INVESTIGATING THE CHARACTERISTICS OF COLLAGEN IN PROTEIN-PROTEIN INTERACTIONS THROUGH NMR RELAXATION EXPERIMENTS AND INDIRECT ELISA BINDING ASSAYS

By

ROBERT T. YOUNG

A thesis submitted to the

Graduate School- New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Computational Biology & Molecular Biophysics

written under the direction of

Dr. Jean Baum

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October 2013

ABSTRACT OF THE THESIS

Investigating the Characteristics of Collagen in Protein-Protein Interactions through NMR Relaxation Experiments and Indirect ELISA Binding Assays

by Robert T. Young

Thesis Director:
Dr. Jean Baum

Triple helical collagen in the most abundant protein found in the human body. It is made of three right-handed polyproline chains that supercoil together with an n-1 stagger to produce a rigid left-handed helix. Collagen has various types based on the combination of chains used to assemble the helix and each type can generate various higher order structures, such as a thick collagen fibril. Collagen's unique G-X-Y repeated sequence and interstrand bonding network contributes to the protein's high overall structural integrity, with the GPO triplet (GPO) being the most rigid structure and, therefore, the most often repeated sequence. A subdomain in the collagen sequence with greater internal motion and flexibility is known to be areas of molecular recognition for proteins such as the transmembrane protein integrin or the MHC-I protein subunit β2-microglobulin (B2m). As full length collagen proteins are approximately 1000 residues in length, collagen model peptides of thirty residues are often used for characterization experiments in nuclear magnetic resonance (NMR) but are not necessary in indirect

ELISA binding assays

NMR relaxation experiments were conducted on a homotrimeric collagen model peptide GFOGER, designed from the high affinity integrin binding sequence in heterotrimeric type I collagen. Previous research indicated key areas of binding between the two proteins, but flexibility information about the six-residue sequence is essential for future binding studies with heterotrimer models. Experiments with longitudinal relaxation and multiple transverse relaxation experiments have indicated the flexible nature of the insert, and further work will be conducted to identify individual trimer residues and their roles in integrin binding.

Indirect ELISA binding assays were conducted between full length type I collagen and B2m to first confirm reports from previous studies, and to work toward potentially identifying the binding domains on collagen that result in the initial binding and subsequent aggregation of B2m. Work in replicating the binding assays did confirm a strong binding interaction between collagen and B2m by experimenting with constant collagen concentrations and various B2m concentrations, ranging from the normal serum level to levels that are what is found in patients undergoing hemodialysis attributed to renal failure.

Acknowledgments & Dedication

There are a lot of people who have helped me get to this particular point in my academic career. Those who have guided me, encouraged me, supported me, and advised me (especially when I didn't want to hear it), I cannot thank you enough. Most importantly, I offer my thanks to the following people.

To those who saw something in me and did their best to prepare me for the rigors of this intellectually-challenging journey, especially Dr. Jennifer Burris, Dr. Leslie Sargent-Jones, Dr. Angela Mead, and Dr. Jennifer Perry-Cecile.

To those who helped me through the time of this thesis, either through content discussions, technique training, or words of encouragement, especially Dr. Jean Baum, Dr. Vikus Nanda, Dr. Darrin York, Dr. David Case, Dr. Joseph Marcotrigiano, Dr. Ana Monica Nunes, Gina Moriarty, Maria Janowska, Jackie Sikora, Maya Lita McCloud, and Amy Mounkes

It is my honor to dedicate this work to my brother, L. Bradley Young, and my sister, Beth Ann Bird. When things got tough, they reminded me of why I wanted to go in this direction and never once wavered on their support. Without them, this would not have happened.

Table of Contents

List of Tables

List of Figures & Illustrations

Introduction: The Triple-Helical Collagen Protein

Biological cells have a sophisticated internal skeletal system that provides not only its overall shape, but a network in which organelles can shift throughout the cytoplasm. Without this ability to provide internal organization, functionality in a cell would suffer. The same can be said for the external environment of a cell or a collection of bonded cells. However, the internal skeletal system cannot be used externally. A separate, more stable and sophisticated environment must be in place to anchor differentiated cells within a specific location as well as provide opportunities to pass information between cellular groups. This is known as the extracellular matrix and in it sits the most abundant protein in the human body: collagen. This rigid, rod-like protein is so critical in an system's organization that collagen accounts for up to 33% of all proteomic content[1,2]. Collagen research has been conducted for several decades by some of the most brilliant thinkers in the early 20th century, such as American chemist Linus Pauling, who proposed a structure of collagen in 1951; Indian physicist G.N. Ramachandran, who presented sequence and composition data in 1955; and English biophysicist Francis Crick, who helped in refining Pauling's original proposal in 1961 to the structure that is accepted today[3,4,5,6] .

This critical protein contains a unique structural motif in that it is made by supercoiling three polypeptide chains together to form its characteristic triple helix[1,7,8] . These proline-rich polypeptide chains (polyproline-II (PP-II) chains) adopt a rigid left-handed configuration and when unbound, have a classical three residue per turn geometry with a 3.1Å rise per residue pattern[7,9,10] . When trimerization occurs during the molecular assembly of collagen, the PP-II chains combine at the C-terminal regions of each strand

with a one-residue stagger between the chains, nucleates at the most rigid areas of the chains, and propagates until a supercoiled, right-handed triple helix is formed, seen in Figure 1[1,11,12]. This supercoiling of the helix results in more residues per turn in the structure (up to 3.33) and, therefore, only a 2.9Å rise per residue[11]. This does not occur in the extracellular matrix, but rather in the endoplasmic reticulum of cells to then secreted out of the membrane, from which further modifications of the triple helix results in a higher order fibril organization of collagen[13].

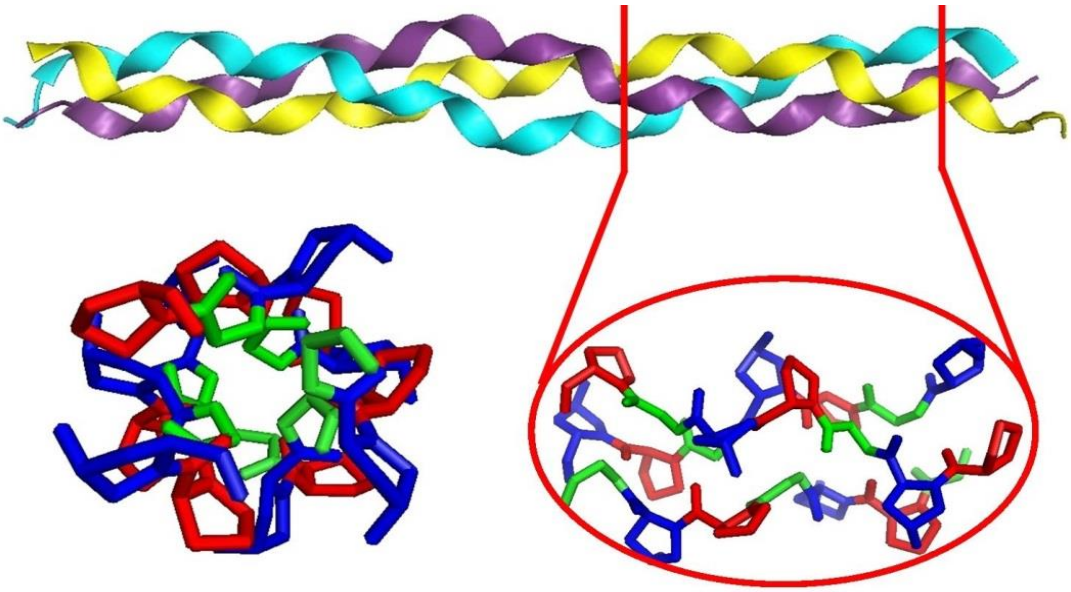

Figure 1 A schematic view of the collagen triple helix where, at top, is the collagen peptide $(POG)_4$-POA-$(POG)_5$ [PDB: 1CGD][14] used to view the three PP-II chains supercoiling together. At bottom left is a cross-section of collagen [PDB: 1V4F][15] where the green residues that are closely packed next to the helical axis are glycine residues, proline residues in red, and hydroxyproline residues in blue. At the bottom right is a zoomed-in view of the triple helical structure, showing the n-1 staggering that occurs between the chains [PDB:1V4F] [15,16]

Another unique characteristic of collagen is that the PP-II chains follow a strict triplet peptide sequence repeat, X-Y-G, where the residues in the X-position are proline (approximately 28% of the protein's amino acid content) and those in the Y-position are frequently hydroxyproline (approximately 38%)[1,10,17]. Every third residue must be a glycine residue as it is requisite for the overall structure and stability of the protein as it does not have a side chain and will allow for close packing of the PP-II chains around a

common helical axis without any steric distortions[13,17]. These glycine residues are therefore essential towards maintaining high level of structural integrity that collagen is known for[17,18]. Sequence analysis in previous experiments have reported that the most repeated triplet is P-O-G at 11% of total triplets in sequence and due to the high imino acid content, it is the most rigid of all triplets possible[1,10,19].

The imino acid content not only provides the rigid structure in the collagen chains, but it also is the bedrock of an extensive hydrogen bonding networking between the neighboring chains. Research conducted by Pauling and Corey[3], Rich and Crick[6], and Bella and Berman, *et al* [20], among others, have reported on how such a network increases the global stability of the protein and has gone so far as to identify the locations of these bonds. This bonding network is based on two hydrogen bonds that are formed for every triplet in sequence between chains: one hydrogen bond is formed between one chain's glycine amide hydrogen and a carbonyl group on a neighboring chain's X-residue (seen in Figure 2); the other hydrogen bond exists between the alpha carbon's hydrogen on either the glycine or the Y-residue and the carbonyl group on the X-residue or glycine[3,6,20]. Further studies by the Helen Berman and Barbara Brodsky's groups have shown evidence of water-mediated hydrogen bonding both intrachain and interchain with carbonyl groups as well as hydroxyproline alcohols with carbonyls intrachain and interchain to the atoms[14].

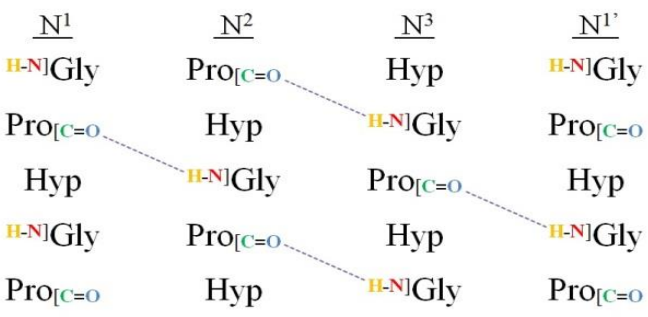

Figure 2 A view of the hydrogen bond between the amide hydrogen from the glycine residue with the carbonyl group on the proline residue on the adjacent chain, where $N^{1'}$ is the first chain to view the interaction between Pro in chain N^3

Due to the POG triplet's relatively low frequency in the PP-II chain, the rod-like collagen proteins, which can stretch up to 1000 residues in length, is neither uniform in stability nor in overall structure[11]. These subdomains within collagen, with greater flexibility and molecular motion because they can unfold cooperatively as an independent unit, are the first to be considered and experimented on as possible molecular recognition and binding sites between collagen and other proteins[2,19]. The natural prevalence of heterotrimeric collagens over homotrimers points researchers in a direction of higher subdomain content with molecular recognition possibilities[1]. Further functional analysis must transcend sequential analysis, much of which has been vastly studied, as the interchain bonding networking and steric interactions between the chains have a strong connection to flexibility and recognition. Even a decrease to the degree of hydrogen bonding within a subdomain can increase or decrease its recognition ability[1,19].

In the human system alone, there are 28 various collagen types made up of close to 50 potential chains, each collagen type creating either higher-order fibrils, helical networks, or other throughout the body in key locations to provide stability and support between cells and the extracellular matrix, where collagen proteins reside[1,19]. For instance, type I collagen is a fibril-forming protein made up of two α1 type I chains and a single α2 type I chain found in tendons and is the major organic matrix in bone, type II collagen, also fibrillar, is a homotrimer of α1 type II chains and is found predominantly in cartilage, and type IV collagen is a networking collagen distributed among basement membranes and has three possible chain compositions[1,9,16]. The subdomains that exist on these specific collagen types allow for increased or decreased rigidity based on external stimuli and even with binding of a specific protein for allow for tissue or barrier repair.

With a large number of different collagen types comes a series of diseases that arise from mutations within the molecular triple helix that has large-scale implications to either the larger collagen fibril or to the basement membrane network that it supports[10]. For instance, with a non-fibrillar collagen, glycine mutations or deletions in the triplet sequence repeats are typically a part of the nature of the protein and these breaks, such as the twenty breaks in type IV collagen, have gone relatively unexplored[21]; however, a single Gly→ Any residue mutation (such as glycine to alanine, an amino acid with a small side side) can result in connective tissue diseases such as osteogenesis imperfecta (OI), known as "brittle-bone" disease, when such a mutation occurs in type I collagen[22]. This single mutation occurs at the molecular level and may occur once, but as mentioned previously, the glycine residue is required to keep that signature stability of the triple helix and once the flawed triple helix assembles into a higher order fibrillar structure, that single kink within the organic matrix in bone can be devastating and even fatal[9,16,17,18].

Research with collagen has limitations due to the sheer size and global structure of the protein. While the use of full length type I and II collagen proteins may be used in exploratory research into protein-protein binding studies, experiments in nuclear magnetic resonance (NMR) require modifications in size[23]. Therefore, a significant amount of collagen experimentation for biophysical characterization requires collagen model peptides: triple helical monomers with a maximum 30 residues per chain. Although this seems rather limiting, collagen model peptides have proved invaluable in studying subdomains of collagen. The models are frequently designed to have the rigid POG triplets flank the subdomains in question. Work with these collagen model peptides have shed light on wild type sequence flexibility and changes due to sequence mutations.

Such works include research conducted on the model peptides T3-785 (seen in Figure 3), designed based on a type III collagen sequence starting at residue 785 and located C-terminal of a collagenase cleavage site, and T1-858, from type I collagen in experiments with Gly→Ser (serine) mutations from OI[24,25].

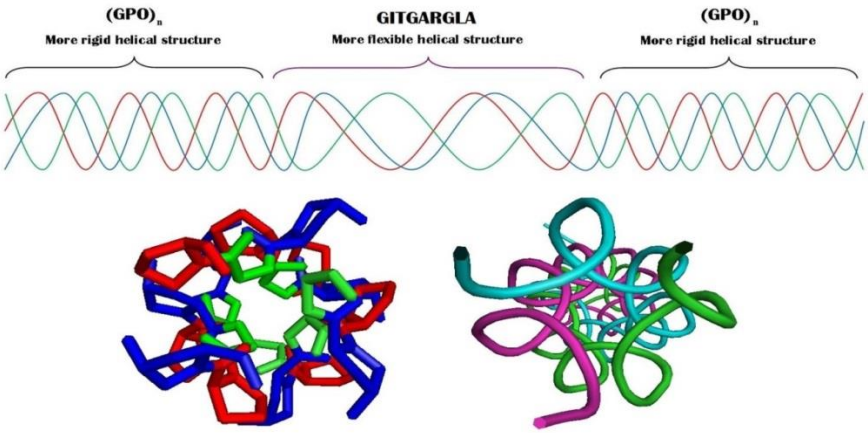

Figure 3 A schematic diagram of the T3-785 collagen model peptide and its more-flexible insert based on the type III collagen sequence of interest compared to the more rigid G-P-O repeat flanking the insert and, at the bottom, a comparable look down the helix between a GPO-repeat region [PDB: 1V4F] and the T3-785 [PDB: 1BKV], which will begin to curve out from the axis due to the internal flexibility brought on by the insert

With this information, two projects were conducted to investigate the versatility of collagen as a binding protein in two ways. The first project focuses on the role of a subdomain's molecular flexibility with binding and how NMR experiments can be used to characterize such properties. A series of NMR experiments relaxation experiments were conducted to study the flexibility of a very powerful subdomain in type I collagen that binds with integrin, a transmembrane protein. The second project was focused on elucidating protein-collagen interactions through binding assays. Using full length type I collagen and the major histocompatibility complex-I subunit known as β2-microglobulin, which catabolizes off the complex at a regular rate, circulates in the blood stream, and causes joint problems with patients undergoing hemodialysis treatments due to renal failure.

Chapter 1: Biophysical Characterization of Collagen Model Peptide GFOGER

Although collagen is a protein that exists in the extracellular matrix, its influence on cellular behavior is staggering. By having a specific domain on either one of its three chains or a collection of residues on multiple chains, collagen's presence can improve a cell's response to its environment. Based on the extracellular protein extending from the cellular membrane interacting with collagen, the cell could improve upon its adhering to the matrix, undergo cellular growth and division, or even express differentiating phenotypes[26,27]. One such protein is the transmembrane protein integrin, the principal family of cell surface proteins that interact with the matrix which is composed of generally conserved structures of α-β heterodimers[27]. Of the many α-subunits, only α1, α2, α10, and α11 can be combined with the β1 domain to constitute the native collagen-binding integrin family, with α2β1 as one of the most widely distributed heterodimers[28]. The α2β1 integrin protein is found on platelets in the circulatory system and is used to adhere platelet to a damaged blood vessel wall, whose damage will result in exposed type I collagen proteins[29,30]. This activity will lead to platelet aggregation and repair of the vessel wall.

A fascinating feature of the α2β1 integrin protein is that it can exist in at least two affinity states, which previous crystallographic experiments have shown to be the two conformations of the α2 I-domain[31]. The I-domain is a 200-residue subunit central to ligand binding due to a very important metal ion-dependent adhesion site (MIDAS site) situated at the upper area of the domain[26,29]. In order for collagen-integrin binding to typically occur, the transmembrane protein must first become activated, as it is typically held at its low affinity resting state.

Integrin proteins bind to a specific six-residue domain: GXX'GEX", where X' is typically hydroxyproline and X" frequently arginine[29,32]. The affinities of this domain can range from low/null to high. An example of a binding domain is the sequence GFOGER, a six-residue sequence found in the α1 chain of type I collagen. The affinity range of integrin to this six-residue sequence can be due to factors such as the heterodimer configuration of integrin, its activation state, or even the state of collagen surrounding the binding domain. With this particular domain, GFOGER has such a high affinity binding affinity that is has the ability to bind to integrin while the transmembrane protein is in its resting state[27]. Experimental work conducted by the Farndale group from the University of Cambridge has shed significant light on the GXX'GEX" collagen domains through the use of toolkits, or large sets of overlapping triple-helical peptide sequences from collagen[32]. One such advancement from the toolkit approach has allowed for the crystallization of a homotrimer model peptide of collagen with GFOGER flanked by GPO-repeats bound to a separated I-domain from α2 (see Figure 4)[27].

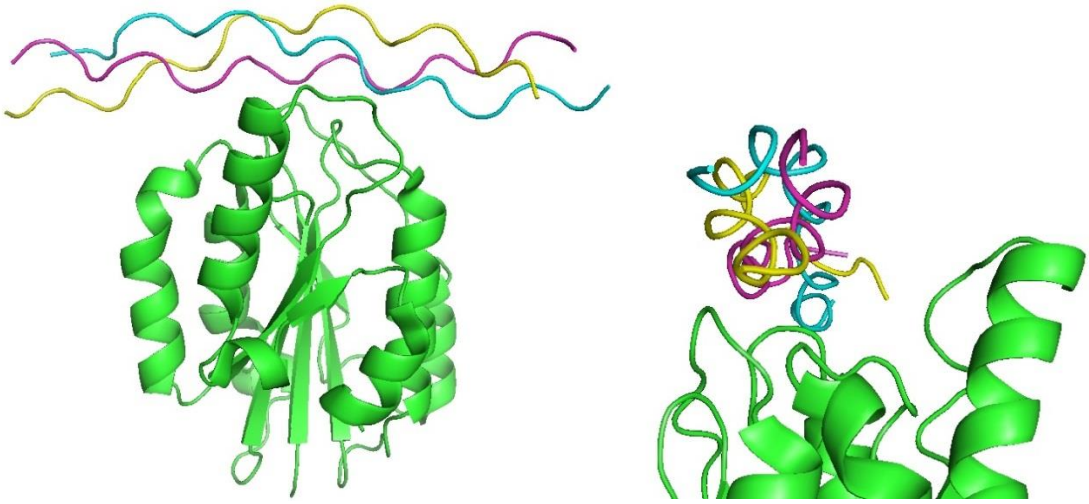

Figure 4 Integrin alpha2-beta1 I-domain (green) bound to a collagen model peptide with GFOGER inserted between GPO-rich regions (with leading chain in yellow, middle chain in purple, lagging in cyan). Image on left shows both the full I-domain and the GFOGER model peptide and the right shows the axial view of the model peptide in relation to the integrin I-domain to show the bend of the collagen's overall structure[PDB: 1DZI][27]

Because of the significance of the high-affinity sequence, closer study into the biophysical nature of GFOGER is paramount. Previous characterization of a smaller GFOGER collagen model peptide unbound and bound to an α2 subunit of integrin showed an increase to the overall angles of the model peptide between three distinct regions. Further analysis showed that of the three collagen strands, the middle strand was responsible for completing the metal ion coordination sphere in the MIDAS region of integrin[26]. Unlike the type I collagen that it is found in, the GFOGER model peptide is homotrimeric and the data previously collected indicates a biologically relevant internal flexibility allowing for protein-protein binding. Therefore, the goal for this project will be to investigate the flexibility of this molecular recognition site by characterizing the $(GPO)_4GFOGER(GPO)_4$ peptide using NMR experiments, where Gly7, Gly13, and Phe14 are labeled with ^{15}N, including a series of longitudinal and transverse relaxation experiments to observe its backbone dynamics.

Methods & Materials

A series of nuclear resonance experiments were conducted in order to biophysically characterize the GFOGER model peptide, whose configuration is seen in Figure 5. Only three residues of the peptide were labeled: Gly7, Gly13, and Phe14. The Gly7 residue sits in the heart of a POG-rich environment while the –GF- neighboring residues sit at the end of the N-terminal POG- cap and are situated N-terminal of a –GER- triplet. Heteronuclear Single-Quantum Coherence (HSQC) experiments were conducted on a Varian 600MHz NMR and also on a Bruker 700Mhz at 10°C to determine the specific ^{1}H-^{15}N spectra for further identification and comparison for additional experiments as spectral data from such experiments are unique to the peptide.

1)	G	P	O	G	P	O	G	P	O	G	P	O	G	F	O	G	E	R	G	P	O	G	P	O	G	P	O	G	P	O	
2)		G	P	O	G	P	O	G	P	O	G	P	O	G	F	O	G	E	R	G	P	O	G	P	O	G	P	O	G	P	O	
3)			G	P	O	G	P	O	G	P	O	G	P	O	G	F	O	G	E	R	G	P	O	G	P	O	G	P	O	G	P	O
1)	G	P	O	G	P	O	G	P	O	G	P	O	G	F	O	G	E	R	G	P	O	G	P	O	G	P	O	G	P	O	

Figure 5 The triple-helix residue map for the model peptide GFOGER, which all isotope-labeled residues in red and the GFOGER insert highlighted

From the HSQC data, a series of relaxation experiments were conducted on the GFOGER peptide, mainly R1, NOE, and R2 relaxation experiments in a Varian 600MHz NMR at 10°C. Two R2 experiments were conducted on the peptide: The first with a delay time of 2.000sec and a τ_d of 0.625msec and the second based on modifying the pulse sequence of the previous R2 experiment ("modR2"), using a τ_d of 0.300msec and a delay of 2.300sec.

After the initial series of experiment, two additional relaxation experiments were conducted as a means to observe large internal motion within the peptide and for details about conformational exchange. Unlike the two previous R2 experiments, whose pulse

sequence causes suppression of conformational exchange by repeatedly perturbing the peptide with radio frequency (rf) pulses, an R2 Hahn-Echo experiment ($R2^{HE}$) uses a pulse sequence with a longer detection time use in obtaining the full conformational exchange of the peptide.

Results

HSQC Experiments

Any information about the flexibility of the insert GFOGER must start from data collected during HSQC experimentation, as seen in Figures 6(a) and 6(b), due to the unique and unchanging spectrum it generates for the peptide. The peptide is a homotrimer with only three residues labeled with ^{15}N, yet peaks are expected for both trimer signals, due to interchain interactions, and for monomer signals, due to the unraveling of the triple helix during experimentation. Figure 6(a) shows the anticipated ten peaks: three monomer peaks (one for each labeled residue), one lone peak for trimer Gly-7 due to the homogeneity of its surroundings, three trimer peaks for Gly-13, and three trimer peaks for Phe-14. The HSQC shows two coupled peaks (E-F and H-I) and further analysis confirmed that they are, in fact, separate residue signals.

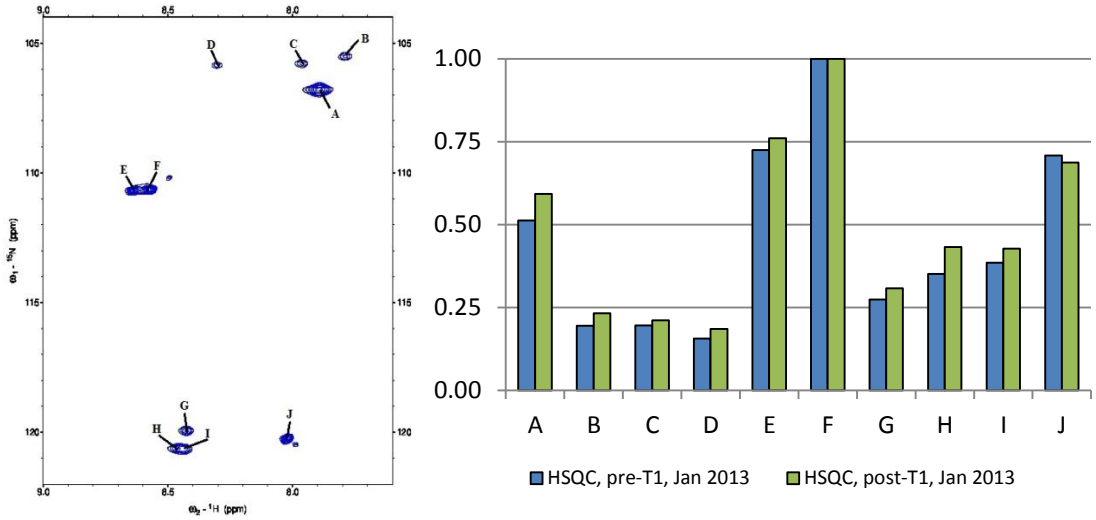

Figure 6 (a, left) An HSQC spectrum of the collagen model peptide GFOGER at 10°C from a Varian 600MHz NMR; (b, right) Normalized HSQC peak intensities of GFOGER taken before and after a T2 relaxation experiment

During processing, normalized HSQC peak intensity values, seen in Figure 6(b), were used to determine the monomer peaks of the peptide due to the difference in signal

strength between a monomer signal and a trimer signal. A second HSQC spectrum was generated, this time overlapping an HSQC of the GFOGER collagen peptide with the highly-documented T3-785 model peptide, seen in Figure 7(a). This was done to determine the location of both the Gly-7 monomer and trimer peaks, as they have the same interchain environment as the Gly-24 residue in T3-785, as well as to differentiate glycine and phenylalanine residue peaks, since the Phe-14 peaks will sit well below the Gly residues on the HSQC. Monomer peaks of the GFOGER peptide and the lone Gly-7 trimer peak can thus be selected and are labeled in Figure 7(b); yet there is not enough evidence to pick specific peaks for the Gly-13 and Phe-14 trimer signals, so a set of arbitrary labels were attached to these peaks to be used during relaxation experimentation for analysis.

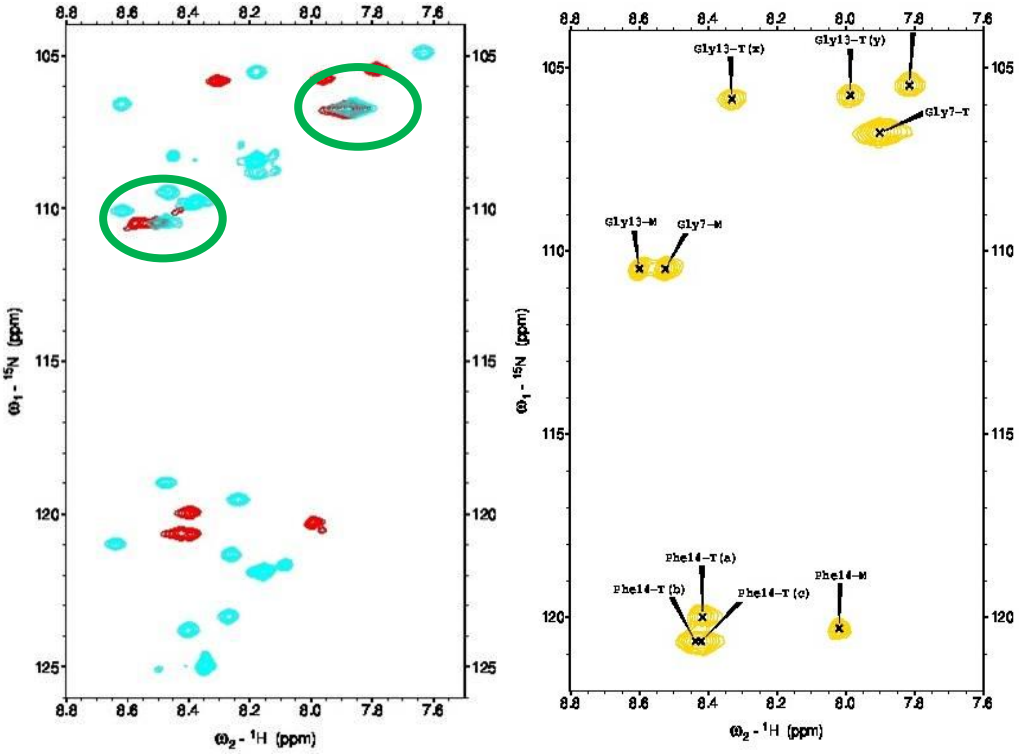

Figure 7 (a, left) An HSQC spectrum overlaying collagen model peptides T3-785 (in cyan) and GFOGER (in red). The green circles indicate glycine residue peaks that the two models have in common; (b, right) The HSQC of GFOGER with peaks either picked or labeled, where M denotes monomer and T for trimer

Relaxation Experiments

Data analyzed from the longitudinal R1 relaxation experiment shows little differences between the Phe-14 trimer peaks and the Gly-13/Phe-14 monomer peaks see in Figure 8. The data shows that the Phe-14 residues in the triple helix are the slowest to return to equilibrium while Gly-13 trimer residues, Gly-13 monomer, and Phe-14 monomer return to equilibrium the quickest. The Gly-13 trimer residues have a higher degree of variability compared to the Phe-14 trimer residues. The monomers returned to equilibrium quickly with the exception of Gly-7, whose R1 value is much closer to the Phe-14 trimers than to the other monomers.

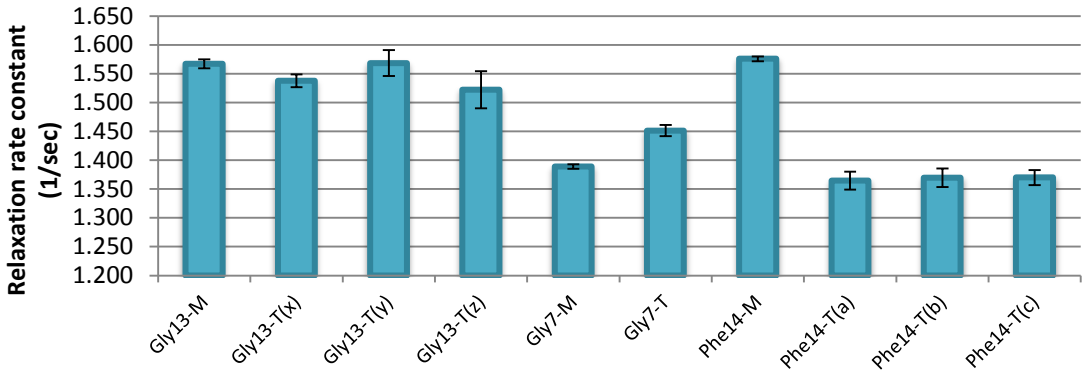

Figure 8 Comparison of R1 values with the Gly7, Gly13, and Phe14 labeled residues from the GFOGER collagen model peptide

Following the longitudinal relaxation experiment, transverse R2 experiments were conducted to observe flexibility of the residues, with collected data seen in Figure 9. The R2 data shows a significant difference between residues in trimer and monomer form: monomeric residues have a rather low R2 value compared to those of the trimeric form, suggesting a greater flexibility with monomeric residues. There also does not appear to be much difference in location, as Gly-7, in the rigid GPO-rich environment, has relaxation rates similar to Gly-13 and Phe-14. Transverse relaxation occurs due to entropic relations,

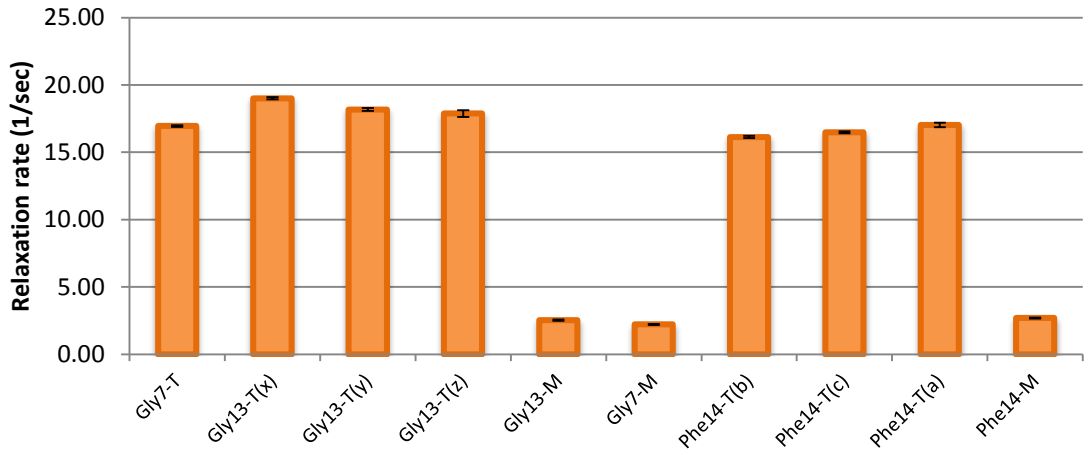

Figure 9 Comparison of R2 values with the Gly7, Gly13, and Phe14 labeled residues from the GFOGER collagen model peptide

through mutual energy swapping between spins, so having greater flexibility with residues with less interchain neighbors would make sense. The relaxation decay curves generated for the Gly-13 and Phe-14 trimer residues (Figures 10 and 11) show a similar decrease in intensity between the two and the Gly-7 timer residue and also a leveling-out of the trending prior to the final time point.

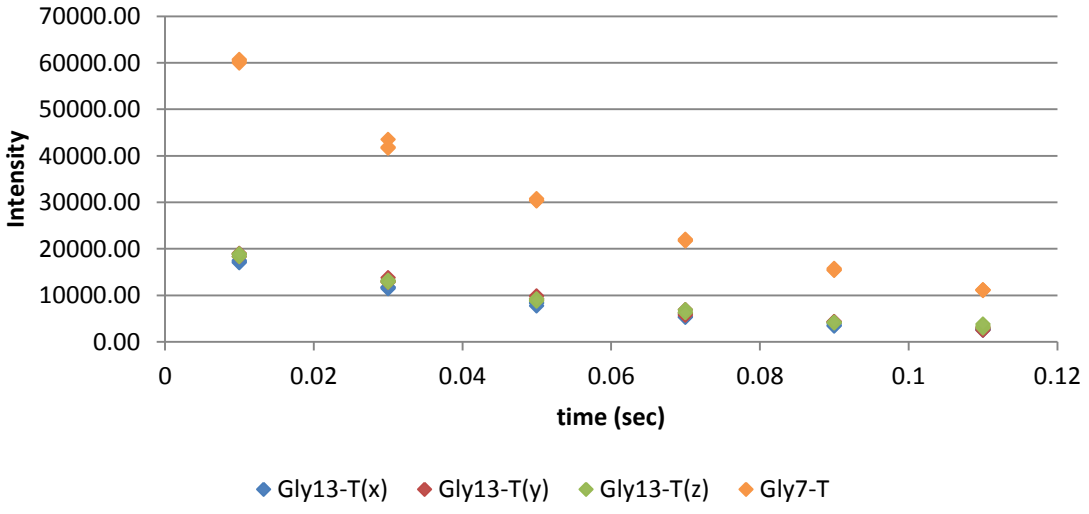

Figure 10 Relaxation R2 experimental decay curves between the three Gly-13 trimer peaks and the Gly-7 timer peak of the GFOGER collagen model peptide

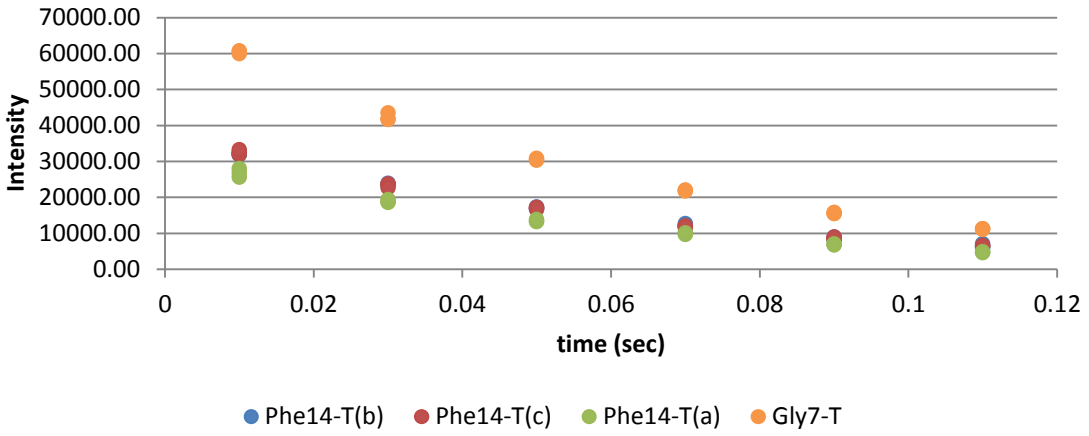

Figure 11 Relaxation R2 experimental decay curves between the three Phe-14 trimer peaks and the Gly-7 timer peak of the GFOGER collagen model peptide

It is because of the leveling-out in the R2 decay profiles that altering the R2 pulse sequence would be the next course of action to observe any changes to the R2 values. Altering the initial R2 relaxation experiment in both changing the pulse sequence and τ_d generated a smaller time scale to view all seven decay curves, from 0.11sec to 0.072sec, as seen in Figure 12, and shows similar data to the original R2 experiment seen in Figure 13. Yet, modR2 reduced the likelihood of the leveling-out of the graph past a particular time point and these new decay profiles have greater error in the intensity points.

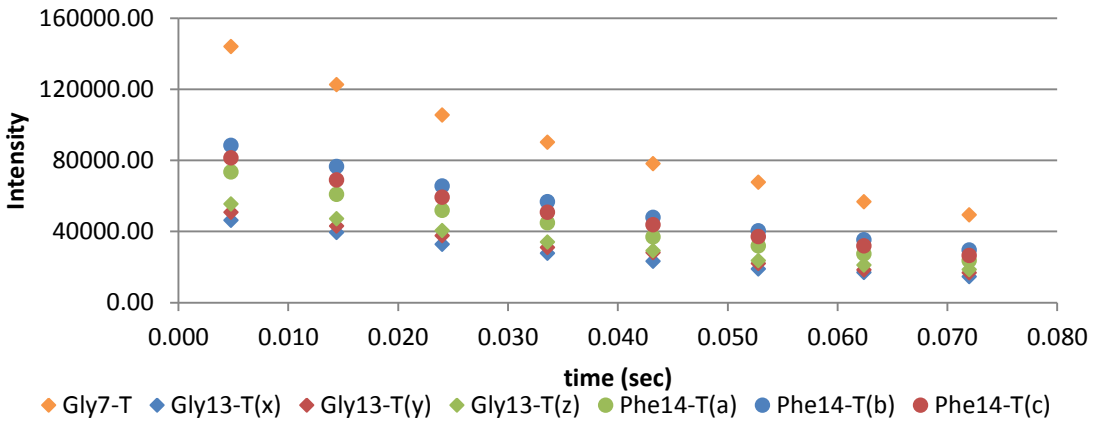

Figure 12 Decay curve generated by the modified R2 experiment for the Gly7 timer peak, the three Gly13 trimer peaks, and the three Phe14 trimer peaks of the GFOGER model peptide

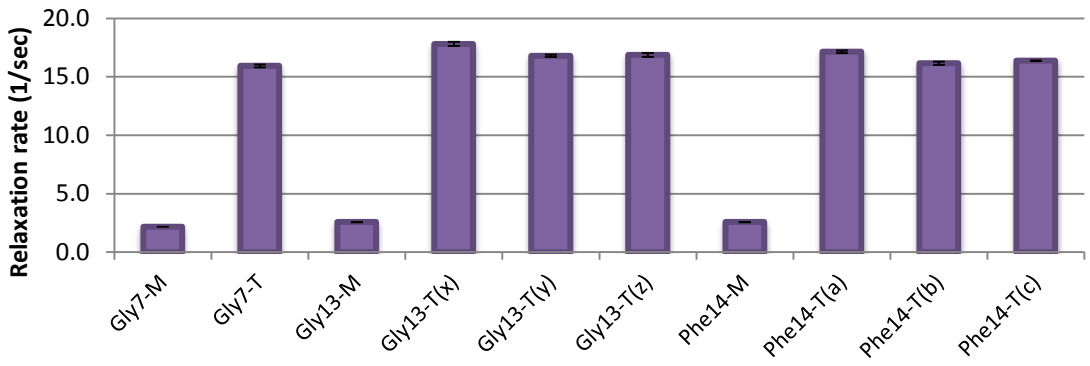

Figure 13 Comparison of R2 values from the modified pulse sequence with the Gly7, Gly13, and Phe14 labeled residues from the GFOGER collagen model peptide

Two additional relaxation experiments were conducted as a means to observe large internal motion within the peptide and for conformational exchange. An NOE relaxation experiment was conducted to determine whether significant internal motion occurs by generating two HSQC spectra: one of a ^{1}H-unsaturated GFOGER peptide and the other of an ^{1}H-saturated peptide. A ratio between the HSQC peak intensities were calculated, tabulated in Table 1, and seen in Figure 14. The Gly-13 and Phe-14 trimer residues show NOE values > 0.60, indicative of no internal motion. The NOE of the Gly-7 trimer residue is 0.547, which is rather close to the other trimer peaks.

Table 1- Peak Intensities of the GFOGER collagen model peptide during unsaturated and saturated proton HSQC experimentation and the NOE ratio produced

HSQC Peak Assignment	Unsaturated HSQC Peak Intensities	Saturated HSQC Peak Intensities	Ratio in Peak Intensities
Gly13-M	51990	1820	0.035
Gly13-T(x)	38388	26060	0.679
Gly13-T(y)	40881	25599	0.626
Gly13-T(z)	41359	25963	0.628
Gly7-M	67769	-22321	-0.329
Gly7-T	116759	63908	0.547
Phe14-M	51695	9796	0.189
Phe14-T(a)	55211	36186	0.655
Phe14-T(b)	68242	48922	0.717
Phe14-T(c)	68612	49906	0.727

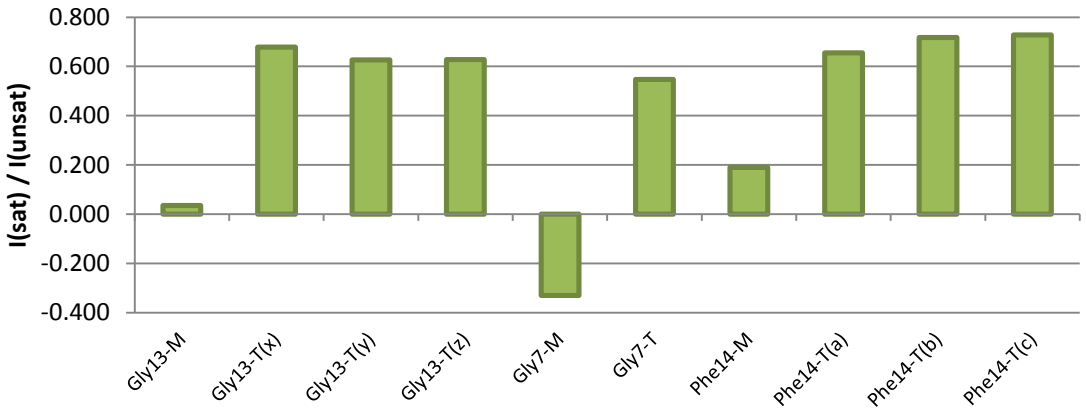

Figure 14 NOE values of each residue from saturated and unsaturated HSQC spectral data from the GFOGER collagen model peptide

The biggest changes occur between the monomers; however, the saturated portion of the NOE experiment will reduce the intensity of the monomer peaks, which is expected since the monomer will not have residues within a distance of 5Å. An interesting feature from the NOE experiment happens to be a negative peak for the Gly-7 monomer residue, generating a negative NOE peak. The GFOGER model peptide is considered a relatively small molecule by weight (~8.74 kDa), which means that there should be all positive NOE peaks.

To view conformational exchange, an R2 Hahn-echo experiment was conducted which, unlike the two previous R2 experiments whose pulse sequence causes suppression of conformational exchange by repeatedly perturbing the peptide with rf pulses, an $R2^{HE}$ experiment's pulse sequence is designed to allow for a longer detection time. This will provide information for the full conformational exchange of the peptide as shown in Figure 15.

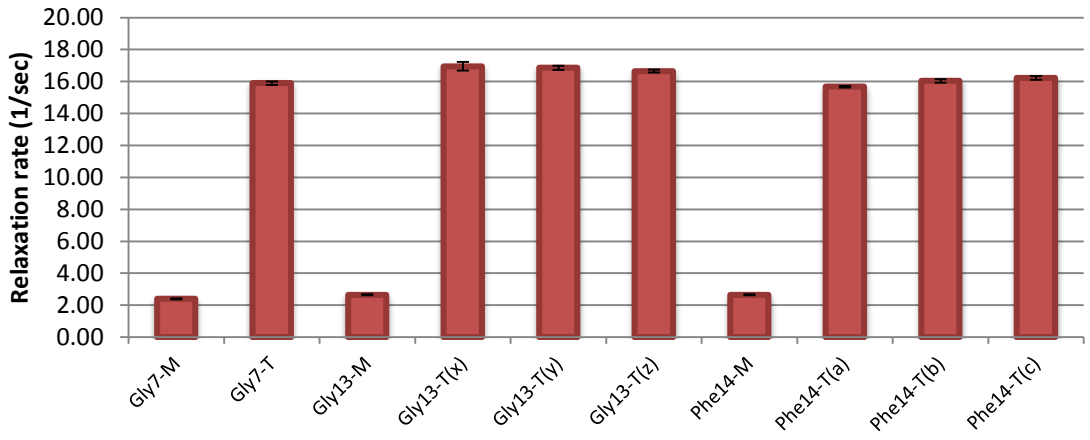

Figure 15 Comparison of R2 values from the Hahn-echo experiment with the Gly7, Gly13, and Phe14 labeled residues for the GFOGER collagen model peptide

To determine the conformational exchange value, or $R2^{EX}$, the R2 values from modR2 and $R2^{HE}$ were subtracted and shown in Figure 16. Based on the range of $R2^{EX}$ values, there does not appear to be minimal conformational exchange, if any at all, although one Phe14 and Gly13 appears to generate the large $R2^{EX}$ values.

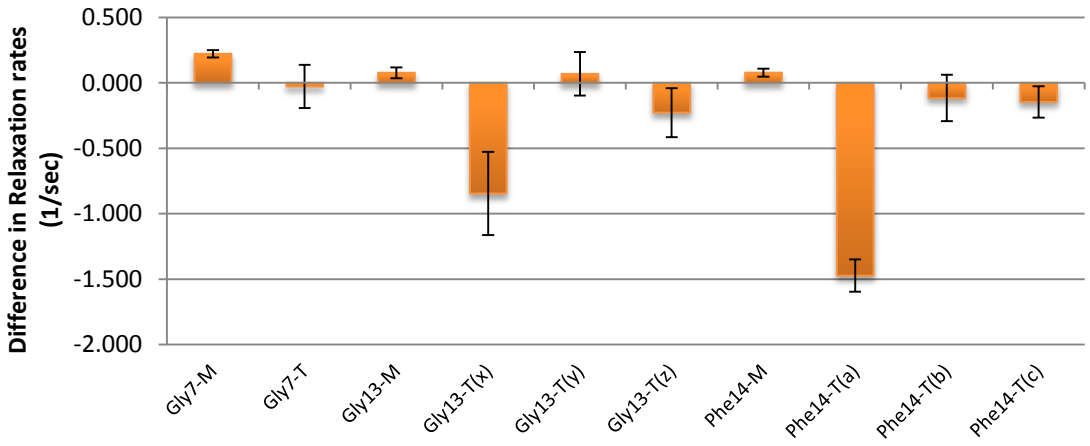

Figure 16 Estimated $R2^{EX}$ values from the modified R2 experiment and the R2 Hahn-echo experiment of the GFOGER collagen model peptide

Discussion

Analysis of the GXX'GEX" sequences has been able to state which particular residues can replace X, X', and X'' and affect integrin binding. Toolkit experimentation has provided a list of low and high affinity binders with X' as hydroxyproline residues and X'' as arginine. This has even lead to a crystal structure of an integrin subunit binding to the high affinity sequence GFOGER. However, more information needs to be collected as to the nature of this binding region. While such valuable binding information has been collected, the question now shifts to what drives the high affinity GFOGER sequence over another, such as GAOGER, of which alanine has a smaller side chain than phenylalanine.

The longitudinal and transverse relaxation experiments confirmed the flexibility of the GFOGER sequence, with Figure 13 showing differences in R2 values between the multiple R2 experiments conducted in this investigation. As shown, the three R2 experiments do not appear to have much in the way of rate value differences, expect between the Gly7 trimer residue, a Phe14 timer residue, and the three Gly13 trimer residues; yet some of the differences seen in Figure 17 are between the initial and modified R2 experiments. Yet this was as expected, as modR2 was designed to reduce the leveling out that occurred in the later time points of the initial transverse pulse sequence.

The relaxation experiments conducted now has led research with GFOGER characterization towards first differentiating the trimer peaks of the Gly-13 and Phe-14 neighbors. Emsley, *et al* (2004)[26], reported that the middle chain's GFOGER sequence closes the MIDAS coordination, so two questions are now raised: does the size of the

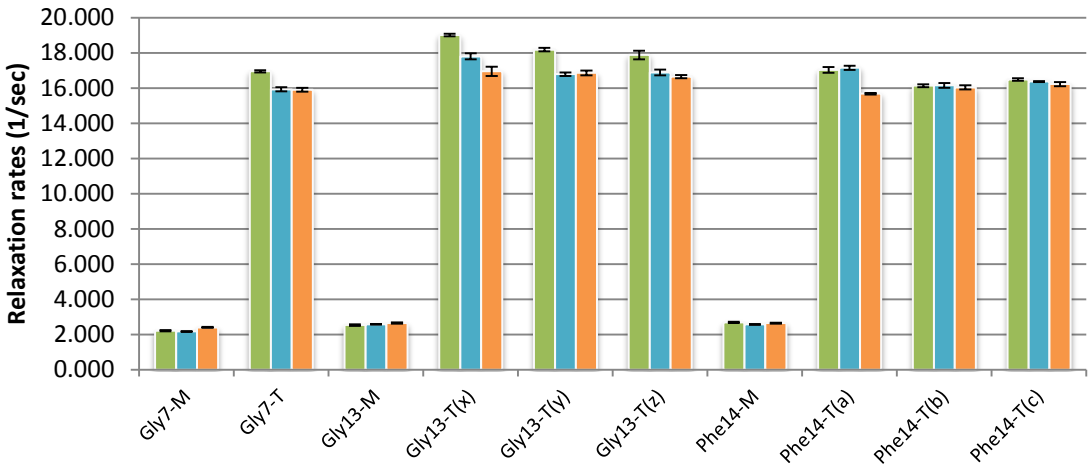

Figure 17 Comparison of all R2 values collected from the original R2 experiment (green), the modified R2 experiment (cyan), and the R2 Hahn-echo experiment (orange)

model play a role in binding with the I-domain and how does this information translate to full length type I collagen, where the GFOGER sequence is only presented twice in a region, as it is only found in the α1 chains and not in the α2? Will the six-residue sequence that sits adjacent to GFOGER in α2 increase or decrease the flexibility and, thus, the molecular recognition between integrin and collagen? Finally, how does a mutation in the α2 six-residue sequence change GFOGER's strength in affinity? These questions can propel integrin-collagen research forward as well as collagen mutation work in computational studies of both structure and functional sequence analysis. Understanding the overall flexibility of the GFOGER peptide will play its part towards improved molecular recognition thermodynamics.

Chapter 2: Binding Study between Type I Collagen and Beta-2 Microglobulin

One of the most critical functions of the immune system is a cell's ability to target antigens and generate an appropriate antibody response, which is done by presenting polypeptide strands to thymus lymphatic cells (T-cells). An integral transmembrane protein complex is used, known as the Major Histocompatibility Complex class I (MHC-I), and is comprised of a 45-kDa heavy-chain (HC) segment, a short antigen peptide that is presented to the extracellular environment, and an 11.8-kDa light-chain segment known as β2-microglobulin (B2m) that is used in chaperoning assembly of the overall complex (Figure 18(a))[33,34]. The 100-residue B2m subunit (Figure 18(b)) is fully exposed to the extracellular environment and has numerous contact points with both the transmembrane HC subunit and antigen-presenting domains, thereby consistently affecting the conformation of the HC subunit[33]. This protein complex has a relatively high turnover rate and as a result, it undergoes catabolism. As B2m is not attached to the cellular membrane, catabolism results in the protein's collection in plasmatic fluids and is circulated through the body until it reaches the kidneys and is excreted as waste[35,36]. The three-dimensional structure of monomeric, wild-type B2m has been solved several times with initial x-ray crystallography data describing the protein, characterized as a "β-sandwich" configuration, comprised of seven anti-parallel β-strands stabilized by a lone disulfide bridge contacting two distant strands (Figure 18(b))[35,37].

The protein's accumulation and subsequent amyloidogenesis, a protofilament seen in Figure 18(c), in a patient undergoing hemodialysis due to renal failure has been the focal point in B2m research[38]. As B2m is dissolved in renal tubules, complications in

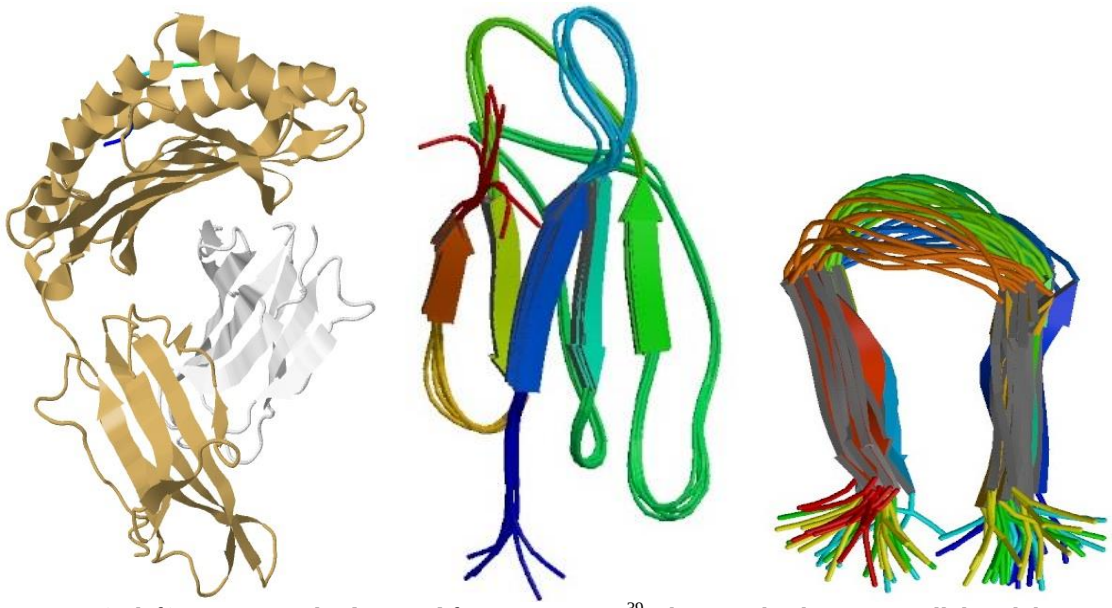

Figure 18 (a, left) MHC-I resolved imaged from PDB 1HSA[39], showing the three extracellular alpha subunits in brown and the B2m subunit in white; (b, middle) ribbon diagram of B2m in solution using NMR from PDB 1JNJ [40]; (c, right) ribbon diagram of B2m amyloid protofilament using solid-state NMR from PDB 2E8D[41]

the kidneys will result is a larger concentration of B2m in the plasmatic fluid, ranging up to 40-50 times greater than normal serum levels (up to 2 μg/mL)[34,36,38]. Dialysis-related amyloidosis (DRA) of B2m has been known to target the osteotendineous system, as such symptoms of carpal tunnel syndrome and arthralgias, among others, have been reported by hemodialysis patients[34,36]. The landmark work reported by Homma, *et al*[38], found from clinical findings that these DRA-related symptoms are occurring specifically in areas of abundant collagen fibers, such as heterotrimeric type I collagen and homotrimeric type II collagen, and noted past research stating the close proximity between generated amyloid fibrils and pre-existing collagen fibers. Research conducted by A. Relini, *et al*[34,36], utilizing atomic force microscopy has shown not only that collagen and B2m interact but that they are closely associated, adding to evidence of collagen's possible role as catalyzer for localized B2m amyloid fibril formation when experimenting with type I collagen[34].

However, there lacks a clear explanation as to what is the primary force driving the interaction between collagen and B2m. Homma, *et al*[38], presented evidence of collagen concentration-dependent binding in the presence of B2m, with binding interaction activity increasing with both increasing B2m concentrations (with constant collagen concentration) and increasing gelatin concentrations (with constant B2m concentration). Experimentation conducted by both Moe and Chen[42] (2001) and S. Giorgetti, *et al*[43] (2005), concluded that binding properties between the MHC-I subunit and type I collagen is higher than that of type II. Relini's atomic force microscopy work discussed the unique role of collagen as an immobilized positively-charged surface and electrostatic interactions between the proteins[34]. Research into the location or specific binding interactions, including a closer look as to whether specific residues in one or both proteins, will provide insights into the mechanism of amyloid formation and ways to target and eliminate DRA symptoms.

Methods & Materials

To determine the potential binding interactions with B2m and collagen, a series of ELISA binding assays were conducted. Materials and objectives were similar to the protocol established by Homma, *et al*[38], with a key difference of utilizing two antibodies over one. This technique, called an indirect ELISA (schematic in Figure 19), was chosen to minimize any cross-binding effects and false-positives that would result between the primary antibody and any non-B2m protein in the assay.

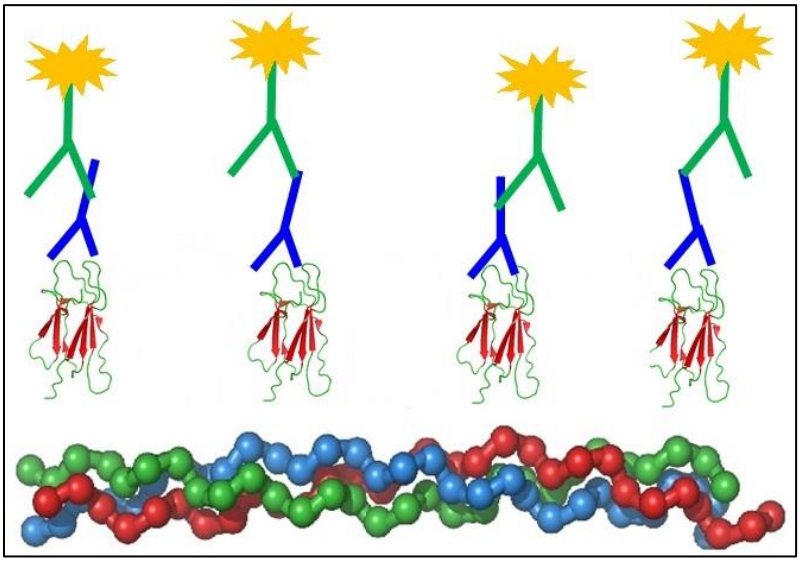

Figure 19 A schematic of a full indirect ELISA binding assay in which B2m (PDB entry 1JNJ)[40] is introduced to type I collagen (PDB entry 1CAG)[20] after it was coated to an untreated polystyrene plate. Primary antibodies (blue) will bind to the B2m proteins and then add secondary antibody

A 5ug/mL sample type I collagen, 100uL/well, was coated to an untreated 96-well polystyrene plate after dialysis from a stock solution of 0.1M acetic acid, pH 4.0, to the bicarbonate coating buffer at 100uL/well. After a 12 hour incubation period at 4°C, the plate was washed with PBS-T four times, blocked with 300uL of 2% bovine serum albumin (BSA) in PBS-T solution, and incubated at 4°C for an hour. For the binding phase of the experiment, a series of 100uL B2m samples, the protein as a kind gift provided by Dr. Sheena Radford (University of Cambridge), in PBS-T with

concentrations used in Homma, *et al*[38], [1.3ug/mL, 10ug/mL, 40ug/mL, and 80ug/mL] were added following a 300uL/well PBS-T wash four times and binding incubation was set for five hours, again at 4°C. Following another 300uL/well PBS-T wash four times, the plate was set at room temperature for the antibody binding stage, 100uL/well of primary antibody [monoclonal mouse anti-human 1° antibody (Fischer Scientific, Hanover Park, IL)] with a dilution factor of 1:50,000 added and incubated for an hour, then repeated the wash procedure and 100uL/well of secondary antibody [polyclonal goat anti-mouse 2° antibody with a TMB substrate attached (GenScript, Piscataway, NJ)] with a dilution factor also of 1:50,000.

Following the one hour incubation of the secondary antibody binding and another round of the wash procedure, a 1:1 mixture of peroxidase and TMB substrate solutions (Fischer Scientific, Hanover Park, IL) were mixed, 100uL aliquots were added to the plate, and the plates were allowed to react for up to an hour or until the color reached a certain high intensity color change. To stop the reaction, 100uL/well of 2.5M H_2SO_4 was added resulting in a blue to yellow color change of the plate. The plate was placed in the OmegaSTAR plate reader for absorbance detection at λ_{450nm}.

Controls utilized during experimentation included equivalent concentrations of BSA rather than collagen, omitting the initial proteins used for coating at the beginning of the assay, and removing B2m protein from the protocol during the binding stage. Changes to the experiment were done to maximize the signal following secondary antibody binding by fine-tuning the secondary antibody's dilution factor and increasing the starting concentration of collagen two-fold and three-fold.

Results

Several experimental conditions needed to be determined prior to the indirect ELISA binding assays. Determining the ideal dilution factors for the mouse anti-human 1° antibody and goat anti-mouse 2° antibody was detrimental for signal strength and also to prevent any precipitation that would occur from loosely-bound 2° antibody that would react with the TMB substrate solution, separate from the 1° antibody, and accumulate as blue particulates in the wells. Based on a combination of absorbance readings and precipitation visual cues, a 1:50,000 dilution factor was selected for both antibodies, as indicated as the red entry in Table 2; yet continual changes of antibody dilution factors occurred to fine-tune the signal once binding assays. These adjustments resulted in an increase to a 1:40,000 dilution factor for either antibody.

Table 2 Absorbance data at 450nm of a primary/secondary antibody optimization assay, with units of OD associated with table value

		Secondary Antibody Dilution Factors					
Primary Antibody Dilution Factors		**1:45K**	**1:46K**	**1:47K**	**1:48K**	**1:49K**	**1:50K**
	1:42K	0.713	0.792	0.785	0.794	0.857	0.814
	1:44K	0.838	0.772	0.773	0.848	0.779	0.839
	1:46K	0.851	0.702	0.721	0.801	0.755	0.796
	1:48K	0.835	0.739	0.750	0.756	0.762	0.742
	1:50K	0.777	0.756	0.773	0.764	0.723	**0.771**
	1:52K	0.701	0.683	0.722	0.775	0.728	0.739
	1:54K	0.740	0.675	0.763	0.865	0.749	0.762
	1:56K	0.735	0.798	0.768	0.779	0.704	**0.063**

Following antibody optimization came determining which coating, blocking, binding, and wash buffers would need to be used to optimize absorbance signal. Although the final buffers that were decided upon were based on previous ELISA assays with B2m, optimization work did provide information about changes in buffer types, such as Tris-based or phosphate-based buffers, and the ideal amount of BSA and Tween-20 in

the blocking buffers can be useful for future assays and can be seen in Figures 20 and 21. The final buffers selected were the bicarbonate coating buffer of pH 8.3, a binding and wash buffer of PBS, pH 7.4, with Tween-20 (PBS-T), and a blocking buffer of PBS-T, pH 7.4, with 2% BSA.

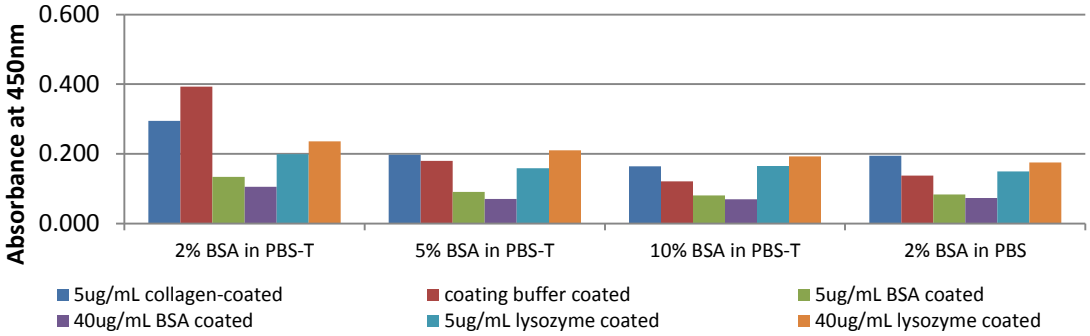

Figure 20 Change in both coating proteins and blocking buffers in combination at various concentrations during collagen-B2m indirect ELISA binding assay optimization

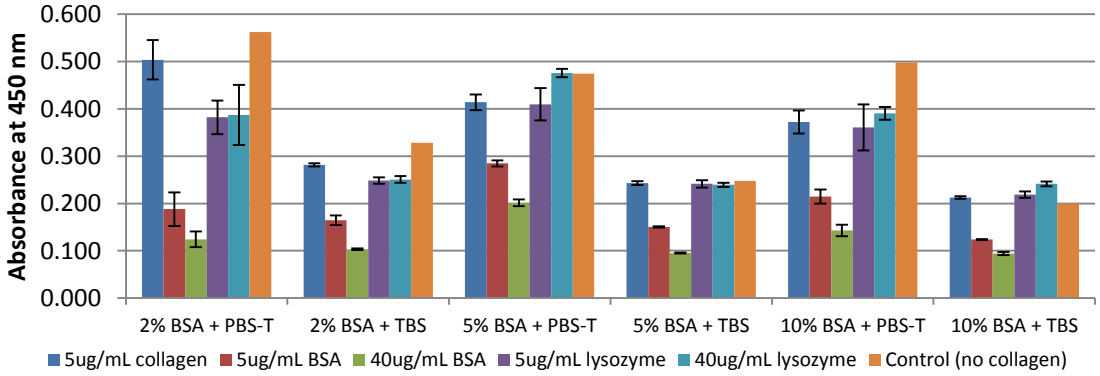

Figure 21 Buffer Optimization data of the PBS-T coating buffer using different coating proteins (in legend and blocking buffers (x-axis)

Once optimized, the experimental objective of the indirect ELISA binding assays was to confirm the findings reported by Homma, *et al*[38], in which ELISA assays showed a relationship between B2m and type I and type II collagen, both in assays of various collagens binding to B2m and in an experiment with a constant type I collagen concentration against various B2m concentrations (1.3ug/mL to 80ug/mL). Initially, the 96-well plates was coated with only a type I collagen of 5ug/mL and allowed to react

against control protein BSA, seen in Figure 22. The results show an increase in absorbance as the concentration of B2m increases, indicative of increased binding activity between collagen and B2m, while the BSA-B2m control interaction did not appear to increase in absorbance.

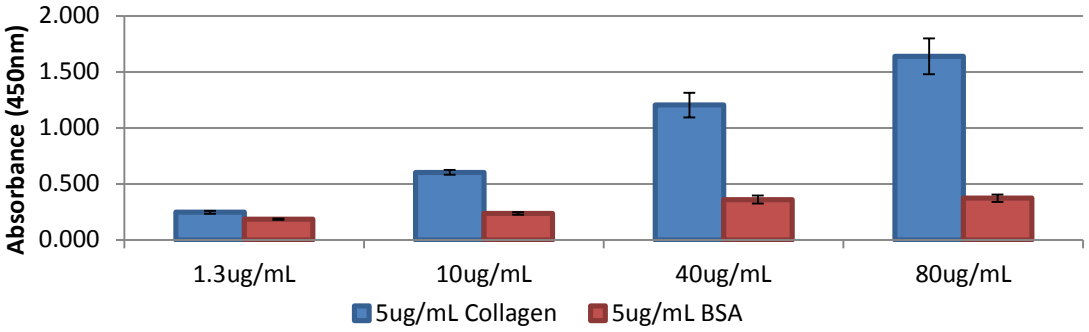

Figure 22 Absorbance data at 450nm for an indirect ELISA binding assay between coating proteins type I collagen and BSA to different concentrations of B2m

Considering the difference in absorbance readings in the above-normal B2m levels as well as the work Homma, *et al*[37] (1989), reported with various gelatin concentrations, two further assays were conducted based on a 2-fold and 3-fold increase to the collagen concentration, in Figures 23(a) and 23(b). In both experiments, there does appear to be the same increase in absorbance at increasing B2m concentrations with collagen, but unlike the 5ug/mL experiment, an increase in B2m-BSA absorbance activity also results.

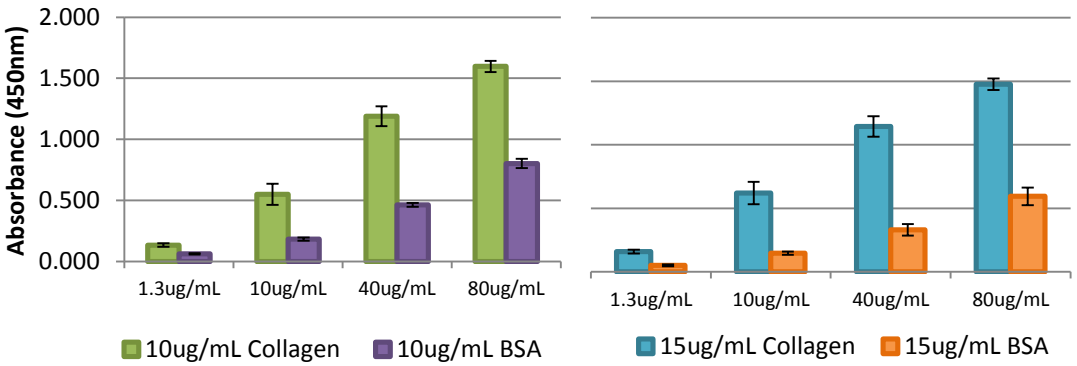

Figure 23 Absorbance data at 450nm for 10ug/mL (a, left) and 15ug/mL (b, right) type I collagen and BSA in an indirect ELISA binding assay with varying concentrations of B2m

Discussion

Research experimentation between collagen and the MHC-I subunit B2m have provided significant insights into protein-protein interactions and fiber-fiber proximity[34,36,38]. However, there still are questions as to the nature of this interaction, including the subdomain on both B2m and collagen where this interaction occurs. Although the further goals within the scope of this work include utilizing collagen model peptides as a coating protein, the primary objective needed was to recreate the experimental findings previously reported. Choosing to conduct an indirect ELISA rather than a direct ELISA[38] or the ^{125}I-labeled B2m binding studies[42] was to reduce any chances of multiple binding events that might occur after the binding protein was allowed to react during incubation.

The data collected from the indirect ELISA experiments for all three collagen concentrations (seen combined in Figure 24) shows an increase in absorbance as the concentration of B2m increases, thereby confirming the findings previously conducted. However, an interesting observation to the data shows that the absorbance data did not increase because of the two-fold and three-fold increases in collagen; that is to say, by

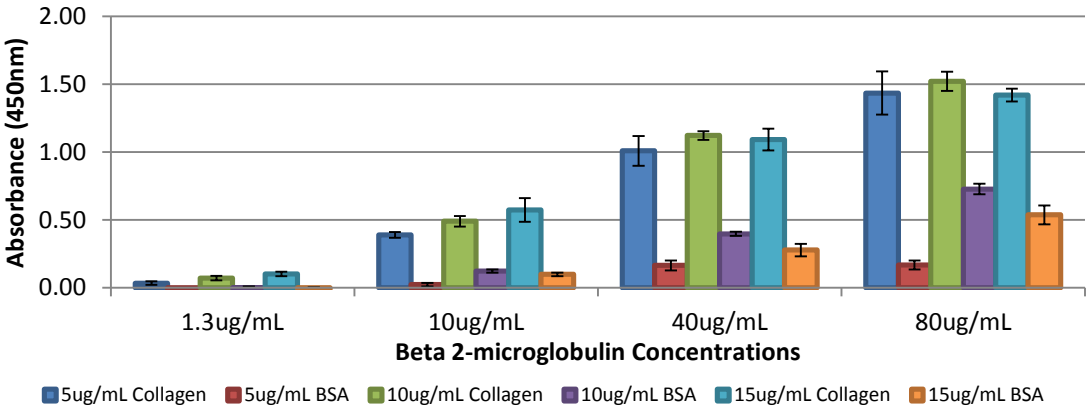

Figure 24 Absorbance data at 450nm for three indirect ELISA binding assays of type I collagen and of BSA with varying concentrations of B2m

doubling or tripling the amount of collagen, the absorbance, and therefore the binding between collagen and B2m, did not increase proportionally. This would indicate that all B2m present would interact with collagen, and this is an occurrence for all four B2m concentrations. However, when considering next steps to the experiment, there was concern as to the nature of collagen. By increasing the concentration, there runs a risk of aggregation into a fibril. Even by conducting the assays in a 0°C cold room and letting all solutions reach that temperature before continuing, there does lie that fibril risk. Although work reported by Relini, *et al*, both in 2006[34] and 2008[36] showed a relationship between the higher-ordered structures of B2m and collagen, such interactions would not help with gathering preliminary information into subdomain binding information between the proteins.

The results shown through indirect ELISA binding assays cements the interaction between collagen and B2m and can allow further work into the question of global structure versus subdomains-specificity between the proteins. By working next with either a different full-length collagen type or by well-studied collagen model peptides, insights into what initiates B2m-collagen binding subsequent B2m aggregation will further work done to help hemodialysis patients minimize side-effects to life-saving treatments.

Works Cited

1. Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **78,** 929–58 (2009).
2. Xu, Y. *Thermal stability of collagen triple helix. Methods Enzymol.* **466,** 211–32 (Elsevier Inc., 2009).
3. Pauling, L. & Corey, R.B. The structure of fibrous proteins of the collagen-gelatin group. *Proc. Natl. Acad. Sci.* **37**:272–81 (1951).
4. Ramachandran, G.N. & Kartha, G. Structure of collagen. *Nature.* **176**:593–95 (1955).
5. Rich, A., & Crick, F.H.C. The structure of collagen. *Nature* **176**:915–16 (1955).
6. Rich, A., & Crick, F.H.C. The molecular structure of collagen. *J Mol Biol.* **3**:483–506 (1961).
7. Persikov, A.V., Ramshaw, J.A., Kirkpatrick, A., & Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **39,** 14960–7 (2000).
8. Beck, K., & Brodsky, B. Supercoiled protein motifs: the collagen triple-helix and the alpha-helical coiled coil. *J. Struct. Biol.* **122,** 17–29 (1998).
9. Persikov, A.V., Ramshaw, J.A., & Brodsky, B. Collagen model peptides: Sequence dependence of triple-helix stability. *Biopolymers* **55,** 436–50 (2000).
10. Mohs, A., Popiel, M., Li, Y., Baum, J. & Brodsky, B. Conformational features of a natural break in the type IV collagen Gly-X-Y repeat. *J. Biol. Chem.* **281,** 17197–202 (2006).
11. Persikov, A.V., Ramshaw, J.A., & Brodsky, B. Prediction of collagen stability from amino acid sequence. *J. Biol. Chem.* **280,** 19343–9 (2005).
12. Makareeva, E., & Leikin, S. Procollagen triple helix assembly: an unconventional chaperone-assisted folding paradigm. *PLoS One* **2,** e1029 (2007).
13. Buevich, A., & Baum, J. Nuclear magnetic resonance characterization of peptide models of collagen-folding diseases. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356,** 159–68 (2001).
14. Bella, J., Brodsky, B. and Berman, H.M.: Hydration structure of a collagen peptide. *Structure.* **3**: 893-906 (1995).
15. Okuyama, K., Hongo, C., Fukushuima, R., Wu, G., Narita, H., Noguchi, K., Tanaka, Y., & Nishino, N. Crystal structures of collagen model peptides with Pro-Hyp-Gly repeating sequences at 1.26Å resolution: implications for proline ring puckering. *Biopolymers*. **80**: 714-5 (2005).
16. Buevich, A.V., Dai, Q. H., Liu, X., Brodsky, B., & Baum, J. Site-specific NMR monitoring of cis-trans isomerization in the folding of the proline-rich collagen triple helix. *Biochemistry* **39,** 4299–308 (2000).
17. Xiao, J., Cheng, H., Silva, T., Baum, J. & Brodsky, B. Osteogenesis imperfecta missense mutations in collagen: structural consequences of a glycine to alanine replacement at a highly charged site. *Biochemistry* **50,** 10771–80 (2011).

18. Xiao, J., Addabbo, R. M., Lauer, J. L., Fields, G. B. & Baum, J. Local conformation and dynamics of isoleucine in the collagenase cleavage site provide a recognition signal for matrix metalloproteinases. *J. Biol. Chem.* **285,** 34181–90 (2010).
19. Fan, P., Li, M. H., Brodsky, B. & Baum, J. Backbone dynamics of (Pro-Hyp-Gly)10 and a designed collagen-like triple-helical peptide by 15N NMR relaxation and hydrogen-exchange measurements. *Biochemistry* **32,** 13299–309 (1993).
20. Bella, J., Eaton, M., Brodsky, B. & Berman, H. Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science.* **266,** 77–81 (1994).
21. Li, Y., Brodsky, B. & Baum, J. NMR shows hydrophobic interactions replace glycine packing in the triple helix at a natural break in the (Gly-X-Y)n repeat. *J. Biol. Chem.* **282,** 22699–706 (2007).
22. Brodsky, B., Li, M. H., Long, C. G., Apigo, J. & Baum, J. NMR and CD studies of triple-helical peptides. *Biopolymers* **32,** 447–451 (1992).
23. Brodsky, B., Thiagarajan, G., Madhan, B. & Kar, K. Triple-helical peptides: an approach to collagen conformation, stability, and self-association. *Biopolymers* **89,** 345–53 (2008).
24. Li, Y., Brodsky, B. & Baum, J. NMR conformational and dynamic consequences of a Gly to Ser substitution in an osteogenesis imperfecta collagen model peptide. *J. Biol. Chem.* **284,** 20660–7 (2009).
25. Brodsky, B. & Ramshaw, J. a. M. The collagen triple-helix structure. *Matrix Biol.* **15,** 545–554 (1997).
26. Emsley, J., Knight, C. G., Farndale, R. W. & Barnes, M. J. Structure of the Integrin α2β1-binding Collagen Peptide. *J. Mol. Biol.* **335,** 1019–1028 (2004).
27. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J. & Liddington, R. C. Structural basis of collagen recognition by integrin alpha2beta1. *Cell* **101,** 47–56 (2000).
28. Siljander, P. R.-M. *et al.* Integrin activation state determines selectivity for novel recognition sites in fibrillar collagens. *J. Biol. Chem.* **279,** 47763–72 (2004).
29. Knight, C. G. *et al.* The Collagen-binding A-domains of Integrins a1b1 and a2b1 Recognize the Same Specific Amino Acid Sequence , GFOGER , in Native (Triple-helical) Collagens *. *J. Biol. Chem.* **275,** 35–40 (2000).
30. Herr, A. B. & Farndale, R. W. Structural insights into the interactions between platelet receptors and fibrillar collagen. *J. Biol. Chem.* **284,** 19781–5 (2009).
31. Raynal, N. *et al.* Use of synthetic peptides to locate novel integrin alpha2beta1-binding motifs in human collagen III. *J. Biol. Chem.* **281,** 3821–31 (2006).
32. Farndale, R. W. *et al.* Cell-collagen interactions: the use of peptide Toolkits to investigate collagen-receptor interactions. *Biochem. Soc. Trans.* **36,** 241–50 (2008).
33. York, I. & Rock, K. Antigen processing and presentation by the class I major histocompatibility complex. *Annu. Rev. Immunol.* 369–396 (1996).

34. Relini, A. *et al.* Collagen plays an active role in the aggregation of beta2-microglobulin under physiopathological conditions of dialysis-related amyloidosis. *J. Biol. Chem.* **281,** 16521–9 (2006).
35. Eichner, T. & Radford, S. E. Understanding the complex mechanisms of β2-microglobulin amyloid assembly. *FEBS J.* **278,** 3868–83 (2011).
36. Relini, A. *et al.* Heparin strongly enhances the formation of beta2-microglobulin amyloid fibrils in the presence of type I collagen. *J. Biol. Chem.* **283,** 4912–20 (2008).
37. Becker, J. W. & Reeke, G. N. Three-dimensional structure of beta 2-microglobulin. *Proc. Natl. Acad. Sci.* **82,** 4225–9 (1985).
38. Homma, N., Gejyo, F., Isemura, M. & Arakawa, M. Collagen-binding affinity of beta-2-microglobulin, a preprotein of hemodialysis-associated amyloidosis. *Nephron* **53,** 37–40 (1989).
39. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. The three-dimensional structure of HLA-B27 at 2.1 A resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* **70,** 1035–48 (1992).
40. Verdone, G. *et al.* The solution structure of human beta2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci.* **11,** 487–499 (2002).
41. Iwata, K. *et al.* 3D structure of amyloid protofilaments of beta2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci.* **103,** 18119–24 (2006).
42. Moe, S. M. & Chen, N. X. The role of the synovium and cartilage in the pathogenesis of beta(2)-microglobulin amyloidosis. *Semin. Dial.* **14,** 127–130 (2001).
43. Giorgetti, S. *et al.* Beta2-Microglobulin isoforms display an heterogeneous affinity for type I collagen. *Protein Sci.* **14,** 696–702 (2005).