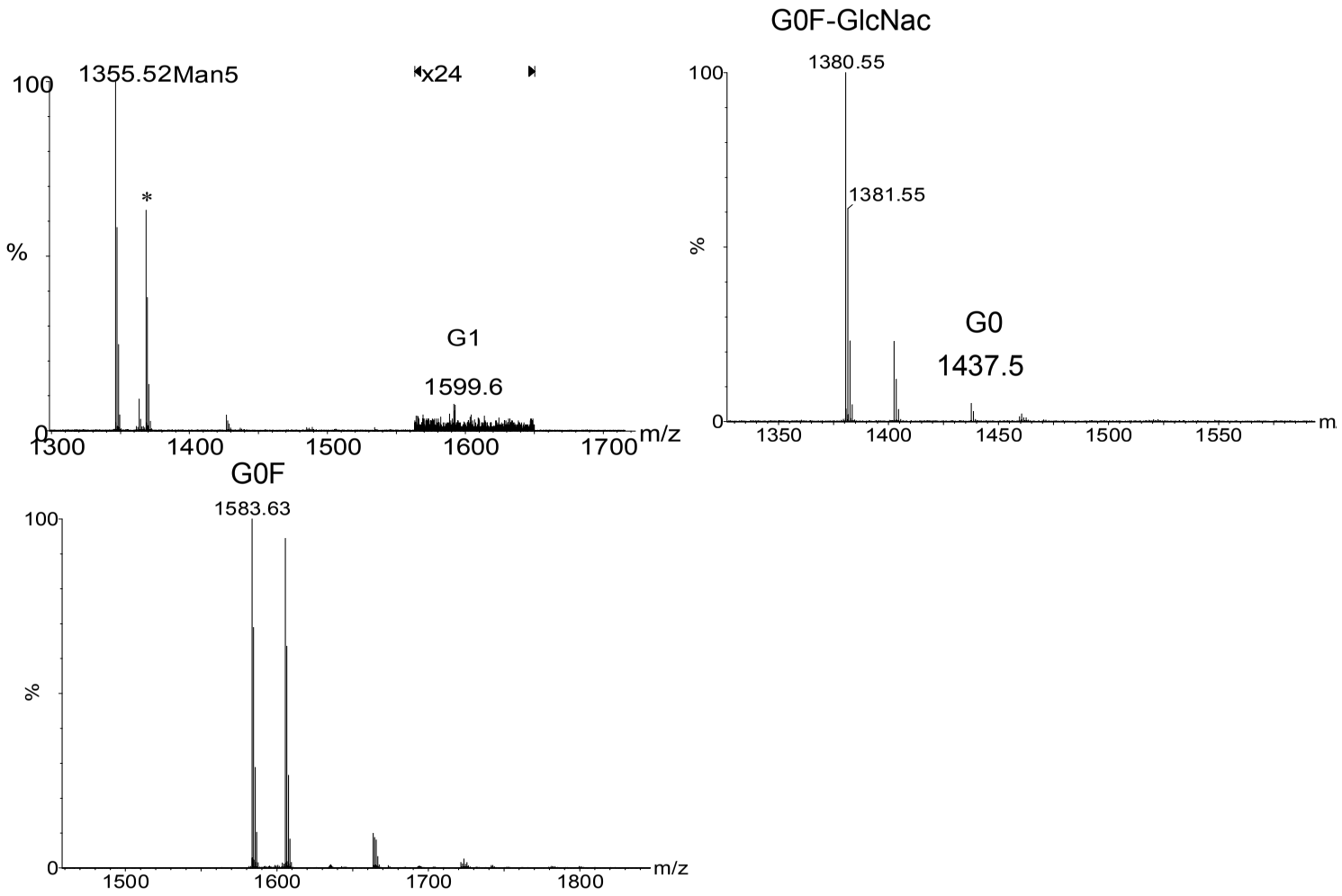
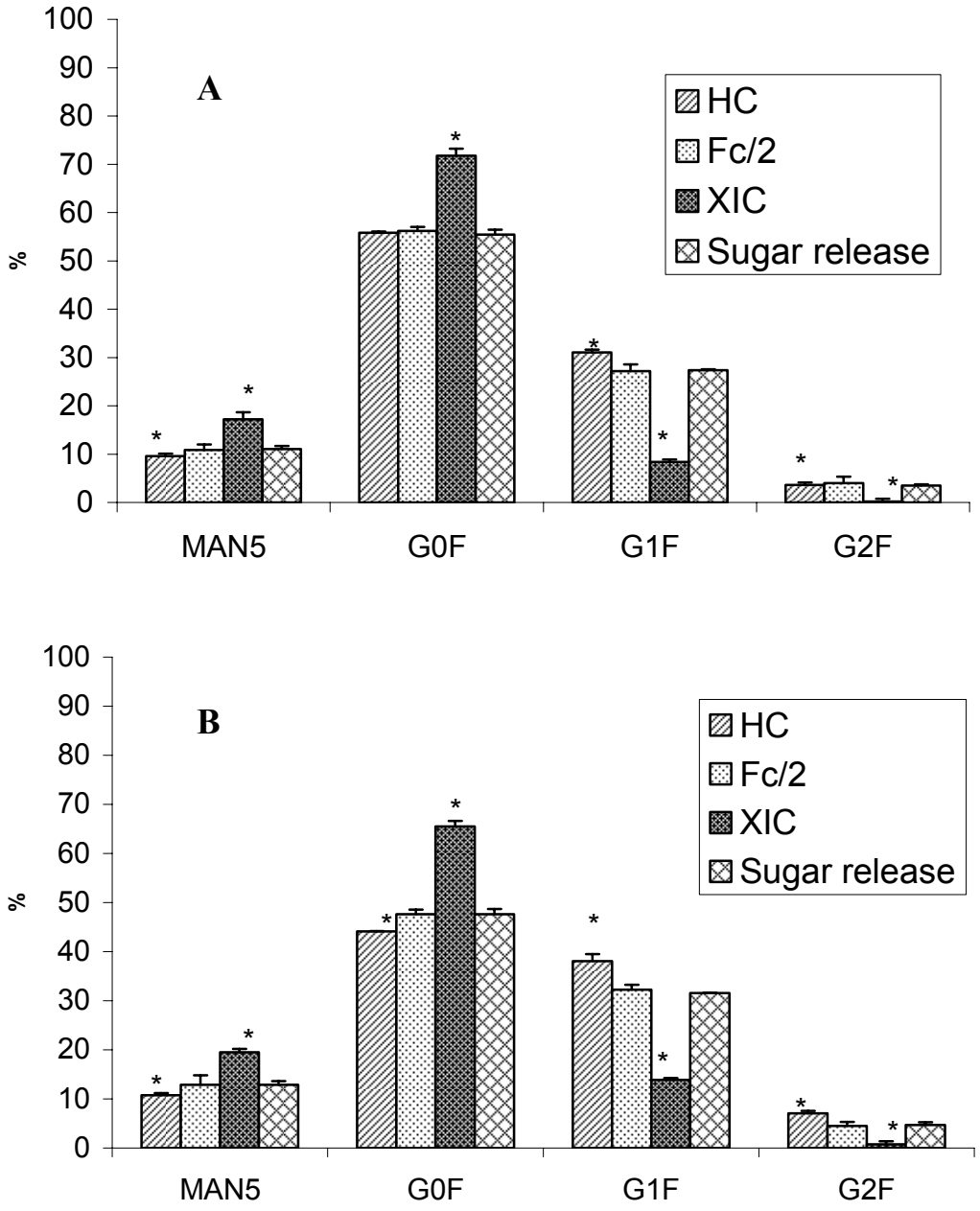


Figure 4.6 B ESI spectra of Man5, G0F and G0F-GlcNAc peaks

accurate identification leading to reassignment of the high mannose peaks. The highly branched nature of the high mannose sugars probably leads to a stronger interaction with the normal-phase column causing these forms to be retained more than the corresponding biantennary structures with higher glucose unit values. LC/MS analysis also allowed the identification of several new peaks such as G0-GlcNac and G0 which were detected but not assigned in the previous study by Hill et al[17] (Fig. 4.6). In addition, the MS analysis showed that the normal-phase method could not fully resolve all the glycoforms: G0F-GlcNac and G0 as well as Man5 and G1 were found to co-elute. The low MS signal for G1 (Fig. 4.6B) may be due to ion suppression of the branched Man5 carbohydrate. Since elution in normal-phase chromatography is generally influenced by the amount of carbohydrate, the G2 form would be expected to co-elute with the Man6 form. However, the spectra for the Man6 form did not show the presence of the G2 form (data not shown). The co-elution of these carbohydrate structures is a major limitation in quantitation using the sugar release assay.

4.3.7. Quantitative comparison of assay results. - The quantitation of glycoforms by four different methods is summarized in Table 4.2 and Figure 4.7. As noted above, LC/MS analysis of intact IgG (Section 4.3.1) and of IgG-Fc (Section 4.3.2) are not included in the quantitative comparison because these methods do not detect individual glycoforms but only glycoform pairs. Table 4.2 summarizes the quantitative analysis of glycosylation by the four methods that detect monomeric (i.e., unpaired) glycoforms. The methods differ in both the number of glycoforms detected and in the quantitative distribution of the glycoforms. Note that in Table 4.2 the total percentages of all glycoforms sum to 100% for each of the methods. This introduces



* Values differ significantly compared to sugar release assay at $\alpha = 0.05$

Figure 4.7. Quantitative comparison of the four major glycoforms Man5, G0F, G1F & G2F as detected by the HC, Fc/2, XIC and sugar release assays for A) IgGA and B) IgGB.

bias in quantitative comparison of methods that detect different numbers of glycoforms. To allow for more accurate comparison, percentages were rescaled to include only the four glycoforms detected by all four methods (i.e., Man5, G0F, G1F, G2F). In addition, isobaric forms that were resolved by the sugar release assay (i.e., Man6 and Man6*, G1F and G1F*) were pooled for this comparison. This normalized comparison is shown graphically in Figure 4.7. The numbers of glycoforms detected by the reduced (HC) and extracted ion (XIC) methods are less than the by the other two methods (Table 4.2). Low abundance glycoforms, accounting for less than ~ 5% of the total, are infrequently detected by the HC and XIC methods. For example, with the exception of G2F, the HC method does not detect any of the glycoforms that are at less than 5% abundance by the sugar release assay. While the XIC method detects some of these low abundance glycoforms (e.g., Man6, G0), low abundance forms with higher mass (e.g., G2) are not detected. In contrast, the limited and reduced (Fc/2) assay detects 11 glycoforms for IgGA and 10 glycoforms for IgGB, comparable to the ten glycoforms detected for each antibody by the sugar release assay (Table 4.2).

Figure 4.7 shows a quantitative comparison of the normalized percentages of the four major glycoforms by each of the four assay methods. For Man5, G1F and G2F, values obtained by the HC and XIC methods differ significantly from those obtained by the sugar release assay. The XIC results also differ significantly from the sugar release assay for the most abundant glycoform, G0F. In contrast, the results of the Fc/2 method are not significantly different from those of the sugar release assay for any of the four major glycoforms. Thus, the Fc/2 assay provides results that are

comparable to the sugar release assay in both the number of glycoforms detected (Table 4.2) and the quantitative values (Fig. 4.7).

Several reasons can be proposed for the quantitative and qualitative differences among the four methods. The poor ability of the HC method to detect glycoforms and to provide quantitative agreement with the standard sugar release assay may reflect poor ionization of the relatively large (~ 50 kD) glycosylated HC molecule. The XIC method may be susceptible to suppression of the glycopeptide signal due to the attached carbohydrate for the relatively small peptide fragments produced by digestion. The good agreement between the Fc/2 assay and the standard sugar release assay may be due in part to the improved ionization of this glycoslated protein, intermediate in size (~ 26 kD) between the HC and XIC fragments.

While the sugar release assay is regarded as a standard in monitoring IgG glycosylation, it is not without limitations. Of the methods studied here, only the sugar release assay could detect and resolve isobaric glycoforms (i.e., Man6 and Man6*, G1F and G1F*, Table 4.2). The sugar release assay also showed high precision as reflected by the low standard deviation. However, online mass spectrometric analysis showed co-elution of some of the carbohydrate structures during the sugar release assay, which greatly restricts the ability of the sugar release assay to quantitate these glycoforms. In particular, values for the Man5 and G0-GlcNAc reported for the sugar release assay in Table 4.2 are not accurate because these carbohydrate structures co-elute with the G0 and G1 forms, respectively, making their quantitation suspect.

Overall, the quantitation obtained with the Fc/2 assay was comparable to that of the sugar release assay and the small differences observed can be attributed to co-elution of certain forms during the sugar release assay. A limitation of the Fc/2 assay, and of any rpLC/MS approach, is that isobaric structures (i.e., Man6 and Man6*, G1F and G1F*) cannot be resolved with this method. The Fc/2 and sugar release assays thus are highly complementary and, when used together, are expected to provide complete characterization of carbohydrates in therapeutic recombinant monoclonal IgG molecules.

4.4. Conclusions

The studies reported here highlight strengths and limitations of LC/ESI-TOF MS assays for the identification and quantitation of glycoforms in IgGs. ESI-TOF analysis of the intact IgG was able to adequately measure the galactose variance in the biantennary sugar structure, but could not resolve the heterogeneity caused by high-mannose carbohydrates. ESI-TOF analysis of the IgG-Fc fragment generated after limited proteolysis enabled detection of both biantennary and high-mannose carbohydrates and was effective in characterizing oligosaccharide pairing caused by the combination of glycans on the two IgG-Fc heavy chains. Neither the intact IgG nor the IgG Fc analysis was found to provide sufficient resolution for quantitation, however. ESI-TOF analysis of the IgG-Fc/2 fragment showed accurate quantitation of various biantennary and high-mannose carbohydrates and was the most effective of the MS based methods evaluated at the identification and quantitation of carbohydrates. Peptide mapping followed by ESI-TOF MS analysis was not effective for absolute quantitation, as the ionization of glycopeptides was influenced by the

size of the carbohydrate. Though the sugar release assay showed high precision, the normal-phase method used for the assay could not fully resolve all the glycoforms. Collectively, the results suggest that MS quantitation based on analysis of Fc/2 (reduced Fc) is accurate and gives results that are both comparable and complementary to the more time-consuming sugar release assay.

4.5. References

1. Zhang J, Wang DI 1998. Quantitative analysis and process monitoring of site-specific glycosylation microheterogeneity in recombinant human interferon-gamma from Chinese hamster ovary cell culture by hydrophilic interaction chromatography 5. *J Chromatogr B Biomed Sci* Appl 712:73-82.

2. Kunkel JP, Jan DC, Butler M, Jamieson JC 2000. Comparisons of the glycosylation of a monoclonal antibody produced under nominally identical cell culture conditions in two different bioreactors 4. *Biotechnol Prog* 16:462-470.

3. Delorme E, Lorenzini T, Giffin J, Martin F, Jacobsen F, Boone T, Elliott S 1992. Role of glycosylation on the secretion and biological activity of erythropoietin. *Biochemistry* 31:9871-9876.

4. Keusch J, Lydyard PM, Delves PJ 1998. The effect on IgG glycosylation of altering beta1, 4-galactosyltransferase-1 activity in B cells. *Glycobiology* 8:1215-1220.

5. Tagashira M, Iijima H, Isogai Y, Hori M, Takamatsu S, Fujibayashi Y, Yoshizawa-Kumagaye K, Isaka S, Nakajima K, Yamamoto T, Teshima T, Toma K 2001. Site-dependent effect of O-glycosylation on the conformation and biological activity of calcitonin. *Biochemistry* 40:11090-11095.

6. Jefferis R 2007. Antibody therapeutics: isotype and glycoform selection 1. *Expert Opin Biol Ther* 7:1401-1413.

7. Edelman GM, Cunningham BA, Gall WE, Gottlieb PD, Rutishauser U, Waxdal MJ 2004. The covalent structure of an entire gamma G immunoglobulin molecule. 1969 5. *J Immunol* 173:5335-5342.

8. Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, Lund J, Jefferis R 2000. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* 37:697-706.

9. Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P 2003. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity 1. *J Mol Biol* 325:979-989.

10. Masuda K, Yamaguchi Y, Kato K, Takahashi N, Shimada I, Arata Y 2000. Pairing of oligosaccharides in the Fc region of immunoglobulin G. *FEBS Lett* 473:349-357.

11. Wuhrer M, Deelder AM, Hokke CH 2005. Protein glycosylation analysis by liquid chromatography-mass spectrometry 18. *J Chromatogr B Analyt Technol Biomed Life Sci* 825:124-133.

12. Gadgil HS, Bondarenko PV, Pipes GD, Dillon TM, Banks D, Abel J, Kleemann GR, Treuheit MJ 2006. Identification of cysteinylation of a free cysteine in the Fab region of a recombinant monoclonal IgG1 antibody using Lys-C limited proteolysis coupled with LC/MS analysis 4. *Anal Biochem* 355:165-174.

13. Gadgil HS, Pipes GD, Dillon TM, Treuheit MJ, Bondarenko PV 2006. Improving mass accuracy of high performance liquid chromatography/electrospray ionization time-of-flight mass spectrometry of intact antibodies 5. *J Am Soc Mass Spectrom* 17:867-872.

14. Gadgil HS, Pipes GD, Dillon TM, Treuheit MJ, Bondarenko PV 2006. Improving mass accuracy of high performance liquid chromatography/electrospray ionization time-of-flight mass spectrometry of intact antibodies 5. *J Am Soc Mass Spectrom* 17:867-872.

15. Gadgil HS, Bondarenko PV, Pipes G, Rehder D, McAuley A, Perico N, Dillon T, Ricci M, Treuheit M 2007. The LC/MS analysis of glycation of IgG molecules in sucrose containing formulations 2. *J Pharm Sci* 96:2607-2621.

16. Mimura Y, Ashton PR, Takahashi N, Harvey DJ, Jefferis R 2007. Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry 1. *J Immunol Methods* 326:116-126.

17. Hills AE, Patel A, Boyd P, James DC 2001. Metabolic control of recombinant monoclonal antibody N-glycosylation in GS-NS0 cells 1. *Biotechnol Bioeng* 75:239-251.

18. Gadgil HS, Bondarenko PV, Pipes G, Rehder D, McAuley A, Perico N, Dillon T, Ricci M, Treuheit M 2007. The LC/MS analysis of glycation of IgG molecules in sucrose containing formulations 2. *J Pharm Sci* 96:2607-2621.

19. Ren D, Pipes G, Xiao G, Kleemann GR, Bondarenko PV, Treuheit MJ, Gadgil HS 2007. Reversed-phase liquid chromatography-mass spectrometry of site-specific chemical modifications in intact immunoglobulin molecules and their fragments 1. *J Chromatogr A.*

20. Ren D, Pipes GD, Hambly DM, Bondarenko PV, Treuheit MJ, Brems DN, Gadgil HS 2007. Reversed-phase liquid chromatography of immunoglobulin G molecules and their fragments with the diphenyl column 2. *J Chromatogr A* 1175:63-68.

21. Masuda K, Yamaguchi Y, Kato K, Takahashi N, Shimada I, Arata Y 2000. Pairing of oligosaccharides in the Fc region of immunoglobulin G. *FEBS Lett* 473:349-357.

22. Boushaba R, Kumpalume P, Slater NK 2003. Kinetics of whole serum and prepurified IgG digestion by pepsin for F(ab')2 manufacture 36. *Biotechnol Prog* 19:1176-1182.

23. Leslie RG, Melamed MD, Cohen S 1971. The products from papain and pepsin hydrolyses of guinea-pig immunoglobulins gamma 1G and gamma 2 G. *Biochem J* 121:829-837.

24. Bongers J, Cummings JJ, Ebert MB, Federici MM, Gledhill L, Gulati D, Hilliard GM, Jones BH, Lee KR, Mozdzanowski J, Naimoli M, Burman S 2000. Validation of a peptide mapping method for a therapeutic monoclonal antibody: what could we possibly learn about a method we have run 100 times? 4. *J Pharm Biomed Anal* 21:1099-1128.

25. Hills AE, Patel A, Boyd P, James DC 2001. Metabolic control of recombinant monoclonal antibody N-glycosylation in GS-NS0 cells 1. *Biotechnol Bioeng* 75:239-251.

26. Bykova NV, Rampitsch C, Krokhin O, Standing KG, Ens W 2006. Determination and characterization of site-specific N-glycosylation using MALDI-Qq-TOF tandem mass spectrometry: case study with a plant protease 4. *Anal Chem* 78:1093-1103.

27. Mirgorodskaya E, Krogh TN, Roepstorff P 2000. Characterization of protein glycosylation by MALDI-TOFMS 8. *Methods Mol Biol* 146:273-292.

Chapter 5

Effect of Secondary Structure on Deamidation in a Tryptic Fragment of the Fc Portion of a Recombinant Monoclonal Antibody

5.1. Introduction

Therapeutic antibodies are one of the fastest growing segments of the pharmaceutical industry[1,2]. To maintain potency and minimize immunogenicity, antibodies and other protein drugs must be protected from physical and chemical degradation during manufacturing and storage. The rational development of stable formulations of these large molecules requires analytical methods capable of detecting chemical changes with amino-acid level resolution.

One of the most common routes of chemical inactivation of proteins is deamidation at asparagine (N) residues. In neutral to basic solution, deamidation proceeds via the nucleophilic attack of the N+1 nitrogen of the protein backbone on the carbonyl group of the N side chain, forming a cyclic imide (succinimide) intermediate. Hydrolysis then occurs at either of the succinimide carbonyl groups to form the aspartate- (D) and isoaspartate- (isoD) containing products, which differ from the parent sequence both in charge (acidic vs. neutral) and mass (+ 1 amu). Protein primary sequence strongly influences deamidation rate. For example, in unstructured model peptides (37 ^{o}C, pH 7.4), deamidation half-lives for the NG sequence are typically on the order of 1 day, while half-lives at NV and NI sequences of more than 200 days have been reported[3]. Secondary structure also influences deamidation rate[4-8], but its effects have not been as clearly elucidated.

Various analytical methods have been used to detect deamidation in proteins[9]. Ion-exchange chromatography (IEC) exploits the difference in charge of the acidic degradation products and the parent molecule, and is routinely used in the pharmaceutical and biotechnology industry. Tryptic digestion with reversed-phase high performance liquid chromatography (rp-HPLC) has been used to identify sites of

deamidation and to quantify deamidation rates. IsoD and D variants can often be resolved from the parent tryptic fragment in tryptic mapping with rp-HPLC, albeit with long analysis times. The enzyme protein isoaspartyl methyl transferase (PIMT) repairs isoD residues in vivo, and has been used to quantify isoD levels in vitro. PIMT detection kits are available commercially (e.g., IsoQuant®, Promega, Madison, WI). PIMT selectively catalyzes the transfer of a methyl group from S-adenosyl methionine to isoD at the α-carbonyl, generating S-adenosyl homocysteine (SAH), which is detected by HPLC.

Recent reports suggest that there may be two deamidation-prone regions in the complementarity determining regions (CDR) of antibodies and several hot spots in the constant regions (Fc). Deamidation in the Fc-domain of a humanized monoclonal antibody was studied recently with tryptic digestion followed by analysis with liquid chromatography and electrospray ionization mass spectrometry (HPLC/+ESI/MS)[10]. The results showed that the reaction is sensitive to antibody structure, with tryptic fragment G_{369}-K_{390}, (GFYPSDIAVEWESNGQPENNYK) showing rapid and structure-dependent deamidation at N_{382} and N_{387}. Interestingly, the authors did not observe several expected deamidation products at these sites, including the isoD product of N_{387} (i.e., $isoD_{387}$), the Asp product at N_{382} (i.e., D_{382}) and the succinimide intermediate at N_{387}. The authors infer that the mechanisms of deamidation differ at the two sites, perhaps due to differences in secondary structure. This is an important observation and a potentially significant contribution to our understanding of the effects of secondary structure on chemical reactivity in proteins. However, it is also possible that the “missing” products were simply poorly resolved chromatographically and so were not detected. This is particularly

problematic if the IsoD and D variants co-eluted, since the two products have identical mass and cannot be resolved by the use of base ion chromatograms, as employed by the authors. As a result, the effect of secondary structure on deamidation in this commercially important protein remains an open question.

The work reported here is a detailed study of deamidation at N_{382} and N_{387} in the Fc portion of a humanized IgG1 antibody and was performed in collaboration with and with support from Merck Research Laboratories, West Point, PA. We have employed ultraperformance liquid chromatography with electrospray ionization mass spectrometry (UPLC/+ESI/MS) in an attempt to achieve chromatographic resolution of deamidation products while maintaining relatively short analysis times. In addition, synthetic peptides corresponding to the tryptic fragment G_{369}-K_{390} have been used as analytical standards and as controls for the effects of structure on deamidation. The results demonstrate that the product profile for the tryptic fragment and the intact protein are indeed different, supporting the hypothesis that higher order structure of the protein plays a significant role in determining both deamidation kinetics and product distribution.

5. 2. Materials and Methods

5.2.1. Materials. – Deamidation was monitored in the Fc portion of a humanized IgG1. Tryptic digests of Fc-IgG and synthetic model peptides corresponding to selected tryptic fragments were used as relatively unstructured controls. Fc-IgG was prepared at Merck Research Laboratories from a humanized monoclonal antibody expressed in NS0 cells, a murine myeloma cell line. The human IgG1 Fc fragment was prepared by digestion with immobilized papain (Pierce, Rockford IL), followed by

protein A purification. Impurities were further removed by cation-exchange high performance liquid chromatography (HPLC) using a ProPac WCX-10 9 x 250 mm column (Dionex, Sunnyvale, CA) and gradient elution (mobile phase A: 10 mM sodium phosphate pH 6.5; mobile phase B: 10 mM sodium phosphate, 0.5 M NaCl pH 6.5; gradient: 3-26% B in 20 min at a flow rate of 2 ml/min). The human Fc IgG1 digestion and purification was conducted by Dr. Josef Vlasak at Merck Research Laboratories, West Point, PA.

Model peptides corresponding to the tryptic fragment G_{369}-K_{390} (i.e., GFYPSDIAVEWESNGQPENNYK) and its deamidated variants were used as synthetic standards and as unstructured controls. The parent peptide GFYPSDIAVEWESNGQPENNYK (hereinafter abbreviated "NNN") and the D_{387}-containing deamidated variant, GFYPSDIAVEWESNGQPE**D**NYK (abbreviated $ND_{387}N$), were synthesized by the Biochemical Research Services Laboratory of the University of Kansas. The $isoD_{382}$-containing variant, GFYPSDIAVEWES(**isoD**)GQPENNYK (abbreviated $isoD_{382}NN$), was purchased from American Peptide Company (Sunnyvale, CA). All other chemicals and reagents were of the highest commercial grade and used without further purification. Milli-Q water was used to prepare all solutions.

5.2.2. Sample preparation and accelerated stability studies. - Accelerated stability studies were conducted for two types of samples at 37 ^{o}C and pH 7.5: (i) the fully-folded, intact Fc IgG (~ 50 kD) ("intact protein") and (ii) tryptic digests of Fc IgG ("digests"). Deamidation in the native Fc IgG (i) was thus compared to deamidation in its tryptic fragments which serve as an unstructured control. For the intact protein, samples contained 50 µL of 0.5 mg/mL Fc IgG in 100mM Tris-HCl buffer (pH 7.5).

Samples were stored in capped microcentrifuge tubes at 37 ^{o}C for two days. Triplicate samples were withdrawn at designated time intervals, digested with trypsin as described below, and stored at -20 ^{o}C prior to LC-MS analysis (see Section 5.3).

To produce digests, the intact Fc IgG was subjected to proteolytic digestion with trypsin prior to accelerated stability testing. Tryptic digestion followed the protocol reported by Chelius et. al[10]. Briefly, an aliquot of intact Fc was first exchanged to 6 M guanidine-HCl, 0.2 M Tris-HCl, 1 mM EDTA (pH 7.5) such that 100 µL of 0.5 mg/ml Fc was obtained. Reduction was performed by addition of 2 µL of 0.5 M dithiothreitol (DTT) at 37 ^{o}C for 40 minutes followed by alkylation by 4 µL of 0.5 M iodoacteamide at room temperature for 40 minutes. The reduced, alkylated Fc portion was then buffer exchanged with a 10,000 MW cut off membrane (Microcon centrifugal filter, Millipore, Billerica, MA) to 100 mM Tris buffer (pH 7.5) such that the Fc concentration was maintained at 0.5 mg/ml. Trypsin (Worthington, MA) was then initially added in a trypsin:protein ratio of 1:50 (w/w) and incubated for 2 hrs at 37 ^{o}C. This was followed by addition of another equal amount of trypsin with incubation for 2 more hours, bringing the final trypsin:protein ratio to 1:25 (w/w). For stability studies, the samples were withdrawn and the reaction quenched by addition of 20 % formic acid such that the final concentration was 2%. For stability studies of digests, a 50 µL aliquot was then stored in capped microcentrifuge tubes and stored at 37 ^{o}C for up to 45 days. Triplicate samples were withdrawn at designated time intervals and stored at -20 ^{o}C prior to LC-MS analysis (see Section 5.2.3). Samples of intact Fc IgG were subjected to tryptic digestion by this protocol following storage. Stability studies of intact protein and digests thus were conducted identically, but with the order of tryptic digestion and storage reversed.

5.2.3. Mass spectrometric analysis (UPLC/+ESI-MS) of Fc-IgG and fragments. - The analytical strategy for monitoring deamidation in Fc-IgG involved tryptic digestion of intact protein (see Section 5.2.2), ultraperformance liquid chromatography with tandem electrospray ionization mass spectrometry (UPLC/+ESI-MS) analysis of the fragments, and quantitation of peptides and their deamidated variants using extracted ion chromatograms (EIC). UPLC is a recent innovation in high performance liquid chromatography (HPLC), in which unique small particles and very low column volumes allow for greater throughput and rapid analysis times. In the G_{369}-K_{390} sequence, deamidation can occur at each of the N residues (N_{382}, N_{387}, N_{388}) to produce the isoD and D variants. Since each of the three sites can exist in three forms (i.e., N, isoD, D), a total of 27 (i.e., 3^3) variants is possible. Identification and quantitation of the product mixture thus requires analytical methods capable of resolving similar peptides in complex mixtures. UPLC is well-suited to the analysis of such mixtures, and was selected as the separation method to minimize possible co-elution of isobaric products.

UPLC was performed using a Waters Acquity UltraPerformance liquid chromatography system (Waters, Inc., Milford, MA) with a BEH C18 column (1 mm × 150 mm, 1.7 µm particle size) at a flow rate of 100 µL/min and a column temperature of 35 ^{o}C. Approximately 10 µg of protein was injected for each analysis (10 µl / injection). Gradient elution was employed, (Solvent A: 99% H_2O, 1% methanol, 0.08% formic acid; Solvent B: 99% acetonitrile, 1% H_2O, 0.06% formic acid), with the profile: 0 to 1 min, solvent B at 1% (v/v); 1 to 2 min, solvent B increased to 10% (v/v); 2 to 19 min, solvent B increased to 18% (v/v); 19 to 45 min, solvent B increased to 22% (v/v); followed by washing and requilibration. Mobile phase solvent

compositions were: Solvent A - 99% H_2O, 1% methanol, 0.08% formic acid; Solvent B - 99% acetonitrile, 1% H_2O, 0.06% formic acid.

The UPLC eluate was coupled to a Micromass® Q-Tof II mass spectrometer (Waters, Inc., Milford, MA) operated in the positive ion mode ($^+$ESI) with a cone voltage of 35V and a collision voltage of 30 eV. The relatively high collision voltage produces partial fragmentation of the parent ions, and was adopted as an alternative to MS/MS analysis since it produced a sufficient number of daughter ions to allow unambiguous identification of deamidation sites. The G_{369}-K_{390} fragment and the corresponding synthetic peptide carry a charge of +2, a theoretical mass of 2544.67 Da and thus an expected m/z value of 1272.57. Data were analyzed using MassLynx® software (Waters, Inc.). Quantitation of native peptides and their deamidated variants was based on the isotopic peak intensities of each form. The theoretical isotope distribution for the parent peptide was employed to separate the contribution of a doubly deamidated product, which eluted as a shoulder with the parent in studies of the undigested protein. The ionization efficiency of the parent peptide was assumed to be equivalent to that of its deamidated forms.

5.2.4. Kinetic modeling. **–** The time-varying concentrations of the N-containing parent species and their deamidated variants were subjected to kinetic analysis to determine the rate constants for deamidation. The native species and the deamidation products form an ensemble that evolves in time based on the microscopic rates of interconversion between species. Assuming that the deamidation process can be well represented as a unimolecular, first-order reaction, the time-dependencies of the population of "n" species can be described by a system of linear differential equations of the form:

$$\frac{d[A_i]}{dt} = (\sum_{j \neq i} k_{ji}[A_j]) - [A_i] \times \sum_{j \neq i} k_{ij} \qquad (5.1)$$

where A_i, A_j, represent molar fractions of any two species in the ensemble, and k_{ji} is the microscopic rate constant describing the conversion of species A_j into species A_i:

$$A_j \xrightarrow{kji} A_i \qquad (5.2)$$

The system of linear differential equations was solved using classical procedures[11] to determine the eigenvalues of the system λ_i (i=1,n). The eigenvalues λ_i represent experimental rate constants and can be expressed as linear combinations of the microscopic rate constants, k_{ji}. For a system comprising "n" distinct species there are at most "n -1" values for λ_i, since one solution is always zero, corresponding to the equilibrium value. The solutions of the system, $[A_i](t)$, were expressed as a sum of exponentials of the form:

$$[A_i](t) = \sum_j \alpha_{ij} \exp(-\lambda_i t) \qquad (5.3)$$

where α_{ij} represent coefficients that are specific to each species, "i". The coefficients α_{ij} are determined from the system of linear equations and from initial (t=0) and equilibrium ($t=\infty$) conditions. Since the system is relatively uncomplicated, the eigenvalues and eigenvectors were calculated ”by hand”, i.e., no particular application was used to solve the system of equations describing the time-dependence of each species. The explicit form of the system of differential

equations, with their corresponding eigenvalues and solutions for the molar fractions of the species, is given in the Appendix.

For a particular data set (i.e., describing either the deamidation of a model peptide, a tryptic fragment or the intact Fc), the microscopic rate constants were determined by performing a global fit, in which the fitting of the experimental points was performed simultaneously for all kinetic traces using the program SigmaPlot (Ver 10.0, Systat Software, Inc., San Jose, CA). Statistical information regarding the fitting results, including standard and relative errors of regression coefficients, was provided by SigmaPlot routines. Kinetic analysis was performed by Dr. Roxana Ionescu at Merck Research Laboratories, West Point, PA.

5.2.5. Molecular dynamics simulations (MDS). - The IgG-Fc structure was visualized using SYBYL (Version 8.0, Tripos, St. Louis, MO) which enabled direct measurement of key interatomic distances as represented in the 1h3u RCSB Protein Data Bank (PDB) crystal structure[12]. Relative solvent accessibility of specific residues was determined in SYBYL by generating Connolly surfaces on those amino acids and evaluating the total area of each resulting surface.

5.3. Results

The studies reported here investigate the effect of higher-order structure on deamidation in the constant (Fc) domain of a humanized IgG1 antibody, focusing on a 22 amino acid tryptic fragment located in the CH3 domain (G_{369}FYPSDIAVEWESNGQPENNYK$_{390}$). To control for the effects of protein structure, deamidation in the intact Fc-protein was compared with deamidation its tryptic digests. The identification of reaction products for each of the sample types is

presented first (Section 5.3.1), followed by data on reaction kinetics (Section 5.3.2) and molecular dynamics simulation (Section 5.3.3). Because the product profiles are considerably more complex in the intact protein than in the digests, results for the digests are presented first within each Section. These are followed by results for the intact protein and then by supporting results on the synthetic peptides.

5.3.1. Product identification

5.3.1.1. Deamidation products in tryptic digests. - The EIC for a representative sample of a tryptic digest of Fc-IgG following 14 hours of storage (37 ^{o}C, pH 7.4) shows three well-resolved peaks eluting at approximately 36, 37 and 39 min (Fig. 5.1A). Figure 1B shows the molecular ion mass spectrum of each of these peaks, presented in the order of elution. The second peak (Fig. 5.1A, ~37 min) shows an m/z of 1273 (Fig. 5.1B(2)) and is thereby identified as the intact parent peptide (i.e., G_{369}-K_{390}). The first peak (Fig. 5.1A, ~36 min) and the third peak (Fig. 5.1A, ~39 min) both shown +1 amu shifts in their molecular isotope envelopes (Fig. 5.1B(1), Fig. 5.1B(3)), consistent with singly deamidated products. These mass envelope shifts cannot identify the site of deamidation, however, nor can they distinguish between the isobaric D and isoD products produced at a single site.

The sites of deamidation were determined using the daughter ions (i.e., b- and y˝-ions) formed during high energy MS1 analysis of each of these peaks (Fig. 5.2). Figures 5.2A and 5.2C show $y˝_{10}$, $y˝_{12}$, $y˝_{13}$ and $y˝_{19}$ ions with a mass increase of +1 amu, consistent with deamidation in these fragments and which may have occurred at the N_{382}, N_{387} or N_{388} sites. However, Figures 5.2A and 5.2C show that there are no mass changes in the $y˝_{6}$, $y˝_{4}$ and $y˝_{3}$ ions, which excludes deamidation at the the N_{387} and N_{388} sites. Thus, using these daughter ions, deamidation at N_{382}

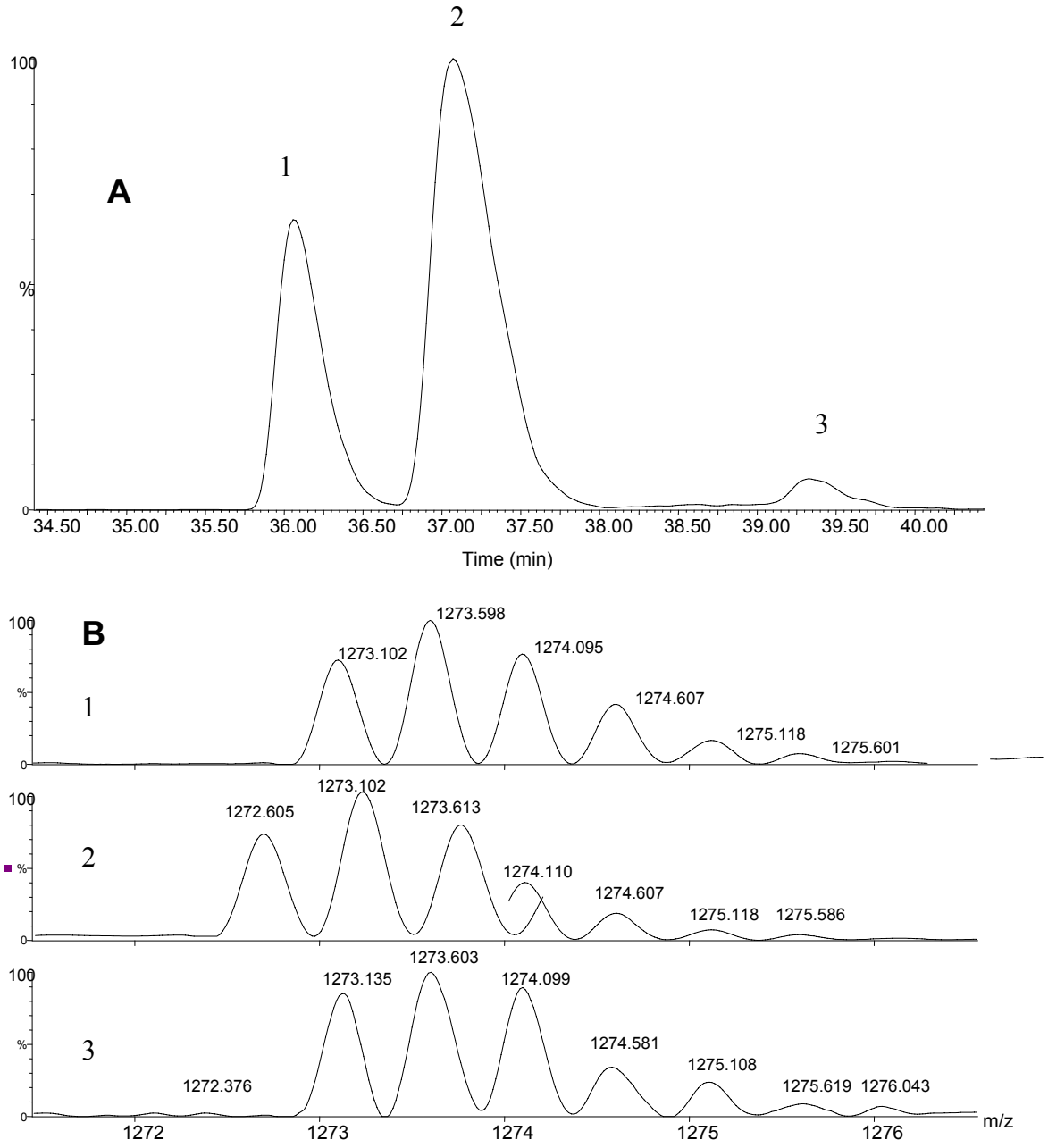


Figure 5.1. Representative A) Extracted ion chromatogram (EIC) and B) molecular ion isotope envelope of 1. IsoDNN, 2. NNN and 3. DNN peaks in the EIC, for a sample stressed for 14 hours at 37 ^{o}C, pH 7.5 after digestion

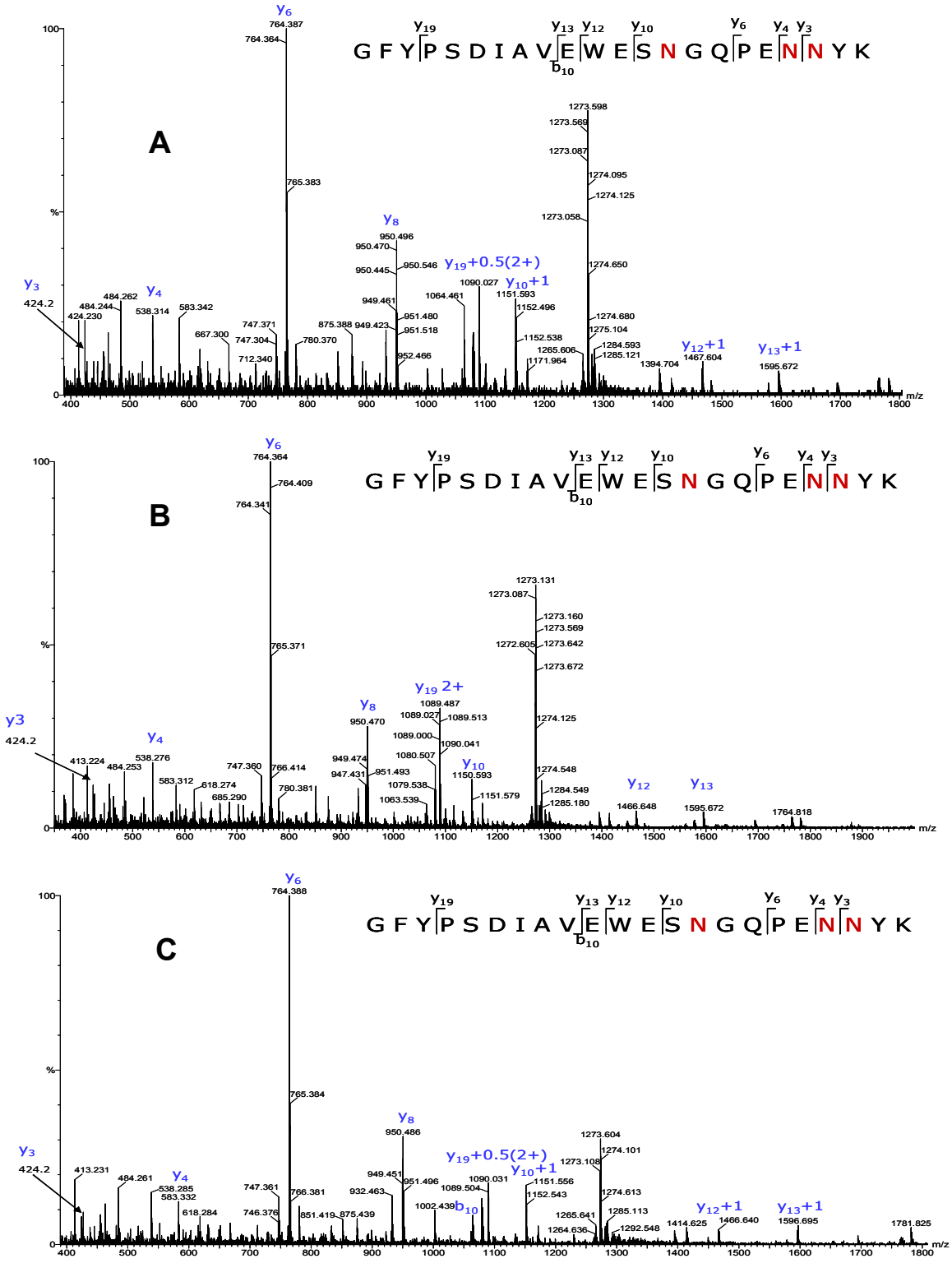


Figure 5.2. High energy MS1 spectra of the peptides eluting at approximately 36 (A), 37 (B) and 39 (C) minutes. The peptides were identified with the y ions as the IsoD382, parent and the D382 forms respectively.

was confirmed for both the first and third peaks of Figure 5.1A, indicating that they are the $isoD_{382}$ and D_{382} products of deamidation at this site. Both the relative peak areas and the elution order further suggest that the first peak corresponds to the $isoD_{382}$NN product. In reverse-phase chromatography, isoD-containing peptides typically elute earlier than their D-containing counterparts[13-16]. In addition, the isoD product is generally favored in unstructured peptides, with a typical isoD:D ratio of 3:1 to 5:1 [6,17]. On this basis, the product peaks in Figure 1A are tentatively assigned as $isoD_{382}$NN (Fig. 5.1A, 36 min) and D_{382}NN (Fig. 5.1A, 39 min). These assignments were later confirmed by comparing the elution pattern of synthetic peptides corresponding to the $isoD_{382}$NN and parent forms (see Section 5.3.1.3, below).

5.3.1.2. Deamidation products in the intact protein. - The intact protein was stressed at 37 ^{o}C, pH 7.4 (100 mM Tris) to allow deamidation to occur in the native (i.e., fully structured) form, then subjected to tryptic digestion prior to UPLC/MS analysis. Figure 5.3A shows the EIC for the G_{369}-K_{390} peptide fragment of intact Fc that had been stressed for 28 hrs before digestion. A total of five species are detected, eluting at 35.7, 36.7, 37.3, 38.2 and 38.8 min. The molecular isotope envelopes (Fig. 5.3B) show that: (i) the first, fourth and fifth peaks correspond to singly deamidated products (Fig. 5.3B(1), 5.3B(4) and 5.3B(5)), (ii) the second peak corresponds to the parent peptide (Fig. 5.3B(2)), and the third peak corresponds to a doubly-deamidated product with +2 amu mass increase relative to the parent (Fig. 5.3B(3)). By comparison with results for the digests (above) and by high energy MS1 analysis

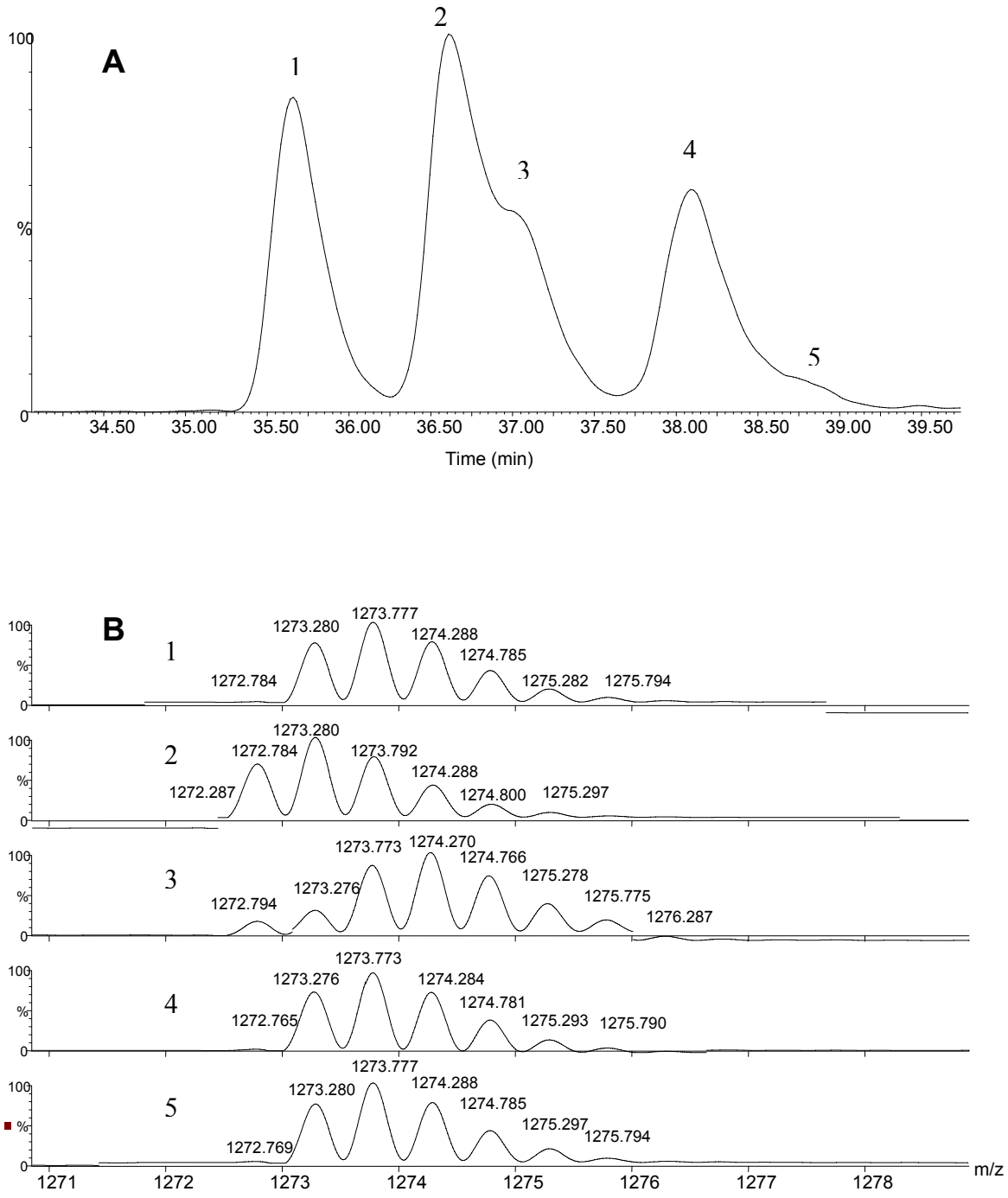


Figure 5.3. Representative A) Extracted ion chromatogram (EIC) and B) molecular ion isotope envelope of 1. IsoDNN, 2. NNN, 3. IsoD/D,IsoD/D,N, 4. NDN and 5. DNN peaks in the EIC, for a sample stressed for 28 hours at 37 ^{o}C, pH 7.4 before digestion.

(not shown), the first and fifth peaks were identified as the isoD$_{382}$NN and D$_{382}$NN variants. The y˝ ions indicate that the deamidation sites for the third and fourth peaks were at N$_{382}$ and N$_{387}$ (Figure. 5.4). By high energy MS1 analysis (Fig. 5.4A) and the molecular isotope envelopes (Fig. 5.3B(3)), the third peak was identified as a doubly deamidated product (Fig. 5.4A). The presence of y˝$_4$, y˝$_6$ and y˝$_8$ daughter ions with a mass change of +1 amu and of a y˝$_3$ daughter ion with unchanged mass indicated deamidation at N$_{387}$. Since this peak eluted after the parent peak and was singly deamidated, it was tentatively assigned as the D form. This result was confirmed with the synthetic peptide studies (see Section 5.3.1.3). A +2 amu mass change for the y˝$_{10}$ and y˝$_{19}$ daughter ions indicated a second deamidation site at N$_{382}$. Thus, the third peak is deamidated at both the N$_{382}$ and N$_{387}$ sites. The identity of the deamidated sites with respect to isoD or D is not determinable with the present data.

5.3.1.3. D and IsoD peak identification with synthetic peptides. – Three synthetic peptides (NNN, ND$_{387}$N, IsoD$_{382}$NN; see Section 5.2.1) were analyzed by UPLC/MS to confirm the tentative peak assignments described above. The elution times of the synthetic peptides agreed with those tentatively assigned to these species in the previous studies, thus confirming their identities (Fig. 5.5). When allowed to deamidate under the same stress conditions used for the digests and intact protein (i.e., 37 ^{o}C, pH 7.4), the synthetic NNN peptide formed the isoD$_{382}$NN and D$_{382}$NN peptides observed in the digest studies (not shown).

5.3.2. Deamidation kinetics - The deamidation products for the Fc IgG digests are consistent with the accepted mechanism for deamidation at a single Asn site[16,18,19]. In the discussion that follows, deamidation kinetics in these samples will be modeled as parallel irreversible first-order reactions to produce the isoD and D products, as

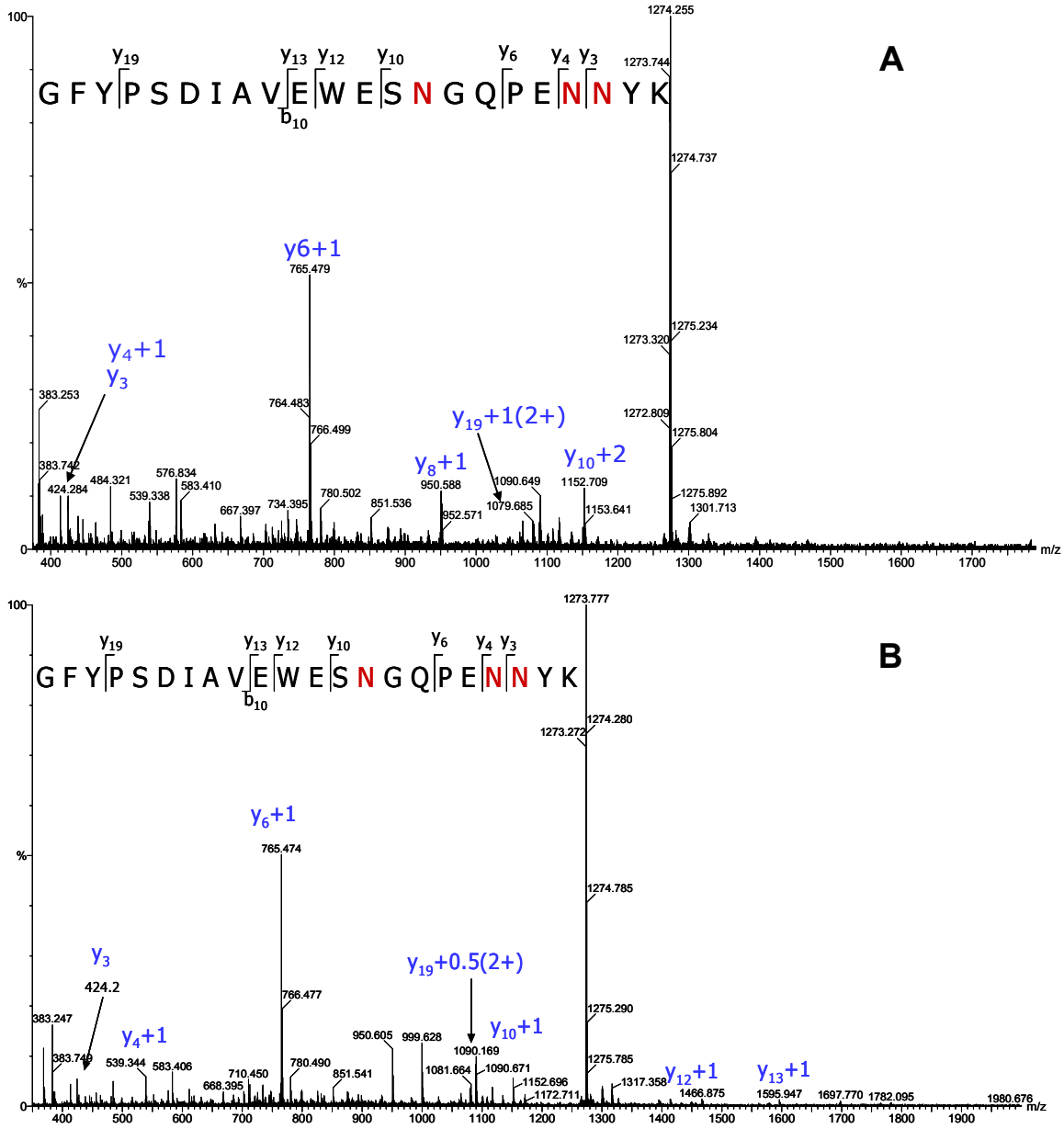


Figure 5.4. High energy MS1 spectra of the peptides eluting as a shoulder a approximately 37.5 minutes (A) and 38 (B) minutes. The peptides were identified with the y ions as the doubly deamidated IsoD/D382, IsoD/D387N and the D387 peptide.

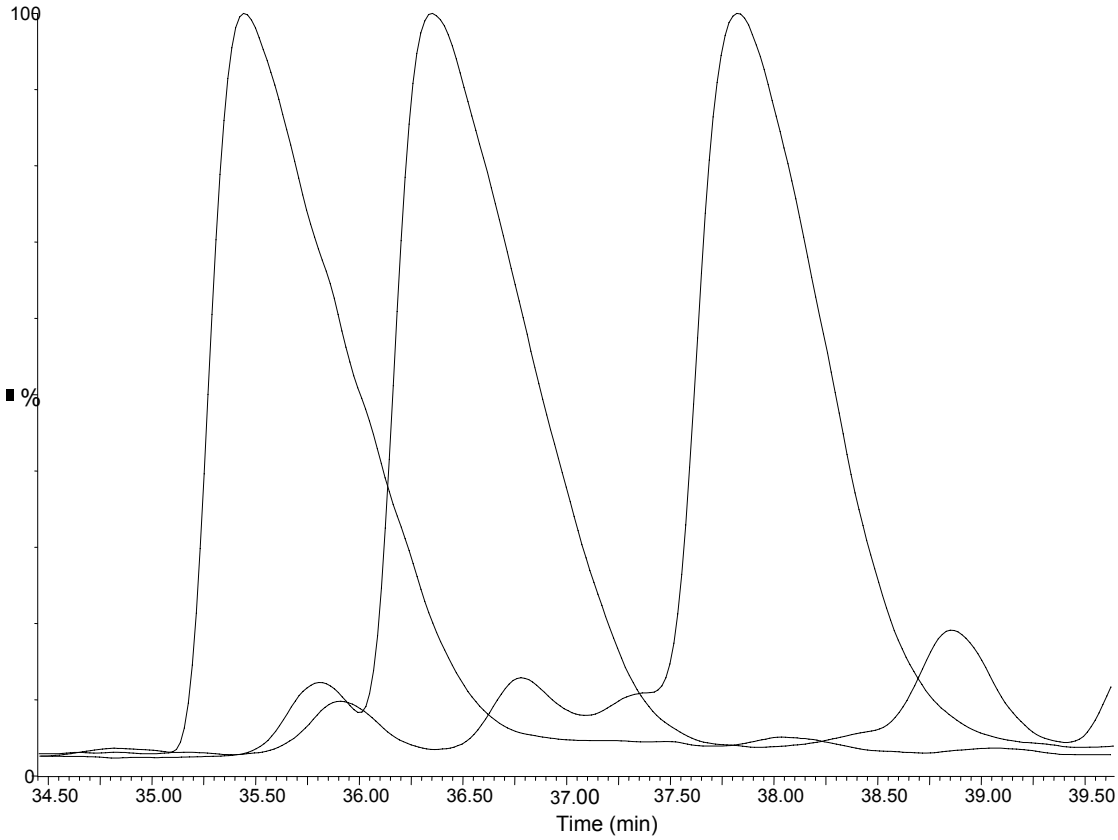


Figure 5.5. Elution times for the synthetic peptides IsoDNN, NNN and NDN. Elution pattern matched with that observed for the digest assay and also the intact protein analysis.

shown schematically in Fig. 5.6A and described in previous reports[17,20-22]. In contrast, deamidation in the intact Fc IgG produced singly deamidated products at two sites (i.e., $isoD_{382}$ and D_{387}) and a doubly deamidated product. As discussed below, reaction kinetics for the intact protein are consistent with parallel irreversible first-order deamidation to produce the singly deamidated products, followed by their further degradation to produce the doubly deamidated product (Fig. 5.6B). In the discussion below, reaction kinetics for the digests are presented first (Section 5.3.2.1) followed by results for the intact protein (Section 5.3.2.2) and supporting results on the synthetic peptides (Section 5.3.2.3).

5.3.2.1. Deamidation kinetics in tryptic digests. – Kinetic profiles for deamidation in the digests are shown in Figure 5.7. The parent fragment (NNN) degrades in a pseudo-first order manner with less than 10% remaining after 42 hrs. Loss of the parent fragment is accompanied by monotonic increase in the isoD- and D-containing products at N_{382} (i.e., $isoD_{382}$NN and D_{382}NN, respectively), which are formed in the ratio of ~ 4:1 (Fig. 5.5). The percentages were calculated assuming that sum of the N-, isoD- and D-containing species is 100%.

The kinetic data for the parent tryptic fragment and its deamidation products were fitted simultaneously to the reaction scheme in Fig. 5.6A to provide values for the rate constants k_{1p} and k_{2p} (Table 5.1). The regression lines are in close agreement with the data (Fig. 5.7) and the coefficients of variation are relatively small (%CV, Table 5.1), indicating a good fit. The calculated half-life for loss of the parent fragment is ~ 14 hours (i.e., 0.693 $(k_{1p} + k_{2p})^{-1}$ = 14.2 h), in good agreement with deamidation rates

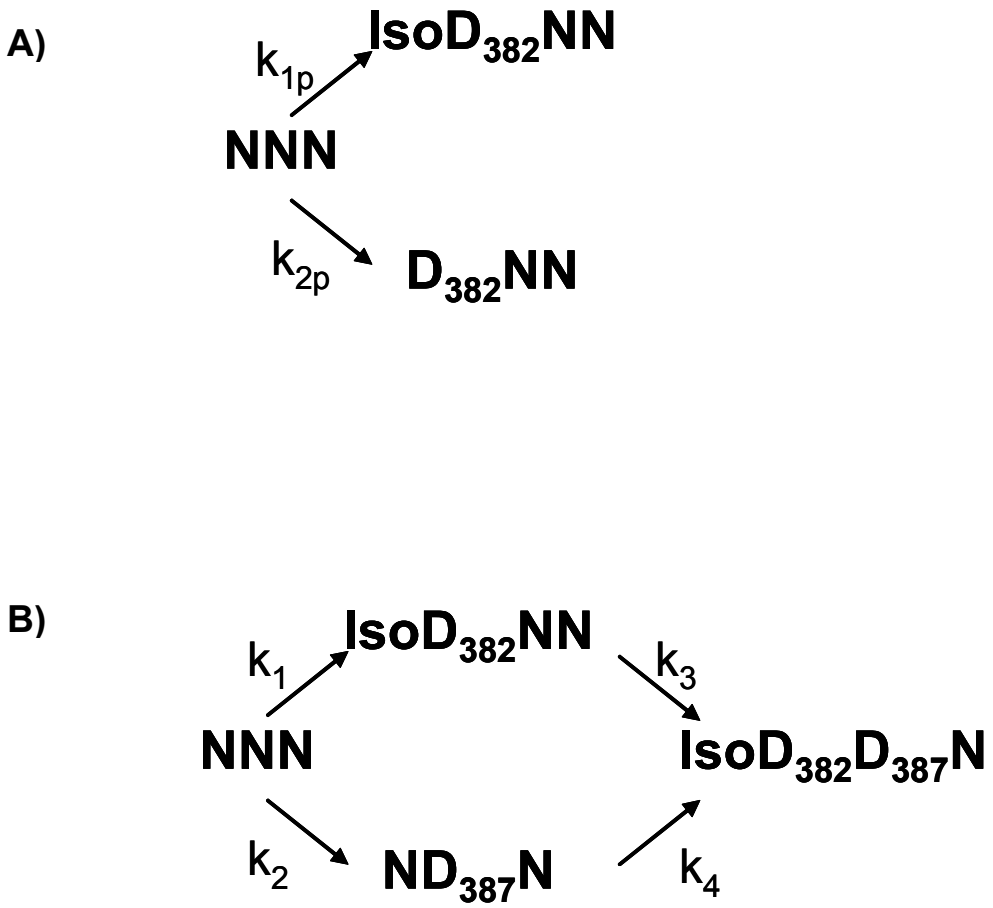


Figure 5.6. Reaction scheme showing the kinetics of deamidation of A) peptide digest assay and B) intact Fc.

Table 5.1. Fitting parameters obtained from global analysis of the data for the 369-390 tryptic peptide of Fc fragment

Parameter	Value	SD	CV (%)
k_{1p} (hours^{-1})	0.0390	0.0017	4.5
k_{2p} (hours^{-1})	0.0097	0.0012	12.3

SD= Standard Deviation, CV=Coefficient of Variance

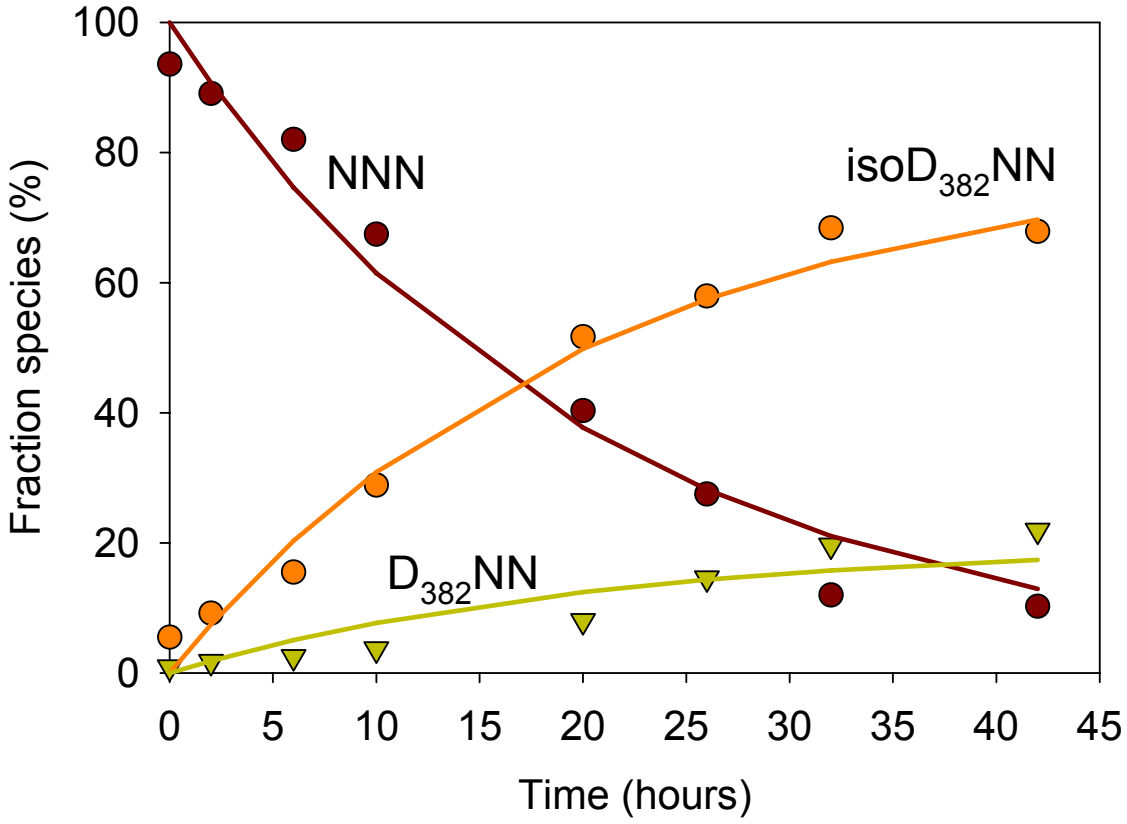


Figure 5.7. Kinetic profile for fragment 369-390, stressed after digestion of the Fc portion at 37°C, pH 7.4. The symbols are experimental data points, the lines represent the fit obtained with the kinetic model described in the text. N=3

predicted for peptides with the NG primary sequence by Robinson et al.[3] and with previous reports on unstructured model peptides with this sequence[20,23]. The half-life is significantly shorter than the value of 110 hours previously reported by Chelius et al. for deamidation in this fragment of an Fc IgG[10], however, a difference of approximately 8-fold. Since the two studies were conducted in solutions of identical pH, storage temperature, buffer, and buffer concentration, the reasons for this difference are unclear, but may involve differences in peptide concentration and/or partial retention of secondary structure in the previous study. Deamidation at N_{387} and N_{388} was not detected during the time course of the studies reported here. Since Tyr is C-terminal to N_{388}, the absence of deamidation products at this site is not surprising as steric hindrance by the bulky Tyr residue typically slows deamidation[3]. The absence of deamidation products at N_{387} in the present study suggests that the presence of a second N, C-terminal to the N of interest, slows deamidation rate. Long-term incubation of the digests at 37 ^{o}C for more than 48 h led to mass balance issues and so, the slower deamidation at N_{387} and N_{388} could not be monitored. Studies with the synthetic peptides (data not shown) also showed mass balance issues after being stressed for 48 h or more. Clipping of the S_{373}-D_{374} bond was observed with the peak for D_{374}-K_{390} peptide increasing over time.

5.3.2.2. Deamidation kinetics in the intact protein. – Kinetic profiles for deamidation in the intact Fc IgG are shown in Figure 5.8. The fully N-containing form of the protein (i.e., NNN) undergoes monotonic loss with corresponding increases in the singly deamidated (i.e., isoD_{382}NN, ND_{387}N) and doubly deamidated (i.e., iso$D_{382}D_{387}$N) products. At time zero, the percentage of the fully N-containing form is less than 100

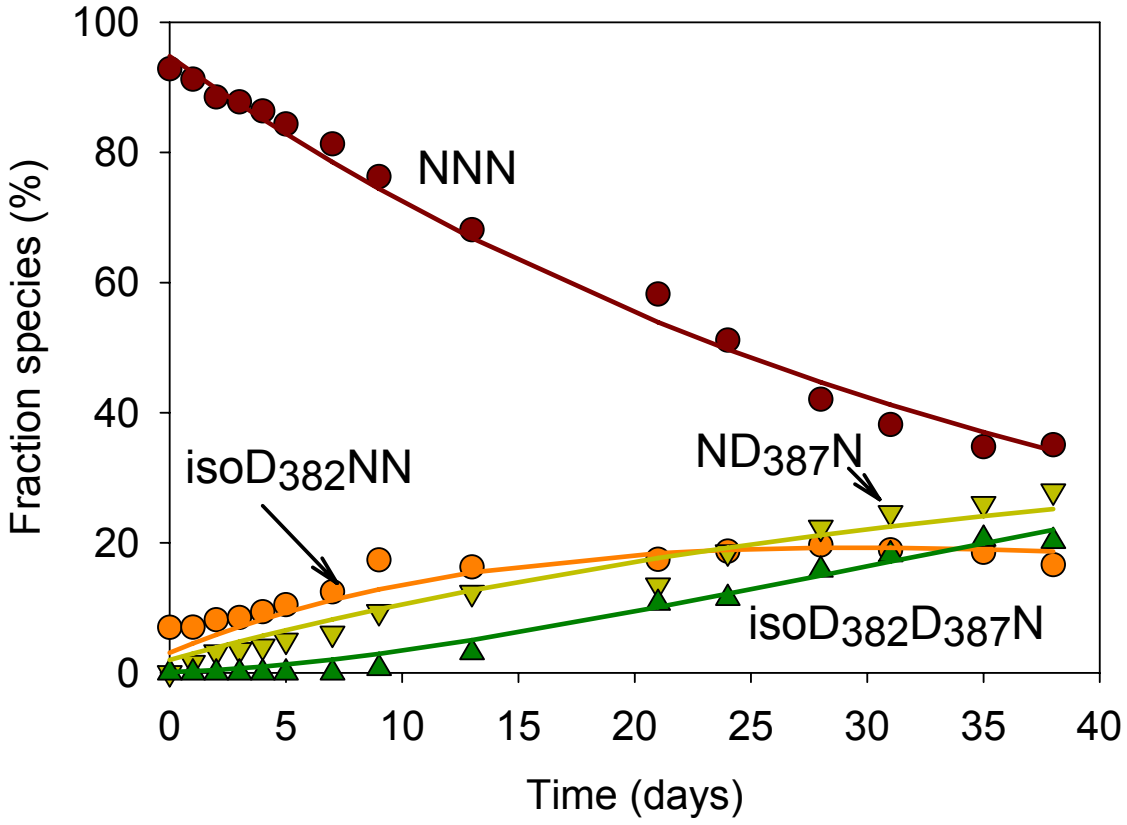


Figure 5.8. Kinetic profile for fragment 369-390, stressed Fc at 37°C, pH 7.4 followed by digestion. The symbols are experimental data points, the lines represent the fit obtained with the kinetic model described in the text. N=3

and the percentage of the isoD$_{382}$NN product is greater than zero, suggesting that partial deamidation of N$_{382}$ occurs during sample preparation or is already present in the protein at time 0. The D$_{382}$ product (i.e., D$_{382}$NN) was also detected throughout the study at a constant level of approximately 2-3% (not shown) consistent with its formation during sample preparation. A similar pattern was also evident in the digest samples (Fig. 5.7), albeit to a lesser extent.

Kinetic analysis of deamidation in intact Fc IgG was performed according to the reaction scheme shown in Figure 5.6B. The D$_{382}$NN product was treated as an artifact of sample preparation and measured amounts of D$_{382}$NN (1.5-2%) were reassigned to the parent form (NNN). Kinetic data for the parent protein and its degradation products then were fitted simultaneously to provide the microscopic rate constants (k_1,k_2,k_3,k_4; Fig. 5.6B). A time-shift parameter, t_0, was included to account for the history of the Fc fragment, i.e., deamidation occurring in the protein prior to initiation of the stability study and/or during sample preparation. Simultaneously fitting the time-dependent profiles for the four species with a model containing five global fitting parameters resulted in good agreement between experimental and theoretical values (not shown). However, values of the rate constant for loss of ND$_{387}$N (i.e., k_4) were small with large standard errors, indicating that the value was not significantly different from zero. Regression was then repeated with the constraint k_4=0, corresponding to a very slow conversion of ND$_{387}$N to isoD$_{382}$D$_{387}$N on the time scale of the experiment. Good agreement between theoretical and experimental values was again observed (Fig. 5.8) and the errors associated with the determination of the regression parameters decreased to acceptable levels (Table

Table 5.2. Fitting parameters obtained from global analysis of the data for the intact Fc fragment

Parameter	Value	SD	CV (%)
t_0 (days)	1.99	0.31	15.6
k_1 ($days^{-1}$)	0.0166	4.8200e-4	2.9
k_2 ($days^{-1}$)	0.0103	3.5780e-4	3.5
k_3 ($days^{-1}$)	0.0369	1.8970e-3	5.1

SD= Standard Deviation, CV=Coefficient of Variance

5.2). If the number of regression parameters was further reduced by requiring equality of the microscopic rate constants for the production and loss of the singly deamidated products (i.e., setting $k_1=k_3$ and $k_2=k_4$), the fit was not satisfactory. Reported model values (Table 5.2) and fitting lines (Fig. 5.8) thus correspond to the schematic shown in Figure 5.6B with k_4 =0 (fixed) and the inclusion of the time-shift parameter, t_0.

The calculated half-life for the overall loss of the parent protein is ~ 26 days (i.e., 0.693 $(k_1 + k_2)^{-1}$ = 25.8 d; Table 5.2), approximately 44-fold slower than the 14-hour half-life observed for the digests. The rate of formation of $isoD_{382}NN$ in the digests (k_{1p}) was approximately 60-fold greater than in the intact protein (k_1) (0.0166 d^{-1} vs. 0.94 d^{-1}), suggesting that reduction in the rate of deamidation at N_{382} is a dominant contribution to the overall reduction in degradation rate. The rates of formation of the two singly deamidated products, $isoD_{382}NN$ (k_1) and $ND_{387}N$ (k_2), were comparable in the intact protein. The rate of formation of $isoD_{382}NN$ (k_1) is less than the rate of loss of this intermediate, reflecting its low accumulation. The $ND_{387}N$ is formed at an approximately equal rate but degrades slowly (k_4 ~ 0) and accumulates through the time course of the study. While we have proposed that the doubly deamidated product ($isoD_{382}D_{387}N$) is produced by parallel reactions from both $isoD_{382}NN$ and $ND_{387}N$ (Fig. 5.6B), the global analysis suggests that this product is formed mainly from $isoD_{382}NN$, since k_4 ~ 0. That $k_3 > k_2$ further suggests that the formation of $isoD_{382}$ at accelerates the formation of D at position 387 (i.e., $ND_{387}N$). The D-product at the 382 position, $D_{382}NN$, was not observed in the intact protein, in agreement with the previous observation by Chelius et al.

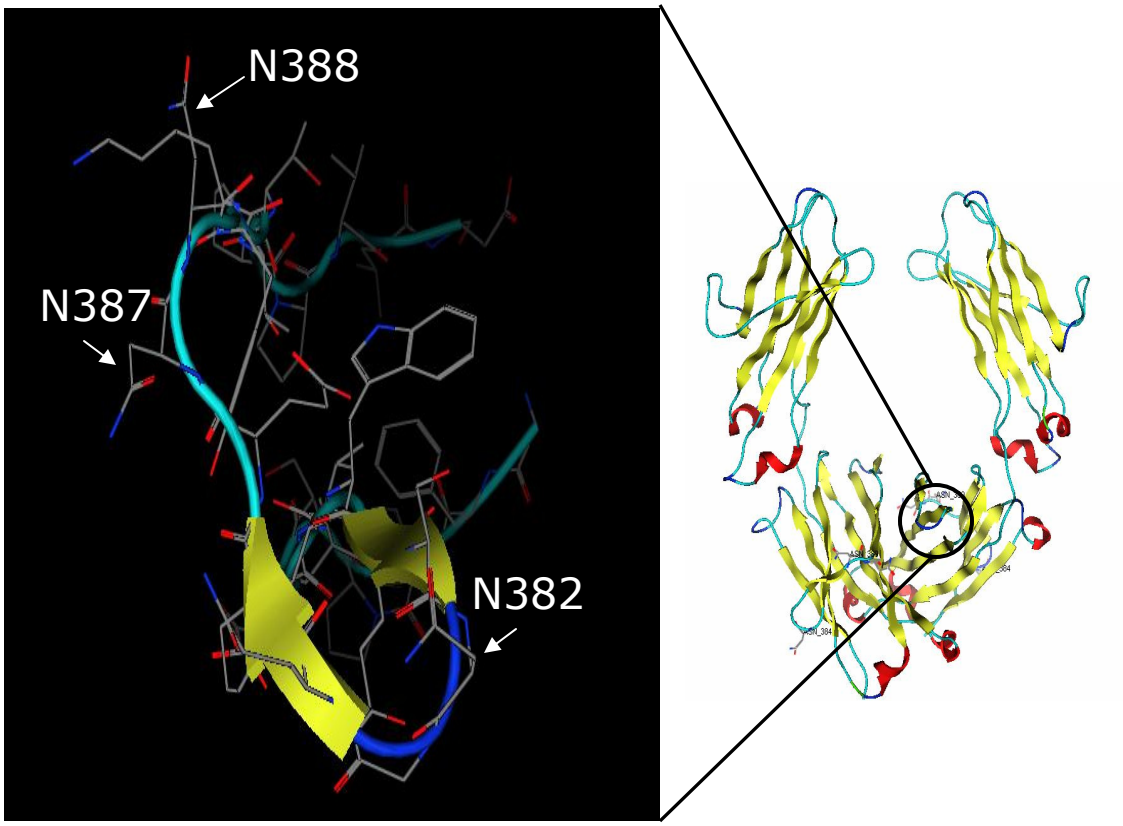


Figure 5.9. Molecular simulation showing the local environment around the 382, 387 and 388 asparagine sites.

5.3.3. Molecular dynamics simulation - Using the PDB crystal structure for human Fc-IgG1 (1h3u), SYBYL MDS revealed that N_{382} is located in a loop between two β-sheet regions (Fig. 5.9), while the N_{387} and N_{388} residues are located in a relatively unstructured loop. Solvent accessibility values and interatomic distances were also calculated for all three N residues (Table 5.3). Previous reports suggest that deamidation is favored by an interatomic distance between the attacking backbone nitrogen and the N-side chain carbonyl of 1.89 Å or less, while distances greater than 4.9 Å are unfavorable for deamidation. Here, the calculated interatomic distances are greater than 1.89 Å and less than 4.9 Å for all three N residues (Table 5.3), suggesting some susceptibility to deamidation based on structure. The smallest simulated interatomic distance corresponds to N_{382}, experimentally observed to deamidate most readily (Table 5.3). Calculated solvent accessibility values decreased in the order $N_{382} > N_{387} > N_{388}$, again corresponding to the measured deamidation rates (Table 5.3).

The MDS results also help to explain the somewhat unusual distribution of deamidation products in the intact protein, in which: (a) only the isoD product is observed at position 382, (b) only the D product is observed at position 387, and (c) no deamidation is observed at position 388. The formation of isoD$_{382}$NN to the exclusion of D$_{382}$NN suggests that hydrolysis of the succinimide occurs preferentially at the carbonyl contributed by the backbone amide and not at the carbonyl contributed by the N side chain. Such preferential hydrolysis could be related to the location of the 382 residue between two β-sheet regions, which may hinder the attack of water from the side-chain side. In contrast, hindered hydrolysis of the

Table 5.3. Solvent accessibility and interatomic distances obtained from molecular simulation studies with SYBYL software

Asparagine residue	Solvent Accessibility (%)	C=O to (N+1)NH Distance (Å)
N382	71.04	3.31
N387	66.82	4.36
N388	40.48	4.13

succinimide is an unlikely explanation for the preferential formation of D at the 387 position, since the site is relatively solvent exposed. The absence of deamidation products at position 388 probably reflects the combined effects of primary sequence, with deamidation hindered by the N_{388}+1 Tyr residue, and the limited solvent exposure of this site (Table 5.3).

5.4. Discussion

The rapid, high-resolution rpUPLC method was able to separate the three N-containing 22-amino acid tryptic fragments and their deamidated variants. The parent peptide was well resolved from its deamidated counterparts, thus emphasizing the utility of UPLC as a high resolution technique. The product profile for the intact protein (structured) differed from that obtained for the tryptic digest (unstructured). The digest deamidated approximately 50-times faster than the intact protein, suggesting that deamidation is inhibited in the structured protein. This hypothesis is supported by the product profiles for the intact and the digest samples. In the digests, the N_{382} site deamidated to form the expected $isoD_{382}$ and D_{382} products in the typical ratios of 4:1, while in the intact protein, only $isoD_{382}$ is observed with no formation of D_{382}. Molecular dynamics studies supported the formation of $isoD_{382}$ since the N_{382} residue was shown to be located in a structurally constrained orientation. In the digests, no deamidation was observed for N_{387} while in the intact protein, only the D_{387} product was observed at this site with no appearance of the corresponding $isoD_{387}$ form. Together, these results indicate a significant role of the local secondary structure in deamidation in this region of Fc IgG.

The product profile is similar to that reported previously by Chelius et al. for Fc IgG[10]. That group also observed deamidation at the N_{382} and N_{387} sites with exclusive formation of the $isoD_{382}$ and D_{387} products. However, the doubly deamidated product was not detected by Chelius et al. for the intact protein, while the succinimide intermediate at N_{382} was observed. The detection of the succinimide at N_{382} supports the hypothesis that the exclusive formation of the isoD product at this site involves the preferential hydrolysis of backbone side of the succinimide, though the succinimide was not detected in the current study. Interestingly, the half-life for loss of the parent tryptic fragment (i.e., G_{369}FYPSDIAVEWESNGQPENNYK_{390}) in the work by Chelius et al. was approximately 9 times greater than the 14-h half life observed here under identical conditions. Also, the intact Fc half-life for the intact protein was reported to be approximately 1.5 times greater than observed in the present study. These differences may be due to differences in pH and/or solution composition in the two studies.

Similar results have also been reported for deamidation of an N-residue located in the CDR1 region of an antibody[24,25]. In that work, the denatured protein formed isoD to D in a ratio of 3.5:1 but the native antibody deamidated via a succinimide intermediate to form only D under neutral and basic conditions. Unpublished work on deamidation in the CDR1 region by Vlasak et al. has shown that the ratio of isoD to D changes over time, presumably due to isomerization of isoD to D as the process tends toward equilibrium. Creighton et al.[26] studied deamidation in RNase A and observed that the structured protein deamidates 30-fold slower at N_{67} than in the reduced and denatured form, a difference attributed to the

local β-turn. Capasso et. al later observed that the isoD:D ratio for this protein was initially as high as in unstructured peptides (~ 3 : 1) but decreased to 1 : 2 as the reaction approached equilibrium[4]. A study by Harris et. al[13] of Herceptin®, a commercial monoclonal antibody drug product, showed that N_{30} in the light chain deamidates to form only the D product. Atypical isoD : D ratios have also been observed for crystallin[7,27]. Capasso et al. [28-30] proposed that the D product will dominate over its isoD counterpart wherever the D configuration has a lower energy, which in turn depends on the protein three dimensional structure, and noted that equilibration between isoD and D via the succinimide occurs more slowly than deamidation. A similar rationale may hold for deamidation at the N_{387} of the Fc IgG studied here, though our inability to detect the succinimide at this site does not provide support for this mechanism.

An alternative deamidation mechanism has been proposed by Clarke et al.[31], in which the backbone carbonyl oxygen rather than the (N+1) nitrogen attacks the side chain carbonyl to form an isonimide ring with release of water. Subsequent hydrolysis of the symmetric isonimide ring from either side yields only the D-containing product, unlike hydrolysis of the succinimide which produces both isoD- and D-containing products. This mechanism has been promoted in several studies in which the formation of the D-containing product was observed to the exclusion of the isoD product, with absence of the succinimide intermediate[32,33]. In the present studies, the formation of the D_{387} to the exclusion of $isoD_{387}$ is consistent with an isonimide mechanism at that site, as is our failure to detect the succinimide. However, since the isonimide intermediate was not detected here, this explanation must be regarded as speculative.

5.5. References

1. Chowdhury PS, Wu H 2005. Tailor-made antibody therapeutics. *Methods* 36:11-24.
2. Gottschalk U 2005. Downstream processing of monoclonal antibodies: from high dilution to high purity. *BioPharm International* 18(6):42-44, 46, 48, 50, 52, 54, 56, 58.
3. Robinson NE, Robinson AB 2001. Molecular clocks. *Proceedings of the National Academy of Sciences of the United States of America* 98(3):944-949.
4. Capasso S, Di Cerbo P 2000. Kinetic and thermodynamic control of the relative yield of the deamidation of asparagine and isomerization of aspartic acid residues. *J Pept Res* 56(6):382-387.
5. Capasso S, Salvadori S 1999. Effect of the three-dimensional structure on the deamidation reaction of ribonuclease A. *J Pept Res* 54(5):377-382.
6. Patel K, Borchardt RT 1990. Deamidation of asparaginyl residues in proteins: a potential pathway for chemical degradation of proteins in lyophilized dosage forms. *J Parenter Sci Technol* 44(6):300-301.
7. Robinson NE, Lampi KJ, McIver RT, Williams RH, Muster WC, Kruppa G, Robinson AB 2005. Quantitative measurement of deamidation in lens betaB2-crystallin and peptides by direct electrospray injection and fragmentation in a Fourier transform mass spectrometer. *Molecular vision* 11:1211-1219.

8. Robinson NE, Robinson AB 2001. Deamidation of human proteins. *Proceedings of the National Academy of Sciences of the United States of America* 98(22):12409-12413.

9. Vlasak J, Ionescu R 2008. Heterogeneity of monoclonal antibodies revealed by charge-sensitive methods. *Current Pharmaceutical Biotech* (In press).

10. Chelius D, Rehder DS, Bondarenko PV 2005. Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. *Analytical chemistry* 77(18):6004-6011.

11. Ikai A, Tanford C 1973. Kinetics of unfolding and refolding of proteins. I. Mathematical analysis. *J Mol Biol* 73(2):145-163.

12. Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P 2003. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 325(5):979-989.

13. Harris RJ, Kabakoff B, Macchi FD, Shen FJ, Kwong M, Andya JD, Shire SJ, Bjork N, Totpal K, Chen AB 2001. Identification of multiple sources of charge heterogeneity in a recombinant antibody. *Journal of chromatography* 752(2):233-245.

14. Johnson BA, Shirokawa JM, Hancock WS, Spellman MW, Basa LJ, Aswad DW 1989. Formation of isoaspartate at two distinct sites during in vitro aging of human growth hormone. *The Journal of biological chemistry* 264(24):14262-14271.

15. Sadakane Y, Yamazaki T, Nakagomi K, Akizawa T, Fujii N, Tanimura T, Kaneda M, Hatanaka Y 2003. Quantification of the isomerization of Asp

residue in recombinant human alpha A-crystallin by reversed-phase HPLC. *Journal of pharmaceutical and biomedical analysis* 30(6):1825-1833.

16. Stephenson RC, Clarke S 1989. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *The Journal of biological chemistry* 264(11):6164-6170.

17. Oliyai C, Borchardt RT 1993. Chemical pathways of peptide degradation. IV. Pathways, kinetics, and mechanism of degradation of an aspartyl residue in a model hexapeptide. *Pharmaceutical research* 10(1):95-102.

18. Clarke S 1987. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. *International journal of peptide and protein research* 30(6):808-821.

19. Geiger T, Clarke S 1987. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *The Journal of biological chemistry* 262(2):785-794.

20. Oliyai C, Borchardt RT 1994. Chemical pathways of peptide degradation. VI. Effect of the primary sequence on the pathways of degradation of aspartyl residues in model hexapeptides. *Pharmaceutical research* 11(5):751-758.

21. Patel K, Borchardt RT 1990. Chemical pathways of peptide degradation. II. Kinetics of deamidation of an asparaginyl residue in a model hexapeptide. *Pharmaceutical research* 7(7):703-711.

22. Patel K, Borchardt RT 1990. Chemical pathways of peptide degradation. III. Effect of primary sequence on the pathways of deamidation of asparaginyl residues in hexapeptides. *Pharmaceutical research* 7(8):787-793.

23. Li B, Gorman EM, Moore KD, Williams T, Schowen RL, Topp EM, Borchardt RT 2005. Effects of acidic N + 1 residues on asparagine deamidation rates in solution and in the solid state. *Journal of pharmaceutical sciences* 94(3):666-675.

24. Xiao G, Bondarenko PV, Jacob J, Chu GC, Chelius D 2007. 18O labeling method for identification and quantification of succinimide in proteins. *Analytical chemistry* 79(7):2714-2721.

25. Chu GC, Chelius D, Xiao G, Khor HK, Coulibaly S, Bondarenko PV 2007. Accumulation of succinimide in a recombinant monoclonal antibody in mildly acidic buffers under elevated temperatures. *Pharmaceutical research* 24(6):1145-1156.

26. Wearne SJ, Creighton TE 1989. Effect of protein conformation on rate of deamidation: ribonuclease A. *Proteins* 5(1):8-12.

27. Takemoto L, Fujii N, Boyle D 2001. Mechanism of asparagine deamidation during human senile cataractogenesis. *Experimental eye research* 72(5):559-563.

28. Capasso S, Balboni G, Di Cerbo P 2000. Effect of lysine residues on the deamidation reaction of asparagine side chains. *Biopolymers* 53(2):213-219.

29. Capasso S, Di Cerbo P 2000. Formation of an RNase A derivative containing an aminosuccinyl residue in place of asparagine 67. *Biopolymers* 56(1):14-19.

30. Capasso S, Di Donato A, Esposito L, Sica F, Sorrentino G, Vitagliano L, Zagari A, Mazzarella L 1996. Deamidation in proteins: the crystal structure of

bovine pancreatic ribonuclease with an isoaspartyl residue at position 67. J *Mol Biol* 257(3):492-496.

31. Clarke S, Stephenson RC, Lowenson JD. 1992. Stability of protein pharmaceuticals. In Ahern TJ, Manning MC, editors. *Chemical and physical pathways of protein degradation*, ed., New York: Plenum press. p 1-29.

32. Athmer L, Kindrachuk J, Georges F, Napper S 2002. The influence of protein structure on the products emerging from succinimide hydrolysis. *The Journal of biological chemistry* 277(34):30502-30507.

33. Napper S, Delbaere LT, Waygood EB 1999. The aspartyl replacement of the active site histidine in histidine-containing protein, HPr, of the Escherichia coli Phosphoenolpyruvate:Sugar phosphotransferase system can accept and donate a phosphoryl group. Spontaneous dephosphorylation of acyl-phosphate autocatalyzes an internal cyclization. *The Journal of biological chemistry* 274(31):21776-21782.

Chapter 6

Conclusions and Recommendations for Future Work

6.1 Summary and Conclusions

The objectives of this dissertation are broadly related to the characterization of protein drugs in solution and in the solid state. The specific objectives were: i) To determine the effects of secondary structure on deamidation in a tryptic fragment in the CH3 domain of an human IgG molecule in solution (Chapter 3), ii) to compare methods for quantifying glycosylation patterns in recombinant human IgG molecules in solution (Chapter 4), and iii) to analyze protein-excipient interactions in lyophilized solids using hydrogen/deuterium exchange (Chapter 5). The summary and conclusions for each of these specific objectives are as listed below.

6.1.1 Protein-Excipient Interactions in Amorphous Solids by Hydrogen/Deuterium Exchange with Mass Spectrometry (Chapter 5) – The objective of this project was to extend and validate a technique developed in our group[1-3] to[1] obtain region-specific information about protein-excipient interactions in the solid state. Hydrogen/deuterium exchange was used in conjunction with mass spectrometry to obtain information about the deuterium uptake capability of the intact protein and its fragments on digestion with pepsin.

A total of six proteins were studied – myoglobin, lysozyme, β-lactoglobulin, ribonuclease A, E casherin 5 and concanavalin A - with α-helical structure decreasing from myoglobin to con A. The HDX exchange for the intact protein was found to be sensitive to the nature of the excipient, with the extent of exchange varying by 2-5 fold for the proteins studied. The use of peptic digests in addition to the intact protein analysis allows excipient effects to be assigned to secondary structural domains of the proteins. The stabilization of proteins by excipients in the solid state occurs in a region-specific manner and not uniformly along the protein

backbone. Effects were primarily exerted in the proteins' α-helical regions, while β-sheet regions were protected to a lesser extent with low molecular carbohydrate excipients, trehalose and raffinose, showing significant protection against exchange. The results demonstrate the utility of H/D exchange with ESI-MS for analyzing protein-excipient interactions in lyophilized samples.

6.1.2 Comparison of LC and LC/MS methods for quantifying glycosylation in recombinant IgGs (Chapter 4) – The objective of this project was to identify a fast and accurate LC/ESI-TOF method to quantify the various glycoforms that would yield comparable values to the standard sugar cleavage assay method. The studies compare six methods for quantifying glycosylation in two production lots of a IgG: (i) LC/ESI-MS analysis of intact IgG ("intact IgG method"), (ii) LC/ESI-MS analysis of the Fc fragment produced by limited proteolysis with Lys-C ("IgG Fc method"), (iii) LC/ESI-MS analysis of the IgG heavy chain produced by reduction ("IgG HC method"), (iv) LC/ESI-MS analysis of Fc/2 fragment produced by limited proteolysis and reduction ("IgG Fc/2 method"), (v) LC/MS analysis of the glycosylated tryptic fragment (293EEQYNSTYR301) using extracted ion chromatograms ("XIC method") and (vi) normal phase HPLC analysis of sugars cleaved from the IgG using PNgase F ("sugar release assay").

The results highlight strengths and limitations of LC/ESI-TOF MS assays for the identification and quantitation of glycoforms in IgGs. ESI-TOF analysis of the intact IgG was able to adequately measure the galactose variance in the biantennary sugar structure, but could not resolve the heterogeneity caused by high-mannose carbohydrates. ESI-TOF analysis of the IgG-Fc fragment generated after limited proteolysis enabled detection of both biantennary and high-mannose carbohydrates

and was effective in characterizing oligosaccharide pairing caused by the combination of glycans on the two IgG-Fc heavy chains. Neither the intact IgG nor the IgG Fc analysis was found to provide sufficient resolution for quantitation, however. ESI-TOF analysis of the IgG-Fc/2 fragment showed accurate quantitation of various biantennary and high-mannose carbohydrates and was the most effective at the identification and quantitation of carbohydrates. Peptide mapping followed by ESI-TOF MS analysis was not effective for absolute quantitation, as the ionization of glycopeptides was influenced by the size of the carbohydrate. Though the sugar release assay showed high precision, the normal-phase method used for the assay could not fully resolve all the glycoforms. Collectively, the results suggest that MS quantitation based on analysis of Fc/2 (reduced Fc) is accurate and gives results that are both comparable and complementary to the more time-consuming sugar release assay.

6.1.3 Effect of Secondary Structure on Deamidation in a Tryptic Fragment of the Fc Portion of a Recombinant Monoclonal Antibody (Chapter 3) – The objective of this study was to develop an assay method to monitor deamidation in a particular tryptic fragment located in the CH3 domain of a recombinant human antibody and to assess the effect of secondary structure on deamidation.

The rapid, high-resolution rpUPLC method used in this study was able to separate the three N-containing 22-amino acid tryptic fragments and their deamidated variants. The parent peptide was well resolved from its deamidated counterparts, thus emphasizing the utility of UPLC as a high resolution technique. The product profile for the intact protein (structured) differed from that obtained for the tryptic digest (unstructured). The digest deamidated approximately 50-times

faster than the intact protein, suggesting that deamidation is inhibited in the structured protein. This hypothesis is supported by the product profiles for the intact and the digest samples. In the digests, the N_{382} site deamidated to form the expected $isoD_{382}$ and D_{382} products in the typical ratios of 4:1, while in the intact protein, only $isoD_{382}$ was observed with no formation of D_{382}. Molecular dynamics studies supported the formation of $isoD_{382}$ since the N_{382} residue was shown to be located in a structurally constrained orientation. In the digests, no deamidation was observed for N_{387} while in the intact protein, only the D_{387} product was observed at this site with no appearance of the corresponding $isoD_{387}$ form. Together, these results indicate a significant role of the local secondary structure in deamidation in this region of Fc IgG.

6.2. Future work

The results of the three projects suggest opportunities for additional research, as described below.

6.2.1 Protein-Excipient Interactions in Amorphous Solids by Hydrogen/Deuterium Exchange with Mass Spectrometry (Chapter 5) **–** The results of this study demonstrated that detailed information about protein-excipient interactions can be obtained by this technique. This suggests that hydrogen/deuterium exchange might be useful in predicting the aggregation propensity of proteins in the solid state. Exposure of aggregation-prone sequences could be measured by hydrogen/deuterium exchange and related to measured aggregation rates. Aggregation prone proteins like myoglobin, human growth hormone, con A etc. could be used as model proteins to validate this theory.

Aggregation can possibly be monitored with respect to temperature or RH which will provide information about the unfolding dynamics of the protein in lyophilized solids. The same study when done for proteins with commonly used excipients like trehalose can provide an insight into protein structure protection in presence of this excipients.

6.2.2 Comparison of LC and LC/MS methods for quantifying glycosylation in recombinant IgGs (Chapter 4) – Several techniques were compared in this study for their ability to quantitate glycosylation in the IgG molecule. The Fc/2 method was concluded to be the best due to its ease of sample preparation and fast analysis time. Its only disadvantage was that it could not resolve the isobaric forms. This limitation could be addressed by the use of a reverse phase column in series with a column that has the capability to separate isobars e.g. the normal phase column used for sugar release assay.

The type glycoform modulates the function of the antibody. For example: (i) antibodies having high mannose sugar content have been observed to be inactive in complement activation[4,5], (ii) glycoforms in IgG in patients with rheumatoid arthritis are known to contain higher G0 sugars[6], and (iii) IgGs with higher F0 sugars are known to have greater binding for human Fcg RIII[7]. The pharmacological effects could be related to a changes in protein higher order structure, since the sugars interact at various contact points with the amino acids in the CH2 region and with sugars from the complementary CH2 domain[8]. Crystal structures of IgG-Fc reveal a distinct conformation for the oligosaccharide resulting from multiple non-covalent interactions with the protein, such that each has a reciprocal influence on the conformation of the other[9-11]. It has also been reported that the deglycosylated

antibody is less thermally stable than the glycosylated variety[12]. Deglycosylation results in a more closed conformation of the IgG-Fc and more importantly leads to an increased internal disorder of the CH2 domains whilst the CH3 domain is unaffected[13]. Therefore, it would be interesting to investigate whether the variability in antibody function with different sugars is related to a change in higher order structure or if it imparts overall stability to the antibody. The study will involve purifying antibodies attached to various glycoforms e.g. Man5/Man5, G0F/G0F, G1F/G1F etc. Structural studies will involve monitoring the structure with Circular Dichroism or FTIR. Differential scanning calorimetry can also be used to determine the thermodynamic parameters for thermal unfolding, which will include a contribution from the intra-molecular oligosaccharide-protein interactions[12]. This will allow an understanding of the importance of the presence of an extra galactose unit or a high mannose glycoform like Man5. To my knowledge no studies on the effect of various glycoforms have been reported.

6.2.3 Effect of Secondary Structure on Deamidation in a Tryptic Fragment of the Fc Portion of a Recombinant Monoclonal Antibody (Chapter 3) – The results reported here demonstrate that secondary structure has a profound effect on deamidation product profile and kinetics. The N_{382} site deamidated to form $isoD_{382}$ while the N_{387} site deamidated to form only the D_{387} product. Although an effort was made to explain the selective formation of these products, additional evidence is needed to test the deamidation mechanism. The succinimide was not detected at either N_{382} or N_{387}, which is in disagreement with the previous report by Chelius et al.[14] and may suggest an alternative pathway for deamidation. An approach to isolating the succinimide would be to perform the study at a lower pH (~ 5). Acidic

conditions will stabilize the cyclic imide if it is formed thereby facilitating detection by mass spectrometry. Recent reports have also used $H_2{}^{18}O$ to identify and quantitate succinimide formation[15,16]., a method which provides an increase of +3 amu in deamidation products and may improve mass resolution. This approach could be taken to prove/disprove the formation of succinimide in the present study and to assess whether isomerization is responsible for detecting D_{387} at the N_{387} site.

6.3. References

1. Li Y, Williams TD, Schowen RL, Topp EM 2007. Characterizing protein structure in amorphous solids using hydrogen/deuterium exchange with mass spectrometry. *Anal Biochem* 366(1):18-28.
2. Li Y, Williams TD, Schowen RL, Topp EM 2007. Trehalose and calcium exert site-specific effects on calmodulin conformation in amorphous solids. *Biotechnol Bioeng* 97(6):1650-1653.
3. Li Y, Williams TD, Topp EM 2007. Effects of Excipients on Protein Conformation in Lyophilized Solids by Hydrogen/Deuterium Exchange Mass Spectrometry. *Pharm Res*.
4. Wright A, Morrison SL 1994. Effect of altered CH2-associated carbohydrate structure on the functional properties and in vivo fate of chimeric mouse-human immunoglobulin G1. *The Journal of experimental medicine* 180(3):1087-1096.
5. Wright A, Morrison SL 1997. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol* 15(1):26-32.
6. Watson M, Rudd PM, Bland M, Dwek RA, Axford JS 1999. Sugar printing rheumatic diseases: a potential method for disease differentiation using immunoglobulin G oligosaccharides. *Arthritis and rheumatism* 42(8):1682-1690.
7. Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, Weikert SH, Presta LG 2002. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *The Journal of biological chemistry* 277(30):26733-26740.

8. Padlan EA 1990. On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins* 7(2):112-124.

9. Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, Lund J, Jefferis R 2000. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Molecular immunology* 37(12-13):697-706.

10. Mimura Y, Ghirlando R, Sondermann P, Lund J, Jefferis R 2001. The molecular specificity of IgG-Fc interactions with Fc gamma receptors. *Adv Exp Med Biol* 495:49-53.

11. Mimura Y, Sondermann P, Ghirlando R, Lund J, Young SP, Goodall M, Jefferis R 2001. Role of oligosaccharide residues of IgG1-Fc in Fc gamma RIIb binding. *J Biol Chem* 276(49):45539-45547.

12. Ghirlando R, Lund J, Goodall M, Jefferis R 1999. Glycosylation of human IgG-Fc: influences on structure revealed by differential scanning micro-calorimetry. *Immunol Lett* 68(1):47-52.

13. Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P 2003. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 325(5):979-989.

14. Chelius D, Rehder DS, Bondarenko PV 2005. Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. *Anal Chem* 77(18):6004-6011.

15. Xiao G, Bondarenko PV, Jacob J, Chu GC, Chelius D 2007. 18O labeling method for identification and quantification of succinimide in proteins. *Anal Chem* 79(7):2714-2721.

16. Terashima I, Koga A, Nagai H 2007. Identification of deamidation and isomerization sites on pharmaceutical recombinant antibody using H(2)(18)O. *Analytical biochemistry* 368(1):49-60.

APPENDIX

Deamidation of 369-390 tryptic peptide of Fc fragment

The analysis of the deamidation of the 369-390 tryptic peptide of Fc fragment was performed according to the reaction shown in Schematic A.I. Using the following notations: NNN=N, $IsoD_{382}NN$=ID, $D_{382}NN$=D, the system of differential equations can be written as follows for 3 species (n=3):

$$\frac{dN}{dt} = -(k_{1p} + k_{2p})N$$

$$\frac{dID}{dt} = k_{1p}N$$

$$\frac{dD}{dt} = k_{2p}N$$

The eigenvalues of the system are $\lambda_1=k_{1p}+k_{2p}$; $\lambda_2=\lambda_3=0$ and the time dependencies of the molar fraction of each species are: $N(t)=\exp(-\lambda_1 t)$, $ID(t)=[k_{1p}/(k_{1p}+k_{2p})]x[1- \exp(-\lambda_1 t)]$, $D(t)=[k_{2p}/(k_{1p}+k_{2p})]x[1- \exp(-\lambda_1 t)]$.

Deamidation of intact Fc fragment

The analysis of the deamidation of the intact Fc fragment was performed according to the reaction shown in Schematic A.II. Using the following notations: NNN=N, $IsoD_{382}NN$=ID, $ND_{387}N$=D, $IsoD_{382}$ $D_{387}N$=DD, the system of differential equations can be written as follows for four species (n=4):

$$\frac{dN}{dt} = -(k_1 + k_2)N$$
$$\frac{dID}{dt} = k_1 N - k_3 ID$$
$$\frac{dD}{dt} = k_2 N - k_4 D$$
$$\frac{dDD}{dt} = k_3 ID + k_4 D$$

The eigenvalues of the system are $\lambda_1=k_1+k_2$, $\lambda_2=k_3$, $\lambda_3=k_4$, $\lambda_4=0$ and the time dependencies of the molar fraction of each species are: $N(t)=\exp(-\lambda_1 t)$, $ID(t)=[k_1/(k_1+k_2-k_3)]x[\exp(-\lambda_2 t)- \exp(-\lambda_1 t)]$, $D(t)=[k_2/(k_1+k_2-k_4)]x[\exp(-\lambda_3 t)- \exp(-\lambda_1 t)]$, $DD(t)=1-[N(t)+ID(t)+D(t)]$.

References

Ikai, A. and Tanford, C (1973) "Kinetics of Unfolding and Refolding of Proteins" *J. Mol. Biol*. 73, 145-163