

Observing Residual Structure in Disordered Peptides with Multidimensional Infrared Spectroscopy

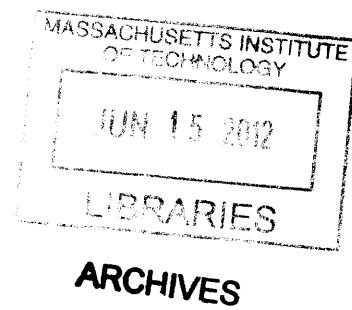
By

Joshua Lessing
ScB. Chemistry
Brown University, 2004

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at the
Massachusetts Institute of Technology
May 2012

© 2012 Massachusetts Institute of Technology. All rights reserved.



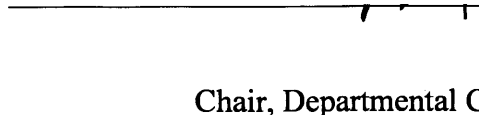
Signature of Author:


Joshua Lessing
May 2, 2012

Certified by:

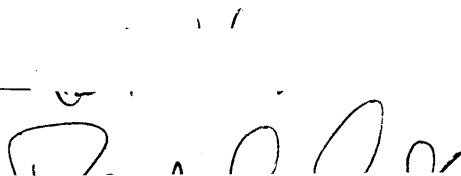

Andrei Tokmakoff
Professor of Chemistry
Thesis Supervisor

Accepted by:


Robert W. Field
Professor of Chemistry
Chair, Departmental Committee on Graduate Studies

This doctoral thesis has been examined by a committee of the Department of Chemistry that included:

Professor Keith A. Nelson
Chair



Professor Robert G. Griffin



Professor Andrei Tokmakoff
Thesis Supervisor



Observing Residual Structure in Disordered Peptides with Multidimensional Infrared Spectroscopy

By

Joshua Lessing

Submitted to the Department of Chemistry on
May 2, 2012
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

An estimated 35% of the human proteome is intrinsically disordered. Disordered proteins play a key role in physiologic and pathologic regulation, recognition, and signaling making protein disorder the subject of increasing investigation. Since disordered samples do not generate x-ray quality crystals and since they have conformations that interconvert faster than the time resolution of NMR or ESR, little is known about their structure or function. By combining isotope-edited two-dimensional infrared spectroscopy (2D IR) with spectral modeling based on molecular dynamics simulations, this work will show that one can measure the residual structure and conformational heterogeneity of a putatively disordered sequence.

This methodology was first used to study the structure and dynamics of the tryptophan zipper 2 (TZ2) peptide. The TZ2 peptide is a 12 residue engineered sequence that employs the “tryptophan zipper” motif to stabilize a β -hairpin fold. For this work five isotopologues of TZ2 were synthesized, the β -turn label K8, the midstrand labels T10 and T3T10, the N-terminal label S1, and the unlabeled peptide UL. Temperature dependent FTIR and 2D IR studies in conjunction with modeling revealed that the TZ2 peptide at low temperatures exists as a folded peptide with a type I' turn and a small previously unobserved subpopulation of a bulged loop structure. Upon heating a fraying of the peptides termini is observed, in addition the population of the type I' turn increases relative to the population of the bulged turn. The TZ2 peptide provided a model hairpin system to test the utility of different forms of analysis and isotope labeling patterns for the subsequent study of disordered hairpin peptides.

This methodology was next employed to gain insight into the structure of the elastin-like peptide GVGVPGVG, a prototypical disordered system. For this work nine isotopologues were studied in addition to size dependent variants based on the VPGVG sequence and point mutants variants of the form GVGXPGVG. The conformational ensemble was found to contain a high

population of irregular β -turns possessing two peptide hydrogen bonds to the proline C=O (see figure below). Further, it was found that this population grows as a function of 1) side-chain volume for the peptide series GVGXPGVG, where $X = \text{Gly, Ala, and Val}$, and 2) polymer size for the peptide series GVG(VPGVG) $_n$, where $n = 1-6$ and 251. These findings provide new insight into the molecular origin of the mechanical properties of biopolymers containing XPG turns including collagen ($X = \text{Pro}$), elastin ($X = \text{Val}$), mussel byssus ($X = \text{Gly}$), dragline spider silk ($X = \text{Gly}$), and wheat glutenin ($X = \text{Gln}$).

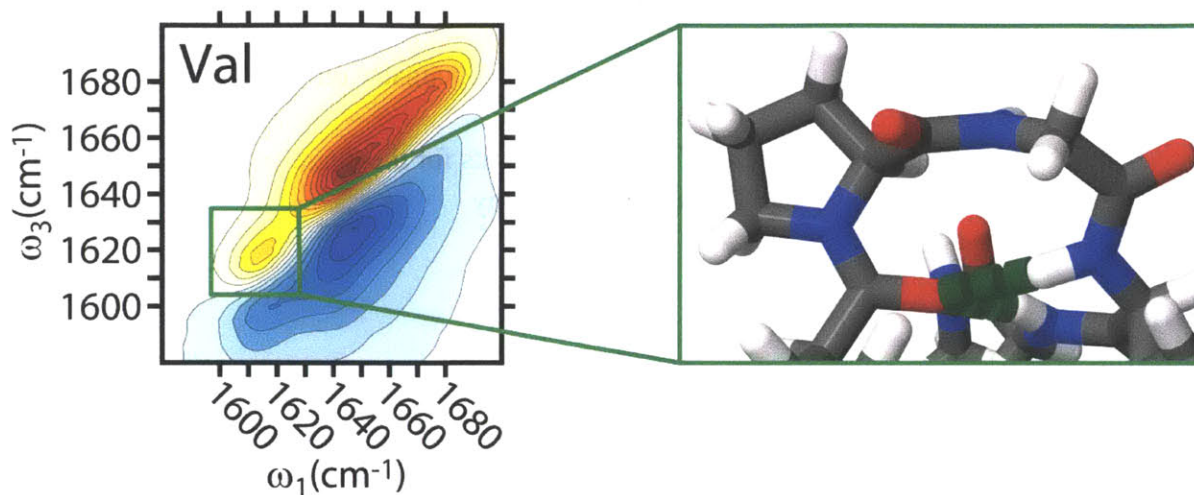


Figure A1. On the left is a 2D IR spectrum of GVGVPGVG at $T = 10^\circ\text{C}$ in pH 1.0 DCl in D_2O . The green box in the spectrum highlights the red shifted amide I' proline peak feature associated with the presence of an irregular β -turn which is depicted on the right. The green dotted lines in this structure indicate the location of the two hydrogen bonds to the proline C=O.

Thesis Supervisor: Andrei Tokmakoff

Title: Professor of Chemistry

The presented work is dedicated to my parents, Judith and Jeffrey Lessing, whose tireless commitment to my education made this thesis possible.

“I love it when a plan comes together.”

Colonel John “Hannibal” Smith

Table of Contents

List of Figures and Tables	9
List of Tables	23
1. Preface	24
1.1 Introduction.....	24
1.2 Acknowledgments.....	27
1.3 References.....	27
2. Synthesis of amide I isotope labeled peptides	32
2.1 ¹⁸ O Isotope labeling.....	32
2.2 Solid Phase Peptide Synthesis.....	36
2.2.1 Choosing a resin.....	38
2.2.2 Deprotection.....	41
2.2.3 Activation and Coupling.....	42
2.2.4 Coupling Reagents.....	44
2.3 Side-chain protection.....	46
2.3.1 Asparagine.....	47
2.3.2 Glutamic Acid.....	48
2.3.3 Tryptophan.....	49
2.3.4 Lysine.....	52
2.3.5 Serine and Threonine.....	54
2.4 Cleavage, Side-Chain Deprotection and Purification.....	56
2.5 Preventing amino acid deletion.....	58
2.6 Acknowledgments.....	61
2.7 References.....	61
3. Melting of a β-Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations	65
3.1 Introduction & Background.....	65
3.2 Spectroscopic Methods.....	69
3.3 Simulation and Modeling Methods.....	69
3.4 Amide I' FTIR Spectroscopy Results.....	71
3.5 Amide I' 2D IR Spectroscopy Results.....	73
3.6 Temperature-Dependent Spectra.....	76

3.7	Simulation Results.....	83
3.8	Discussion and Conclusions.....	95
3.9	Acknowledgements.....	98
3.10	References.....	99
4.	Identifying Residual Structure in Intrinsically Disordered Systems: A 2D IR Spectroscopic Study of the GVGXPGVG Peptide.....	105
4.1	Introduction.....	105
4.2	Experimental spectra.....	106
4.3	Modeling and Simulations.....	109
4.4	Conclusions.....	120
4.5	Acknowledgments.....	121
4.6	References.....	122
5.	Isotope-Edited 2D IR Study of the GVGVPGVG Peptide.....	124
5.1	Introduction.....	124
5.2	Background.....	127
5.3	Current models for elastin.....	129
5.4	Experimental overview.....	141
5.5	Experimental methods.....	143
5.6	FTIR and 2D IR spectra.....	144
5.7	Single site analysis.....	149
5.7.1	Single site analysis: V1.....	149
5.7.2	Single site analysis: G3 and V4.....	153
5.7.3	Single site analysis: G5 [†]	156
5.7.4	Single site analysis: G3 [†]	157
5.8	Multiple site analysis.....	158
5.9	Temperature and salt dependence.....	162
5.10	Size dependent FTIR.....	169
5.11	Conclusion.....	175
5.12	Acknowledgments.....	176
5.13	References.....	176
6.	Appendix.....	184

List of Figures

A.1	On the left is a 2D IR spectrum of GVGVPGVG at $T = 10^{\circ}\text{C}$ in pH 1.0 DCl in D_2O . The green box in the spectrum highlights the red shifted amide I' proline peak feature associated with the presence of an irregular β -turn which is depicted on the right. The green dotted lines in this structure indicate the location of the two hydrogen bonds to the proline C=O.....	4
2.1	N-terminal Fmoc protected amino acid.....	32
2.2	Steps involved in performing Fmoc SPPS.....	35
2.3	Common linkers used for SPPS. Blue circle – solid polymer support; Green square – peptide attachment point.....	39
2.4	Mechanism of Fmoc deprotection with piperidine. Purple – carbamate; Blue – Dibenzofulvene.....	41
2.5	Reaction mechanism for the activation of an amino acid by the reaction of HBTU with an unprotected carboxylic acid to form a HOBt ester. Blue – Urea; Green – Triazolol; Purple – O-Acylisourea; Yellow – HOBt ester.....	42
2.6	Mechanism for the acetylation of an amine by an activated Fmoc amino acid.....	43
2.7	Potential pathways for the racemization of an amino acid under basic conditions including the deprotonation of the alpha carbon (top) and the formation of a cyclic azlactone intermediate (bottom). The rate of both mechanisms depends on the acidity of	

	the amino acids α -carbon which varies as a function of the electron-withdrawing effects of the side-chain and N-terminal protecting group.....	45
2.8	High yeild coupling agents for the activation of carboxylic acids. Historically HATU and HBTU were believed to exist as an uronium salt (O form), as depicted above for HATU and HBTU, though the current consensus is that there true structure is a guanidinium salt (N form), depicted for HCTU. ²⁶ This finding was significant because the uronium isomer is a more effective coupling agent then the guanidinium isomer. These isomers are discussed here for future reference since these coupling agents are described as both uronium or guanidinium salts in the references provided for this chapter.....	46
2.9	Commonly used side-chain protecting groups for Fmoc SPPS. Benzyl (Bzl) protects carboxylic acids, alcohols, thiols, and imidazoles. Tert-Butyl (tBu) protects carboxylic acids, alcohols, and thiols. Benzyloxy-carbonyl (Boc) protects amines. Triphenylmethyl (trt) protects carboxamides, imidazoles, and thiols. 2,2,5,7,8-pentamethyl-chromane-6-sulfonyl (Pmc), 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf), 4-methoxy-2,3,6-trimethylbenzene-sulfonyl (Mtr), and 4-toluenesulfonyl (Tos) protect guanidinium.....	47
2.10	Mechanism of asparagine dehydration to create a cyano group. Purple – carboxamide; Green – cyano.....	47
2.11	Cyclization and ring opening reaction of a Glu side-chain in the presence of piperidine to generate α -Glu peptide with a γ -piperidide side-chain (Left product) and a γ -Glu peptide	

	with an α -piperidide side-chain (Right product). Green – piperidinedione; Purple – piperidide.....	48
2.12	Alkylation of a Trp side-chain followed by migration of the alkyl group. Green – indole.....	49
2.13	Acid induced dimerization of Trp side-chains.....	50
2.14	Two stage orthogonal deprotection of Fmoc-Trp(Boc-Nmbu)-OH. Blue – Nmbu.....	52
2.15	Acetylation of an unprotected Lys to generate a branched product.....	53
2.16	Polymerization of Lys as a result of coupling in the absence of a side-chain protecting group. To avoid these side-reactions in the synthesis of TZ2 the primary amine of Lys was protected with Boc using Fmoc-Lys(Boc)-OH as the incorporated Lys precursor.....	53
2.17	Acetylation of the Ser hydroxyl group under basic conditions.....	54
2.18	Silylation of the hydroxyl and carboxylic acid group of Ser followed by the removal of the silyl group from the carboxylic acid. Here the EtOH selectively removes the silyl group from the carboxylic acid due to its increased susceptibility of the silyl ester to nucleophilic attack over the silyl ether.....	55
2.19	Creation of Ruhemanns purple by addition of ninhydrin to compounds containing exposed amines.....	59

3.1	Structure of TZ2 highlighting the S1 (green), TT (blue), T10 (purple), and K8 (pink).....	66
3.2	Equilibrium FTIR spectra are shown for each TZ2 isotopologue at 25 °C and pH 2.5. Each spectrum is baseline corrected with a linear subtraction, and area normalized for comparison. Difference data (bottom) are obtained by subtracting the unlabeled spectrum (TZ2-UL) from each of the spectra. The S1 difference spectra is presented at 2×magnification for clarity. Arrows have been drawn to emphasize the isotope labeled peaks.....	72
3.3	2D IR Spectra of TZ2 isotopologues at 25 (top) and 75 °C (bottom).....	74
3.4	Temperature-dependent FTIR spectra of each TZ2 isotopologue taken from 10 to 85 °C in 15 °C increments (pH 2.5). Line color shows progression from cold (blue) to hot (red).....	76
3.5	Second component amplitudes from SVD analysis of the temperature-dependent FTIR spectra. The inset shows the second component spectra. (Bottom) FTIR peak absorbance versus temperature. The ¹² C peak intensities are obtained from the absorbance within 5 cm ⁻¹ of the amide I' peak maximum. The integration regions for other traces are ν_{K8-1} , 1588-1598 cm ⁻¹ ; ν_{K8-2} , 1612-1619 cm ⁻¹ ; ν_{TT} , 1603-1611 cm ⁻¹ . Traces are scaled to overlap between 70 and 95 °C.....	77
3.6	Temperature-dependent 2D-IR spectra of isotope-shifted peaks in K8, T10, and TT. Contour lines are plotted as a percentage of the amide I' maximum at the respective	

	temperature. Twenty seven evenly spaced contours are drawn from -18 to +18% (K8), -80 to +80% (T10), and -35 to +35% (TT).....	80
3.7	Lineshape analysis of temperature-dependent 2D IR spectra of K8 and TT. Antidiagonal line width (FWHM) for (A) K8 and (B) TT. (C) Ratio of K8-2 to K8-1 peak volumes. (D) Ellipticity of isotope-shifted peaks. The data for K8-2 has been upshifted by 0.2 for clarity.....	81
3.8	(A) Folded (FO , 250851) is a well-ordered ensemble that has an average of four backbone hydrogen bonds, which correspond to those in the NMR structure. ¹ The S1, T3, K8, and T10 carbonyls are all oriented for cross strand hydrogen bonds, and all four tryptophan side chains are packed to one face of the peptide.....	84
	(B) Bulged Turn (BT , 214369) is a compact coiled structure with a well-solvated, bulged turn region and an average of one cross-strand hydrogen bond that appears near the strand termini. This state contains S1, T3, K8, and T10 peptide carbonyls that are solvent exposed, and tryptophan side chains make rare contacts.....	84
	(C) Disordered (SR , 271154) is a disordered ensemble whose common feature is a bulged turn and an average of one backbone hydrogen bond with a misaligned registry near the C terminus. Tryptophans are poorly organized.....	85
	(D) Frayed (FR , 11131) has a compact properly formed type I' turn region with an average of two backbone hydrogen bonds but disordered strand termini. The T3 and T10 are oriented toward the opposite strand, and tryptophan side chains are disordered.....	85

	(E) Extended Disorder (ED , 64336) is an extended and well solvated state with substantial disorder and rare backbone hydrogen bonds.....	86
3.9	2D IR spectral simulations of all five TZ2 isotopologues studied for five different conformational ensembles. One representative structure for each conformer macrostate is presented. Simulations of K8 for the spectral region $\omega_1 < 1650 \text{ cm}^{-1}$ have the amplitudes scaled by a factor of 10x.....	87
3.10	(Top) Isotope-labeled TZ2 amide I' site energy distributions for the five Markov states. (Bottom) Distribution of calculated amide I' site energies grouped by number of hydrogen bonds. A linear fit to the mean value gives $\omega = 1645.7 \text{ cm}^{-1} - 16.3 \times (\text{no. of CO H-bonds})$	88
3.11	(A) Analysis of simulated site energy variation by conformer.....	89
	(B) Site energy mean values showing error.....	89
3.12	Simulated FTIR and FTIR difference spectra for isotopologues of all Markov States.....	91
3.13	2D IR spectral simulations of TZ2 isotopologues for specific conformations.....	92
4.1	(Top) FTIR and (Bottom) 2D IR spectrum of GVGX ¹ PGVG at T = 10°C in pH = 1.0 DCl in D ₂ O. The proline resonance frequencies are Gly: 1625 cm ⁻¹ , Ala: 1619 cm ⁻¹ , and Val: 1614 cm ⁻¹ . Maximum optical densities were 0.23-0.25. Contours for all 2D IR	

surfaces presented in this paper are plotted from -1.0 to 1.0 in equally spaced steps of 0.08.....107

4.2 Val peptide 2D IR collected in ZZYY polarization at $T = 10^{\circ}\text{C}$ in a (Left) $\text{pH} = 1.0$, DCl in D_2O solution using $\tau_2 = 150$ fs, and in a (Right) $\text{pH} = 7.0$, 50 mmol phosphate buffer solution using $\tau_2 = 0$ fs. Changing the solvent conditions produces the appearance of on-diagonal broadening for the proline peak, a blue-shift of the main amide I' band, and a reduction in the off-diagonal line width. The change in the appearance of the proline peak results from the generation of the COO moiety upon changing the pH, creating a peak at 1590 cm^{-1} from the COO asymmetric stretch. The blue shift of the main amide I' band and the reduction in the off-diagonal line width observed at $\text{pH} = 7$ are likely due to the addition of buffer salts which produces a similar effect in $\text{pH} = 1.0$ samples.....108

4.3 The images above show representative structures from the 1/0 and 2/0 hydrogen bonding bins selected from the Val MD trajectories used in this work. Hydrogen bonds are labeled for a single selected structure which is fully represented, with an ensemble of similar structures (peptide backbone only) is shown in gold for comparison.....111

4.4 Ramachandran plots for the four central turn residues ($\text{X}_5\text{P}_6\text{G}_7\text{V}_8$) for the 1/0 and 2/0 hydrogen bond bins. Each plot was generated by culling structures based on hydrogen bonding criteria from 22 nsec of MD simulations for each of the three peptides GVGVPGVG, GVGAPGVG, and GVGPGGVG. Culling was performed using the `g_hbond` utility implemented in GROMACS 4.5 using an acceptor--donor--H cutoff angle of 30° and a donor-acceptor cutoff distance of 3.5 \AA . Ramachandran angles were likewise extracted using the GROMACS `g_chi` utility for each trajectory.....112

4.5	(Top) Visualization of representative turn structure in the 0/0, 1/0 and 2/0 bins. (Bottom) Simulated FTIR spectra for the Gly peptide in the 0/0, 1/0 and 2/0 bins.....	113
4.6	Comparison of experimental and simulated FTIR spectra for the Gly, Ala and Val peptides in the six hydrogen bond bins. Note the non-zero offset of the blue side of the spectrum is the result of a broad –COOD stretching vibration centered at 1722 cm ⁻¹	114
4.7	Comparison of the experimental data (Top row) to the six simulated surfaces generated for each peptide.....	115
4.8	(a) Simulated 2D IR spectra of the Val peptide as a function of peptide-peptide hydrogen bond number. (b) Simulated 2D IR spectra of the Val peptide as a function of peptide-peptide vs. peptide-water hydrogen bonds.....	116
4.9	Average frequency of the simulated proline peak maximum for the 0/0, 1/0 and 2/0 bins. The linear fit to this data produced an equation relating the frequency of the proline peak and the number of peptide-to-proline hydrogen bonds: $\omega_p(\text{cm}^{-1}) \approx - 13.8 \cdot n_{\text{PP}} + 1642$	117
4.10	Average frequency of the simulated proline peak maximum for all six bins plotted as a function of the total number of peptide and water hydrogen bonds to proline. The linear fit to this data produced an equation relating the frequency of the proline peak and the total number of hydrogen bonds to proline: $\omega_p(\text{cm}^{-1}) \approx - 12.6 \cdot n_{\text{THB}} + 1642$	117

4.11 (Top row) Experimental 2DIR data, (Bottom row) - Fitted spectra, generated by fitting the experimental data over the frequency range 1575 to 1650 cm^{-1} with the simulated spectra for all six peptide bins using a genetic algorithm containing a least squares fitness function and no additional constraints. Since the amide I' band maximum is located at $\sim 1645\text{cm}^{-1}$ this fit range allowed for fitting both the proline peak shift and its intensity relative to the main amide band. Listed to the right of each surface is the population of individual HB bin's predicted from the fitting routine. Contours for all 2D IR surfaces presented in this paper are plotted from -1.0 to 1.0 in equally spaced steps of 0.08.....118

4.12 Comparison of the experimental spectrum for the Val peptide to simulated spectra using the bin populations generated for the fitting of the experimental data and populations calculated from a 200 ns MD trajectory.....120

5.1 Structure of elastin proposed by Partridge in 1966, as well as Weis-Fogh and Anderson in 1970. They described the elastin fiber as a cross linked network of collapsed peptide globules, represented by circles and ellipses, with water in the spaces in between. Image (A) represents the extended state of the fiber and image (B) represents the contracted relaxed state of the fiber. This figure was reproduced from the work of Weis-Fogh and Anderson.....130

5.2	Proposed oiled coil model for elastin taken from the work of Gray, Sandberg, and Foster. Here the elastin fiber consists of alternating segments of cross linking α -helical regions and elastic oiled coils.....131
5.3	The β -spiral structure of elastin proposed by Dan W. Urry and coworkers. (A) Schematic representation of the poly(GVGVP) sequence demonstrating how the PG unit creates periodic β -turns. (B) PG turn structure from the crystal structure of the cyclo-(GVGVP) ₃ peptide. (C and D) Helical representations of the poly(GVGVP) sequence with and without the PG β -turns. (E) Stereo plot of the β -spiral structure. (F) Proposed structure for the association of β -spirals to form twisted filaments drawn with a space-filling united residue representation. This figure was reproduced from the work of Dan W. Urry.....132
5.4	MD structures for the (VPGVG) ₁₈ peptide at temperatures from 7 to 42 °C, 6 ns after starting from a β -spiral conformation. The peptides main chain is drawn in red and side chain atoms in green. This image was reproduced from the work of Daggett and coworkers.....136
5.5	Depiction of the three phase model proposed by Debelle et. al. In the lower portion of the image is a cross linked (crosses) network of globular tropoelastin proteins (circles) with bulk water (B) in the interstitial space. Above is a representation of the structure with in a single tropoelastin globule. Here α -helical cross linking domains are at the surface of the globule with a semi ordered peptide rich in β turns and strands with hydration water molecules solvating these structures. This image is reproduced from the work of Debelle et. al.....138

- 5.6 Library of peptides used for assigning GVGn1 structure. Boxes indicate the location of isotope labeled amide sites. All labeled amide sites were labeled with $^{13}\text{C}^{18}\text{O}$ except for the G3 site in V1G3V4 which was labeled with $^{13}\text{C}^{16}\text{O}$ for increased peak resolution. The monomer unit VPGVG is drawn in black in the WT structure with amide sites numbered to demonstrate the naming convention used in this work.....141
- 5.7 FTIR spectrum for the GVGn1 peptide library collected at 10 °C on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). The green and circle are placed at 1635 and 1675 cm⁻¹ to indicate the location of the β-sheet ν_{\perp} and ν_{\parallel} modes respectively.....145
- 5.8 2DIR spectrum for the library of GVGn1 peptides used for assigning the peptides structure. Spectrum were collected in ZZYY polarization at 10 °C and $\tau_2 = 150$ fs on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). Note that due to the desalting of the peptides these 150 mmol spectrum contain line-shapes similar to spectrum collected for peptides prepared without a desalting step and dissolved in pH = 1.0 DCl in D₂O with no additional salt added.....146
- 5.9 [(Isotope Label) – WT] FTIR difference spectrum for the GVGn1 peptide library collected at 10 °C on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄.....148
- 5.10 Experimental V1 FTIR and ZZYY 2DIR spectrum collected at 10 °C on a desalted peptide at $\tau_2 = 150$ fs in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄. The FTIR spectrum was fit with 2 Gaussian peaks giving populations for the 2/0 and 1/0 turns of

	82% and 18%. In the upper right hand corner of the figure is a schematic representation of the GVGn1 peptide in a 2/0 turn with hydrogen bonds depicted as red lines.....	149
5.11	Waiting time 2DIR data for non-desalted V1 and N-Acetyl-Pro-COOD collected in ZZZZ polarization at 10 °C in pH = 1.0 DCl in D ₂ O with no additional salt added. The black lines in the figure are the result of fitting the fundamental peak maximum.....	152
5.12	Waiting time 2DIR data for non-desalted V4 and G3 collected in ZZZZ polarization at 10 °C in pH = 1.0 DCl in D ₂ O with no additional salt added. Each figure has two perpendicular lines, one indicating the fit line for the fundamental peak maximum and a perpendicular line in the anti-diagonal direction provided as a guide for the eye.....	154
5.13	[(Isotope Label) – WT] FTIR difference spectrum for G3 [†] (Red), G5 [†] (Cyan), G3 (Green), and V4 (Blue), collected at 10 °C on desalted peptides in pH = 1.0 DCl in D ₂ O with 150 mmol of KD ₂ PO ₄	155
5.14	FTIR spectrum for G3 [†] (Red) and WT (Black) collected at 10 °C on desalted peptides in pH = 1.0 DCl in D ₂ O with 150 mmol of KD ₂ PO ₄	157
5.15	FTIR and 2DIR spectrum for G5 [†] , G3, and G5 [†] G3 collected at 10 °C on desalted peptides in pH = 1.0 DCl in D ₂ O with 150 mmol of KD ₂ PO ₄ . 2DIR spectrum were collected in ZZZY polarization at $\tau_2 = 150$ fs.....	159
5.16	FTIR and 2DIR spectrum for V1, V4, and V1V4 collected at 10 °C on desalted peptides in pH = 1.0 DCl in D ₂ O with 150 mmol of KD ₂ PO ₄ . 2DIR spectrum were collected in ZZZY polarization at $\tau_2 = 150$ fs.....	160

- 5.17 2DIR spectrum for G3, V1V4, and V1G3V4 collected at 10 °C on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄. 2DIR spectrum were collected in ZZYY polarization at $\tau_2 = 150$ fs. Green circles indicate the location of isotope peaks and the corresponding green lines are provided to indicate the location of cross peaks. In the G3 spectrum the green dots are placed over the peak maximum of the G3 and unlabeled V1 peaks with $\omega_{\max} = 1581.0$ and 1614.0 cm^{-1} respectively. In the V1V4 spectrum the green dots are placed over the labeled V1 2/0 and 1/0 peaks with $\omega = 1554.4$ and 1572.2 cm^{-1} along with a dot over the V4 peak maximum at $\omega_{\max} = 1584.1 \text{ cm}^{-1}$. The V1G3V4 peptide has green dots over the V1 2/0 and 1/0 peaks with $\omega = 1554.4$ and 1572.2 cm^{-1} along with a dot over the V4 and G3 peak maxima at $\omega_{\max} = 1584.1$ and 1610.3 cm^{-1} respectively.....162
- 5.18 FTIR spectrum for non-desalted TZ2 in pH = 2.5 DCl in D₂O with no salt added and non-desalted V1 in pH = 1.0 DCl in D₂O with no salt.....163
- 5.19 FTIR spectrum for non-desalted WT, V1, G3, and G5[†]G3 collected as a function of temperature from 10-60 °C in 5 °C steps. Spectrum in the left column were prepared in pH = 1.0 DCl in D₂O and in the right column in pH = 1.0 DCl in D₂O with 150mmol of KD₂PO₄ added.....164
- 5.20 FTIR spectrum for non-desalted WT, V1, G3, and G5[†]G3 collected at T = 10 and 60 °C. Data collected in pH = 1.0 DCl in D₂O is displayed with red and blue solid lines and data collected in pH = 1.0 DCl in D₂O with 150mmol of KD₂PO₄ added is displayed with red

	and blue dashed lines. The FTIR difference spectrum (No salt – Salt) at constant temperature is displayed in black.....	165
5.21	FTIR spectrum collected on the non-desalted V1 peptide. (Left) Temperature dependent FTIR spectrum collected from $T = 10-60\text{ }^{\circ}\text{C}$ in $5\text{ }^{\circ}\text{C}$ steps, (Center) Fit to the $10\text{ }^{\circ}\text{C}$ spectrum, and (Right) Fit to the $60\text{ }^{\circ}\text{C}$ spectrum.....	166
5.22	FTIR spectrum collected on the non-desalted V1 peptide plotted for the COOD region.....	167
5.23	2DIR spectra of the non-desalted V1 peptide as a function of temperature and salt concentration. All spectra were collected in a ZZZZ polarization geometry with $\tau_2 = 1.5\text{ps}$	168
5.24	Temperature dependent FTIR spectra for non-desalted GVG(VPGVG) _n where $n = 1-6$ and the (GVGVP) ₂₅₁ polymer in $\text{pH} = 1.0$ DCl in D ₂ O with 150 mmol of KD ₂ PO ₄	170
5.25	First and Second SVD component spectrum and corresponding melting curves for GVG(VPGVG) _n where $n = 1-6$ and the (GVGVP) ₂₅₁ polymer in $\text{pH} = 1.0$ DCl in D ₂ O with 150 mmol of KD ₂ PO ₄	171
5.26	(Left) Temperature dependent FTIR spectrum for (GVGVP) ₂₅₁ in $\text{pH} = 1.0$ DCl in D ₂ O with 150 mmol of KD ₂ PO ₄ . (Right) A comparison between the (GVGVP) ₂₅₁ FTIR spectrum with the simulated 2/0, 1/0, and 0/0 spectrum for the GVGn1 peptide.....	172

5.27 ZZZY 2DIR spectrum for the GVGn1 (Left column) and (GVGVP)₂₅₁ (Right column) peptides. Desalted GVGn1 was collected at T = 10 °C and $\tau_2 = 150$ fs in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). (GVGVP)₂₅₁ was collected at T = 15 °C and $\tau_2 = 150$ fs in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄. Dots are placed to mark the location of the antiparallel β -sheet modes, $\nu_{\perp} = 1640$ and $\nu_{\parallel} = 1670$ cm⁻¹, as well as the location of the GVGn1 and (GVGVP)₂₅₁ proline peak maximum at $\omega_{\max} = 1614$ and 1608.3 cm⁻¹ respectively.....174

List of Tables

4.1 Bin populations calculated from a 200 ns MD trajectory, 27.5 ns MD trajectory, and the 2D IR experimental surface fitting for the Val peptide.....119

5.1 Center frequency (ω), on diagonal FWHM (Δ), off diagonal FWHM (Γ), and ellipticity (E) for the isotope labeled sites of GVGn1. The values for ω , Δ , and Γ are presented here in cm⁻¹ and the value for the ellipticity is given by $E = (\Delta^2 - \Gamma^2)/(\Delta^2 + \Gamma^2)$. Peak metrics for all compounds with the exception of G3[†] were measured by fitting their 2D IR spectra. Since the G3[†] isotope peak is partially overlapped in frequency with the proline resonance it was not possible to accurately fit its 2D IR surface. As a result the value for ω was measured from the [G3[†] - WT] FTIR difference spectrum presented in figure 9.....147

Chapter 1

Preface

1.1 Introduction

Protein conformational dynamics lie at the heart of numerous biophysical processes, including protein folding, substrate binding, molecular recognition, and self-assembly. Historically, such processes were often conceived in terms of rigid protein molecules with minimal conformational degrees of freedom. However, structure formation in such processes involves a complex interplay between the favorable energy of contact formation and conformational entropy. Protein flexibility and conformational heterogeneity are intrinsic to all systems, but this disorder and its role in biophysical processes remains difficult to study and describe experimentally. This stems from a paucity of methods that can explore these questions. Conformational dynamics occur over picosecond and longer timescales, and atomistic structural resolution in solution is only routinely available from computer simulations. The most detailed experimental observations have emerged from magnetic resonance spin relaxation rate measurements that characterize inherently disordered proteins,¹ residue-level protein folding,² transient protein-protein complexes,^{3,4} and microsecond structural fluctuations.⁵⁻⁸ Two-dimensional infrared spectroscopy (2D IR) is an emerging method that is being used to directly

measure structure and conformational dynamics in proteins and peptides.⁹⁻¹⁴ Analogous to multidimensional NMR, 2D IR observes structure sensitive vibrational couplings through cross peaks in a 2D spectrum.¹⁵ In addition, 2D line shapes provide information on molecular flexibility and solvent exposure through their peak frequency and ellipticity. Time-dependent changes to cross peaks and line shapes reveal chemical exchange processes and fluctuations. Since the measurement time scale for vibrational spectroscopy is in the picosecond range, such studies provide an avenue to characterize conformational variation and ultrafast fluctuations. The technique has been used to measure protein structure and dynamics with particular focus on the amide I vibrations, which are sensitive to secondary structure and the local hydrogen bonding geometry of the peptide groups.^{9-12, 16-20}

Extensive study of small peptides has revealed the local forces that contribute to protein secondary structure.²¹⁻³⁰ In contrast, far less is known about the forces that contribute to the structure of disordered proteins where local effects dominate.³¹⁻³⁷ One striking example of a poorly understood disordered protein is elastin whose structure, despite extensive investigation, remains unassigned. Elastin is a ubiquitous component of the extracellular matrix of animals and is responsible for the elasticity of bodily tissues such as the skin, arteries, and lungs.³⁸⁻⁴¹ The elastin protein has been the subject of interest in biology, chemistry, and material science because of its strength, durability, phase transition behavior, and ability to self-assemble.⁴²⁻⁴⁵ To date the molecular origin of these properties is unknown owing to a dearth of structural tools that can perform measurements on the elastin fiber. Elastin protein cannot be crystallized, and X-ray diffraction or scattering studies provide limited structural information. Although NMR is selective to long lived conformational states, since conformational exchange for elastin occurs on

picosecond to nanosecond time scales, NMR cannot be used to resolve its conformational heterogeneity.^{35,46}

Some clues in to the mechanism of the proteins elasticity have been revealed through amino acids sequence analysis. The amino acids sequence of elastin consists of cross-linking domains with a high percentage of lysine with alanine or proline and hydrophobic elastic domains with a high percentage of alanine, glycine, leucine, isoleucine, proline, and valine.⁴⁷⁻⁴⁹ In addition these elastic domains are found to consist of small repeating blocks of PGV, PGGV, PVGG, PGVGV, PGVGVA, and PGFGVGAGV generating a simple biopolymer. One dominant and extensively studied elastin repeat sequence is PGVGV, as a result the canonical elastin mimic GVGVPGVG (GVGn1) was used as a model system.

The investigation of elastin presented in this work is broken down into 4 main chapters which will focus on the synthesis, vibrational spectroscopy, and structure of two amide I' isotope labeled peptides, Tryptophan Zipper 2 (TZ2) and GVGn1. Chapter 2 will present the steps taken to synthesize the peptides studied. An explanation of Fmoc solid phase peptide synthesis (SPPS) will be provided including reaction mechanisms, guidelines for designing a peptide synthesis, methods for avoiding unwanted side reactions and peptide cleavage. This chapter will also discuss the synthesis of Fmoc amino acids with ¹⁸O amide I' isotope labels for use in SPPS. In Chapter 3 a 2D IR study of the 12 residue TZ2 is presented. TZ2 is an engineered peptide that contains four indole rings from the peptides tryptophan side-chains whose interaction stabilizes the formation of a β -hairpin fold. TZ2 was chosen because it has been the subject of numerous spectroscopic and theoretical investigations and as a result is an ideal model system for understanding vibrational couplings in small hairpin peptides. This work will act as a template for the approach used in Chapter 5 for studying the GVGn1 peptide by demonstrating the ways

in which isotope labels can be employed to understand peptide structure and dynamics. Next in Chapter 4 the GVGXPGVG peptide will be introduced where X = Gly, Ala, and Val. This point mutation was found to significantly alter the turn structure for the peptide with an increase in the population of closed turns as a function of increasing side-chain volume. Finally in Chapter 5 more subtle aspects of the GVGn1 peptides structure were explored using a library of 9 isotopologues as well as oligomers based on the PGVGV sequence. These peptides were studied as both a function of temperature and salt concentration providing information on the local structural changes that occur as the peptide transitions from its unfolded to folded state.

1.2 Acknowledgments

The introduction presented in this chapter was reproduced in part from the article “Melting of a β -Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations. Adam W. Smith, Joshua Lessing, Ziad Ganim, Chunte Sam Peng, Santanu Roy, Thomas L. C. Jansen, Jasper Knoester and Andrei Tokmakoff; *J. Phys. Chem. B* 2010, 114, 10913–10924.”

1.3 References

- (1) Dyson, H. J.; Wright, P. E. *Chem. Review* **2004**, *104*, 3607.
- (2) Sadqi, M.; Fushman, D.; Munoz, V. *Nature* **2006**, *442*, 317.
- (3) Junji Iwahara, G. M. C. *Nature* **2006**, *440*, 1227.
- (4) Chun Tang, J. I., G. Marius Clore *Nature* **2006**, *444*, 383.
- (5) Anthony Mittermaier, L. E. K. *Science* **2006**, *312*, 224.
- (6) III, A. G. P. *Curr. Opin. Struct. Biol.* **1997**, *7*, 732.

- (7) Mikael Akke, J. L., John Cavanagh, Harold P. Erickson, Arthur G. Palmer III *Nat. Struct. Biol.* **1998**, *5*, 55.
- (8) III, A. G. P. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 129.
- (9) Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. *Acc. Chem. Res.* **2008**, *41*, 432.
- (10) Yung Sam Kim, R. M. H. *J. Phys. Chem. B* **2009**, *113*, 8231.
- (11) Mukherjee, P.; Kass, I.; Arkin, I.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3528.
- (12) Sengupta, N.; Maekawa, H.; Zhuang, W.; Toniolo, C.; Mukamel, S.; Tobias, D. J.; Ge, N.-H. *J. Phys. Chem. B* **2009**, *113*, 12037.
- (13) Zhuang, W.; Hayashi, T.; Mukamel, S. *Angew. Chem. Int. Ed.* **2009**, *48*, 3750.
- (14) Jeon, J.; Yang, S.; Choi, J.-H.; Cho, M. *Acc. Chem. Res.* **2009**, *42*, 1280.
- (15) Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258.
- (16) Hoi Sung Chung, Z. G., Kevin C. Jones, Andrei Tokmakoff *Proc. Nat. Acad. Sci. U.S.A.* **2007**, *104*, 14237.
- (17) Jianping Wang, W. Z., Shaul Mukamel, Robin Hochstrasser *J. Phys. Chem. B* **2008**, *112*, 5930.
- (18) Yung Sam Kim, L. L., Paul H. Axelsen, Robin M. Hochstrasser *Proc. Nat. Acad. Sci. U.S.A.* **2008**, *105*, 7720.
- (19) Hiroaki Maekawa, M. D. P., Claudio Toniolo, Nien-Hui Ge *J. Am. Chem. Soc.* **2009**, *131*, 2042.
- (20) Sang-Hee Shim, R. G., Yun L. Ling, David B. Strasfeld, Daniel P. Raleigh, Martin T. Zanni *Proc. Nat. Acad. Sci. U.S.A.* **2009**, *106*, 6614.

- (21) H. Jane Dyson, P. E. W. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519.
- (22) Francisco J. Blanco, M. A. J., José Herranz, Manuel Rico, Jorge Santoro, José L. Nieto *J. Am. Chem. Soc.* **1993**, *115*, 5887.
- (23) Mark S. Searle, D. H. W., Leonard C. Packman *Nature Structural Biology* **1995**, *2*, 999.
- (24) Faisal A. Syud, H. E. S., Samuel H. Gellman *J. Am. Chem. Soc.* **2001**, *123*, 8667.
- (25) Juan F. Espinosa, F. A. S., Samuel H. Gellman *Protein Sci.* **2002**, *11*, 1492.
- (26) Jovencio Hilario, J. K., Timothy A. Keiderling *J. Am. Chem. Soc.* **2003**, *125*, 7562.
- (27) Tamas Blandl, A. G. C., Nicholas J. Skelton *Protein Sci.* **2003**, *12*, 237.
- (28) Adam W. Smith, A. T. *Angew. Chem. Int. Ed.* **2007**, *46*, 7984.
- (29) Adam W. Smith, A. T. *J. Chem. Phys.* **2007**, *126*, 045109.
- (30) Adam W. Smith, J. L., Santanu Roy, Ziad Ganim, Chunte Sam Peng, Thomas L. C. Jansen, Jasper Knoester, Andrei Tokmakoff *J. Phys. Chem. B* **2010**, *114*, 10913.
- (31) Urry, D. W. *J. Phys. Chem. B* **1997**, *101*, 11007.
- (32) David S. Wishart, C. G. B., Arne Holm, Robert S. Hodges, Brian D. Sykes *J. Biomol. NMR* **1995**, *5*, 67.
- (33) Eduard Schreiner, C. N., Björn Ludolph, Revanur Ravindra, Nikolaj Otte, Axel Kohlmeyer, Roger Rousseau, Roland Winter, Dominik Marx *Phys. Rev. Lett.* **2004**, *92*.
- (34) Marcel Baer, E. S., Axel Kohlmeyer, Roger Rousseau, Dominik Marx *J. Phys. Chem. B* **2006**, *110*, 3576.
- (35) Rachel Graves, M. B., Eduard Schreiner, Raphael Stoll, Dominik Marx *ChemPhysChem* **2008**, *9*, 2759.

- (36) Roger Rousseau, E. S., Axel Kohlmeyer, Dominik Marx *Biophys. J.* **2004**, *86*, 1393.
- (37) Zeeshan Ahmed, J. P. S., Sanford A. Asher *Biopolymers* **2008**, *91*, 52.
- (38) Bernadette Vrhovski, A. S. W. *Eur. J. Biochem.* **1998**, *258*, 1.
- (39) Dean Y. Li, B. B., Elaine C. Davis, Robert P. Mecham, Lise K. Sorensenk, Beth B. Boak,; Ernst Eichwald, M. T. K. *Nature* **1998**, *393*, 276.
- (40) Daniel P. Wendel, D. G. T., Kurt H. Albertine, Mark T. Keating, Dean Y. Li *Am. J. Respir. Cell Mol. Biol.* **2000**, *23*, 320.
- (41) Richard L. Kornberg, S. S. H., Aarne I. Oikarinen, Lois Y. Matsuoka, Jouni Uitto *N. Engl. J. Med.* **1985**, *312*, 771.
- (42) Urry, D. W. *Angew. Chem. Int. Ed.* **1993**, *32*, 819.
- (43) Steven D. Shapiro, S. K. E., Michael A. Province, John A. Pierce, Edward J. Campbell *J. Clin. Invest.* **1991**, *87*, 1828.
- (44) Bernadette Vrhovski, S. J., Anthony S. Weiss *Eur. J. Biochem.* **1997**, *250*, 92.
- (45) M. Manno, A. E., V. Martorana, P. L. San Biagio, D. Bulone, M. B. Palma-Vittorelli, D. T. McPherson, J. Xu, T. M. Parker, D. W. Urry *Biopolymers* **2001**, *59*, 51.
- (46) Jürgen Graf, P. H. N., Gerhard Stock, Harald Schwalbe *J. Am. Chem. Soc.* **2007**, *129*, 1179.
- (47) William R. Gray, L. B. S., Judith A. Foster *Nature* **1973**, *246*, 461.
- (48) Zena Indik, H. Y., Norma Ornstein-Goldstein, Paul Sheppard, Noel Anderson, Joan C. Rosenbloom, Leena Peltonen, Joel Rosenbloom *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 5680.

(49) Sandberg, L. B.; Leslie, J. G.; Leach, C. T.; Alvarez, V. L.; Torres, A. R.; Smith, D. W. *Pathologie-biologie* **1985**, 33, 266.

Chapter 2

Synthesis of amide I isotope labeled peptides

2.1 ^{18}O Isotope labeling

Amide I ^{18}O isotope labeling was performed via acid hydrolysis in mixed H_2^{18}O and organic solvent. This synthesis was designed based on the procedure of Murphy et. al.¹ for the ^{18}O labeling of the carboxyl moiety of an unprotected amino acid. In their simple procedure a solution of an unlabeled amino acid was dissolved in acidic H_2^{18}O and refluxed at 60-70°C for several days generating over 90% isotope enrichment without racemization. This procedure was modified for the present work in order to isotope label amino acids with an N-terminal fluorenylmethoxycarbonyl (Fmoc) protecting group (Figure 1). The Fmoc moiety was

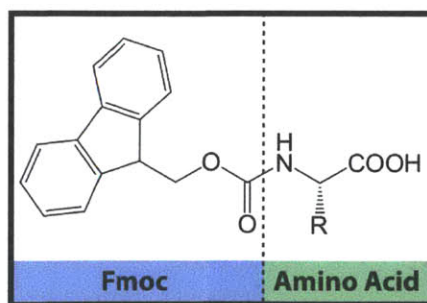
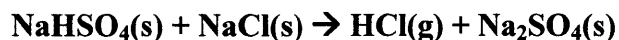


Figure 1. N-terminal Fmoc protected amino acid.

necessary because it is needed for solid phase peptide synthesis (SPPS) which will be discussed later in this chapter. Since the Fmoc group is hydrophobic an organic co-solvent was added to

aid in dissolving the protected amino acids in H_2^{18}O . The desired co-solvent for the labeling reaction would dissolve Fmoc protected amino acids, would be easy to remove by lyophilization, and would not act as a source of ^{16}O during labeling. Acetonitrile was originally used for this purpose but was replaced with 1,4-dioxane² due to the undesired hydrolysis of acetonitrile to generate an isotope labeled acetamide.³ The first amino acid to be ^{18}O labeled was serine (Ser or S) for incorporation at the N-terminal Ser1 position in the Tryptophan Zipper 2 (TZ2) peptide $\text{NH}_2\text{-SWTWENGKWTWK-CONH}_2$. The procedure presented below details the synthesis of the isotope labeled Ser using acetonitrile as the co-solvent, though this procedure can be generalized to the labeling of other none hydrolysable amino acids using 1,4-dioxane as a co-solvent including aspartic acid (Asp or D), glutamic acid (Glu or E), threonine (Thr or T), proline (Pro or P), alanine (Ala or A), valine (Val or V), leucine (Leu or L), isoleucine (Ile or I), phenylalanine (Phe or F) and tryptophan (Trp or W). Though in the case of Asp and Glu subsequent side-chain protection is difficult because it requires the selective protection of the carboxylic acid side-chain while keeping the main chain carboxylic acid deprotected. Labeled Ser was synthesized directly from Fmoc-Ser-OH (Novabiochem) using acid catalyzed hydrolysis. First the glassware for a reflux reaction was assembled, which primarily consisted of a two-neck round-bottom flask and a condenser, and heated under vacuum for 20 min at 110°C for the purpose of removing ambient H_2^{16}O water. The reflux setup was then cooled under positive N_2 pressure to prevent contamination from ambient moisture. Next 0.131 g of Fmoc-Ser-OH was dissolved in a minimal amount of acetonitrile (~3.5 ml) and injected into the reaction vessel followed by the addition of 97% isotope enriched H_2^{18}O (Isoplex USA) in a 1:4 ratio of H_2^{18}O to acetonitrile. It should be noted that the ratio of water to organic solvent will change dramatically depending on one's choice of amino acid and co-solvent.

Next the solution was acidified either by bubbling anhydrous HCl gas through the solution or generation of HCl in situ through the addition of acetyl chloride. HCl gas was generated by heating a powdered mixture of anhydrous sodium bisulfate and sodium chloride.



To do this 0.717 g of NaCl and 1.137 g of NaHSO₄ were finely ground, mixed and placed in a small septum sealed round bottom flask. The flask was then purged with nitrogen following which two syringes with stopcocks were inserted into the septum. Next the first stopcock was opened and the flask was heated with a hand held propane torch until the syringe contained 20-30 ml of HCl gas. Since the first aliquot of evolved gas is mixed with N₂ this gas was disposed of by bubbling it through a flask containing H₂¹⁶O water. Next the second stopcock was opened and the flask was again heated until an additional 10 ml of gas was collected. This gas was then bubbled through the dissolved Fmoc-Ser-OH solution until the pH = 1.0. Alternatively HCl can be generated in situ by injecting a solution of 1M acetyl chloride in dichloromethane (DCM) (aldrich) into the reaction vessel which hydrolyzes to form HCl and acetic acid.²

The solution was then refluxed at 100°C for 72hrs and lyophilized generating a ~75% isotopically enriched sample. Alternatively the H₂¹⁸O/organic solvent mixture can be removed by distillation thereby providing material for future low isotopic enrichment of amino acids. For the Ser1 and Thr10 ¹⁸O isotope labeled TZ2 peptides described in chapter 4 the amino acids were incorporated into the peptide at ~75% enrichment though for the GVGn1 peptides discussed in chapters 5 and 6 the labeling procedure was repeated to achieve ~97% isotopic enrichment. Next a sample was extracted from the solution for electrospray mass spectroscopy to determine the percent isotope incorporation (these mass spectra and HPLC data are presented in the appendix).

Fmoc Solid Phase Peptide Synthesis

X = Main chain protecting group

Y = Side-chain protecting group

A = Carboxy activating group

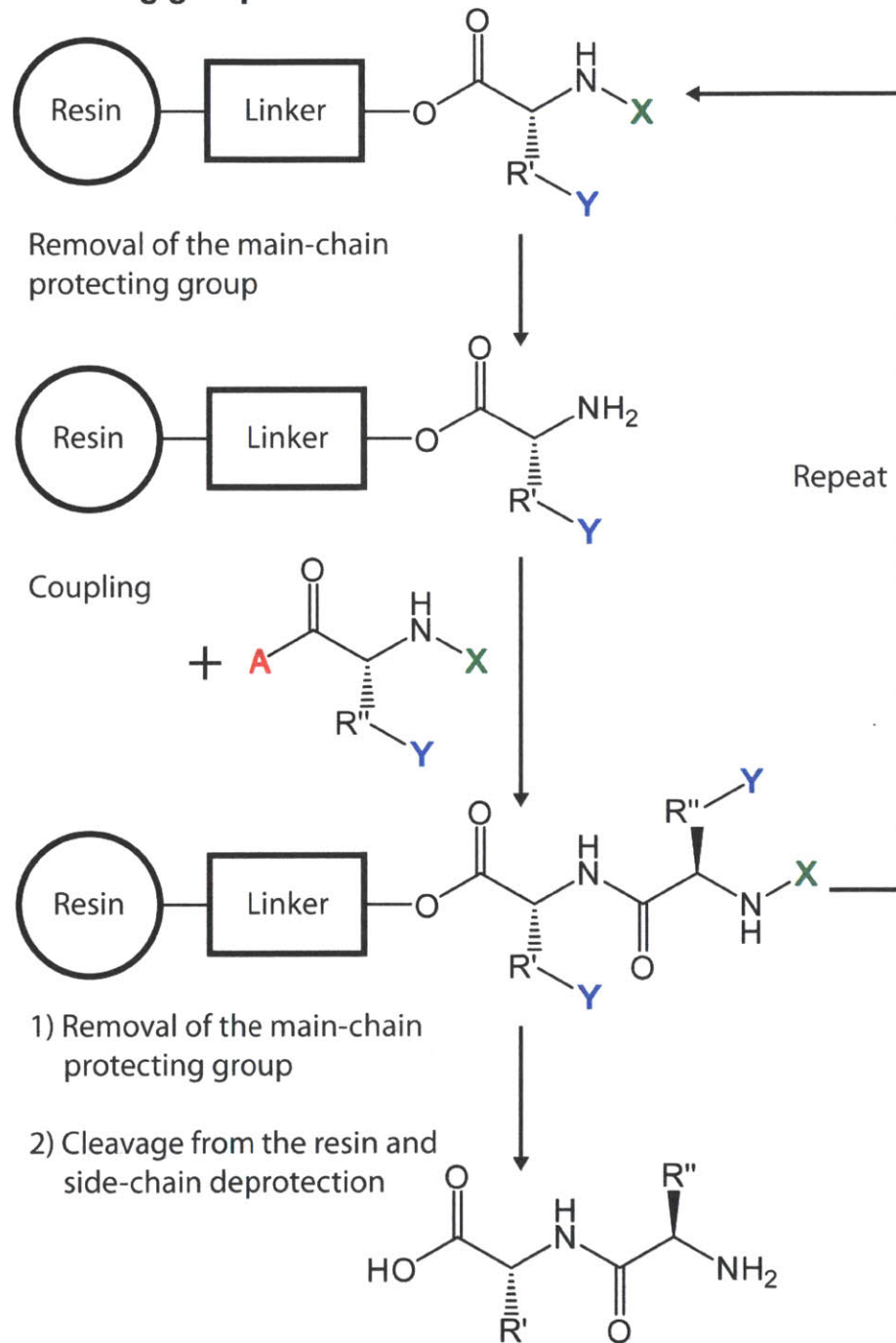


Figure 2. Steps involved in performing Fmoc SPPS.

2.2 Solid Phase Peptide Synthesis

SPPS is a process for generating peptides and small proteins by sequential tethering of amino acids to a linker on a solid support. The main advantage of SPPS is the use of a heterogeneous synthesis approach thereby simplifying purification between reactions to rinsing and filtration of the solid support. The development of SPPS by Robert B. Merrifield⁴ represented a paradigm shift in biochemistry and pharmacology, by making peptide synthesis simple, practical, and automatable.

The steps of an SPPS synthesis can be broken down into deprotection, coupling, and cleavage. During “deprotection” the growing resin bound peptides main-chain protecting group is removed exposing the C or N-terminus for the benzyloxy-carbonyl (Boc) or Fmoc procedures (Figure 2) respectively. Next a “coupling” step is performed which adds the subsequent amino acid in the peptide sequence. These steps of deprotection and coupling are then repeated for each amino acid addition until the desired sequence is generated. Finally a “cleavage” step is performed removing side-chain protecting groups and main-chain linkers to generate the peptide product.

The two common methods for solid phase peptides synthesis are acid labile Boc synthesis and base labile Fmoc synthesis.⁵⁻⁸ Although each method requires different resins, side-chain protecting schemes, reagents, and synthesis directions the underlying principle remains the same. The significant difference between the Boc and Fmoc procedures is the choice of main-chain and side-chain protecting scheme. For Boc SPPS the peptide is synthesized from the N to the C-terminus using a Boc protecting group on the C-terminus which is labile under mildly acidic conditions typically using trifluoroacetic acid (TFA). In contrast the side-chain protecting groups and the linker are cleavable under harsh acidic conditions typically using hydrofluoric

acid. As a result the side-chain and main-chain protecting groups as well as the linker are cleavable under acidic conditions resulting in a slow loss of side-chain protecting groups and cleavage during sequential steps of amino acid coupling thereby lowering the final yield of the peptide. Since peptide synthesis requires multiple steps a small reduction in yield at each step propagates to generate a larger reduction in the yield of the final product. For example the longest and shortest peptides studied in this work are GVGVPGVG requiring 7 coupling steps and GVG(VPGVG)₆ which requires 32 coupling steps. If the yield for each coupling step decreases from 99% to 97% the final yield for GVGVPGVG would decrease from 93% to 81% and for GVG(VPGVG)₆ would decrease from 72% to 38%. As a result optimizing the yield of each step in a peptide synthesis is of paramount importance. An easier and safer alternative to main-chain protection with Boc protecting groups is the use of the orthogonal Fmoc based protecting scheme. For this approach the peptide is synthesized in the opposite direction from the Boc procedure, i.e. from the N to the C-terminus, using a main-chain protecting group on the N-terminus. Here the side-chain protecting groups and main-chain linker are labile under mildly acidic conditions, and the main-chain is protected with the Fmoc group which is labile under basic conditions. This approach allows for the deprotection of the main-chain functional group at each sequential step of amino acid addition without deprotecting the side-chain groups or cleaving the peptide and is therefore termed an orthogonal synthesis. This orthogonal synthesis approach in conjunction with the use of more mild reagents makes Fmoc SPPS a natural choice for the synthesis of the peptides studied in this work. Below is a detailed description of the steps involved in performing an Fmoc synthesis.

2.2.1 Choosing a resin

SPPS synthesis begins with the selection and preparation of a resin. Many factors are considered when choosing a resin such as the chemical composition of the polymer support, degree of cross-linking, degree of swelling upon solvent addition, resin size, resin size distribution, choice of linker group and binding capacity.^{9,10}

Typical polymer support materials include polystyrene, polyacrylamide, or a copolymer of polyethylene glycol (PEG) grafted to polystyrene. When choosing a polymer matrix it is important to match the hydrophobicity of the polymer with the hydrophobicity of the solvent system in use and the sequence being synthesized. Matching the hydrophobicity of the polymer to that of the solvent increases the swelling of the polymer thereby increasing the rate of molecular diffusion within the resin. Matching the hydrophobicity of the polymer support and the solvent to that of the sequence being synthesized improves both coupling yields and yield upon cleavage. To increase coupling yields, especially for the synthesis of long sequences, it is important that the peptide exist in an extended conformation while attached to the resin in order to keep the peptide terminus accessible for the coupling of sequential amino acids. Matching the hydrophobicity of the sequence to that of the polymer and solvent allows the peptide to remain in an extended state during synthesis. Finally a common cause of poor yields in SPPS is destruction of the peptide upon cleavage. One cause for this is resin nucleated aggregation of the peptide product after cleavage. Using a polymer with a similar degree of hydrophobicity reduces the probability of polymer mediated aggregation.

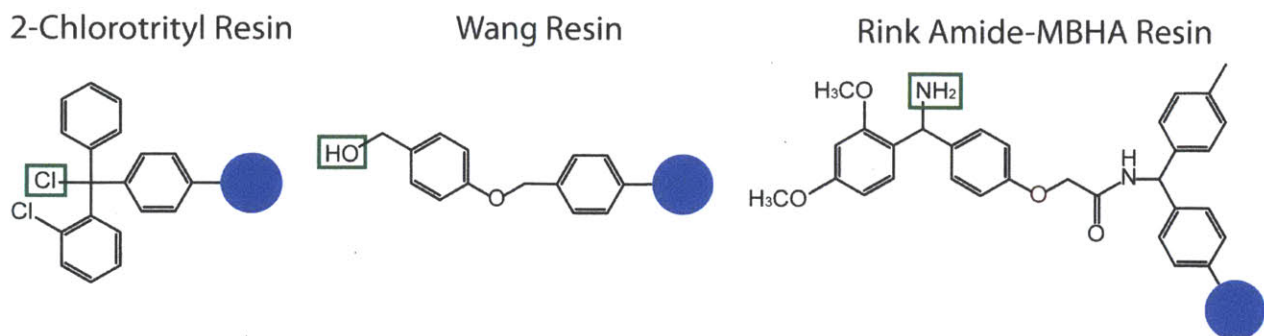


Figure 3. Common linkers used for SPPS. Blue circle – solid polymer support; Green square – peptide attachment point.

The hydrophobic cross-linked polystyrene was used for the synthesis of all peptides presented in this work. All resins had a 1% divinylbenzene (DVB) cross-linking which is a common choice for SPPS because it provides the resin with mechanical stability while still allowing for resin swelling which is required for fast chemical diffusion. All resins had a size distribution of 100-200 mesh (75-150 μm) where both the resins size and size distribution effects an SPPS processes yield. The absolute size of a resin particle determines the surface area to volume ratio of the resin which in turn determines its swelling property's whereas the size distribution determines the degree of uniformity of reaction conditions in the ensemble and in turn the degree to which the reaction can be optimized. The linkers used in this study were 4-(2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetamido-norleucyl-MBHA resin (Rink amide MBHA resin), 4-benzyloxbenzyl alcohol (Wang resin)¹¹, and 2-chlorotrityl chloride (Barlos resin)¹², which were purchased from Novabiochem (Figure 3). Rink amide MBHA resin was used for the synthesis of the TZ2 peptide because upon cleavage it generates the desired C-terminal amide group $-\text{CONH}_2$. In addition the rink amide MBHA resin does not require a separate synthetic step for the coupling of the first amino acid to the resins linker¹³, termed preloading. Wang resin was used preloaded with glycine for the synthesis of elastin-like peptides in order to generate the desired C-terminal carboxylic acid group. Finally 2-chlorotrityl

chloride resin was used for the synthesis of the ELP polymer precursor Boc-GVGVP for two reasons 1) the steric bulk of the 2-chlorotrityl linker prevents the creation of a diketopiperazines^{12,14,15} which commonly occurs for Pro-Wang resin and 2) the 2-chlorotrityl linker can be cleaved under mild conditions^{12,14,15} (see section on peptide cleavage) allowing for the removal of the peptide from the resin while still retaining the desired N-terminal Boc protecting group.

Since the 2-chlorotrityl resin was purchased without pre-loading an additional step was necessary to attach the first amino acid Fmoc-Pro-OH to the resin. This was accomplished by dissolving 1.2 equivalents of Fmoc-Pro-OH relative to the mols of linker, which is expressed as (mmol of linker)/(grams polymer) which gives the linker substitution, and 4 equivalents of *N,N*-Diisopropylethylamine (DIPEA) in dry DCM. This solution was then added to the resin and stirred for 120 min. Next the resin was washed with 3x DCM/MeOH/DIPEA (17:2:1), 3x DCM, 2x DMF, 2x DCM and dried under vacuum.

After resin selection 0.1 mmol of resin was weighed out for the peptide synthesis where the scale of the synthesis is limited by the size of the reaction vessel and the degree of swelling of the resin. For this work Rink amide MBHA resin was used at a linker substitution of 0.69 mmol/g, Fmoc-Gly-Wang resin was used at a linker substitution of 0.78 mmol/g, and 2-chlorotrityl chloride resin was used with a linker substitution of 1.20 mmol/g. It should be noted that the substitution levels used in this work are skewed towards higher levels of substitution where the typical range is from 0.5-0.8mmol/g. A higher degree of resin substitution was chosen because of the small size of the peptides being synthesized. For the synthesis of difficult sequences or sequences longer than 25 amino acids a linker substitution of 0.1-0.2mmol/g is more appropriate.¹⁶

2.2.2 Deprotection

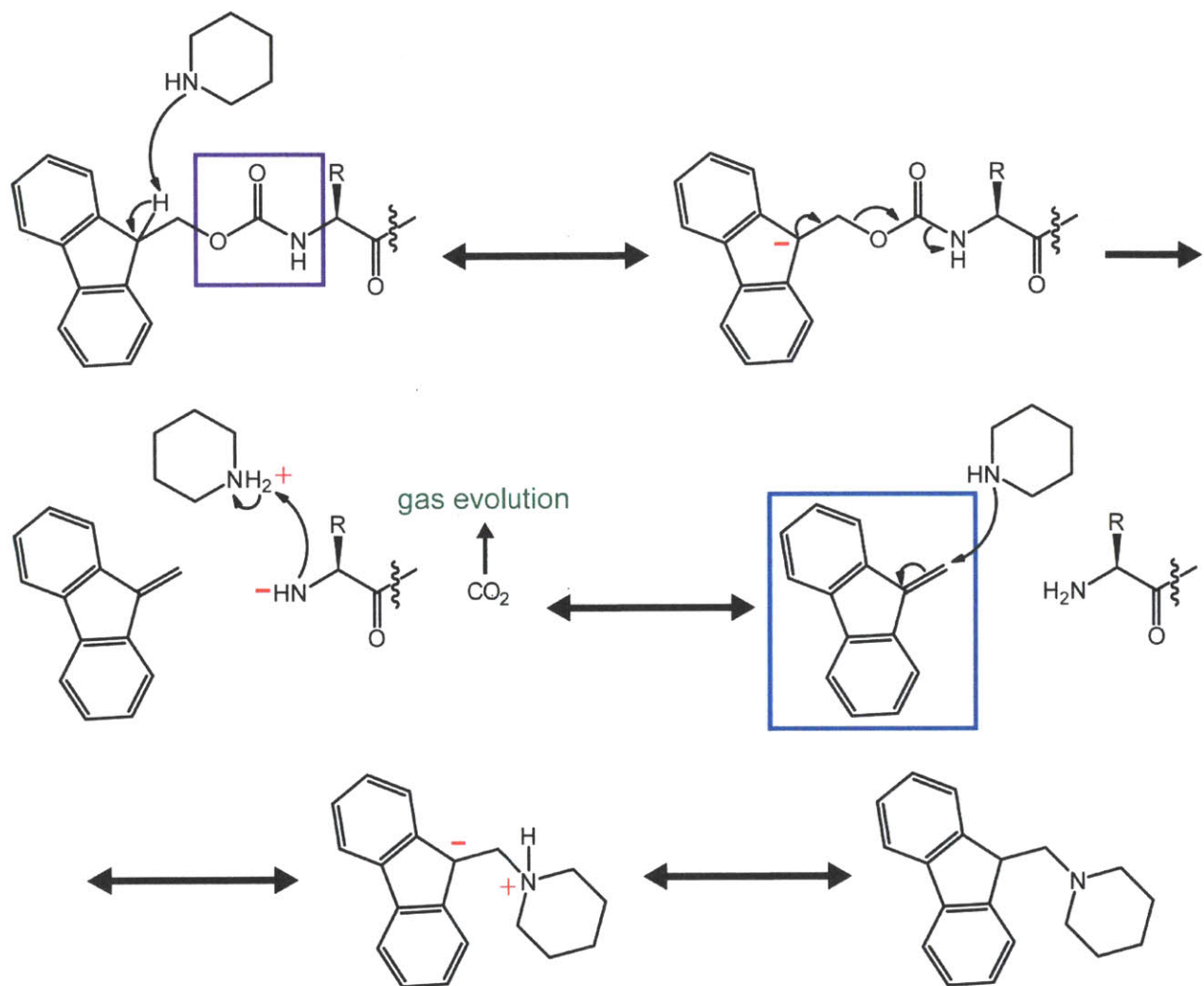


Figure 4. Mechanism of Fmoc deprotection with piperidine. Purple – carbamate; Blue – Dibenzofulvene.

The first automated step in the SPPS procedure is the cleavage of the Fmoc protecting group from the resin in order to unblock amine sites for the addition of the first amino acid. To accomplish this, the resin is suspended in dry DMF for 40 min to allow the resin to swell before application of reagents. This step and all subsequent steps up to peptide cleavage were

performed under dry N_2 to avoid reductions in yield due to hydrolysis. Next the resin was suspended 2x in a 20% piperidine in DMF solution for 7 min. First the piperidine removes the Fmoc group via base-induced β -elimination followed by evolution of the carbamate linkage as CO_2 gas (Figure 4). Finally the piperidine scavenges the highly reactive dibenzofulvene byproduct to form a stable adduct^{5,6}. The resin was then cleaned by rinsing with 6x with DMF for 30 sec to remove piperidine and the dibenzofulvene/piperidine adduct leaving behind the unprotected product.

2.2.3 Activation and Coupling

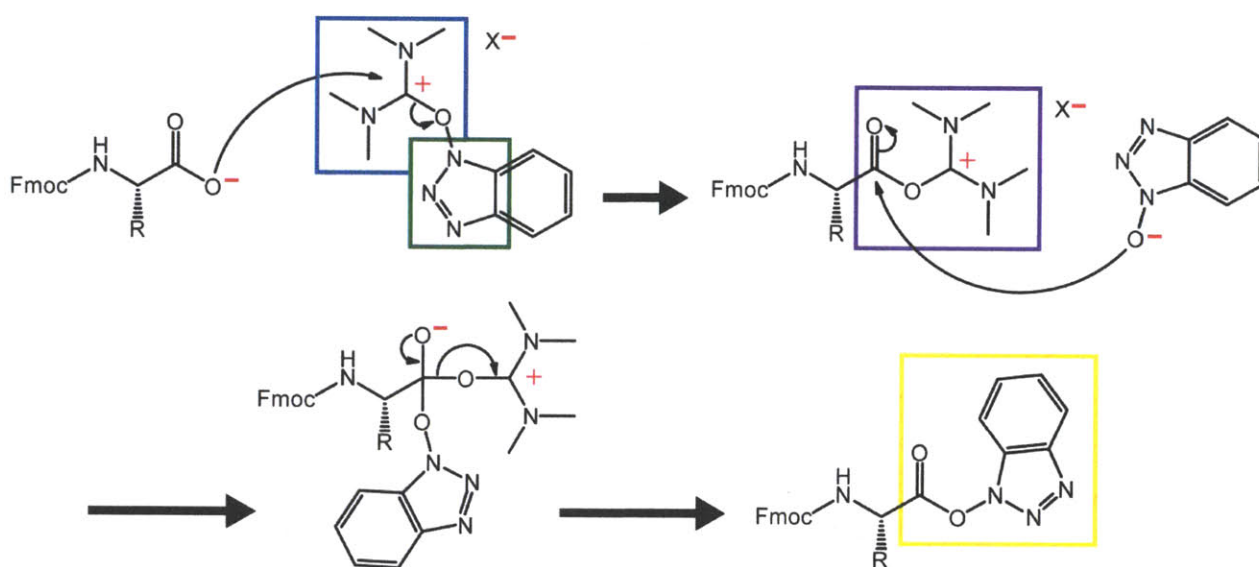


Figure 5. Reaction mechanism for the activation of an amino acid by the reaction of HBTU with an unprotected carboxylic acid to form a HOBt ester. Blue – Urea; Green – Triazolol; Purple – O-Acylisourea; Yellow – HOBt ester.

To perform the first coupling step the carboxylic acid of the Fmoc amino acid being coupled to the peptide is activated using a triazolol based salt as a coupling reagent before being added to the growing peptide. To accomplish this a solution of 0.4 M *N*-methylmorpholine in DMF was added to a vial containing a dry powdered mixture of Fmoc amino acid and coupling

reagent in a 4x excess relative to the number of linkers on the resin. This solution was mixed for 30 sec before being injected into the reaction vessel to create a suspension of resin and activated amino acid solution. The mechanism for the formation of the activated Fmoc amino acid using the coupling reagent O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU) is shown in Figure 5. In the first step nucleophilic attack of the Fmoc amino acids carboxyl group oxygen on the central carbon of HBTU's urea moiety generates the O-acylisourea intermediate. The urea moiety is then replaced by hydroxybenzotriazole (HOBT) to form the more stable HOBT ester which is the activated form of the Fmoc Amino acid. This activation step dramatically increases the rate of acylation though caution should be used before using a large excess of coupling agent to prevent overactivation. Overactivation occurs when the acylating agent is strong enough that it is no longer selective and acylates both the primary amine on the backbone and weaker nucleophiles such as hydroxyl groups. Next the suspension was then mixed for 40 min at room temperature for standard coupling steps and 1.5 hr for difficult coupling steps or for the incorporation of isotope labeled amino acids to allow for a higher coupling yield (Figure 6). In this step the deprotected amine on the N-terminus of the growing

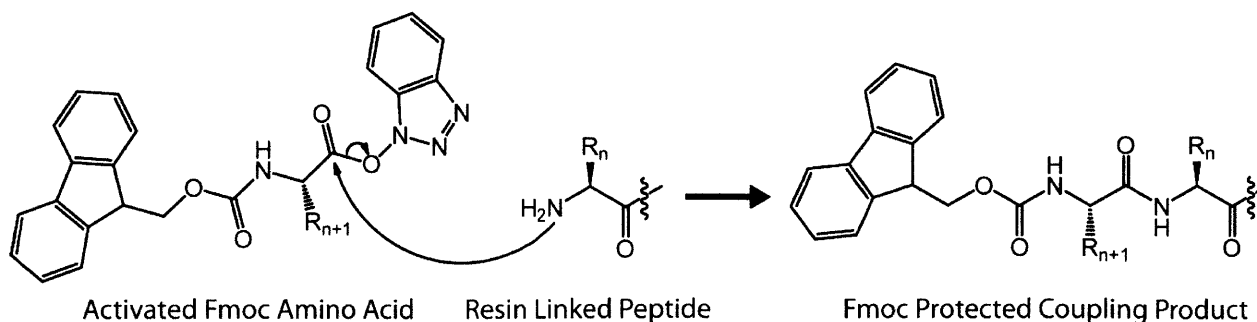


Figure 6. Mechanism for the acylation of an amine by an activated Fmoc amino acid.

resin linked peptide bonds to the carbon of the Fmoc amino acids HOBt ester linkage generating the desired amide bond. Note that the resulting product of this reaction is an Fmoc protected resin linked peptide thereby preventing further addition of the activated amino acid and in turn polymerization. After the reaction was complete the activated amino acid solution was removed by rinsing the resin in DMF 3x for 30 sec. This cycle of resin deprotection, Fmoc amino acid activation, and coupling is then repeated for the addition of each amino acid in the peptide sequence. Though since the resin has already been exposed to DMF the 40 min resin swelling step used for the first deprotection can be replaced with a DMF 3x for 30 sec resin rinsing for subsequent deprotections.

2.2.4 Coupling Reagents

Many activating agents have been developed for the coupling of amino acids including azide, carbodiimide, and triazolol based reagents.^{7,17} For these reactions racemization is a frequently encountered problem. There are three pathways for the racemization of an amino acid: 1) β -elimination which occurs in amino acids with β -side-chains such as cysteine (Cys or C), lysine (Lys or K), Phe, Ser, and Thr, 2) elimination of the amino acids α -hydrogen, and 3) generation of an azlactone intermediate (Figure 7).¹⁸⁻²⁰ For this reason the triazolol based salt activating agents 2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium-hexafluorophosphate methanaminium (HATU), HBTU, and 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) (Figure 8) were chosen for their high yield and lower rate of racemization relative to older carbodiimide activating agents such as

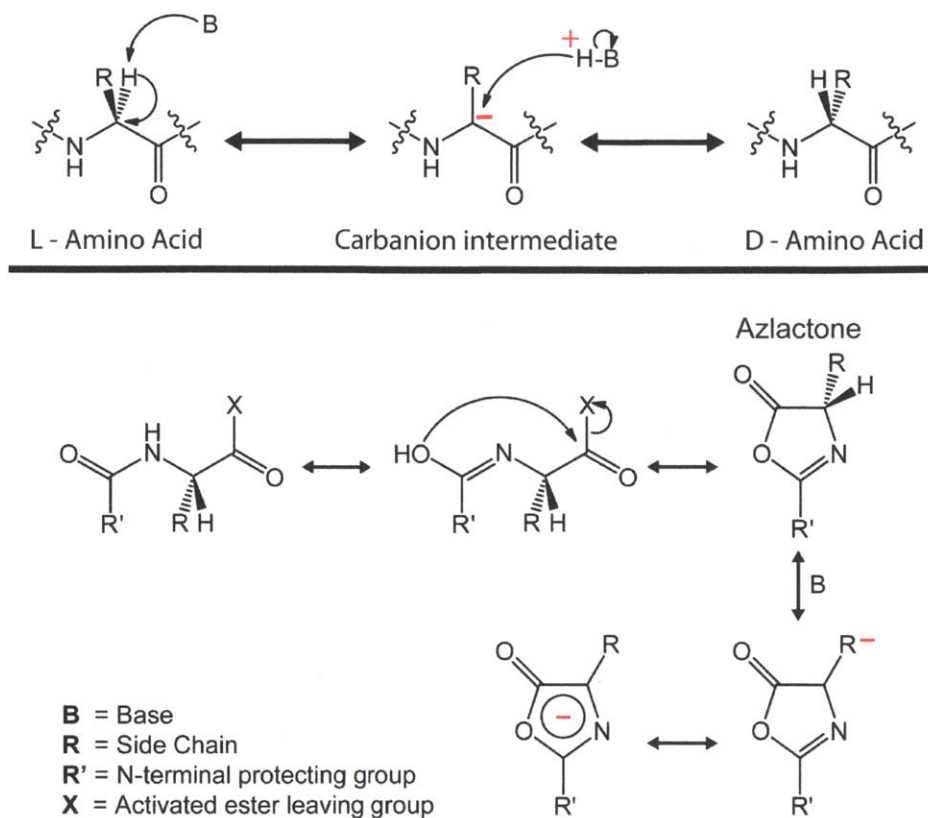


Figure 7. Potential pathways for the racemization of an amino acid under basic conditions including the deprotonation of the alpha carbon (top) and the formation of a cyclic azlactone intermediate (bottom). The rate of both mechanisms depends on the acidity of the amino acids α -carbon which varies as a function of the electron-withdrawing effects of the side-chain and N-terminal protecting group.

dicyclohexylcarbodiimide (DCC) and diisopropylcarbodiimide (DIC). Of these coupling reagents HATU²¹⁻²³ provides the fastest and most efficient coupling with the least racemization followed by HCTU²⁴ and then HBTU^{22,23}. Though due to the cost of HATU and HCTU there use was reserved for difficult coupling steps and for the incorporation of isotope labeled amino acids. Unfortunately HBTU has become difficult to obtain due to the United Nations reclassifying of HBTU's byproduct HOBT as a class 1c explosive.²⁵ In addition since it is the triazole moiety that is explosive both HATU and HCTU are also potential explosive hazards.

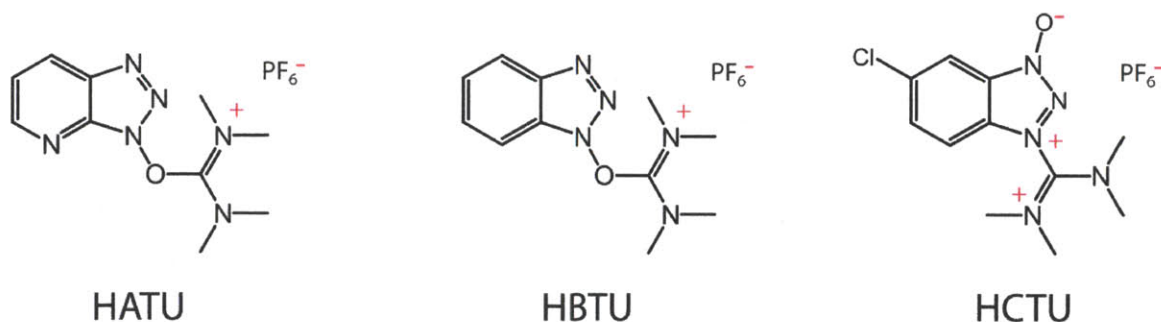


Figure 8. High yield coupling agents for the activation of carboxylic acids. Historically HATU and HBTU were believed to exist as an uronium salt (O form), as depicted above for HATU and HBTU, though the current consensus is that their true structure is a guanidinium salt (N form), depicted for HCTU.²⁶ This finding was significant because the uronium isomer is a more effective coupling agent than the guanidinium isomer. These isomers are discussed here for future reference since these coupling agents are described as both uronium or guanidinium salts in the references provided for this chapter.

For future work it will be necessary to replace these coupling agents with other chemicals that offer comparable reactivity and low rates of racemization. One potential family of compounds for this purpose is ethyl 2-cyano-2-(hydroxyimino) acetate (Oxyma) based coupling agents.²⁷

2.3 Side-chain protection

Choosing an appropriate side-chain protecting scheme is arguably the most significant component to a successful SPPS synthesis. This step is significant because side-chain protecting groups prevent unwanted side reactions that diminishes the synthesis's yield, the purity of the crude product, and may even render a synthesis intractable.^{8,28} Of the amino acids used in this work Asparagine (Asn or N), Glu, Lys, Ser, Thr, and Trp required side-chain protection. A set of commonly used functionalities for protecting amino acids side-chains are presented in Figure 9. In the proceeding discussion is a description of the protection scheme used for each amino acid and a brief discussion of the considerations that need to be taken in to account when incorporating these amino acids into a peptide.

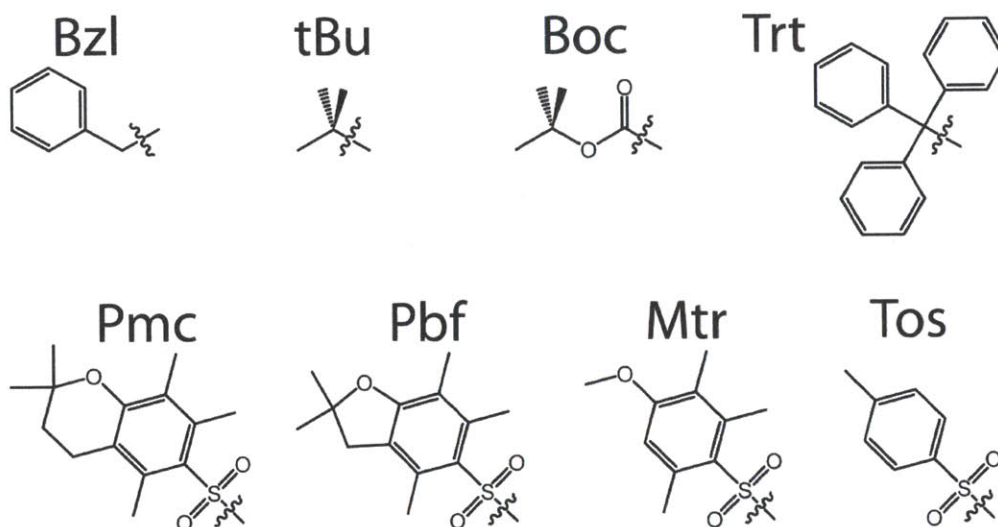


Figure 9. Commonly used side-chain protecting groups for Fmoc SPPS. Benzyl (Bzl) protects carboxylic acids, alcohols, thiols, and imidazoles. Tert-Butyl (tBu) protects carboxylic acids, alcohols, and thiols. Benzyloxy-carbonyl (Boc) protects amines. Triphenylmethyl (trt) protects carboxamides, imidazoles, and thiols. 2,2,5,7,8-pentamethyl-chromane-6-sulfonyl (Pmc), **2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl** (Pbf), 4-methoxy-2,3,6-trimethylbenzene-sulfonyl (Mtr), and 4-toluenesulfonyl (Tos) protect guanidinium.

2.3.1 Asparagine

Asn is known to undergo numerous side reactions the most common of which is the dehydration reaction of the Asn β -carboxamide functional group with a carboxyl group to form β -cyano-alanine (Figure 10).²⁹ This side reaction can be avoided by coupling using an equimolar quantity of HBTU and HOBT in the absence of a side-chain protecting group³⁰ or by

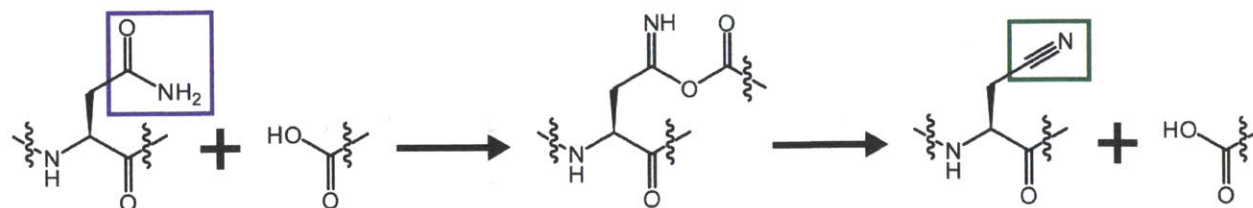


Figure 10. Mechanism of asparagine dehydration to create a cyano group. Purple – carboxamide; Green – cyano.

a second method, which is used in the present work, of protecting the carboxamide functional group with an triphenylmethyl (Trt) group.

2.3.2 Glutamic Acid

Glu and Asp both contain a carboxylic acid side-chain and as a result require similar protecting schemes. In both cases a cyclization followed by reopening of the ring generates two undesired modified peptide species. As a result the tert-Butyl group is commonly used to protect

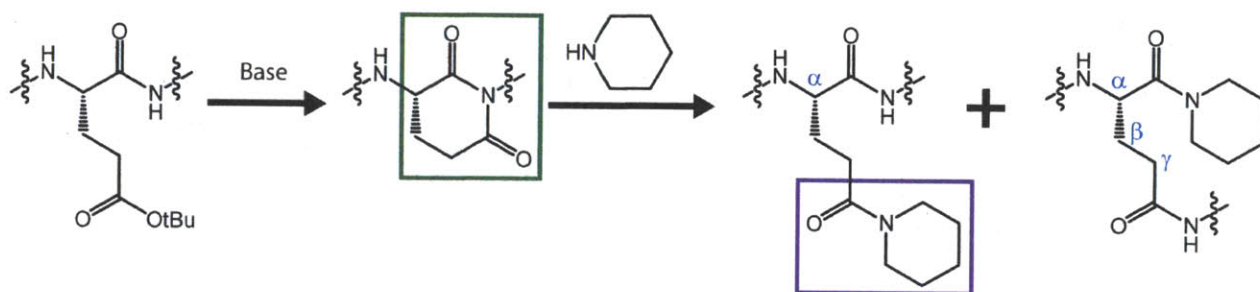


Figure 11. Cyclization and ring opening reaction of a Glu side-chain in the presence of piperidine to generate α -Glu peptide with a γ -piperidide side-chain (Left product) and a γ -Glu peptide with an α -piperidide side-chain (Right product). Green – piperidinedione; Purple – piperidide.

the Glu and Asp carboxyl group. Though the work of Fields and coworkers,³¹ has shown that Asp cyclization was especially common for the sequence Asp(tBu)-Asn(Trt). This is similar to the TZ2 coupling sequence used for this work Glu(tBu)-Asn(Trt), and although the formation of the 5 member ring in Asp is more probable than the formation of the 6 membered ring in Glu this side reaction should still be considered when attempting to optimize the yield for the synthesis of isotope labeled TZ2 peptides. This reaction is shown in Figure 11 where Glu cyclizes to form a piperidinediones followed by a ring opening reaction with piperidide to form both an α -Glu peptide with a γ -piperidide side-chain and a γ -Glu peptide with an α -piperidide side-chain. In the case of Asp this side reaction was minimized by the addition of 2,4-dinitrophenol or HOBt to

the piperidine/DMF reaction solution or using a N-2-hydroxy-4-methoxybenzyl (HbM) side-chain protecting group on the Asp side-chain.³² As a result similar strategies should be explored for the coupling of the Glu residue in the synthesis of cost prohibitive isotope labeled TZ2 peptides.

2.3.3 Tryptophan

Since the amine of the Trp indole side-chain is a secondary amine it typically does not interfere with the coupling of primary amines on the peptide back bone. Never the less the indole ring presents pathways for alkylation, oxidation, dimerization and aggregation. As a result great care should be placed in choosing an appropriate protecting scheme for this residue that takes into account the neighboring amino acids, composition of the resin, and potential radical species generated upon cleavage of the peptide from the resin. In the case of TZ2 the Fmoc-Trp(Boc)-OH was used for coupling Trp residues though for the synthesis of longer sequences a more careful choice of protecting scheme will be required.

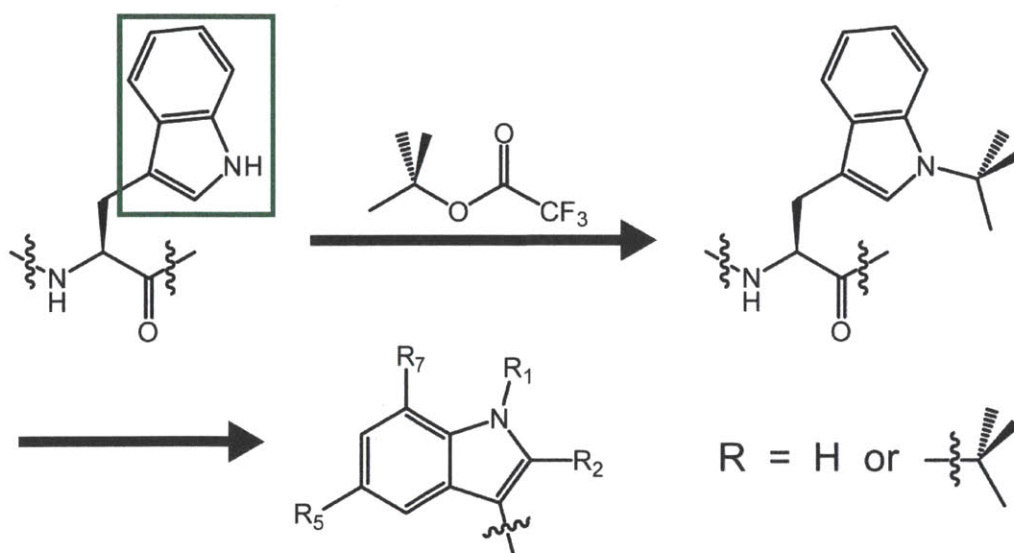


Figure 12. Alkylation of a Trp side-chain followed by migration of the alkyl group. Green – indole.

One potential cause for side-reactions in the synthesis of TZ2 is that the nucleophilic Trp side-chain captures reactive electrophilic species generated during cleavage with TFA.³³⁻³⁷ An example for this class of side reactions is demonstrated for the formation of a tertiary amine, a commonly encountered side-reaction for Trp, by addition of a tert-butyl group to the indole nitrogen (Figure 12). Here the nitrogen of the indole interacts with tert-butyl-trifluoroacetate, a strong alkylating agent that is a byproduct in the TFA cleavage of Boc, to form the a Nⁱⁿ-tert-butyl Trp derivative. This alkyl group can then migrate from the nitrogen to the 2, 5, or 7 position of the indole ring followed by further alkylation to generate a multitude of undesired products.³⁷ To avoid chemical modification of Trp what is required is a protecting group that can draw π -electron density out of the indole ring to prevent electrophilic attack and the use of scavengers to eliminate the reactive byproducts of cleavage (see the section on cleavage).

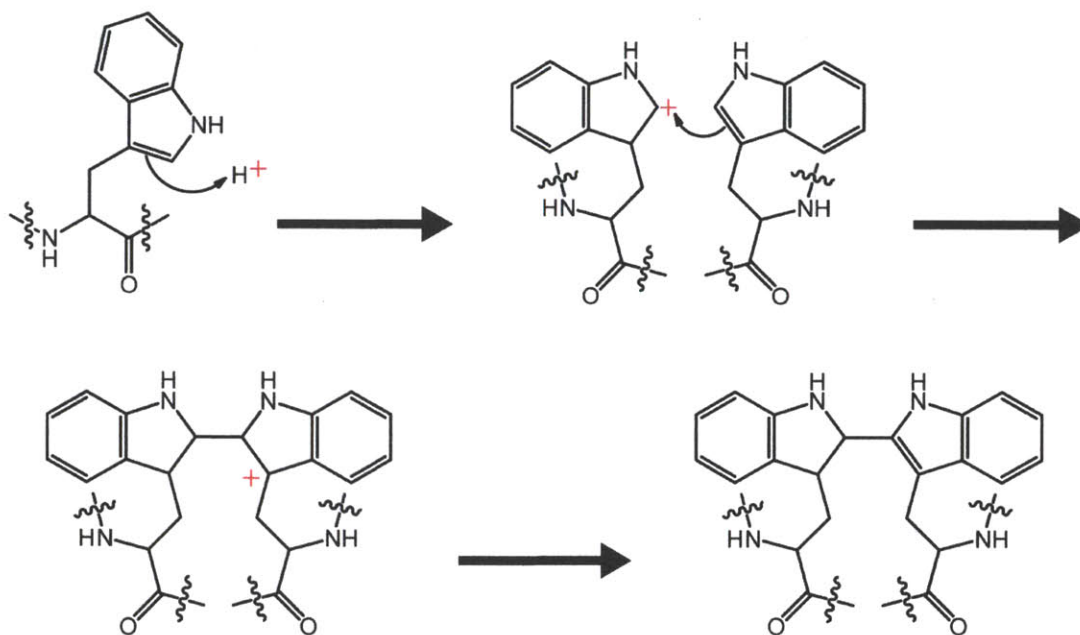


Figure 13. Acid induced dimerization of Trp side-chains.

Since the hydrophobic core of TZ2 is comprised of four tightly packed Trp side-chains another alkylation reaction worth noting is Trp dimerization (Figure. 13). Dimerization can

occur from the acidic conditions in the cleavage cocktail that will be discussed in the next section. Here the 3 carbon of the indole ring can be protonated generating a positively charged carbon at the 2 position that undergoes electrophilic attack on a neighboring Trp side-chain. The resulting charged dimerized species then undergoes an elimination reaction to form the stable dimer product. As a result minimizing the exposure of the unprotected indole rings in TZ2 to the harsh acidic conditions found during cleavage will improve the synthetic yield.

Another concern for synthesizing the TZ2 peptide is the oxidation of its indole rings. In acidic media the indole ring can be oxidized to generate a pink, purple, or yellow solid depending on the oxidation. Since the TZ2 peptide has a tendency to form a yellow solid care should be taken in the future to minimize potential oxidation by performing acid cleavage (see section on cleavage cocktails) under an inert atmosphere and the peptide should be stored cooled after being lyophilized from neutral solvent.

Finally the Trp indole ring is hydrophobic and as a result the TZ2 peptide which contains a quarter Trp residues is likely to aggregate during cleavage and HPLC purification. To prevent this aggregation as well as electrophilic addition a two-step side-chain protection scheme is required (Figure 14). Depicted in this figure is an example of a doubly protected Trp side-chain based on the work of Undén and co-workers.³⁸ In the first step the N-terminus of the Nmbu protecting group is protonated by TFA which then drives the removal of the acid labile Boc group. This generates CO₂ gas, a tert-butyl cation, and a cationic Nmbu protected Trp. This charged Nmbu both prevents aggregation by increasing the residues solubility and prevents electrophilic attack by the tert-butyl cation and other reactive species generated during cleavage. After HPLC purification the peptide is dissolved in a basic solution to remove the Nmbu group

via an intramolecular cyclization reaction to generate N-methylpyrrolidone and the desired un-protected Trp side-chain.

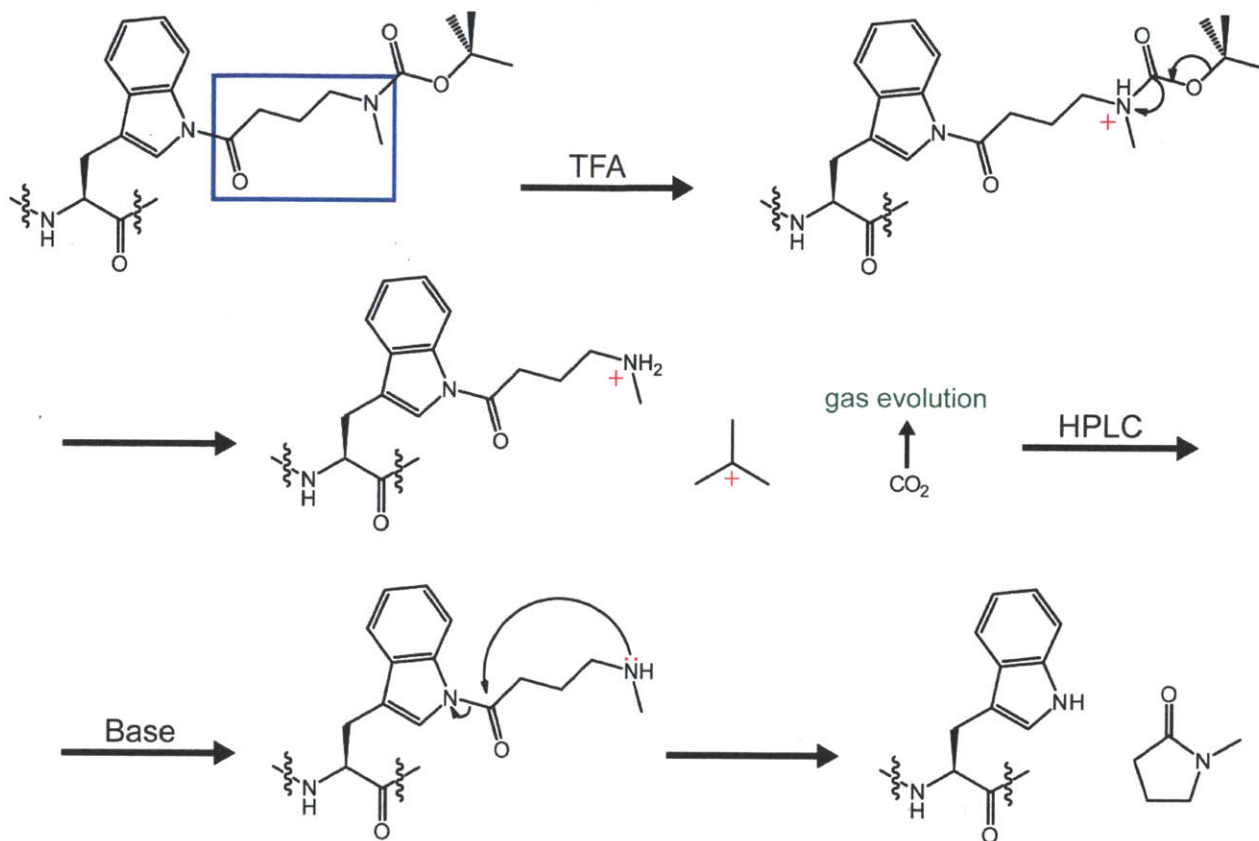


Figure 14. Two stage orthogonal deprotection of Fmoc-Trp(Boc-Nmbu)-OH. Blue – Nmbu.

2.3.4 Lysine

Of the proteinogenic amino acids only Lys has a primary amine as a side-chain functional group. As a result an unprotected Lys side-chain will compete with the primary amine on the peptide backbone during the coupling of activated amino acids (Figure 15). In this reaction

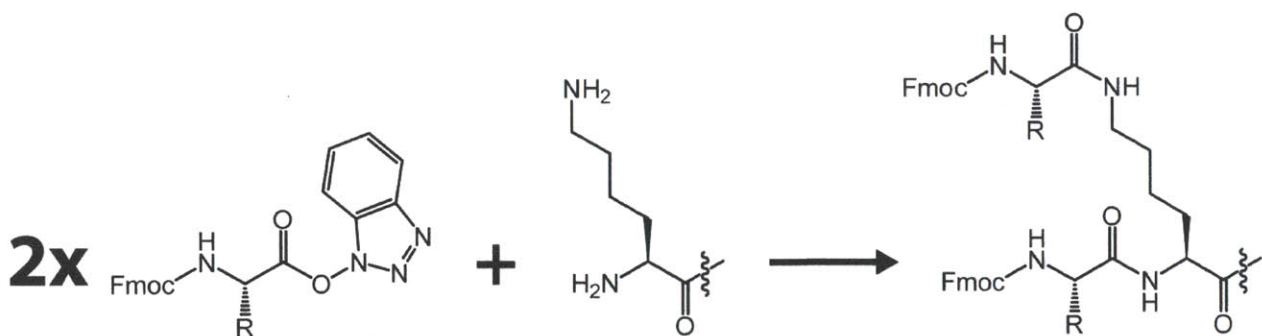


Figure 15. Acetylation of an unprotected Lys to generate a branched product.

an activated amino acid with side-chain and Fmoc protecting groups couples to both the main-chain and unprotected Lys side-chain of the tethered peptide generating a branched product that grows along both chains with each additional coupling step. An undesirable side reaction is also generated in the reverse case where a Fmoc-Lys-OH without a side-chain protecting group is

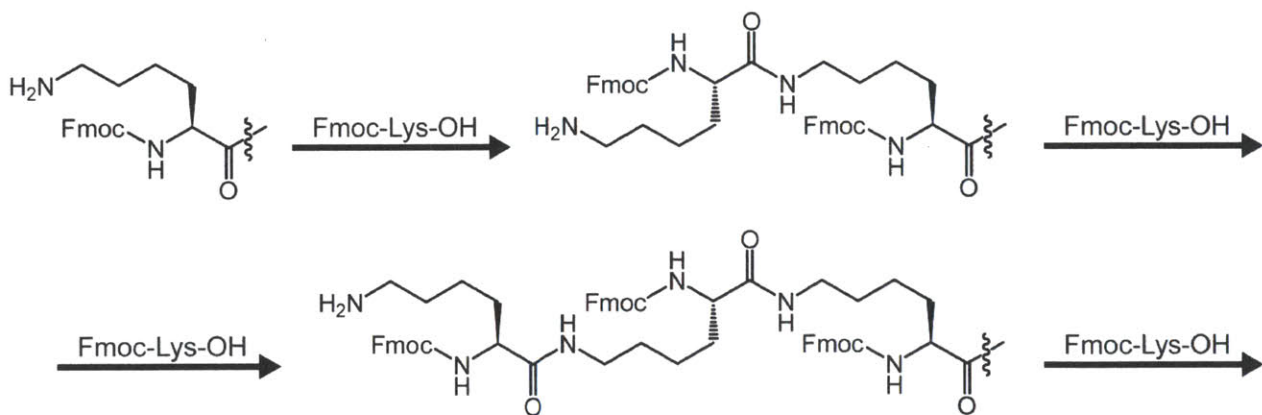


Figure 16. Polymerization of Lys as a result of coupling in the absence of a side-chain protecting group. To avoid these side-reactions in the synthesis of TZ2 the primary amine of Lys was protected with Boc using Fmoc-Lys(Boc)-OH as the incorporated Lys precursor.

coupled to a peptide. Here the activated Lys continues to be coupled to the resin bound peptide because of the availability of the unprotected primary amine on the Lys side-chain generating a polymerized product (Figure 16).

2.3.5 Serine and Threonine

Ser and Thr side-chains were protected with either a *tert*-Butyl (tBu) group for unlabeled amino acids or a silyl group for isotope labeled amino acids. A side-chain protecting group is employed to prevent acylation of the side-chain which happens readily for the primary alcohol of Ser and to a lesser extent to the secondary alcohol of Thr during Fmoc SPPS. Thr is found to be less reactive because its hydroxyl oxygen is sterically hindered and its nucleophilicity is small relative to the backbone amino group. On the other hand the hydroxyl group of Ser is not sterically hindered and as a result can compete with the backbone amine for the coupling of amino acids under strong acylating conditions like those found in Fmoc SPPS (Figure 17). Here

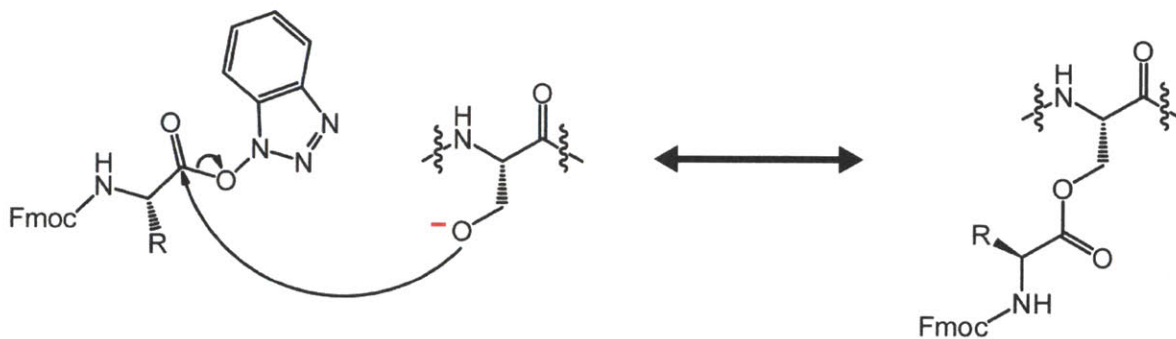
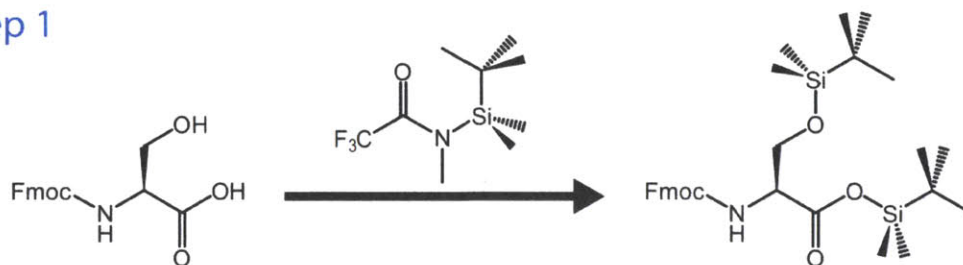


Figure 17. Acetylation of the Ser hydroxyl group under basic conditions.

the deprotonated oxygen of Ser attacks the central carbon of the ester to form a branched peptide through the formation of an *o*-acyl Ser linkage. As a result of this reactivity the isotope labeled Ser and Thr that were prepared as $-C^{18}O^{18}OH$ labels both required an additional synthetic step to protect their hydroxyl side-chain. For these residues a side-chain protecting scheme that would not act as a source of ^{16}O contamination was required. To do this these side-chains were silylated using *N*-methyl-*N*-(*tert*-butyldimethylsilyl)-trifluoroacetamide (MTBSTFA) (Pierce

Protein Research) following the procedure of Madson et al.³⁹ In the case of Ser, 0.139 g of ¹⁸O labeled Fmoc-Ser-OH was dissolved in 15 mL of dry dimethylformamide (DMF) or acetonitrile and the solution was degased using N₂. Next 3 mL of MTBSTFA was injected into the reaction vessel and the reaction was allowed to run for 10 min and then was immediately frozen using liquid N₂ and lyophilized. This step generated a Ser derivative containing both a silylated side-chain hydroxyl group and the carboxylic acid (Figure 18). The dried product was then

Step 1



Step 2

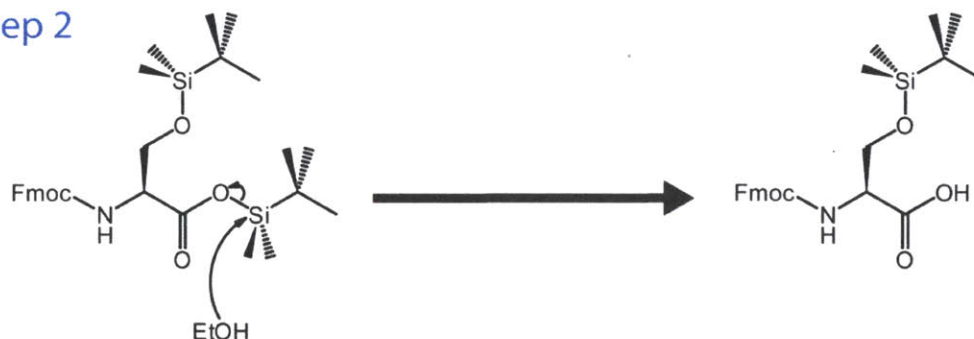


Figure 18. Silylation of the hydroxyl and carboxylic acid group of Ser followed by the removal of the silyl group from the carboxylic acid. Here the EtOH selectively removes the silyl group from the carboxylic acid due to its increased susceptibility of the silyl ester to nucleophilic attack over the silyl ether.

suspended in hexane (20 mL) and a 1.1 mol excess of dry ethanol was added to selectively desilylate the carboxylic acid with minimal cleaving of the silyl protecting group. The solution was again lyophilized and the resulting solid was subsequently purified via recrystallization in DMF/hexane to generate the final product (see mass-spec in appendix). Finally the silylated product was then used immediately for peptide synthesis due to the fact that it on average is only stable for 6 months at room temperature.³⁹ It should be noted that although this side-chain

protecting procedure was used for Ser and Thr it is generally applicable to the protection of primary and secondary amines, primary and secondary alcohols, and primary thiols. In addition since the silyl group is labile under mildly acidic conditions it can be cleaved selectively when used in the presence of the less acid sensitive tBu, Boc, and Trt. As a result this protection scheme is useful for the generation of cyclic peptides and the synthesis of larger protein constructs through a convergent peptide synthesis approach.

2.4 Cleavage, Side-Chain Deprotection and Purification

The yield from a Fmoc SPPS synthesis can be significantly reduced by side reactions and peptide aggregation which can occur during resin cleavage and side-chain deprotection.⁴⁰ Both side-chain protecting groups and resin linkers can form reactive radicals and cationic species that can attack nucleophilic side chains like those of Cys, methionine (Met or M), tyrosine (Tyr or Y), and Trp. As a result it is crucial to choose appropriate scavengers and solvent conditions to optimize yield. Given the variability in the behavior of different peptide and resin combinations it is best to first test small portions of the peptide bound resin to find the ideal cleavage cocktail and reaction time.

Before cleavage the resins used for this work were transferred to a filter-frit reaction vessel and washed with 3x cycles of the solvent series DCM to swell the resin followed by DMF to clean the resin, and diethyl ether to remove the DMF. Since DMF does not evaporate easily this final diethyl ether step is required in order to prevent the acid hydrolysis of residual DMF into formic acid and dimethylamine during cleavage. Once the ether had fully evaporated the resin was suspended in 4 mls of a cleavage cocktail and was occasionally stirred for 2-3 hrs.

In the case of Wang and rink amide resin a cleavage cocktail of 95% TFA and 5% triisopropylsilane (TIPS) was used. For both of these resins this cleavage cocktail simultaneously cleaves the peptide and removes the side-chain protecting groups. TIPS was used as a scavenger to prevent side reactions caused by species generated during the removal of side-chain protecting groups and from the resin linkers. Other commonly used scavengers are water, 1,2-ethanedithiol (EDT), and triethylsilane (TES) where here the different scavengers are chosen based on the scavengers efficacy at removing the reactive species generated in a particular peptide synthesis.

In contrast the 2-chlorotrityl resin was removed under more mild conditions using acetic acid/2,2,2-trifluoroethanol (TFE)/DCM (1:8:8, v/v/v) which only cleaves the peptide leaving any protecting groups like tBu, Boc, and Trt attached. In cases where the addition of acetic acid is undesirable hexafluoroisopropanol (HFIP)/DCM (1:4, v/v) was used as a replacement. Since the *2-chlorotrityl* resin allows for selective cleavage and deprotection it was used to synthesize the protected polymer precursor Boc-GVGVP-OH described by Urry and co-workers.⁴¹ It should be noted that the selective cleavage offered by 2-chlorotrityl resin is also useful for generating precursors for peptide cyclization or convergent synthesis.

After cleavage the cleavage cocktail was removed and the resin was rinsed with another 1ml of cleavage cocktail. This solution was then cooled in an ice water bath before adding 250mls of cold diethyl ether to cause the peptide to precipitate out of solution. This peptide suspension was then centrifuged at >5 kHz at 0°C to remove the supernatant. Next the recovered crude peptide was dissolved in dilute HCl solution and lyophilized 3x to remove residual TFA.^{42,43} In cases where this approach did not yield a precipitate the ether was removed in a rotary evaporator followed by dissolving the peptide in DCM and removing the solvent repeatedly in a

rotary evaporator until the smell of TFA was removed. Finally all peptides were purified using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.1% TFA in H₂O and (buffer B) 80% acetonitrile, 0.085% TFA in H₂O. The product was then dried and lyophilized 3x in HCl to again remove TFA followed by lyophilization 2x from D₂O (Cambridge Isotope *Laboratories*) to remove H₂O and acidic protons which is required for subsequent IR studies.

2.5 Preventing amino acid deletion

Amino acid deletion occurs when an amino acid during synthesis fails to be incorporated into the peptide resulting in the creation of a peptide that is missing amino acids relative to the desired sequence. Amino acid deletion is problematic because these deletions lower the yield of the desired peptide and these deleted peptides can be difficult to separate from the desired peptide using HPLC. As a result monitoring for deletions is crucial for designing the synthesis of a new peptide.

The two common causes for deletions are incomplete deprotection and incomplete coupling which originates in both cases from lack of access to the peptides terminus due to steric hindrance, peptide folding, or peptide hydrophobicity. This risk should be taken into consideration when synthesizing TZ2 because of its propensity to form a β -hairpin thereby blocking amine terminus from being exposed to incoming reagents. This risk should also be considered when synthesizing peptide sequences longer than 30 amino acids as was the case for the synthesis of GVG(VPGVG)₆.

An effective method for detecting incomplete deprotection is by flowing the deprotection waste from the peptide synthesizer through a UV spectrometer or by separating the waste using an HPLC with UV detection. This allows for the quantitative determination of the

dibenzofulvene/piperidine adduct concentration, generated during Fmoc deprotection, by measuring its absorption at 301 nm. In the event of low concentrations of the dibenzofulvene/piperidine adduct the reaction time for deprotection can be increased to allow the reaction to go to completion. Though because of the risk of alkylation during deprotection it is best to initially deprotect the N-terminus using a short reaction time followed by rinsing the resin to remove reactive species before starting a second deprotecting reaction. The process of optimizing the deprotection time also provides insight into the ideal reaction conditions for the subsequent coupling reaction since the same issues of steric hindrance, peptide folding, or peptide hydrophobicity that effect deprotection can also effect coupling. As a result in the event that a slow deprotection reaction is detected one should then determine the efficacy of the subsequent coupling reaction.

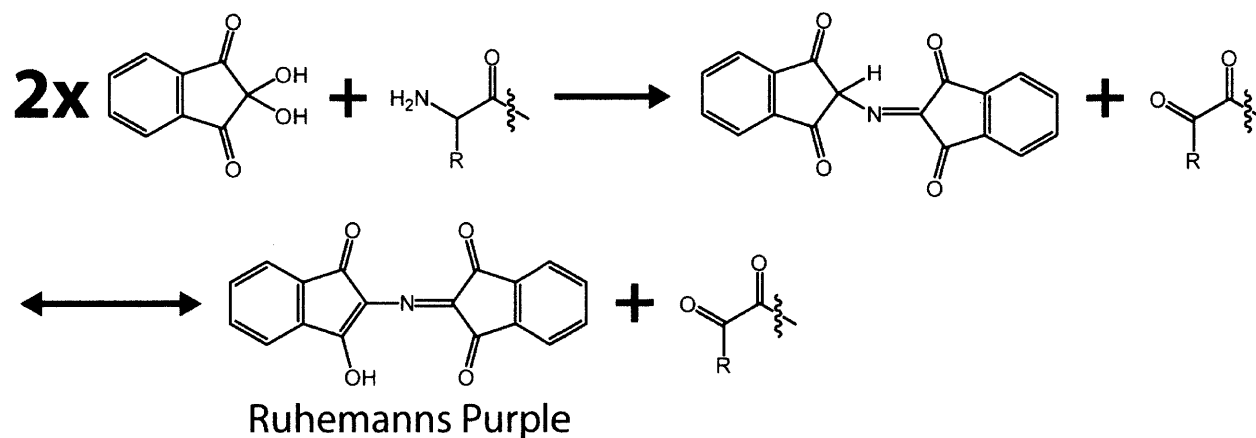


Figure 19. Creation of Ruhemanns purple by addition of ninhydrin to compounds containing exposed amines.

An easy method for both qualitatively and quantitatively identifying the extent of coupling is through the use of a ninhydrin test (also known as a Kaiser test).^{44,45} The ninhydrin test detects the presents of free primary amines through the abstraction of the amine nitrogen to generate the Ruhemanns purple chromophore (Figure 19). Since any successfully coupled site

will contain the Fmoc protecting group on the N-terminus of the newly coupled amino acid only unsuccessfully coupled sites will contain a free amine. Therefore by measuring the intensity of the purple/blue color it is possible to identify the extent of coupling. The ninhydrin test was performed using a ninhydrin test kit (sigma-aldrich) consisting of three solutions: 80% phenol in ethanol (w/v), 6% ninhydrin in ethanol (w/v), and 2% 0.001M aqueous KCN in pyridine (v/v). To perform the test first a small number of resin beads were removed from the reaction vessel and washed 3 times with dry ethanol. Next these beads were suspended in a mixture containing 3 drops of each solution and heated at 120°C for 5 min. After 5 min the solution was removed and examined by eye for a qualitative measure of coupling or alternatively it was examined with a UV-Vis spectrometer to quantitatively measure the extent of coupling. In the case of proline, a chloranil test is required to detect its secondary amine.⁴⁶ The chloranil test consists of a 2% chloranil in DMF solution and a 2% acetaldehyde in DMF solution which are mixed in equal parts to generate a blue/green color in the presents of free primary and secondary amines. Since poor solvent exposure and structuring are commonly the cause of difficult coupling steps, for future work it maybe informative to perform FTIR measurements on resin bound peptides to detect these potential causes for a failed coupling reaction.

Difficult coupling steps can be driven to completion by either extending the reaction time or by washing followed by the injection of a fresh solution of activated amino acid known as double coupling. Alternatively peptide chains with uncoupled N-termini can be capped with an unreactive functional group to prevent the growth of a deletion peptide. Capping was performed by washing the resin 3x with DMF for 30 sec followed by suspending the resin in 2 ml of acetic anhydride for 20 min to acetylate the peptide and finally again washing the resin 3x with DMF for 30 sec. This capping reaction typically proceeds smoothly because of the small size of the

acetylating agent allowing it to access hidden amino groups. Finally the capped peptides are removed during HPLC purification which has now been simplified by removing deletion peptides which can have similar elution times as the full sequence. Though it should be noted that capping should be reserved for difficult coupling because indiscriminant capping can lead to poor yields.

2.6 Acknowledgments

I would like to acknowledge the contributions of Jongjin Kim for his assistance in synthesizing the isotope labeled GVGn1 peptides, Sam Chunte Peng for his assistance in synthesizing GVGn3-6 peptides, Khaled A. Nasr and Dan Kemp for providing advice for designing the isotope labeling procedure and side-chain protecting scheme, and Dan W. Urry and Karshad Parry for helpful discussions on the synthesis of elastin-like peptides.

2.7 References

- (1) Robert C. Murphy, K. L. C. *Biological Mass Spectrometry* **1979**, 6, 309.
- (2) James Marecek, B. S., Scott Brewer, Jenifer Belyea, R. Brian Dyer, Daniel P. Raleigh *Org. Lett.* **2007**, 9, 4935.
- (3) Luis A. M. M. Barbosa, R. A. v. S. *J. Catal.* **2000**, 191, 200.
- (4) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, 85, 2149.
- (5) Louis A Carpino, G. Y. H. *J. Org. Chem.* **1972**, 37, 3404.
- (6) Carpino, L. A. *Acc. Chem. Res.* **1987**, 20, 401.
- (7) Albericio, F. *Curr. Opin. Chem. Biol.* **2004**, 8, 211.
- (8) Gregg B. Fields, R. L. N. *Int. J. Peptide Protein Res.* **1990**, 35, 161.

- (9) Bodanszky, M. *Peptide Chemistry: A Practical Textbook*; Springer-Verlag, 1988.
- (10) Virender K. Sarin, S. B. H. K., R. B. Merrifield *J. Am. Chem. Soc.* **1980**, *102*, 5463.
- (11) Wang, S.-S. *J. Am. Chem. Soc.* **1973**, *95*, 1328.
- (12) Kleomenis Barlos, O. C., Dimitrios Gatos, George Stavropoulos *Int. J. Peptide Protein Res.* **1991**, *37*, 513.
- (13) Michael S Bernatowicz, S. B. D., Hubert Köster *Tetrahedron Lett.* **1989**, *30*, 4645.
- (14) Kleomenis Barlos, D. G., Georgios Papaphotiu, Wolfram Schüfer *Liebigs Ann. Chem.* **1993**, *1993*, 215.
- (15) Kleomenis Barlos, D. G., John Kallitsis, Giorgos Papaphotiu, Petros Sotiriu, Yao Wenqing, Wolfram Schüfer *Tetrahedron Lett.* **1989**, *30*, 3943.
- (16) Muriel Amblard, J.-A. F., Jean Martinez, Gilles Subra *Mol. Biotech.* **2006**, *33*, 239.
- (17) Eric Valeur, M. B. *Chem. Soc. Rev.* **2009**, *38*, 606.
- (18) Murray Goodman, K. C. S. *J. Org. Chem.* **1962**, *27*, 3409.
- (19) Kemp, D. S. *Peptides* **1979**, *1*, 317.
- (20) Bodansky, M. *Principles of peptide synthesis*; Springer-Verlag, New York, 1984.
- (21) Louis A. Carpino, A. E.-F., Fernando Albericio *Tetrahedron Lett.* **1994**, *35*, 2279.
- (22) Fernando Albericio, J. M. B., Ayman El-Faham, Steven A. Kates *J. Org. Chem.* **1998**, *63*, 9678.
- (23) So-Yeop Han, Y.-A. K. *Tetrahedron* **2004**, *60*, 2447.
- (24) Oleg Marder, F. A. *Chim Oggi: Chem Today* **2003**, *21*, 35.

- (25) K.D. Wehrstedt, P. A. W., D. Heitkamp *J. Haz. Materials* **2005**, *A126*, 1.
- (26) Louis A. Carpino, H. I., Ayman El-Faham, Fernando J. Ferrer, Chongwu Zhang, Yunsub Lee, Bruce M. Foxman, Peter Henklein, Christiane Hanay, Clemens Mügge, Holger Wenschuh, Jana Klose, Michael Beyermann, Michael Bienert *Angew. Chem. Int. Ed.* **2002**, *41*, 441.
- (27) Ramon Subirós-Funosas, R. P., Rafael Barbas, Ayman El-Faham, Fernando Albericio *Chem. Euro. J.* **2009**, *15*, 9394.
- (28) George Barany, N. K.-C., Daniel G. Mullen *Int. J. Peptide Protein Res.* **1987**, *30*, 705.
- (29) H. Gausepohl, M. K., R. W. Frank *Int. J. Peptide Protein Res.* **1989**, *34*, 287.
- (30) Paolo Rovero, S. P., Fabio Bonelli, Antonio Triolo *Tetrahedron Lett.* **1993**, *34*, 2199.
- (31) Janelle L. Lauer, C. G. F., Gregg B. Fields *Lett. Pept. Sci.* **1994**, *1*, 197.
- (32) Packman, L. C. *Tetrahedron Lett.* **1995**, *36*, 7523.
- (33) Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 1637.
- (34) Hiroshi Ogawa, T. S., Hiroshi Irie, Haruaki Yajima *Chem. Pharm. Bull.* **1978**, *26*, 3144.
- (35) Cynthia G. Fields, G. B. F. *Tetrahedron Lett.* **1993**, *34*, 6661.
- (36) Jones, J. *The Chemical Synthesis of Peptides*; Oxford University Press, 1991; Vol. 23.
- (37) M. Löw, L. K., E. Jaeger, P. Thamm, S. Knof, E. Wönsch *Hoppe Seyler's Z. Physiol. Chem.* **1978**, *359*, 1637.
- (38) Karolina Wahlström, A. U. *Tetrahedron Lett.* **2009**, *50*, 2976.

- (39) Thomas P. Mawhinney, M. A. M. *J. Org. Chem.* **1982**, *47*, 3336.
- (40) David S. King, C. G. F., Gregg B. Fields *Int. J. Peptide Protein Res.* **1990**, *36*, 255.
- (41) K. Okamoto, D. W. U. *Biopolymers* **1976**, *15*, 2337.
- (42) Valery V. Andrushchenko, H. J. V., Elmar J. Prenner *J. Pept. Sci.* **2007**, *13*, 37.
- (43) Stéphane Roux, E. Z., Bernard Rousseau, Maïté Paternostre, Jean-Christophe Cintrat, Nicolas Fay *J. Pept. Sci.* **2008**, *14*, 354.
- (44) E. Kaiser, R. L. C., C. D. Bossinger, P. I. Cook *Anal. Biochem.* **1970**, *34*, 595.
- (45) *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Weng C. Chan, P. D. W., Ed.; Oxford University Press: Oxford, 2000.
- (46) Vojkovsky, T. *Pept. Res.* **1995**, *8*, 236.

Chapter 3

Melting of a β -Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations

3.1 Introduction & Background

This chapter describes the investigation of the structure of a β -hairpin peptide using isotope-edited 2D IR to site-specifically investigate the conformation of its backbone as a function of temperature. In this mode, the isotope-shifted amide I' band can be used to infer local structure and dynamics through variations in its vibrational frequency or couplings to other vibrations. The approach combines 2D IR spectroscopy on a series of peptide isotopologues with spectroscopic modeling of the peptide's conformers. This pair of tools is used to develop a picture of conformational heterogeneity across the thermal unfolding transition.

The system that is studied is a 12-residue hairpin peptide tryptophan zipper 2 (TZ2) designed by Cochran et al.¹ It has received significant attention in the literature, with multiple experimental studies and numerous theoretical treatments.²⁻⁸ Within the last several years, TZ2 has become a benchmark system which is now used in the development of new simulation techniques.^{3,7,9-11} The trpzip peptides were engineered with four tryptophan residues arranged with two side chains on each of the opposing β -strands that lead to a remarkably stable system through the favorable perpendicular stacking of the indole rings.¹ Single-domain peptides like β -hairpins display conformational dynamics that mirror the folding of larger globular proteins.^{12,13}

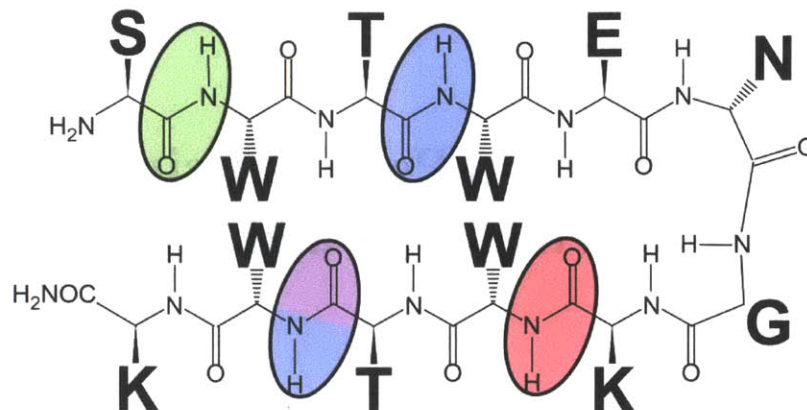


Figure 1. Structure of TZ2 highlighting the S1 (green), TT (blue), T10 (purple), and K8 (pink).

Discussion of hairpin folding mechanisms are often guided by two simple pictures: (1) the kinetic zipper model, in which nucleation of the β -turn is followed by “zipping” of cross-strand backbone contacts,¹⁴⁻¹⁶ and (2) hydrophobic collapse, in which initial clustering by a core of hydrophobic side chains precedes formation of native backbone contacts.^{17,18} Extensive fluorescence^{19,20} and IR absorption²¹⁻²³ measurements have monitored the kinetics of hairpin unfolding on 1-10 μ s time scales but have not reached agreement on which model best describes the folding.^{21,23,24} Often, folding of this peptide is described within the framework of a two-state folding model, although its validity and the nature of the two states is unclear.²⁵

Molecular dynamics (MD) simulations of TZ2 folding have provided intricate details that enrich and challenge these two models, such as partial hydrophobic collapse, fraying of the terminal residues, ordering of backbone hydrogen bonds, salt bridging, and rearrangements of the hydrophobic core.^{11,26,27} To date, however, the complex set of possible folding pathways observed by simulations have not led to a consensus view of the folding mechanism. Experimental evidence has substantiated this complexity and supports a rough free energy landscape in which there are a number of populated minima.¹¹ To date, the structural details of

this heterogeneity have not been experimentally observed, particularly in the peptide backbone. The key information missing from these studies is specific details of hydrogen bond and side chain contact formation at a level of detail that can distinguish conformers.

To obtain residue-level structural insight in to the thermal denaturation of TZ2, amide I' 2D IR was used in combination with ^{13}C and ^{18}O isotope-labeling methods.²⁸⁻³⁵ Inserting these labels into the peptide amide group red shifts the amide I' vibrational frequency away from the main band and provides a spectroscopic probe of the labeled site. Four isotopologues are used to probe cross-strand contacts and isolate peptide units in the turn, termini, and central region of the folded structure (see Figure 1).

In the first isotopologue, S1, an ^{18}O isotope is introduced into the N-terminal serine to probe the stability and fraying of the terminal peptide groups. In the second, T10, an ^{18}O label is placed on the T10 site to monitor the midstrand contacts. The third, TT, also monitors midstrand contacts using ^{13}C isotope labels at threonines T3 and T10 of the peptide. The two labeled amides of TT are hydrogen bound across the β -strands in the folded structure, and vibrational coupling between these sites provides sensitivity to the interstrand contacts in this region.³⁶ Finally, the isotopologue K8 has a ^{13}C label incorporated into the lysine-8 residue to probe the thermal disordering of the β -turn.

The interpretation of 2D IR spectra uses a structure-based spectroscopic model that draws on MD simulations of the TZ2 peptide in water. The spectral simulations make use of the extensive set of TZ2 conformers obtained from the MD simulations of Swope and co-workers,^{11,27} and accompanying Markov state analysis of Chodera et al.¹⁰ Experimental amide I' frequency shifts and 2D lineshapes are assigned and interpreted on the basis of calculated 2D spectra for these Markov states. Spectra are calculated using the local amide I' Hamiltonian that

describes the amide I' frequency shifts of the peptide groups and vibrational couplings between these groups.³⁷⁻⁴⁰ The amide I' site frequencies are mapped onto the local molecular electric field.^{38,39,41,42} This describes the red-shifting of the amide I' frequency upon hydrogen bonding to the peptide unit.

The disorder and local fluctuations of each conformer is reflected in the 2D line shape of the labeled site.^{2,43} The diagonal line width of the 2D IR peak is sensitive to the degree of conformational disorder, whereas the antidiagonal width reflects the fluctuations and time scale of interchange. Therefore the ellipticity of the line shape is inversely related to the time scale over which the vibration samples the possible configurations. A large value of the ellipticity indicates significant disorder in the corresponding local structure. Fast fluctuations in structure and hydrogen bonded solvent lead to round peaks.⁴⁴

This chapter establishes a detailed picture of the conformational changes upon temperature-induced unfolding of TZ2. The isotopologues provide evidence that the termini disorder as the temperature is raised, and the central region retains cross-strand contacts. Further this chapter will provide explicit evidence for distinct conformers of the turn, and that the population of the expected type I' turn increases with temperature. Taken together, these observations reveal a temperature varying heterogeneous ensemble, whose conformers all experience thermally induced fraying, but do not unfold into a random coil structure. In addition to this agreement with previous results, the work shown here gives a level of detail unattainable with other methods. Isotope-editing provides structural detail at the level of individual amide groups with higher spectral resolution than FTIR spectroscopy, while resolving picosecond dynamics that report on peptide flexibility and solvent exposure.

3.2 Spectroscopic Methods

Temperature dependent infrared absorption spectra of each peptide isotopologue were collected from 5 to 95 °C in 5 °C increments. The sample cell used in these measurements consists of ~25 μL of the peptide solution sandwiched between two 1 mm thick CaF_2 windows that are separated by a 50 μm Teflon spacer. For temperature control, the cell was mounted in a brass housing whose temperature was regulated with water from a recirculating chiller. Spectra were acquired on a Nicolet 380 FTIR spectrometer at 1.0 cm^{-1} spectral resolution by averaging 64 one-second scans.

2D IR spectra were collected as described in previous publications,⁴⁵ with $\tau_2 = 100$ fs and the polarization geometry all parallel (ZZZZ) unless otherwise noted. The evolution time, τ_1 , was scanned to 1.5 ps for the nonrephasing spectrum and 2.0 ps for the rephasing spectrum at 4 fs time steps, and a rectangular window function was used for the Fourier transform.

3.3 Simulation and Modeling Methods

To simulate 2D IR spectra, five kinetically metastable macrostates were chosen from the decomposition made by Chodera et al.¹⁰ Their work distilled the 3.23 μs of TZ2 simulation data by Pitera et al.²⁷ into 40 Markov states that displayed patterns in backbone hydrogen bonds and side chain packing. Each macrostate, which was comprised of 30 structurally related microstates with atomistic coordinates, was chosen to represent a different structural model of TZ2.

All MD simulations were performed with GROMACS 3.3.1.^{46,47} The simulations of Pitera et al.²⁷ considered the unblocked peptide, which was fully protonated here (+3 charge) with chloride counterions. Each of the atomistic structures comprising the macrostate was subjected to ~5000 steps of a steepest descent energy minimization to conform to the OPLS/AA

force field⁴⁸⁻⁵⁴ using the particle mesh Ewald method to treat long-range electrostatics.⁵⁵ The peptides were solvated in a box of rigid SPC/E water⁵⁶ that extended 1.5 nm from the solute, and the water was allowed to equilibrate for 10 ps in the NPT ensemble (300 K, 1 atm) while the peptide was fixed.⁵⁷⁻⁶⁰ The simulation was continued with no constraints, except on the bond lengths, for 1 ns of equilibration and 100 ps of data acquisition. Structures were saved every 20 fs to yield 5000 frame trajectories.

For each frame, a Hamiltonian was constructed using the ab initio based models given in the references^{40,41}

$$H_s = \sum_i \left(\omega_i B_i^\dagger B_i - \frac{1}{2} \Delta_i B_i^\dagger B_i^\dagger B_i B_i \right) + \sum_{i,j} J_{i,j} B_i^\dagger B_j$$

where B_i^\dagger and B_i are bosonic creation and annihilation operators for a vibration at site i . The frequency of each site and the coupling between sites are given by ω_i and $J_{i,j}$ respectively. The anharmonicity, Δ_i , lowers the energy for double excitation of site i , and was fixed at 16 cm⁻¹. The frequencies and coupling values were calculated as previously described.^{40,41} The transition dipoles for each site had a unit magnitude and were oriented 20° off the CO bond vector in the CN direction. Where applicable, the sites were isotope-labeled by shifting their frequency 41 cm⁻¹ to the red.

The 2DIR spectra were calculated using the numerical integration of the Schrödinger equation (NISE) scheme.^{61,62} In this procedure the time-evolution matrices needed in the nonlinear response functions that govern the 2DIR spectra are calculated by solving the time-dependent Schrödinger equation. This is done numerically by dividing time into small intervals during which the Hamiltonian can safely be assumed to be constant. The time-evolution matrix for the time-independent Schrödinger equation can then be found for each interval and the time-

evolution matrices for longer time periods are found by multiplying the matrices for the small intervals in a time ordered fashion. The resulting spectra were summed over each structure in the macrostate to produce spectra that reflected the disordered ensembles.

3.4 Amide I' FTIR Spectroscopy Results

FTIR spectra of the TZ2 isotopologues at 25 °C are shown in Figure 2. The peak maximum of the unlabeled (UL) amide I' band is 1636 cm^{-1} , and there is a prominent shoulder at 1674 cm^{-1} . On the basis of the NMR structure, these bands have previously been assigned to ν_{\perp} and ν_{\parallel} excitonic transitions, delocalized vibrations of the amide backbone in which cross-strand peptide bonds oscillate in-phase or out-of-phase with one another.³⁶ Twisting of the β -strands leads to two nearly degenerate ν_{\perp} vibrations located at the turn and midstrand regions of the peptide, while the ν_{\parallel} vibration is largely localized at the turn. Because the spectra shown are taken at pH 2.5, the glutamic acid side chain carbonyl stretch contributes at 1710 cm^{-1} .⁶³

Changes to the IR spectrum upon isotope substitution are also plotted in Figure 2 for K8, T10, TT and S1. Isotope-labeling leads to small shifts in intensity and frequency of the ν_{\perp} and ν_{\parallel} bands, and an increase of amplitude to the red of the bands from the isotope-labeled peptide carbonyl. As seen in difference spectra relative to UL, K8 has a broad ($\sim 35\text{ cm}^{-1}$) increase in intensity from $\nu_{\text{K8}} = 1585\text{-}1615\text{ cm}^{-1}$, indicating significant structural heterogeneity at this site. The ν_{\perp} band of K8 narrows and shifts slightly to the blue (1638 cm^{-1}), suggesting that isotopic substitution significantly rearranges the underlying ν_{\perp} eigenstate composition.

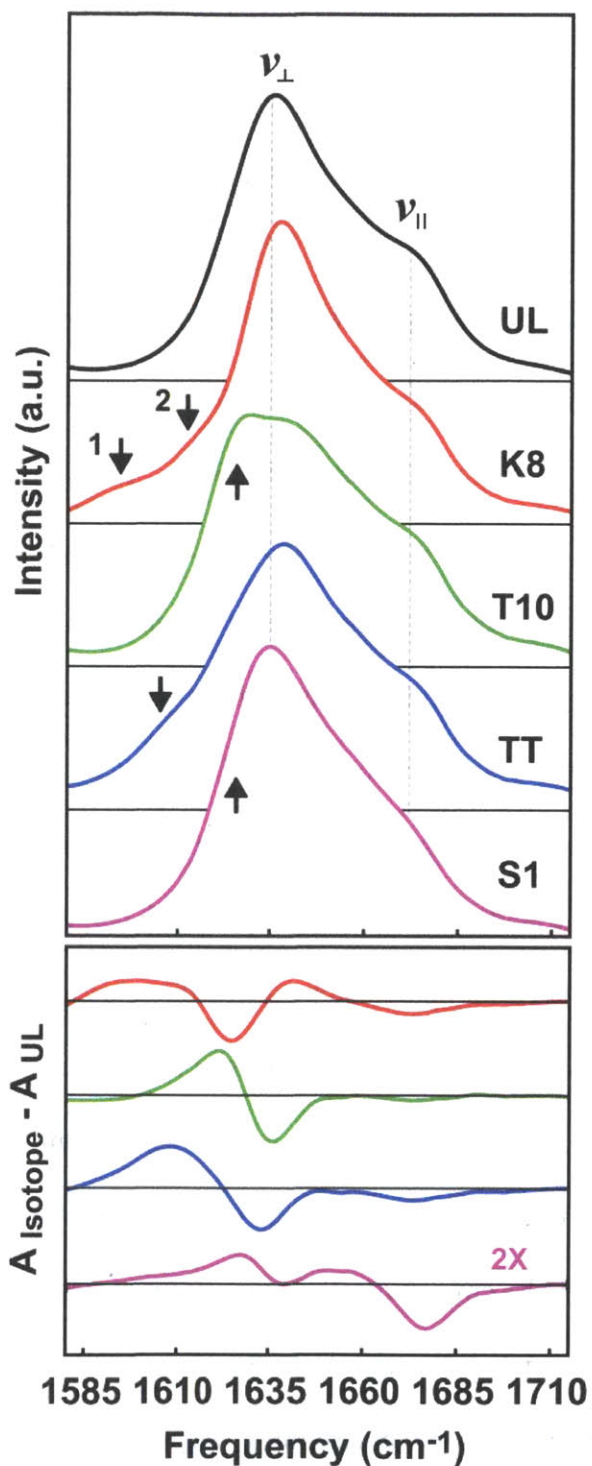


Figure 2. Equilibrium FTIR spectra are shown for each TZ2 isotopologue at 25 °C and pH 2.5. Each spectrum is baseline corrected with a linear subtraction, and area normalized for comparison. Difference data (bottom) are obtained by subtracting the unlabeled spectrum (TZ2-UL) from each of the spectra. The S1 difference spectra is presented at 2×magnification for clarity. Arrows have been drawn to emphasize the isotope labeled peaks.

In addition to a drop of intensity in the ν_{\perp} band, the T10 spectrum shows a pronounced shoulder at 1627 cm^{-1} due to the labeled peptide group. This is a relatively high frequency, indicative of a single moderately strong hydrogen bond to the oxygen as expected from the NMR structure. Upon labeling both threonines (TT), spectral loss is observed at 1640 and 1675 cm^{-1} , and a broad absorbance gain is observed at $\nu_{\text{TT}} = 1610\text{ cm}^{-1}$. The additional red shift of ν_{TT} relative to ν_{T10} reflects the strength of coupling between the cross-strand labeled oscillators of the ν_{TT} band. Labeling both of the threonine residues redshifts the midstrand ν_{\perp} vibration and also perturbs the high frequency, ν_{\parallel} , band.⁶⁴ A similar but narrower band was observed at 1608 cm^{-1} for the TZ2 mutant TZ2-T3A*T10A*.⁶

The FTIR spectrum of S1 is qualitatively similar to the UL spectrum. The peak absorption frequency is red-shifted by 2 cm^{-1} relative to UL, and the difference spectrum reveals a slight intensity increase between 1590 and 1625 cm^{-1} . A distinct ν_{S1} band is unresolved in the FTIR spectrum, which indicates that the 39 cm^{-1} frequency shift from the ^{18}O label moves intensity from the blue side of the amide band to the red side, where it remains hidden underneath the ν_{\perp} band. There is also loss from 1670 to 1685 cm^{-1} and from 1637 to 1650 cm^{-1} in the difference spectrum.

3.5 Amide I' 2D IR Spectroscopy Results

Amide I' 2D IR spectra of the TZ2 isotopologues at $25\text{ }^{\circ}\text{C}$ are shown on the top row of Figure 3. The spectra show on and off-diagonal features in the amide I' main band between 1630 - 1690 cm^{-1} , as well as red-shifted features along the diagonal for the isotope-labeled species. Positive features arise from fundamental ($\nu = 0$ -1) vibrational transitions, whereas

negative features arise from $\nu = 1-2$ transitions. As described previously,⁵ the UL spectrum shows the Z-shaped contour profiles of the excitonic amide I' bands expected for antiparallel contacts between two strands. The ν_{\parallel} band is resolved as a separate peak rather than a shoulder as in FTIR, and the peak splitting between ν_{\perp} and ν_{\parallel} peaks is 37 cm^{-1} in the ω_1 dimension. Coupling between ν_{\perp} and ν_{\parallel} vibrations is observed in the off-diagonal region as ridges that extend along $\omega_3 = 1636$ and 1674 cm^{-1} .

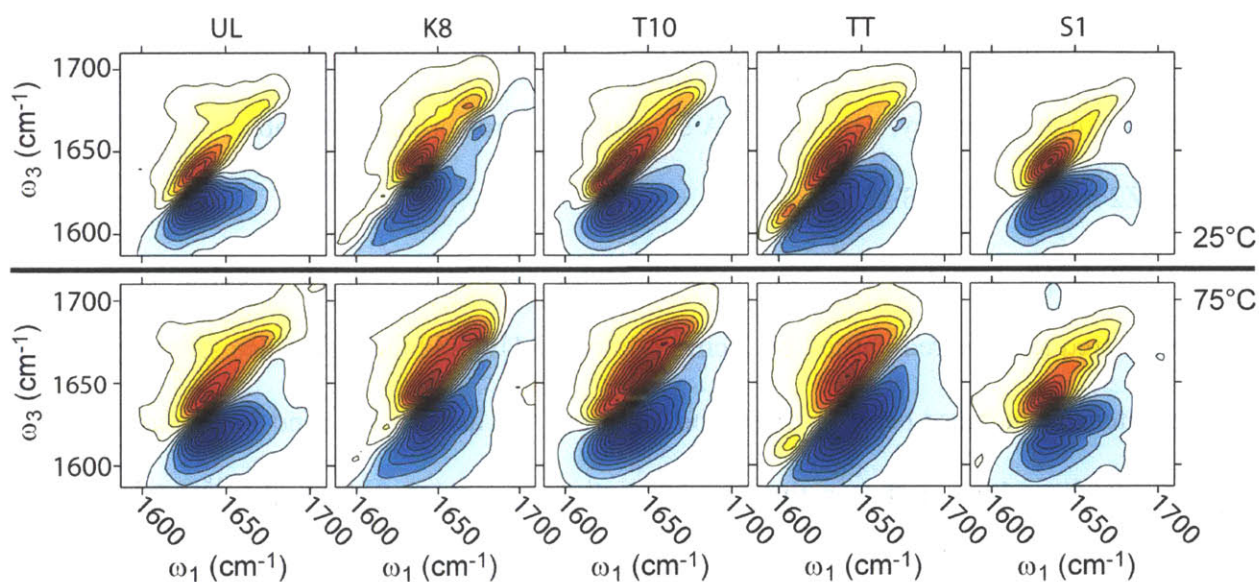


Figure 3. 2D IR Spectra of TZ2 isotopologues at 25 (top) and 75 °C (bottom).

For K8, the most striking feature in the 2D IR spectrum is the observation that the ν_{K8} band is composed of two distinct diagonal peaks: one at 1593 cm^{-1} (ν_{K8-1}) and the other at 1611 cm^{-1} (ν_{K8-2}), with no detectable cross peak between them. (See also Figure 6.) Also, the diagonal line width of ν_{K8-1} ($\sigma = 6.9 \text{ cm}^{-1}$) is broader than ν_{K8-2} ($\sigma = 4.5 \text{ cm}^{-1}$). Strong coupling of the ν_{K8-2} to the ν_{\perp} band is observed as a ridge extending to $\omega_1 = 1612 \text{ cm}^{-1}$ and $\omega_3 = 1650 \text{ cm}^{-1}$ (Figure 6). The presence of two shifted peaks for a single K8 isotope-label is an indication of two structurally distinct subensembles. The 18 cm^{-1} peak splitting between these features indicates

that the conformers differ considerably in hydrogen bonding configuration to the K8 peptide oxygen. The relatively high frequency of ν_{K8-2} indicates a single hydrogen bond to oxygen. The most reasonable explanation for this observation is that the ν_{K8-2} peak originates in a properly folded type I' turn in which the labeled oxygen has a single cross-strand hydrogen bond. The ν_{K8-1} peak frequency is characteristic of a peptide group with two or more hydrogen bonds and a higher degree of configurational freedom. This suggests a solvent exposed carbonyl that might be present in a disordered or bulged turn.

For T10, the main excitonic band retains the Z-shaped contours of UL, but an additional sharp peak from the labeled site is observed at $\omega_1 = \omega_3 = 1630 \text{ cm}^{-1}$. Like ν_{K8-2} this high frequency is suggestive of a single hydrogen bond to oxygen as seen in the NMR structure. In the 2D IR spectrum of the double labeled species, TT, the ν_{TT} band is well-resolved at 1602 cm^{-1} , and the splitting between ν_{\perp} and ν_{TT} in the ω_1 dimension is 28 cm^{-1} . The additional red-shift relative to T10 arises from strong coupling between the Thr3 and Thr10 amides forming a midstrand ν_{\perp} -like mode.⁶⁴ Changes to the 2D IR ν_{\perp} band relative to the other isotopologues are dramatic, including loss of off diagonal structure and a diagonal line width increase of approximately 21 cm^{-1} . This line width change suggests that the eigenstate distribution of TZ2-TT is much more heterogeneous than in the other peptides.

The 2D IR spectrum of S1 does not show a peak corresponding to the ^{18}O label in the serine group; however, there are changes relative to UL along the diagonal. The ν_{\parallel} band of the S1 2D IR spectrum, which has a resolvable peak at 1674 cm^{-1} in the UL spectrum, now appears as a diagonal ridge to the blue of the ν_{\perp} band. In addition there is a 2 cm^{-1} red-shift to 1634 cm^{-1} and line-narrowing of the ν_{\perp} peak. These observations indicate that the frequency of the UL S1 site energy is relatively high ($\sim 1675 \text{ cm}^{-1}$) and that labeling moves the oscillator into resonance

with the ν_1 exciton band. This would also explain the lack of both a resolvable peak and interstrand coupling observed between labeled residues in the TZ2 mutant TZ2-S1A*,T10A*.⁶

3.6 Temperature-Dependent Spectra

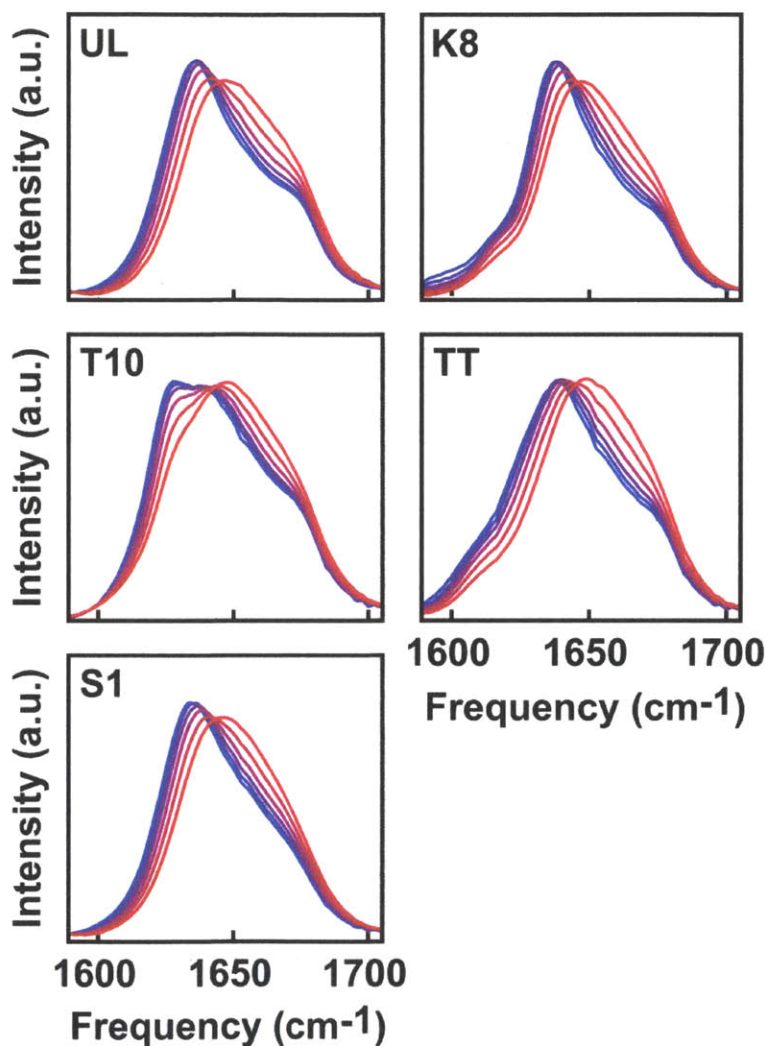


Figure 4. Temperature-dependent FTIR spectra of each TZ2 isotopologue taken from 10 to 85 °C in 15 °C increments (pH 2.5). Line color shows progression from cold (blue) to hot (red).

Temperature-dependent amide I' FTIR spectra of each isotope-labeled hairpin are shown in Figure 4. Each spectrum shows similar trends with temperature, including the loss of

amplitude in the ν_{\perp} region and rise of amplitude in the 1660 cm^{-1} region commonly attributed to disordered chains. To characterize the overall melting behavior, singular value decomposition (SVD) of the entire amide I' band is carried out for each data set. For each isotopologue, the data is well described with the first two spectral components. Spectral changes with temperature are characterized by the second SVD component, and are shown as melting curves in Figure 5.

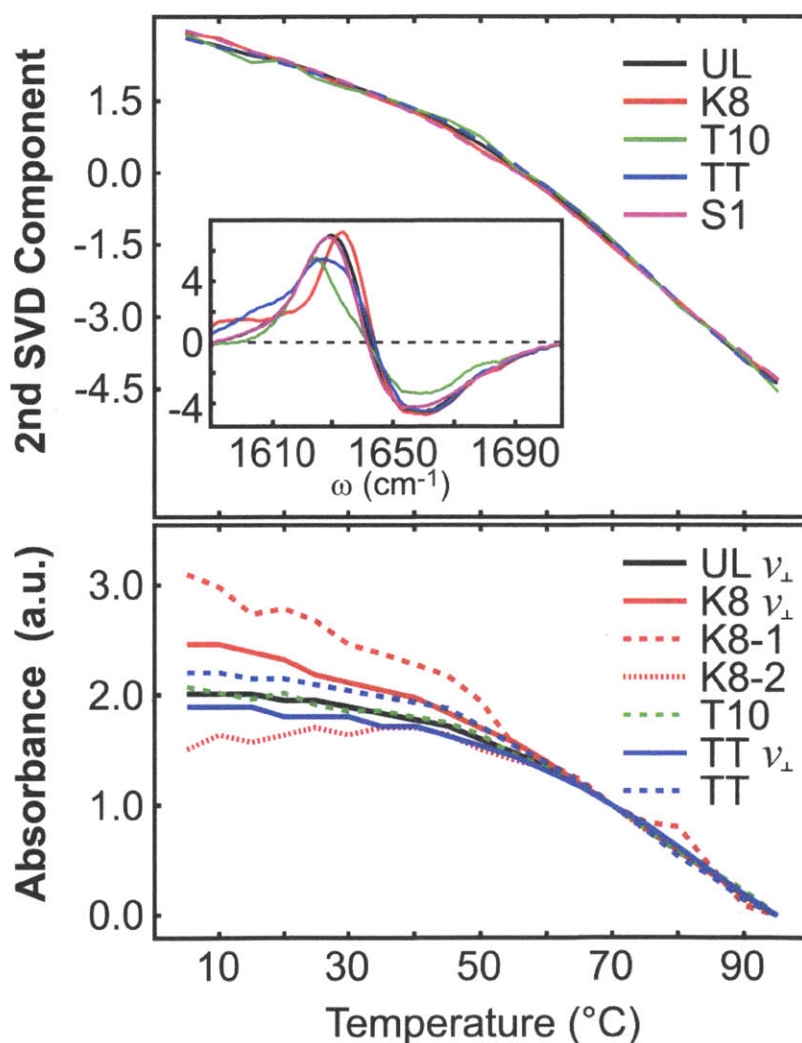


Figure 5. (Top) Second component amplitudes from SVD analysis of the temperature-dependent FTIR spectra. The inset shows the second component spectra. (Bottom) FTIR peak absorbance versus temperature. The ^{12}C peak intensities are obtained from the absorbance within 5 cm^{-1} of the amide I' peak maximum. The integration regions for other traces are $\nu_{\text{K8-1}}$, $1588\text{-}1598\text{ cm}^{-1}$; $\nu_{\text{K8-2}}$, $1612\text{-}1619\text{ cm}^{-1}$; ν_{TT} , $1603\text{-}1611\text{ cm}^{-1}$. Traces are scaled to overlap between 70 and $95\text{ }^{\circ}\text{C}$.

The melting curves are nearly identical and each show slight curvature without a distinct cooperative unfolding transition. This data is consistent with previous IR, fluorescence, and UVCD measurements^{2,4-6} and is identical to the earlier UL data at pH 7.⁶⁴

In addition to SVD characterization, it is also useful to observe specific spectral regions within the band. The relative intensities of the isotope-shifted peaks of K8, T10 and TT are compared with the ν_{\perp} peak of UL in Figure 5. Intensities are obtained by averaging the absorbance around the peak positions ($\nu = \pm \sim 5 \text{ cm}^{-1}$), and are compared by overlaying the curves between 70-95 °C. This demonstrates that the melting behavior follows the same trends for all spectroscopic features between 55 and 95 °C. The melting curves for ν_{\perp} peak intensities of all isotopologues have profiles matching the second SVD component melting curves within 10%. Similarly, the intensity of the T10 and TT labels both track the SVD melting curves, suggesting that these curves report on the interstrand contacts in the central region of the peptide. For K8, the ν_{K8-1} region changes quasi-linearly with increasing temperature. In contrast, the ν_{K8-2} region is largely unchanged for temperatures <40 °C. This melting behavior suggests that the ν_{K8-2} peak reports on peptides with a natively structured turn, while the ν_{K8-1} peak reports on non-native turn configurations that show a melting curve similar to that expected for a less-stable or disordered peptide. The high apparent melting temperature of the ν_{K8-2} peak suggests that the natively structured turn region is in fact more stable than the rest of the peptide and could account for the residual backbone structure previously reported over similar temperatures.

Temperature-dependent changes to the 2D IR spectra of each isotopologue can be seen by comparing the 25 and 75 °C spectra shown in Figure 3. Along the diagonal axis ($\omega_1 = \omega_3$), spectral shifts track those observed in the FTIR data, and spectral features broaden along the

antidiagonal axis with increasing temperature. This analysis focuses on the center frequency, amplitude, and 2D lineshapes of the isotope-labeled features. The line shape can be characterized by their widths parallel and perpendicular to the diagonal axis. The diagonal line width (σ) reflects inhomogeneous broadening that originates in static structural disorder on the time scale of the experiment (~ 2 ps). Antidiagonal widths (Γ) report on rapid frequency shifts that originate in subpicosecond peptide-solvent interactions and structural fluctuations. The relative magnitudes of these linewidths, or ellipticity $E = (\sigma^2 - \Gamma^2)/(\sigma^2 + \Gamma^2)$,^{44,65} measured as a function of temperature can be used to interpret changes from static disorder to dynamic disorder due to increasing structural flexibility and solvent exposure.⁴⁴

Spectral changes to the UL 2D IR data are similar to those reported earlier at pH 7, which were interpreted as fraying of the termini and an intact turn at all temperatures. The antidiagonal line width grows with temperature for both of the β -sheet modes in the 2D IR spectrum (data not shown). The full width at half-maximum (FWHM) of the ν_{\parallel} band is only ~ 2.0 cm^{-1} larger than that of the ν_{\perp} band at 15 $^{\circ}\text{C}$, but the slope of the line width increase versus temperature of ν_{\parallel} is twice that of the ν_{\perp} band.

Thermally induced spectral changes to the ν_{K8} peaks can be seen more clearly in Figure 5. As the temperature of the system increases, the spectrum responds both to structural rearrangements and shifting populations between native and misfolded ensembles. The center frequency of the $\nu_{\text{K8-1}}$ and $\nu_{\text{K8-2}}$ bands does not change with temperature, but the amplitudes of both peaks decrease. In addition, the $\nu_{\text{K8-1}}$ intensity decreases with respect to the $\nu_{\text{K8-2}}$ peak, indicating that the native turn is more stable than the disordered turn. Linewidth changes to the TZ2-K8 spectra are consistent with those of the unlabeled hairpin (Figure 7). The slope of the line fit to the FWHM of the ^{12}C ν_{\perp} band (0.058 ± 0.012 $\text{cm}^{-1}/^{\circ}\text{C}$) is twice as large as the

corresponding slope of TZ2-UL ($0.020 \pm 0.006 \text{ cm}^{-1}/^{\circ}\text{C}$). For the $\nu_{\text{K8-1}}$ and $\nu_{\text{K8-2}}$ peaks, the FWHM measurement is noisier due to lower signal strength. However, the line width of the high frequency peak, $\nu_{\text{K8-2}}$, does not display a significant slope with temperature, indicating that the homogeneous broadening does not increase significantly with temperature. The limited resolution of the ν_{K8} peaks prevents a quantitative description of the diagonal line width. However, in the spectra shown in Figure 6 there does not seem to be any clear sign of diagonal peak broadening with temperature.

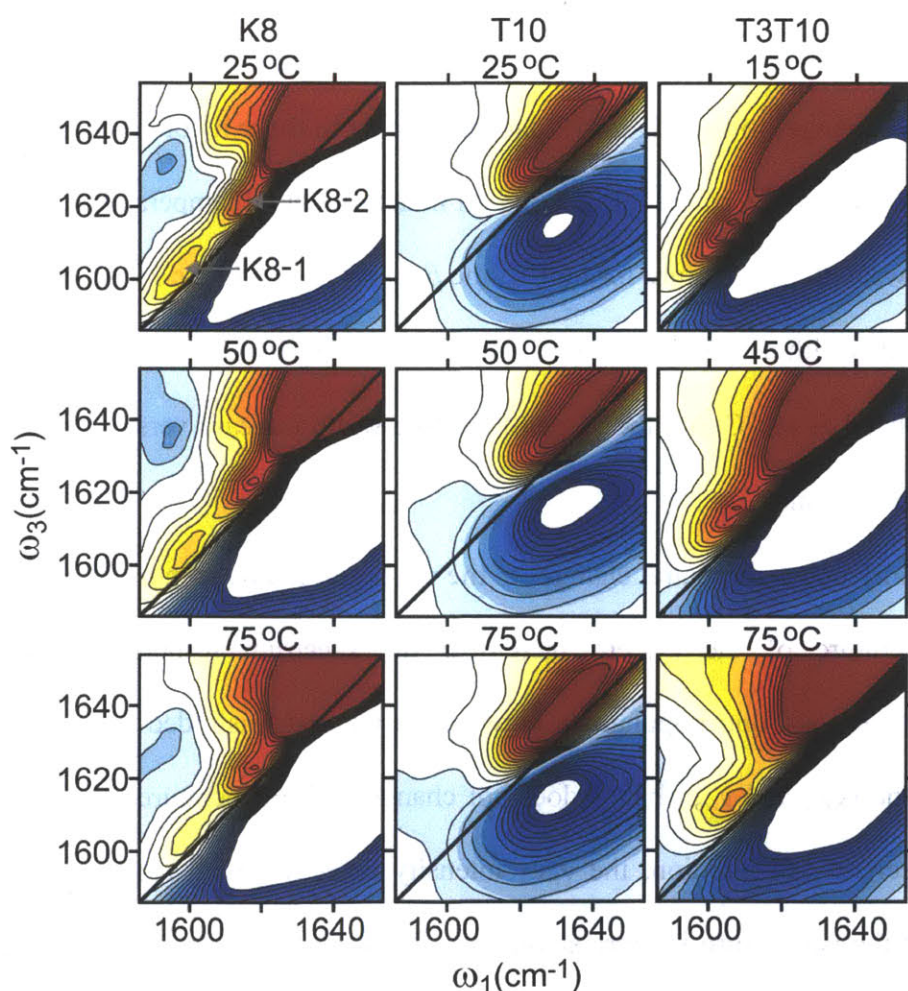


Figure 6. Temperature-dependent 2D-IR spectra of isotope-shifted peaks in K8, T10, and TT. Contour lines are plotted as a percentage of the amide I' maximum at the respective temperature. Twenty seven evenly spaced contours are drawn from -18 to +18% (K8), -80 to +80% (T10), and -35 to +35% (TT).

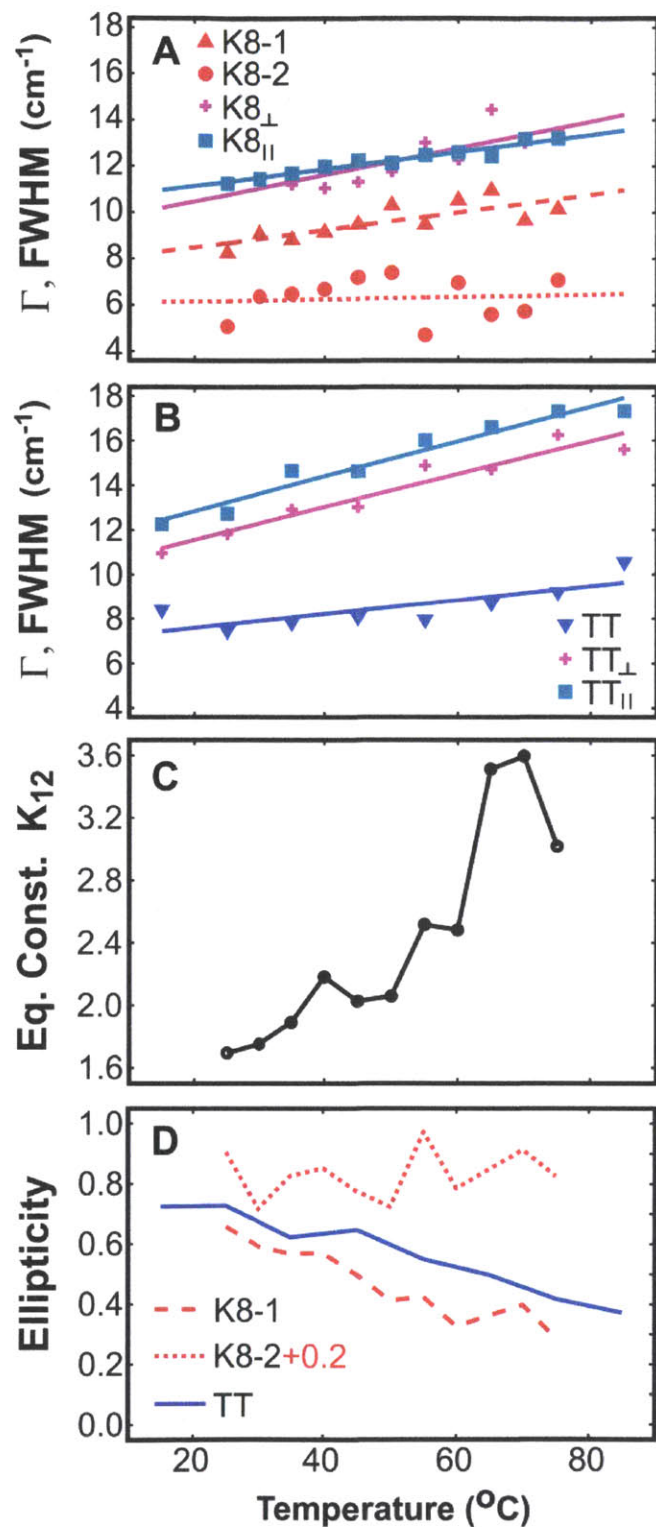


Figure 7. Lineshape analysis of temperature-dependent 2D IR spectra of K8 and TT. Antidiagonal line width (FWHM) for (A) K8 and (B) TT. (C) Ratio of K8-2 to K8-1 peak volumes. (D) Ellipticity of isotope-shifted peaks. The data for K8-2 has been upshifted by 0.2 for clarity.

Together, these observations support the idea that two bands are reporting on subensembles of peptides for which the relative population shifts with temperature, but the line width properties remain relatively constant. The presence of two isotope-shifted peaks in the turn may reflect two distinct conformational states of the peptide; however, from the K8 measurement alone one cannot rule out additional states that have the same type of solvent exposure in the turn. Of course, conformational dynamics will be present that interconvert these turn structures. Not surprisingly, if a two state analysis is applied to this exchange of conformers (K8-1 to K8-2), one finds that the temperature dependence does not follow van't Hoff behavior (Figure 7). However, if one restricts the analysis to the high temperature range between 45 and 75 °C for which most spectral features shift in concert, one finds that $\Delta H_{1-2} = 20.2 \pm 2.8$ kJ/mol.

Thermally induced spectral changes to the ν_{TT} peak are emphasized in Figure 6. Similar to the ν_{K8} peaks, the maximum of the positive diagonal ν_{TT} peak does not shift appreciably with temperature. Instead, the $\nu_{TT} - \nu_{\perp}$ peak splitting increases from 26 cm^{-1} at 15 °C to 34 cm^{-1} at 85 °C, which results from the blue-shift of the ν_{\perp} peak. The relative intensity of the isotope-shifted peak also drops, going from 40% of the band maximum at 15 °C to 20% at 85 °C. This decrease in intensity is partially ascribed to a loss of coupling between isotope-labeled amide groups and the ^{12}C amides. The shape of the ν_{TT} peak also changes drastically with temperature, going from a diagonally elongated peak to a symmetrically distributed peak as the temperature is raised. From 15 to 75 °C, the antidiagonal ν_{TT} FWHM increases more than 2.0 cm^{-1} , and the ν_{TT} diagonal width decreases by a similar amount over the equivalent temperature range. These line width shifts show that the region around T3 and T10 are relatively rigid at low temperature but experience increased configurational fluctuations of the peptide groups and surroundings as the temperature is raised. Yet, on average the frequency of the ν_{TT} peak and magnitude of

crossstrand coupling remains unchanged from lower temperatures. These disordering pathways are consistent with the fraying transition described in previous hairpin simulations.^{11,66,67}

3.7 Simulation Results

To better interpret how the observed amide I' spectral features report on peptide conformation, structure-based modeling of 2D IR spectra was applied to five sets of TZ2 conformers drawn from the Markov state analysis of Chodera et al.¹⁰ Each configuration chosen for the spectral simulations should be viewed as an ansatz from which structural features that play an important role in determining the spectroscopy can be identified. Conformational disorder for each of these Markov states is shown in Figure 8 and one representative structure from each state chosen is shown at the top of Figure 9. In Figure 8 a visualization of these states is presented showing 30 overlapped structures of the peptide. These backbone visualizations serve to characterize the amide group orientations and conformational disorder of each state, and the lower right panel illustrates the packing of tryptophans for each state. The set of five representative Markov States, shown in the figures below, were used to create structure/spectrum correlations. Here the number identifying each state refers to the designation in appendix for McDonald & Jorgensen 1998.⁶⁸

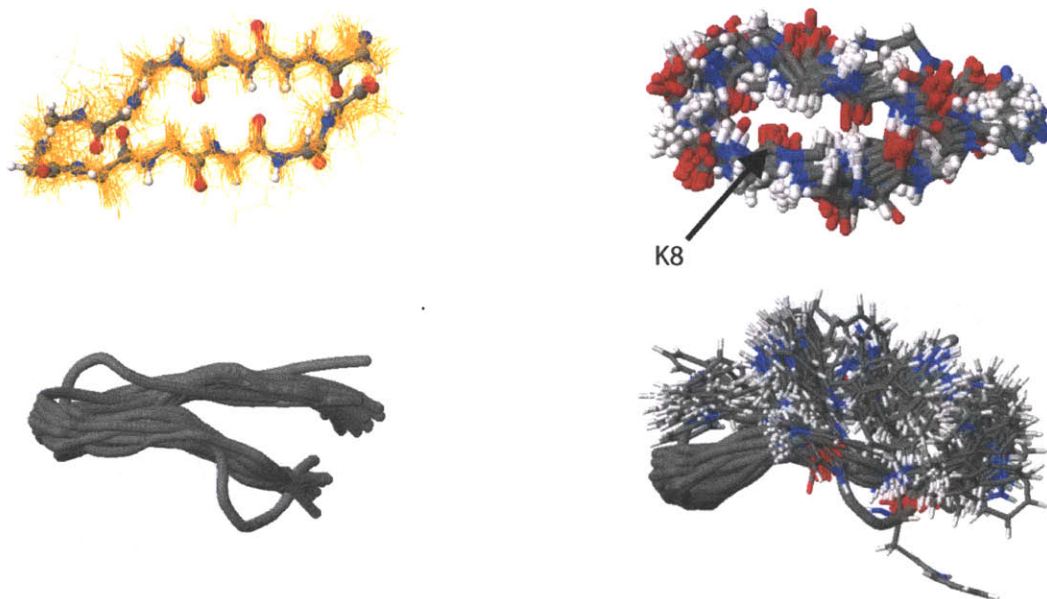


Figure 8a. Folded (FO, 250851) is a well-ordered ensemble that has an average of four backbone hydrogen bonds, which correspond to those in the NMR structure.¹ The S1, T3, K8, and T10 carbonyls are all oriented for cross strand hydrogen bonds, and all four tryptophan side chains are packed to one face of the peptide.

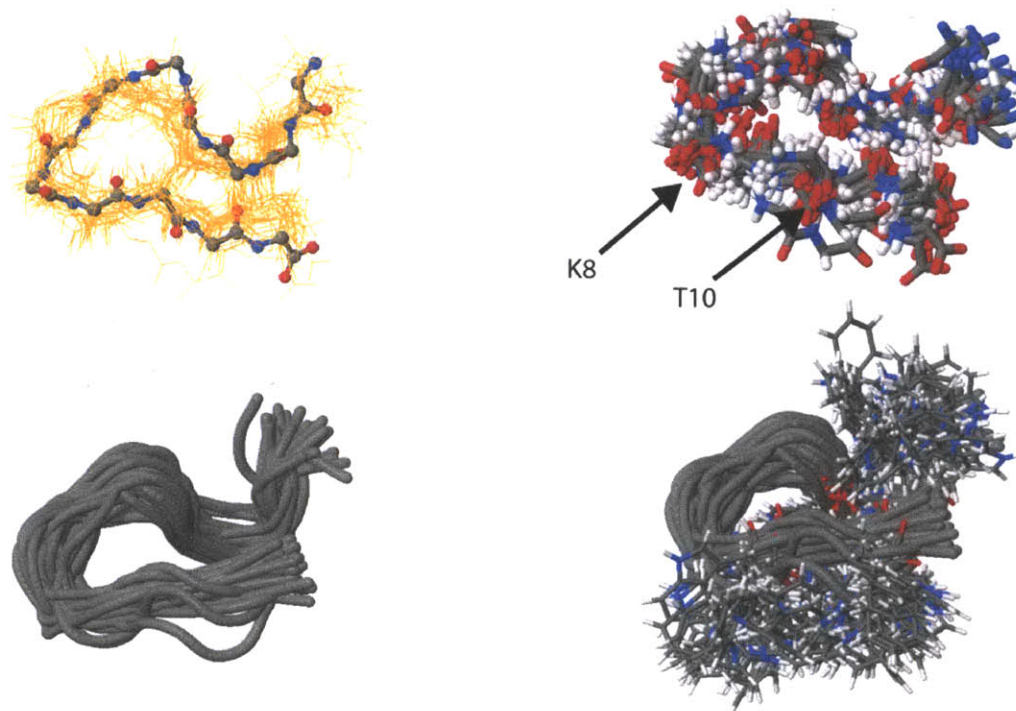


Figure 8b. Bulged Turn (BT, 214369) is a compact coiled structure with a well-solvated, bulged turn region and an average of one cross-strand hydrogen bond that appears near the strand termini. This state contains S1, T3, K8, and T10 peptide carbonyls that are solvent exposed, and tryptophan side chains make rare contacts.

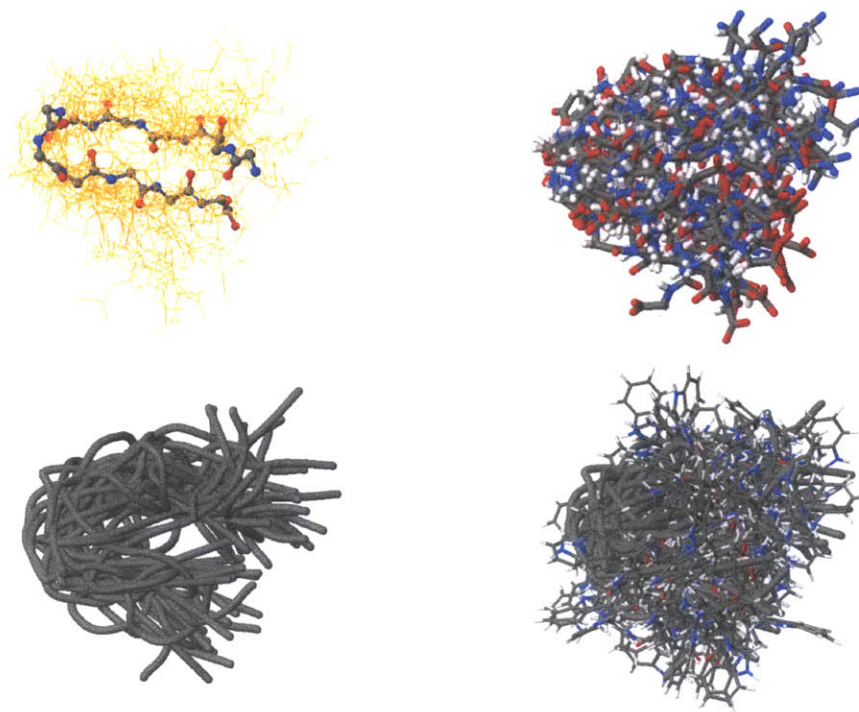


Figure 8c. Disordered (SR, 271154) is a disordered ensemble whose common feature is a bulged turn and an average of one backbone hydrogen bond with a misaligned registry near the C terminus. Tryptophans are poorly organized.

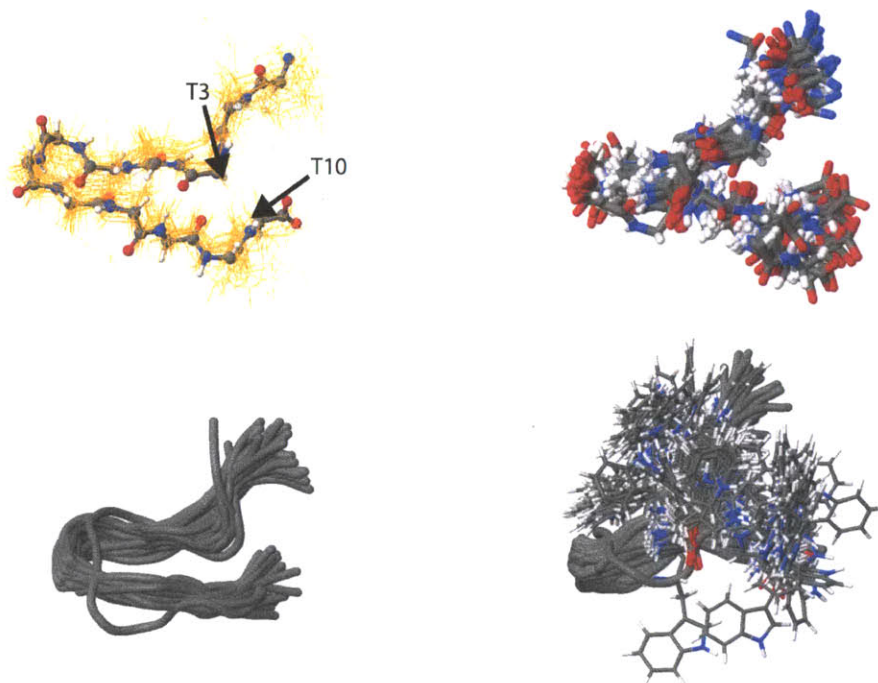


Figure 8d. Frayed (FR, 11131) has a compact properly formed type I' turn region with an average of two backbone hydrogen bonds but disordered strand termini. The T3 and T10 are oriented toward the opposite strand, and tryptophan side chains are disordered.

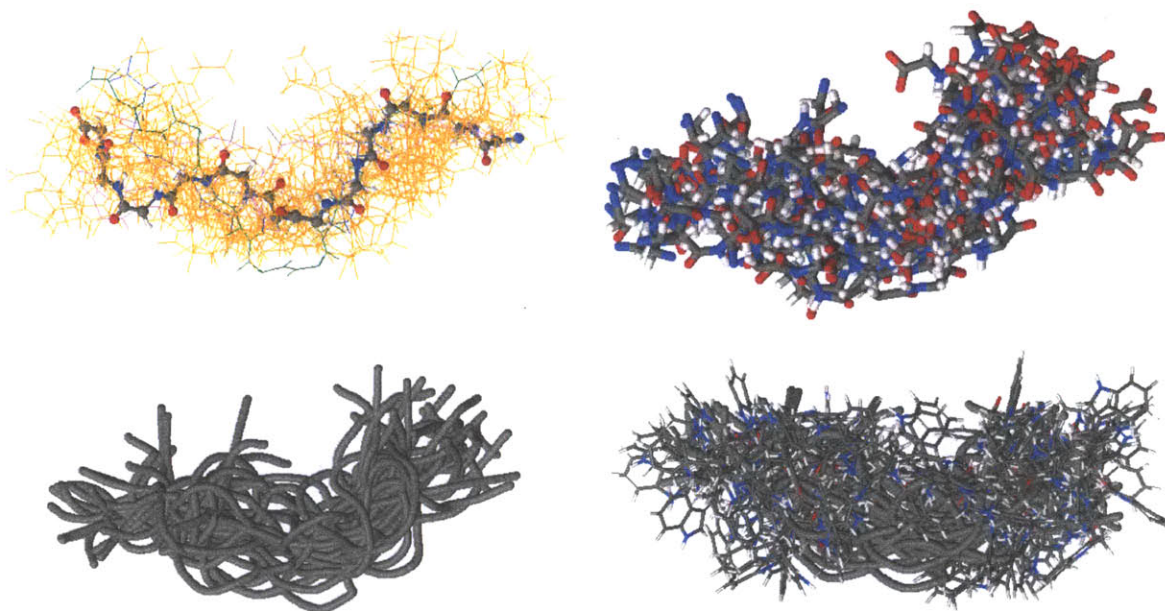


Figure 8e. Extended Disorder (ED, 64336) is an extended and well solvated state with substantial disorder and rare backbone hydrogen bonds.

As a first point of contact to experiment, electric field mapping was used to calculate histograms of amide I' site energies for the 11 peptide units of the backbone. The red-shift is primarily sensitive to hydrogen bonding to the oxygen site. Histograms for sites S1, T3, K8, and T10 are shown in Figures 10 and 11. It is apparent that the distribution of site energies for site T3 and K8 vary in such a way that the mean frequency is unique for the FO and BT conformers. On the other hand, S1 and T10 appear to be insensitive to conformation. Time-correlation functions for the site energy showed biphasic behavior, with a picosecond component attributed to solvent induced dephasing, and a longer component assigned to sampling of backbone configurations. The longer decay time varied from ~ 50 ps for S1, T3, and T10 to >200 ps for W4 and K8.

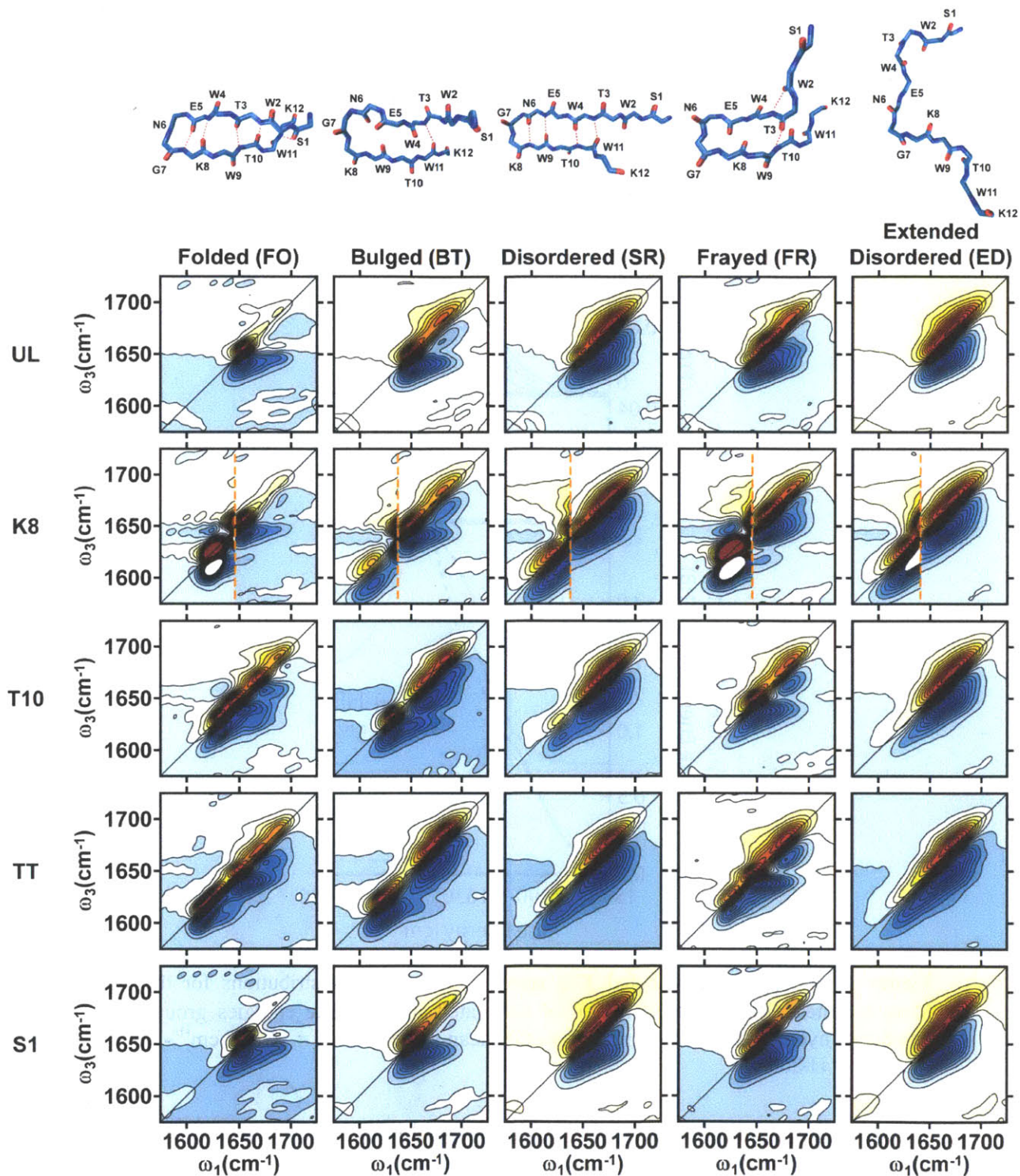


Figure 9. 2D IR spectral simulations of all five TZ2 isotopologues studied for five different conformational ensembles. One representative structure for each conformer macrostate is presented. Simulations of K8 for the spectral region $\omega_1 < 1650 \text{ cm}^{-1}$ have the amplitudes scaled by a factor of 10x.

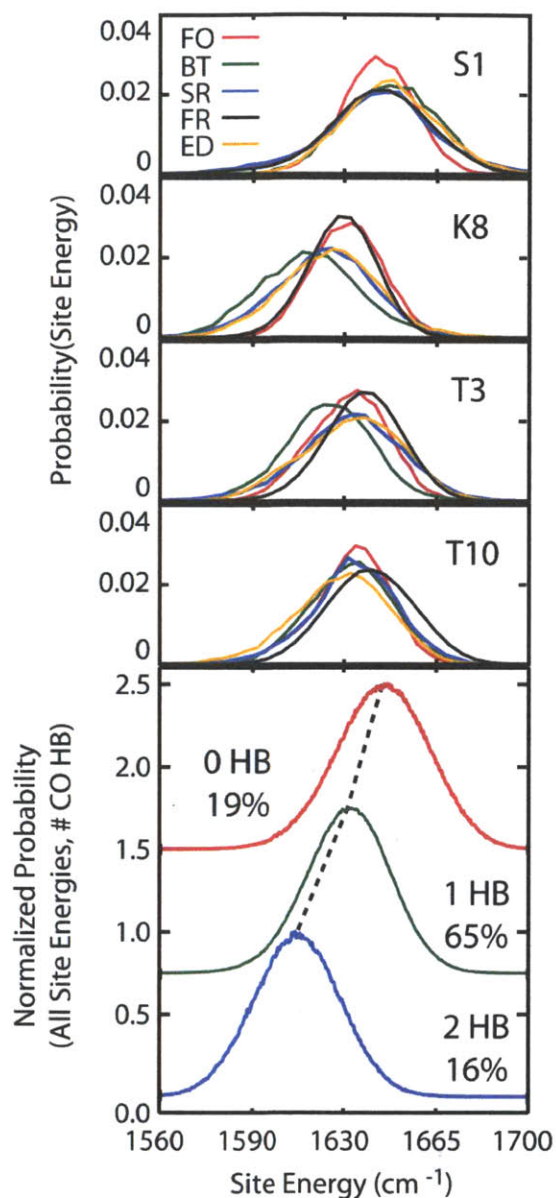


Figure 10. (Top) Isotope-labeled TZ2 amide I' site energy distributions for the five Markov states. (Bottom) Distribution of calculated amide I' site energies grouped by number of hydrogen bonds. A linear fit to the mean value gives $\omega = 1645.7 \text{ cm}^{-1} - 16.3 \times (\text{no. of CO H-bonds})$.

The modeling of all conformers provides a general relationship between the amide I' site frequency and the hydrogen bonds it participates in. This modeling found that a hydrogen bond (HB) of an amide carbonyl to protein gives about the same red shift as a hydrogen bond to water.

(For this chapter, a HB is defined by the hydrogen bond distance and angle: $r_{X \cdots O} \leq 4.5 \text{ \AA}$, $\theta_{X \cdots O} < 35^\circ$). However, when an amide group is fully solvated by water, it has the ability to form

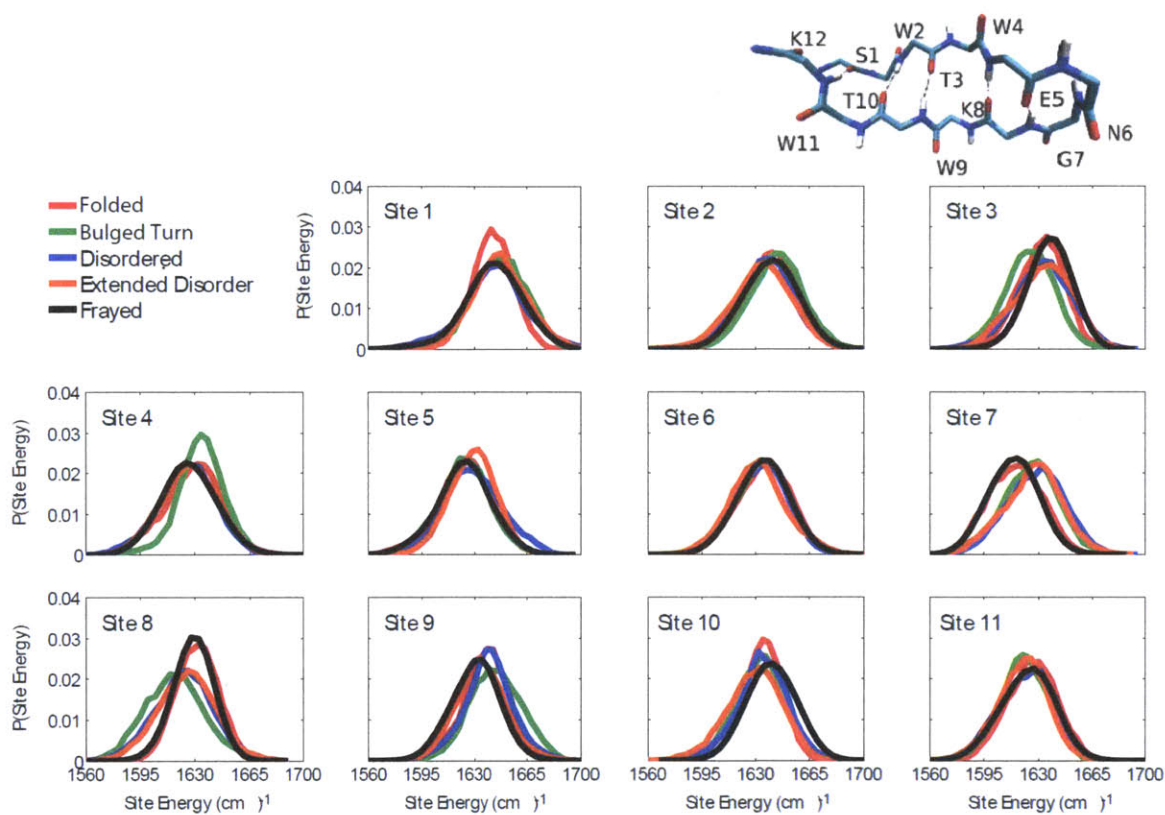


Figure 11a. Analysis of simulated site energy variation by conformer

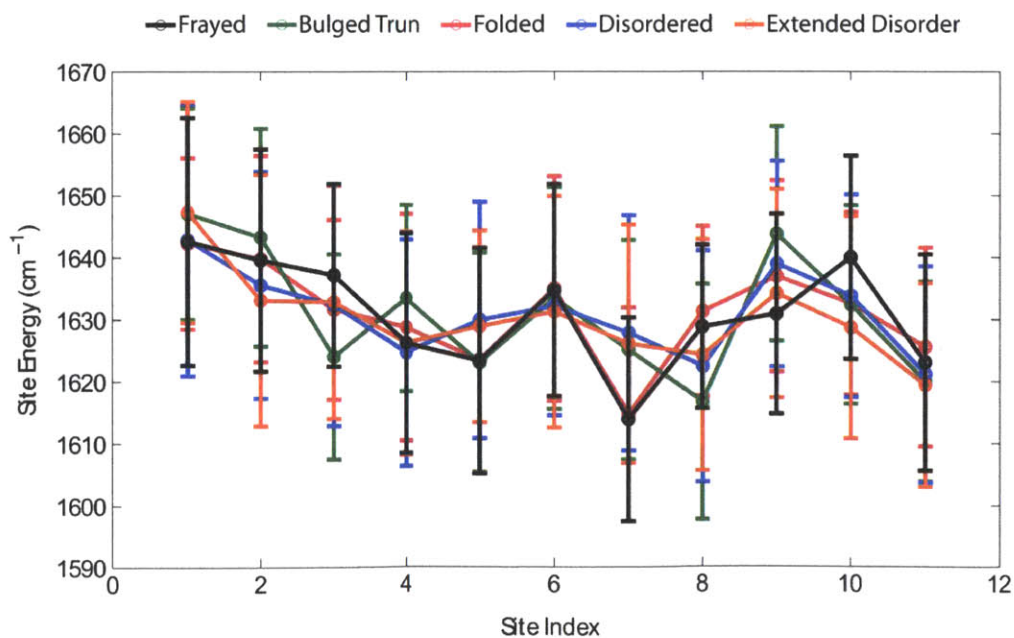


Figure 11b. Site energy mean values showing error

multiple HBs, giving a larger average red shift. From analysis of both types of HBs, on average, one HB leads to an 11 cm^{-1} red shift, however the dispersion in this relationship ($\sigma = 25 \text{ cm}^{-1}$) leads only to a modest correlation ($\rho = -0.50$). These results can be compared with the results of ab initio calculations on isolated clusters, which reveal a 20 cm^{-1} red-shift for each HB to oxygen and an additional 10 cm^{-1} red shift for hydrogen bonds donated by the peptide N-H.⁶⁹

For each of these five conformers, infrared spectra were simulated for all five isotopologues studied. Simulated FTIR and FTIR difference spectra for each structure are shown in Figure 12. In general qualitative similarities exist between several of the experimental and simulated spectra, which are discussed here in terms of peak positions, amplitudes, linewidths, and lineshapes. The simulated 2D IR spectra of the unlabeled peptide for FO, BT, and FR conformers each display two diagonal peaks and Z-shaped contours that resemble the experiment at low temperatures. For the disordered SR state, two bands are also resolved although in this case, the high-frequency band carries most of the intensity. Limiting the simulation only to a specific structure with +1C slipped registry and 4 hydrogen bonds leads to a spectrum similar to BT, see Figure 13. For the disordered ED state, the amide I' band has little structure and reflects the diagonally stretched resonance of an inhomogeneous line shape.

As expected, each of the five K8 simulations gives rise to one resonance for the labeled site, although the peak position and line width vary with conformation. In particular, there is a correlation between the orientation and solvent exposure of the K8 amide group and the frequency of the isotope-shifted peak. In the case of FO and FR states each have a K8 amide group oriented so that its carbonyl oxygen forms a cross-strand hydrogen bond with the W4

amide proton, as expected for a type I' β turn. In their calculated spectra, the ν_{K8} peak is split from the amide I' maximum by 25-30 cm^{-1} , and the diagonal line width is a compact 5-6 cm^{-1} . In

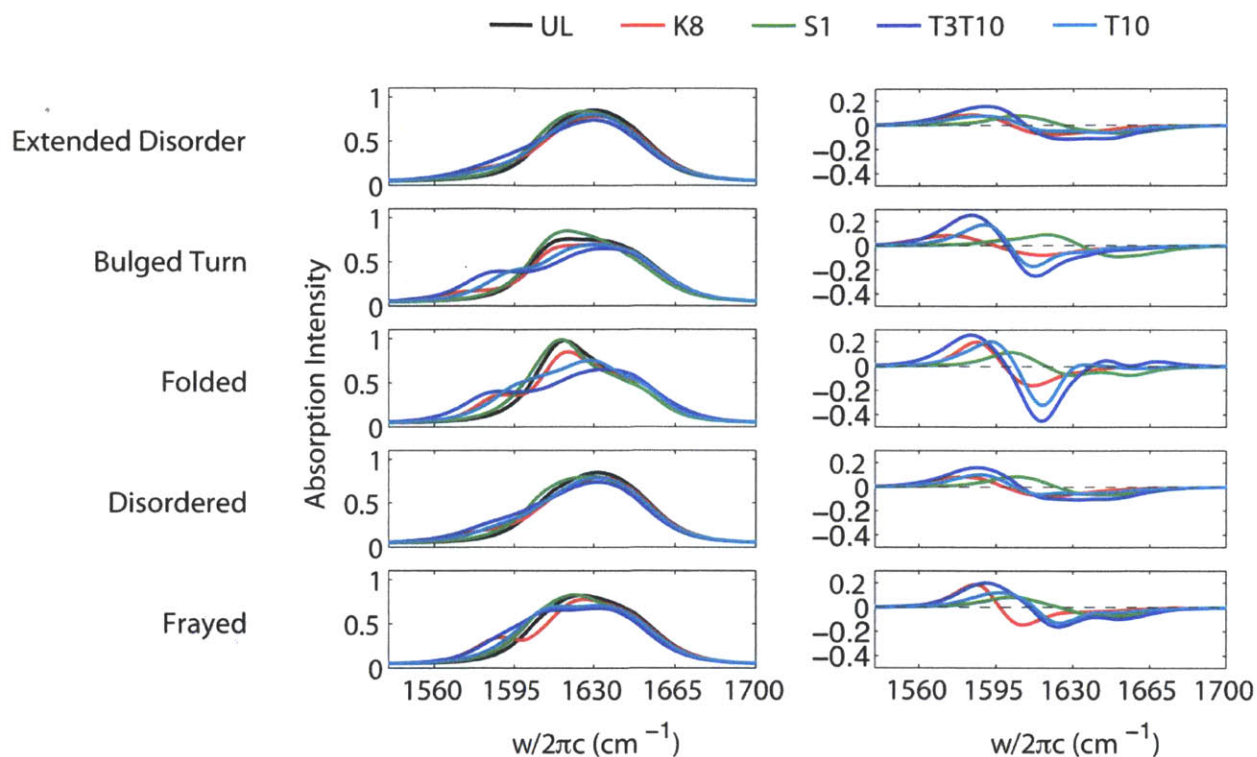
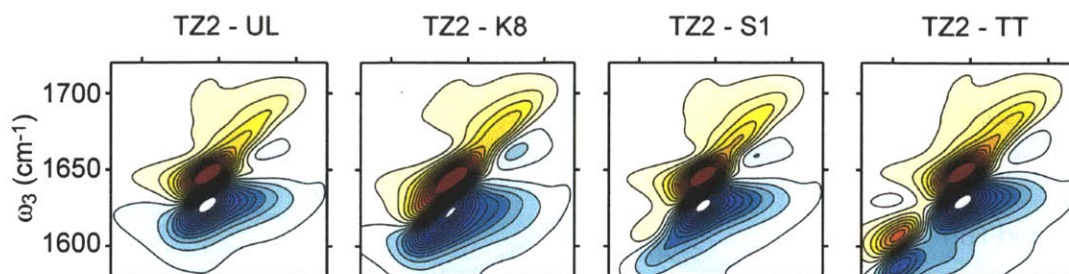


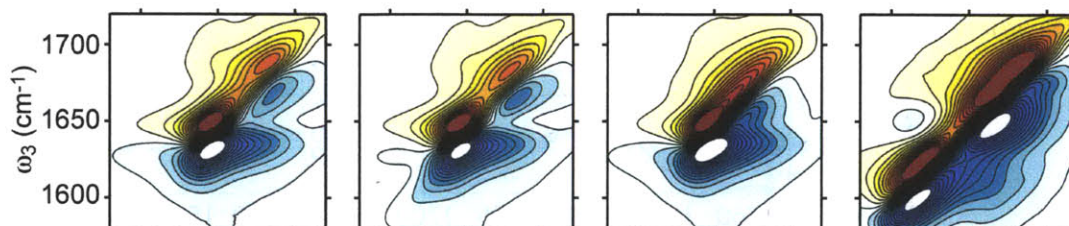
Figure 12. Simulated FTIR and FTIR difference spectra for isotopologues of all Markov States.

addition, a distinct cross peak is observed between the ν_{K8} and ν_{\perp} bands in the calculated spectra for these states. For each of the other structures (SR, BT, and ED) the K8 oxygen has higher water exposure on average, although with higher variance in hydrogen bonds formed. The $\nu_{K8}-\nu_{\perp}$ amide I' peak splitting for each of the corresponding spectra is 30-36 cm^{-1} , and the label resonance has distinctly broader diagonal linewidths of 10-14 cm^{-1} . The calculated BT spectrum shows the largest shift and line width. The observed differences can be attributed to variation in hydrogen bonding to the K8 carbonyl. Fully solvated carbonyls will on average participate in two hydrogen bonds to water, as opposed to one cross-strand hydrogen bond in the case of properly folded β turns. For the K8 carbonyl, it was found that each hydrogen bond contributes

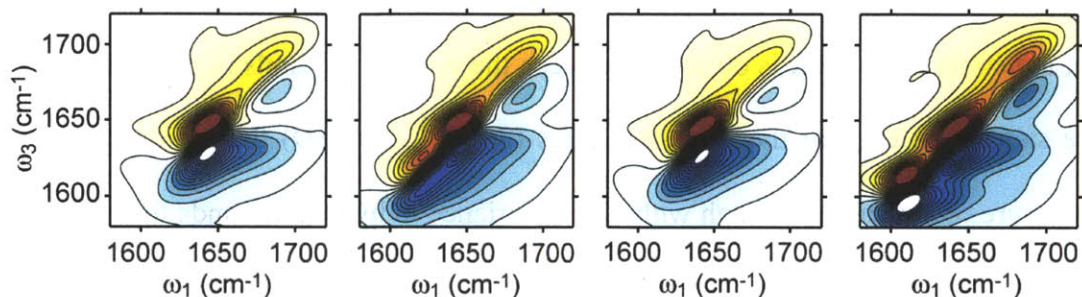
on average a 16 cm^{-1} red shift in amide I' frequency; however, the correlation between hydrogen bond number and red-shift is mild ($\rho = -0.53$). Therefore the solvent-exposed K8 conformers



NL simulation for single structure with intact S1-W12 hydrogen bond.



Slipped registry simulation for a single +1C slipped registry conformer with 4 hydrogen bonds.



Simulations of the 2D IR spectra for the NMR structure of TZ2.

Figure 13. 2D IR spectral simulations of TZ2 isotopologues for specific conformations.

should exhibit an additional red shift relative to FO, but the conformational disorder and fluctuations in peptide backbone and water hydrogen bonds will also lead to a larger inhomogeneous width. Based on these observations, one can conclude that the ν_{K8-1} peak

represents peptides with non-native turns, and the high frequency ν_{K8-2} peak reports on structures with native type I' β -turns.

During simulations of the IR spectra for the NMR structure at pH 2.5, a correlation between the K8 site frequency and two distinct E5 side chain configurations was observed. Specifically, two configurations of the glutamic acid side chain and backbone dihedrals that differed in $O(K8) \cdots \delta C(E5)$ separation by 6.1 and 6.8 Å were predicted to be observed with frequency splitting of 12 cm^{-1} . The observation of these two conformers required a protonated form of the glutamic acid (pK_a 4.6). This was ruled out as an explanation of the two K8 peaks in experiments because the amide I' resonances in FTIR and 2D IR spectra exhibit no pH dependence between pH 2.5 and 7.

The T10 peptide carbonyl adopts a variety of conformations in the Markov states simulated. In the case of the FO and FR states it is oriented for a cross-strand hydrogen bond, for BT it is entirely solvent exposed, and for SR and ED it samples a variety of configurations. In simulated spectra conformers with solvent exposed T10 carbonyls give rise to a structured excitonic spectrum and a red-shifted and spectrally distinct T10 peak, whereas the internally hydrogen bonded site gives a peak merged with the excitonic band, as observed in experiment. For the SR and ED states, the intensity of the labeled transition is greatly diminished over the intense band seen in the completely water-exposed state. Of all simulations, the FR state corresponds best to the low temperature experimental spectra. These observations demonstrate that the T10 experimental spectra provide evidence of a T10 carbonyl that participates in a single cross-strand hydrogen bond over the temperature range sampled.

For 2D simulations of TT, it was found that all conformers except ED have a resolvable red-shifted peak arising from the TT label. In the case of the disordered states SR and ED, there

is little change in the spectrum between the T10 and TT simulations, which reflects the limited correlation between the positions of their T3 and T10 peptide units. For the FO and FR conformers, which have close T3-T10 contacts, a distinct ν_{TT} peak is observed although the T10 spectrum showed no distinct T10 isotope-shifted feature. This is a result of the additional redshift of the ν_{TT} band as a result of T3-T10 amide I' coupling, which varies between 6-7 cm^{-1} for the FO conformers and leads to a $\nu_{\text{TT}} - \nu_{\perp}$ splitting of $>30 \text{ cm}^{-1}$. Comparing the T10 and TT 2D spectra for the BT state shows few differences besides additional inhomogeneous broadening of the ν_{TT} transition. Each of the simulated spectra containing proximal T3 and T10 sites bear similarity to the experimental spectra, but those that compare favorably to the calculated T10 and TT spectra are FR and FO. Finally, the intensity of the simulated ν_{TT} feature scales with the strength of coupling between the two threonine sites. These observations show that contacts between strands in the central region of the peptide exist throughout the temperature range studied. In addition, out-of-registry structures are not significantly populated in the equilibrium conditions observed here.

Simulations of the S1 spectrum were used to address the lack of a distinct peak for the ^{18}O peptide site, and the shifts in intensity and line shape seen in FTIR and 2D IR spectra. In general, for the Markov states studied, the S1 site is protected from the solvent, because the serine side chain and terminal $-\text{NH}_3^+$ both prefer to be oriented to the solvent, and because tryptophan side chains interfere with water penetration to the S1 carbonyl. Thus, the S1 site energies are generally high regardless of conformation. The FO state has many intact S1-W11 hydrogen bonds, the SR has S1 carbonyls that experience a range of solvent and side chain environments, and the FR is the most solvent exposed.

The experimental (S1-UL) FTIR difference spectrum indicates that intensity is lost at 1675 and 1640 cm^{-1} upon S1 labeling. Simulations of this spectrum for various conformers shows that this characteristic double peak loss feature matches those observed in the FO state, whereas other states show a single loss feature (see Figure 12). Cross-strand hydrogen bonding between the S1 and W11 in the FO state leads to strong coupling across the β -strand and a low frequency ν_{\perp} -like mode which is decoupled from the rest of the band and shifts upon ^{18}O isotope-labeling. However, the high frequency loss features in FTIR and 2D IR are also present in BT, SR, and FR, which have high site energies and weak coupling to the remaining peptide units. ^{18}O labeling shifts the S1 vibrational frequency for each system from very high energies ($\geq 1700 \text{ cm}^{-1}$) in the unlabeled spectrum to frequencies just to the blue of the ν_{\perp} . The distribution of S1 configurations available in these conformers suggests that the S1 site is generally disordered at all temperatures.

3.8 Discussion and Conclusions

Spectral changes to the amide I' band of TZ2, using four isotopic substitutions to probe the β -turn (K8), the N-terminus (S1), and the midstrand region (T10 and TT) of the hairpin structure, provide a site-specific description of the thermal denaturation of this β hairpin. The $^{13}\text{C}'$ label at K8 reveals two unique turn geometries, the type I' β -turn observed in the NMR structure ($\nu_{\text{K8-2}}$), and a less populated disordered or bulged loop ($\nu_{\text{K8-1}}$). Structural disorder within the $\nu_{\text{K8-2}}$ band does not increase significantly with temperature, which shows that the β -turn structures remain well-ordered. The $\nu_{\text{K8-1}}$ population in contrast does decay with temperature, indicating a relatively unstable loop. The hairpin retains interstrand contacts in the midstrand T3-T10 region over the entire temperature range studied as evidenced by the consistently high

frequency of the T10 resonance and coupling-shifted TT resonance. Unlike ν_{K8-2} , the 2D ν_{TT} line shape changes from a heterogeneously broadened peak to a more symmetric homogeneously broadened peak, indicating that this region experiences an increase in the magnitude of picosecond fluctuations in the peptide, its internal contacts, and immediate surroundings. Spectroscopy of the N-terminal serine group amide group puts constraints on the possible conformations of the termini. From these observations, this chapter has shown that S1 is free to explore a large configurational space, relative to the other labeled residues. These vary from structures in which the S1 amides are hydrogen bonded to the W11 amide to structures in which there is no interaction with the other peptide amides.

Additional information on the TZ2 turn is available from recent experiments on turn labeled TZ2-G7* and spectroscopic simulations based on the NMR structure.^{2,70} An isotope-shifted peak was well-resolved in the room temperature 2D IR spectrum, and shown to be consistent with a solvent exposed amide groups. In addition, 1 ps waiting-time 2D IR data revealed cross peak structure to the G7* site that was attributed to different solute-solvent substates. These observations are consistent with the assignments presented in this here, and can be explained with the conformer simulations. The dominant isotope-shifted peak is at 1595 cm^{-1} , reflecting the high degree of solvent exposure of the G7 carbonyl observed in the FO structure. Other conformers studied have lower solvent exposure and considerable disorder, leading to higher transition frequencies, as observed in experiment.

The picture that emerges from this data provides a detailed view of a heterogeneous thermal fraying process. The 25 °C state includes roughly 60% conformers that retain most of the cross-strand contacts seen in the NMR structure, whereas others have midstrand hydrogen bonding contacts and a bulged loop. As temperature is raised, the ends fray for all molecules,

becoming more solvent exposed, and introducing more disorder into the central contacts of the hairpin. Contacts remain within the central region of the hairpin, but on average the turn is more ordered. No idealized unfolded state with fully solvated backbone is observed. These results add structural insight to the building body of evidence that TZ2 folding is heterogeneous.^{11,19,71} Further, the data presented in this chapter provides a point of reference for testing the conformational variation and energy landscape observed by TZ2 MD simulations.^{7,10,11,27,72}

Curiously, the type I' turn appears to be favorable at high temperature. One can rationalize the origin of this apparent annealing by comparison of Markov states. The frayed state provides the possibility of packing W4 and W9 side chains about the turn which is not possible with the bulged turn. As such it appears that the desolvation entropy for water about tryptophans plays a key role in determining the relative stability of these conformers. This counterintuitive shift of turn stability could be compared with cold-denaturing transitions attributed to hydrophobic hydration.⁷³

What can be said about the TZ2 and the folding pathways of β hairpins? While it was originally assumed that there is a choice to be made between zipping²⁰ and hydrophobic collapse,⁷⁴ it appears that both of these processes play crucial roles for the folding of TZ2. For TZ2, the simulations seem to be in agreement that folding proceeds from the turn to the termini, but that this “zip-out” process involves both contacts between W4 and W9 and turn hydrogen bonding contacts.^{9,72,75} This hybrid zipper type model^{26,76} explains the varying experimental views that both the nature of the turn and the hydrophobic packing⁴ are critical to hairpin folding.

The observations presented here closely match those of Gao and co-workers,³ who recently performed MD simulations of the sequence dependent temperature-dependence of β

hairpin folding. Temperature-dependent studies of TZ2 unfolding indicated that there was a propensity for the HB closest to the turn (E5-G7) to form as the temperature is raised, even though other HBs tended to rupture. These authors used clustering analysis of states during TZ2 folding to argue that the folding mechanism from an extended state is dominated by a zip-out mechanism in which the turn forms first, although they do identify bulged configurations along the folding path.³ Hydrophobic clustering and cross-strand HBs were favorable at low temperature in the central region, whereas the frayed state was favored at high temperature.

Recently there have been several studies of peptide structure with isotope-labeled 2D IR spectroscopy that demonstrate the capabilities of this method for insight into backbone conformation and hydrogen bonding environment.^{28,30,31,77} This work indicates the potential of the method to unravel systems with conformational heterogeneity and disorder. With an appropriately large set of trial conformers and increased improvements in spectroscopic modeling, one should be able to use self-consistent modeling of multiple isotopologues to provide a detailed picture of the global range of conformations adopted by the peptide. Experimentally, one can also imagine that for cases such as TZ2, 3D IR spectroscopy on multiply labeled peptides can provide a higher level of conformational detail by correlating the local structural environment between multiple sites of the same peptide. Such experiments provide an avenue to characterize conformational heterogeneity of rapidly exchanging conformers of proteins and peptides in folded and disordered states.

3.9 Acknowledgements

The work presented in this chapter was reproduced with limited alterations from the article “Melting of a β -Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and

Simulations. Adam W. Smith, Joshua Lessing, Ziad Ganim, Chunte Sam Peng, Santanu Roy, Thomas L. C. Jansen, Jasper Knoester and Andrei Tokmakoff; *J. Phys. Chem. B* 2010, 114, 10913–10924.” 2D IR data was collected by Adam W. Smith and J.L., 2D IR line-shape analysis was performed by Adam W. Smith and Chunte Sam Peng. MD simulations and 2D IR spectral modeling was performed by Ziad Ganim, Santanu Roy, and Thomas L. C. Jansen, FTIR data collection and analysis was performed by J.L., synthesis of the UL, K8, T10, and S1 labeled peptides was performed by J.L., synthesis of the 18O labeled Thr and Ser was performed by J.L., and synthesis of the TT peptide was performed by Anaspec Inc. (San Jose, CA). The primary investigators for this work were Jasper Knoester and Andrei Tokmakoff.

3.10 References

- (1) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5578.
- (2) Jianping Wang, W. Z., Shaul Mukamel, Robin Hochstrasser *J. Phys. Chem. B* **2008**, *112*, 5930.
- (3) Lijiang Yang, Q. S., Yi Qin Gao *J. Phys. Chem. B* **2009**, *113*, 803.
- (4) Ling Wu, D. M., Rong Huang, Timothy A. Keiderling *Biochemistry* **2009**, *48*, 10362.
- (5) Smith, A. W.; Chung, H. S.; Ganim, Z.; Tokmakoff, A. *J. Phys. Chem. B* **2005**, *109*, 17025.
- (6) Karin Hauser, C. K., Rong Huang, Ling Wu, Timothy A. Keiderling *J. Am. Chem. Soc.* **2008**, *130*, 2984.
- (7) Nina Singhal, C. D. S., Vijay S. Pande *J. Chem. Phys.* **2004**, *121*, 415.

- (8) Thomas la Cour Jansen, J. K. *Biophys J.* **2008**, *94*, 1818.
- (9) Yi Xiao, C. C., Yi He *Int. J. Mol. Sci.* **2009**, *10*, 2838.
- (10) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* **2007**, *126*, 155101.
- (11) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Bio.* **2004**, *336*, 241.
- (12) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76.
- (13) Searle, M. S.; Ciani, B. *Current Opinion in Structural Biology* **2004**, *14*, 458.
- (14) Martin Karplus, D. L. W. *Protein Science* **1994**, *3*, 650.
- (15) Kim, P. S.; Baldwin, R. L. *Annual Review of Biochemistry* **1982**, *51*, 459.
- (16) Ptitsyn, O. B. *Journal of Protein Chemistry* **1987**, *6*, 273.
- (17) Christopher M. Dobson, A. Š., Martin Karplus *Angewandte Chemie, International Edition* **1998**, *37*, 868.
- (18) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83.
- (19) Wei Yuan Yang, M. G. *J. Am. Chem. Soc.* **2004**, *126*, 7758.
- (20) Victor Muñoz, P. A. T., James Hofrichter, William A. Eaton *Nature* **1997**, *390*, 196.
- (21) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15915.
- (22) Du, D.; Tucker, M. J.; Gai, F. *Biochemistry* **2006**, *45*, 2668.
- (23) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. *Biochemistry* **2004**, *43*, 11560.

- (24) Olsen, K. A.; Fesinmeyer, R. M.; Stewart, J. M.; Andersen, N. H. *Proceedings of the National Academy of Sciences* **2005**, *102*, 15483.
- (25) Anthony Mittermaier, L. E. K. *Science* **2006**, *312*, 224.
- (26) Tsai, J.; Levitt, M. *Biophys. Chem.* **2002**, *101*, 187.
- (27) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* **2006**, *124*, 141102.
- (28) Mukherjee, P.; Kass, I.; Arkin, I.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3528.
- (29) Yung Sam Kim, L. L., Paul H. Axelsen, Robin M. Hochstrass *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 7720.
- (30) Hiroaki Maekawa, M. D. P., Alessandro Moretto, Claudio Toniolo, Nien-Hui Ge *J. Phys. Chem. B* **2009**, *113*, 11775.
- (31) Sang-Hee Shim, R. G., Yun L. Ling, David B. Strasfeld, Daniel P. Raleigh, Martin T. Zanni *Proc. Nat. Acad. Sci. U.S.A.* **2009**, *106*, 6614.
- (32) Fang, C.; Wang, J.; Charnley, A. K.; Barber-Armstrong, W.; Smith, A. B.; Decatur, S. M.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2003**, *382*, 586.
- (33) Chong Fang, R. M. H. *J. Phys. Chem. B* **2005**, *109*, 18652.
- (34) Yung Sam Kim, L. L., Paul H. Axelsen, Robin M. Hochstrasser *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 17751.
- (35) Joshua Manor, P. M., Yu-Shan Lin, Hadas Leonov, James L. Skinner, Martin T. Zanni, Isaiah T. Arkin *Structure* **2009**, *17*, 247.
- (36) Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2004**, *126*, 7981.
- (37) Ganim, Z.; Tokmakoff, A. *Biophysical Journal* **2006**, *91*, 2636.

- (38) Tomoyuki Hayashi, W. Z., Shaul Mukamel *J. Phys. Chem. A* **2005**, *109*, 9747.
- (39) Choi, J.-H.; Ham, S.; Cho, M. *J. Phys. Chem. B* **2003**, *107*, 9132.
- (40) Thomas la Cour Jansen, A. G. D., Tim M. Watson, Jonathan D. Hirst, Jasper Knoester *J. Chem. Phys.* **2006**, *125*, 044312.
- (41) Thomas la Cour Jansen, J. K. *J. Chem. Phys.* **2006**, *124*, 044502.
- (42) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. *J. Chem. Phys.* **2004**, *121*, 8887.
- (43) Y.-S. Lin, J. M. S., P. Mukherjee, M. T. Zanni, J. L. Skinner *J. Phys. Chem. B* **2009**, *113*, 592.
- (44) Roberts, S. T.; Loparo, J. J.; Tokmakoff, A. *J. Chem. Phys.* **2006**, *125*, 084502.
- (45) Munira Khalil, N. D., Andrei Tokmakoff *J. Phys. Chem. A* **2003**, *107*, 5258.
- (46) Erik Lindahl, B. H., David van der Spoel *J. Mol. Model* **2001**, *7*, 306.
- (47) H. J. C. Berendsen, D. v. d. S., R. van Drunen *Comput. Phys. Commun.* **1995**, *91*, 43.
- (48) William L. Jorgensen, D. S. M., Julian Tirado-Rives *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (49) William L. Jorgensen, N. A. M. *J. Mol. Struct.* **1998**, *424*, 145.
- (50) Nora A. McDonald, W. L. J. *J. Phys. Chem. B* **1998**, *102*, 8049.
- (51) Robert C. Rizzo, W. L. J. *J. Am. Chem. Soc.* **1999**, *121*, 4827.
- (52) Melissa L. P. Price, D. O., William L. Jorgensen *J. Comput. Chem* **2001**, *22*, 1340.
- (53) Edward K. Watkins, W. L. J. *J. Phys. Chem. A* **2001**, *105*, 4118.
- (54) George A. Kaminski, J. T.-R., Richard A. Friesner, William L. Jorgensen *J. Phys. Chem. B* **2001**, *105*, 6474.

- (55) Ulrich Essmann, T. D., Lalith Perera, Hsing Lee, Lee G. Pedersen, Max L. Berkowitz *J. Chem. Phys.* **1995**, *103*, 8577.
- (56) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
- (57) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. *R. J. Chem. Phys.* **1984**, *81*, 3684.
- (58) Shuichi Miyamoto, P. A. K. *J. Comput. Chem* **1992**, *13*, 952.
- (59) David Eisenberg, A. D. M. *Nature* **1986**, *319*, 199.
- (60) Berk Hess, H. B., Herman J. C. Berendsen, Johannes G. E. M. Fraaije *J. Comput. Chem.* **1997**, *18*, 1463.
- (61) Thomas la Cour Jansen, J. K. *Acc. Chem. Res.* **2009**, *42*, 1405.
- (62) Thomas la Cour Jansen, J. K. *J. Phys. Chem. B* **2006**, *110*, 22910.
- (63) Pinchas, S.; Laulicht, I. *Infrared spectra of labelled compounds*; Academic Press (London and New York), 1971.
- (64) Adam W. Smith, A. T. *J. Chem. Phys.* **2007**, *126*, 045109.
- (65) Lazonder, K.; Pshenichnikov, M. S.; Wiersma, D. A. *Optics Letters* **2006**, *31*, 3354.
- (66) Bolhuis, P. G. *Biophysical Journal* **2005**, *88*, 50.
- (67) Bolhuis, P. G. *Proc. Nat. Acad. Sci., USA* **2003**, *100*, 12129.
- (68) Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. *Acc. Chem. Res.* **2008**, *41*, 432.
- (69) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491.
- (70) Kim, Y. S.; Hochstrasser, R. M. *J. Phys. Chem. B* **2007**, *111*, 9697.

- (71) Carsten Krejtschi, R. H., Timothy A. Keiderling, Karin Hauser *Vib. Spectrosc.* **2008**, *48*, 1.
- (72) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4077.
- (73) Urry, D. W.; Hugel, T.; Seitz, M.; Gaub, H. E.; Sheiba, L.; Dea, J.; Xu, J.; Parker, T. *Phil. Trans. R. Soc. Lond. B* **2002**, *357*, 169.
- (74) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 9068.
- (75) Wei Yang, H. N., Huan-Xiang Zhou, Bernd Berg, Rafael Brüschweiler *J. Comput. Chem* **2008**, *29*, 668.
- (76) Zhou, R.; Berne, B. J.; Germain, R. *Proc. Nat. Acad. Sci. U.S.A.* **2001**, *98*, 14931.
- (77) Hiroaki Maekawa, M. D. P., Claudio Toniolo, Nien-Hui Ge *J. Am. Chem. Soc.* **2009**, *131*, 2042.

Chapter 4

Identifying Residual Structure in Intrinsically Disordered Systems: A 2D IR Spectroscopic Study of the GVGXPGVG Peptide

4.1 Introduction

In this chapter the role that the amino acid preceding a PG turn plays in determining the turn structure of an intrinsically disordered peptide is investigated. The XPG sequence is commonly found in fibrous and elastomeric proteins, in which X influences their structural and mechanical properties. Examples include collagen (X=Pro)¹, elastin (X=Val)², mussel byssus (X=Gly)³, dragline spider silk (X=Gly)⁴, and wheat glutenin (X=Gln)⁵. To better understand the structure and folding of a XPG peptide turn, the side chain size of proline's N-terminal amino acid was modulated for the disordered peptide GVGX¹PGVG, where X = Gly, Ala, or Val. These peptides were characterized with amide I' 2D-IR in combination with structural modeling of spectra⁶⁻⁸.

Unlike other proteinogenic amino acids, the tertiary amide structure of proline changes the amide-I vibration of the peptide linkage preceding proline in sequence. The three heavy atoms bound to its nitrogen give proline's amide I' vibration a larger reduced mass, shifting it to a lower frequency. This frequency shift largely decouples proline's amide I' vibration from the amide I' vibrations of its neighbors, moving its resonance to a predominantly background free region on the red side of the absorption spectrum. As a result, proline amide I' spectroscopy can also be used as a local probe of protein structure. Because the proline amide I' vibration has almost all C=O stretch character, a frequency

shift of the proline amide I' vibration is either predominantly the result of a change in hydrogen bonding to the amide oxygen or other changes in the electric field experienced by the amide C=O. (For spectroscopic reasons, the main chain amide or carbonyl preceding proline will be referred to as the proline C=O, although it derives from the preceding amino acid). With the help of a recently parameterized structural model for proline IR spectroscopy,⁸ it is now possible to exploit this intrinsically local vibration and use it as a sensitive probe of proline turn structure.

4.2 Experimental spectra

The series of model peptides studied here differ in the volume of the side-chain on the N-terminal side of proline. The peptide with the smallest volume in the series is Gly with a volume increase of $\Delta(\text{X-Gly}) = 25.1$ and 75.3 \AA^3 for $\text{X} = \text{Ala}$ and Val respectively.⁹ The resulting change in turn conformation is detected in the experimental infrared spectra as a red shift of the low frequency proline amide I' vibration with increasing side-chain size, from Gly = 1625 cm^{-1} to Val = 1614 cm^{-1} (Figure 1). The features in β -turn peptides at 1670 cm^{-1} and 1640 cm^{-1} are commonly assigned to the parallel and perpendicular bands of an antiparallel β sheet.¹⁰ Since proline's amide I' vibration is only weakly coupled to the remaining amide units of the peptide, the observed frequency shift primarily reflects variation in hydrogen bond configurations to the oxygen of the C=O moiety preceding proline.

As a rule of thumb for a time-averaged observation, the amide frequency red shifts by roughly 16 cm^{-1} for each hydrogen bond accepted, and on average the oxygen can accept up to two.¹¹ Therefore the shifts in Figure 1 encode considerable changes in turn hydrogen bond configuration and possible changes in solvent exposure.

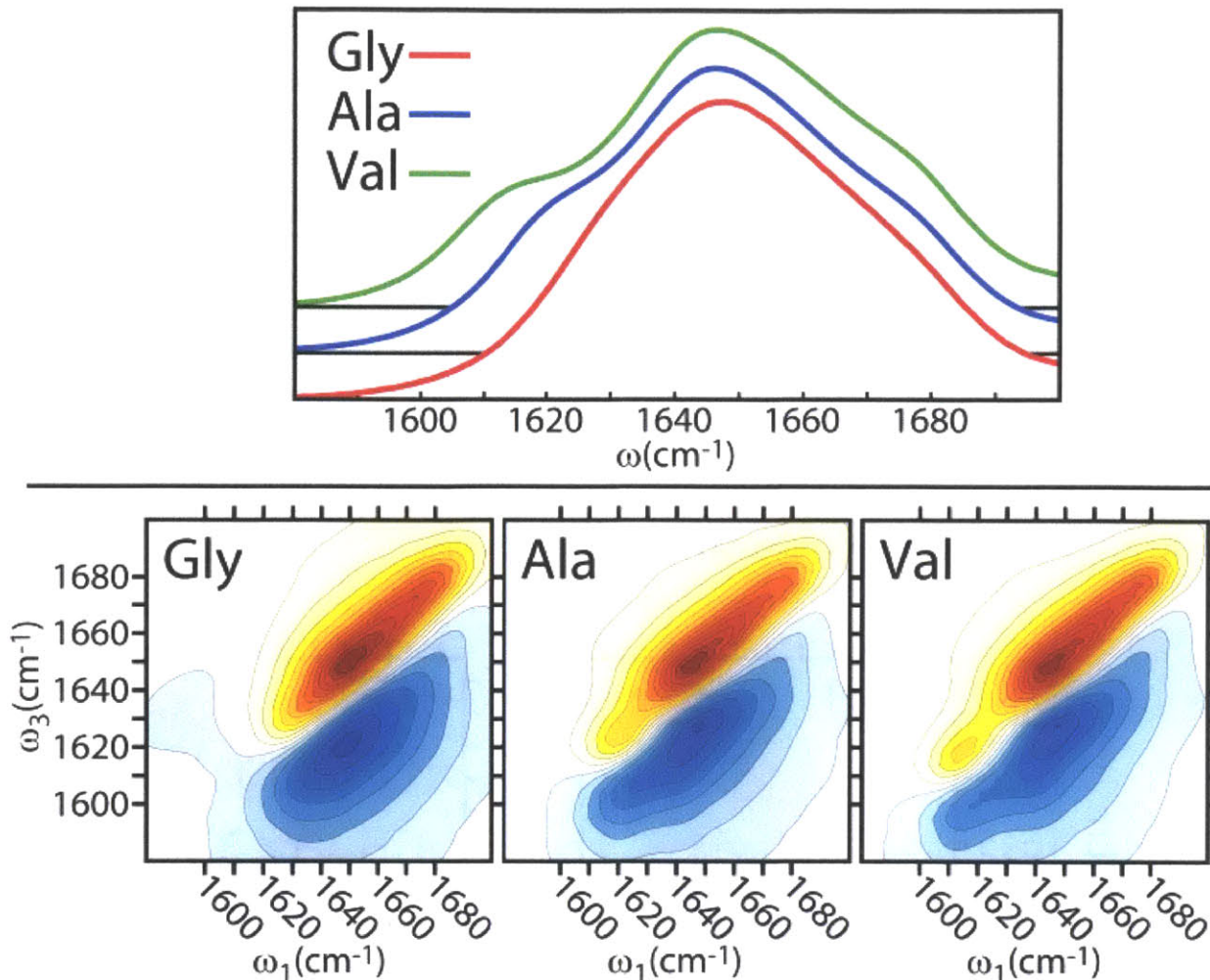


Figure 1. (Top) FTIR and (Bottom) 2D IR spectrum of GVGX¹PGVG at T = 10°C in pH = 1.0 DCl in D₂O. The proline resonance frequencies are Gly: 1625 cm⁻¹, Ala: 1619 cm⁻¹, and Val: 1614 cm⁻¹. Maximum optical densities were 0.23-0.25. Contours for all 2D IR surfaces presented in this paper are plotted from -1.0 to 1.0 in equally spaced steps of 0.08.

To perform these experiments samples for FTIR and 2D IR spectroscopic measurements were first dissolved in D₂O and lyophilized in order to exchange the acidic hydrogen's of the peptide for deuterium. The resulting dry deuterated peptide was then dissolved in a pH = 1.0 DCl in D₂O solution at a concentration of 18.2 mg/ml. No aggregation was observed in the 2D IR spectra at this concentration, in addition for the Val peptide no spectral changes were observed over a range of 10 to 50 mg/ml. Next the sample was then placed between two 1mm thick CaF₂ windows separated by a 50µm Teflon spacer creating a ~25 µL sample volume. The FTIR spectra were collected on a dry air purged Nicolet 380

FTIR spectrometer at a spectral resolution of 2.0 cm^{-1} and averaged for 64 one-second scans. The 2D IR spectra were collected as described previously using 100 fs broad band mid-infrared pulses and a ZZZZ polarization geometry¹². This data was collected using a waiting time of $\tau_2 = 150\text{ fs}$ and a τ_1 evolution time that was scanned to 3.1 ps for the rephasing spectrum and 2.5 ps for the non-rephasing spectrum in 4 fs time steps. All experiments were performed at pH = 1.0 in order to eliminate the $-\text{COO}^-$ stretch at 1590 cm^{-1} by deuterating the C-terminus creating the $-\text{COOD}$ stretching vibration at 1722 cm^{-1} . It should be noted that the findings presented here are consistent with spectra measured at neutral pH (Figure 2). In the case of Figure 2 the 2D IR data was collected in a ZZYY polarization geometry at both pH = 1.0 and 7.0 for reference with no other alterations to experimental conditions.

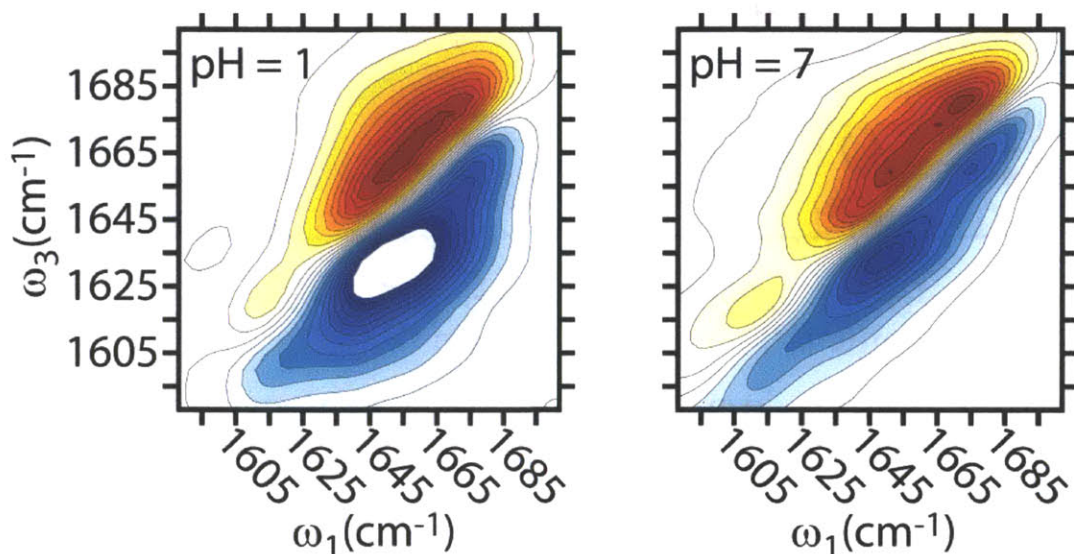


Figure 2. Val peptide 2D IR collected in ZZYY polarization at $T = 10^\circ\text{C}$ in a (Left) pH = 1.0, DCl in D_2O solution using $\tau_2 = 150\text{ fs}$, and in a (Right) pH = 7.0, 50 mmol phosphate buffer solution using $\tau_2 = 0\text{ fs}$. Changing the solvent conditions produces the appearance of on-diagonal broadening for the proline peak, a blue-shift of the main amide I' band, and a reduction in the off-diagonal line width. The change in the appearance of the proline peak results from the generation of the COO moiety upon changing the pH, creating a peak at 1590 cm^{-1} from the COO asymmetric stretch. The blue shift of the main amide I' band and the reduction in the off-diagonal line width observed at pH = 7 are likely due to the addition of buffer salts which produces a similar effect in pH = 1.0 samples.

4.3 Modeling and Simulations

To extract information on peptide structure from the infrared data, FTIR and 2D IR spectra were calculated from molecular dynamics (MD) simulations of the three peptides using a similar mixed quantum-classical model as was presented in chapter 3. A vibrational Hamiltonian was constructed for every MD snapshot by employing a Hamiltonian of the form¹³:

$$H(t) = \sum_{i=1}^N \left[\omega_i(t) b_i^\dagger b_i - \frac{\Delta_i(t)}{2} b_i^\dagger b_i^\dagger b_i b_i \right] + \sum_{ij} J_{ij}(t) b_i^\dagger b_j + \sum_{i=1}^N \mu_i(t) \cdot E(t) [b_i^\dagger + b_i]$$

Here the summations run over the amide I' sites in the peptide where for any amide site i , b_i^\dagger and b_i are the Bosonic creation and annihilation operators for the amide I' vibration, $\omega_i(t)$ is the frequency, $\Delta_i(t)$ is the anharmonicity, $J_{ij}(t)$ is the coupling between the i^{th} and j^{th} site, $E(t)$ is the applied laser field and μ_i is the transition dipole. For this scheme the Hamiltonian was considered to be constant for the 10 fs time intervals between successive saved structures of a MD trajectory.

Amide I' site frequencies (ω_i) for secondary amide groups were calculated with a Stark effect based approach using the electric field and its gradient on the C, O, N and D atoms⁶. For the proline amide unit the field components were considered on the C, O, N and C $_{\delta}$ atoms⁸. Next, frequency shifts and couplings due to through bond effects with neighboring amide units were determined using Ramachandran angle based look up tables^{7,8}. The non-nearest-neighbor couplings were then calculated using a transition charge coupling scheme^{7,8}. The anharmonicity needed for describing doubly excited states was again set to a value of 16 cm⁻¹ for all amide units¹³. A 14 cm⁻¹ systematic red-shift of the frequencies of all amide units except proline was used in this model because the frequency maps developed for the secondary amide units overestimate their frequency by the same amount.⁸ Finally an *ad hoc* vibrational life time (1 ps) was used in the spectral simulations¹⁴. The total time interval for

calculating a sample for an FTIR spectrum was 2.5 ps and for a 2D IR spectrum it was 5 ps ($\tau_1=2.5$ ps, $\tau_2=0$ ps, $\tau_3=2.5$ ps, where τ_1 , τ_2 , and τ_3 are the time delays between the pulses).

These spectral calculations were performed on eleven structurally distinct initial configurations that were selected from a 200 ns MD trajectory of the Val peptide carried out previously.¹⁵⁻¹⁸ These initial configurations are representative of the main structural and dynamical motifs observed, including extended and collapsed states, and a variety of PG turn configurations. This judicious choice of initial conformations makes it possible to use rather short simulations for computing 2D IR spectra of local structural motifs since the picosecond time scale of the experiments is far shorter than the (multi-)nanosecond relaxation times of the peptide. These eleven configurations were then mutated to obtain the initial configurations of Ala and Gly variants. Next, MD simulations in NVT ensembles at 300 K were performed on these conformers using Gromacs-4.0.7 and an all atoms OPLS force field for the peptide which was solvated in a box of TIP3P water¹⁹⁻²¹. D₂O was used to reproduce the experimental conditions and all acidic protons of the peptides were deuterated. These simulations were carried out with 2 fs time steps and the LINCS algorithm was used to constrain all bond lengths²². For temperature coupling the Nose-Hoover thermostat was used^{23,24}. The reaction field method was used to treat the long range electrostatic interactions with a cutoff of 1.4 nm. The length of each trajectory was 2.5 ns, where snapshots were saved every 10 fs yielding 2750000 conformations for each peptide studied. The simulations of these disordered peptides sampled a variety of configurations, including variations in turn structures, fraying or extension, and side-chain packing. Although these simulations may not provide a complete sampling of all configurations, it is extensive enough to provide a diversity of trial structures by which to create structure/spectrum correlations.

Turn conformers were distinguished based on hydrogen bonding patterns to the **X** carbonyl preceding the proline. Six hydrogen-bond bins were investigated: 0/0, 1/0, 0/1, 1/1, 2/0 and 0/2, where the first digit corresponds to the number of hydrogen bonds the proline CO accepts from other amide groups

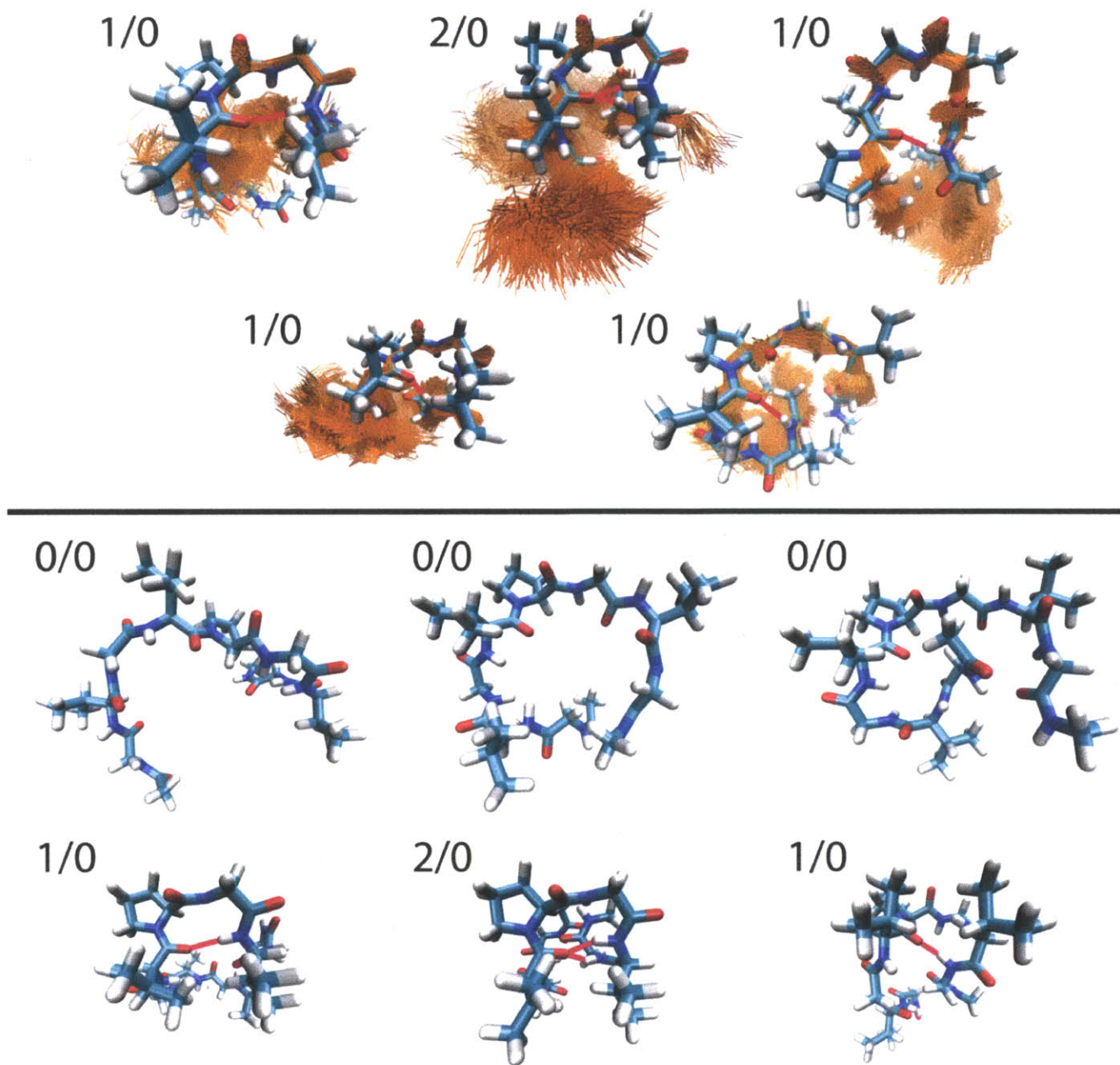


Figure 3. The images above show representative structures from the 1/0 and 2/0 hydrogen bonding bins selected from the Val MD trajectories used in this work. Hydrogen bonds are labeled for a single selected structure which is fully represented, with an ensemble of similar structures (peptide backbone only) is shown in gold for comparison.

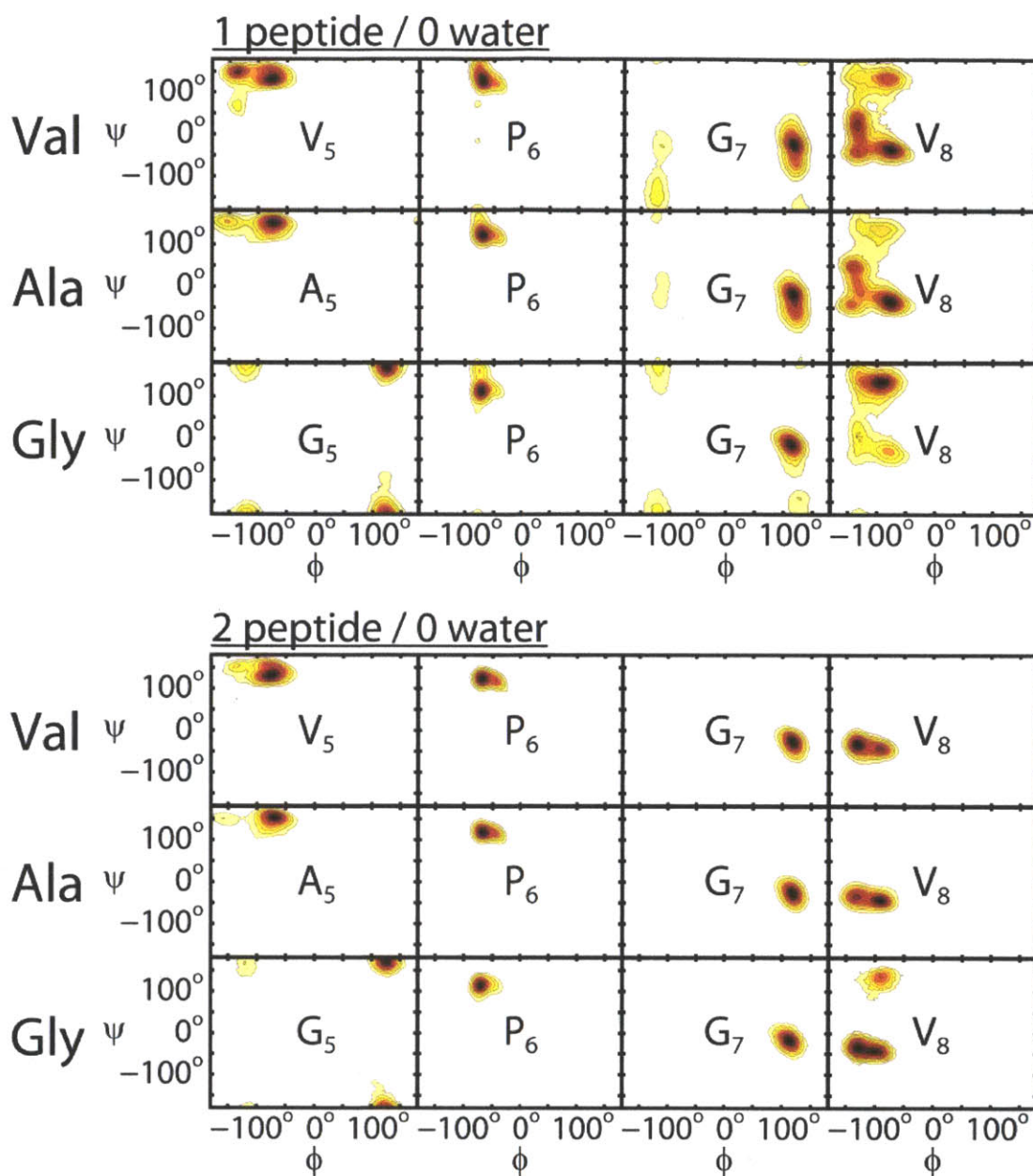


Figure 4. Ramachandran plots for the four central turn residues ($X_5P_6G_7V_8$) for the 1/0 and 2/0 hydrogen bond bins. Each plot was generated by culling structures based on hydrogen bonding criteria from 22 nsec of MD simulations for each of the three peptides GVGVPGVG, GVGAPGVG, and GVGPGGVG. Culling was performed using the `g_hbond` utility implemented in GROMACS 4.5 using an acceptor--donor--H cutoff angle of 30° and a donor-acceptor cutoff distance of 3.5 Å. Ramachandran angles were likewise extracted using the GROMACS `g_chi` utility for each trajectory.

in the peptide, and the second digit corresponds to the number of hydrogen bonds the proline CO accepts from water (Figure 3 and 4). For this chapter a peptide-peptide hydrogen bond is defined as having an

$O_{\text{peptide}} \cdots N_{\text{peptide}} - D_{\text{peptide}}$ angle $\leq 30^\circ$ and a $O_{\text{peptide}} \cdots N_{\text{peptide}}$ distance $\leq 3.5 \text{ \AA}$ and a peptide-water hydrogen bond is defined as having an $O_{\text{peptide}} \cdots O_{\text{water}} - D_{\text{water}}$ angle $\leq 30^\circ$ and a $O_{\text{peptide}} \cdots O_{\text{water}}$ distance $\leq 3.5 \text{ \AA}$. Dynamic FTIR and 2D IR spectra were calculated for each hydrogen-bond bin. This was performed by ensemble averaging only the parts of a MD trajectory belonging to a particular bin. To do this a trajectory was checked for whether a structure corresponding to a particular bin was found in the middle of the next 5 ps of the simulation. If it was found the spectrum was calculated. As expected, calculations of FTIR spectra show that the lowest frequency amide I' peak red-shifts with increasing number of hydrogen bonds to the proline (Figures 5 and 6). The simulated FTIR spectra of Gly in Figure 5

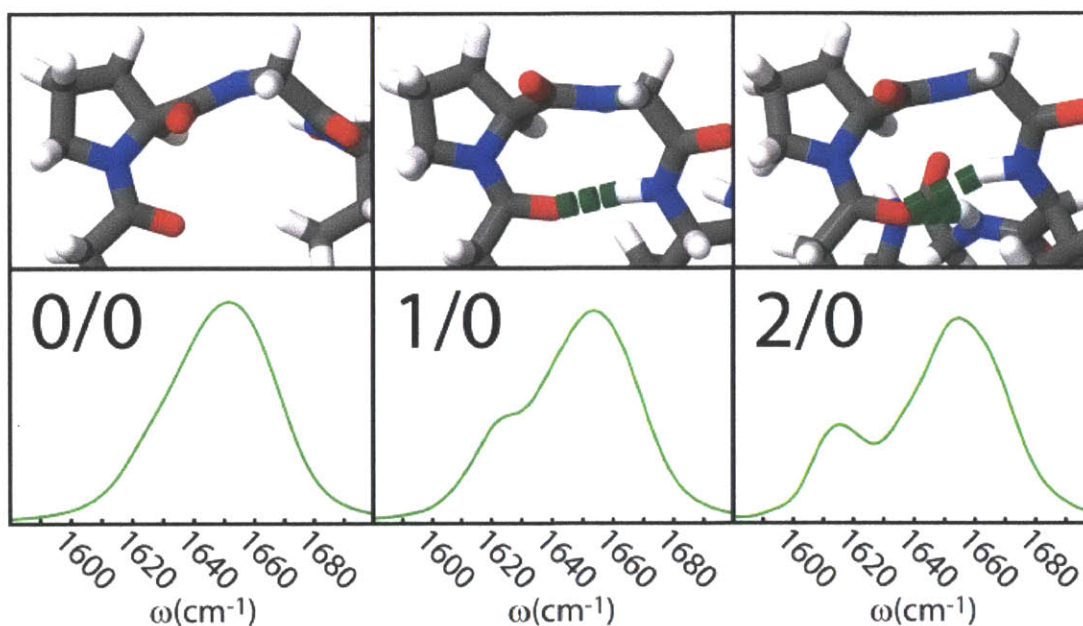


Figure 5. (Top) Visualization of representative turn structure in the 0/0, 1/0 and 2/0 bins. (Bottom) Simulated FTIR spectra for the Gly peptide in the 0/0, 1/0 and 2/0 bins.

shows that: 1) the proline resonance is not resolved for the 0/0 bin, which consists of disordered turns or condensed states that do not accommodate hydrogen bonds, 2) the proline resonance has a weakly red shifted tail and a small increase in intensity in the 1/0 bin, which predominantly consists of type II β -turns ($i \rightarrow i+3$ hydrogen bond) or bulged turns ($i \rightarrow i+4$ hydrogen bond), 3) the proline resonance has both a large red-shift and increase in intensity in the 2/0 bin, which predominantly consists of an irregular β turn

with dual hydrogen bonds ($i \rightarrow i+3$ and $i \rightarrow i+4$) corresponding to ϕ/ψ angles for the VPGV sequence of i : $-79^\circ/151^\circ$, $i+1$: $-59^\circ/127^\circ$, $i+2$: $118^\circ/-29^\circ$, and $i+3$: $-128^\circ/-30^\circ$. These simulated proline

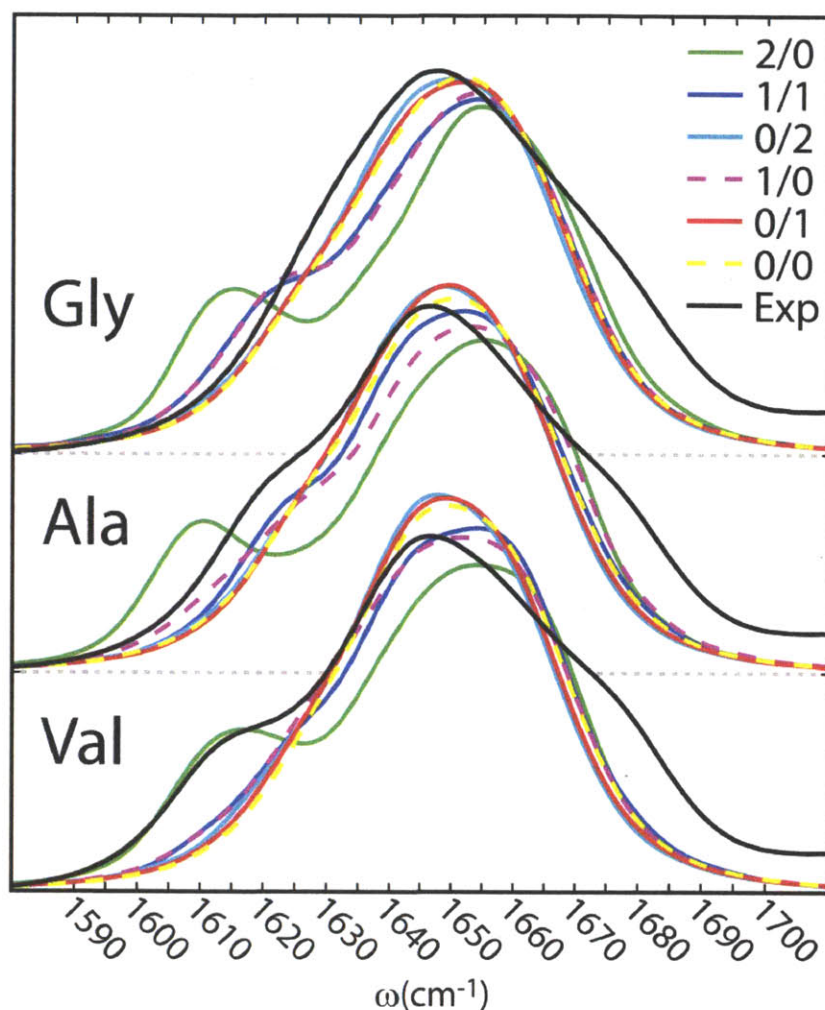


Figure 6. Comparison of experimental and simulated FTIR spectra for the Gly, Ala and Val peptides in the six hydrogen bond bins. Note the non-zero offset of the blue side of the spectrum is the result of a broad $-\text{COOD}$ stretching vibration centered at 1722 cm^{-1} .

peak shifts were also observed for the lowest frequency amide I' peak in the experimental spectra. For example a high population of the irregular β turn is observed in the experimental spectra for the Val peptide which has a proline peak with a red-shift relative to the band maximum of $(\omega_{\text{max}} - \omega_{\text{p}}) = 31 \text{ cm}^{-1}$. This model reproduces the features in the experimental spectra with the exception of the antiparallel β

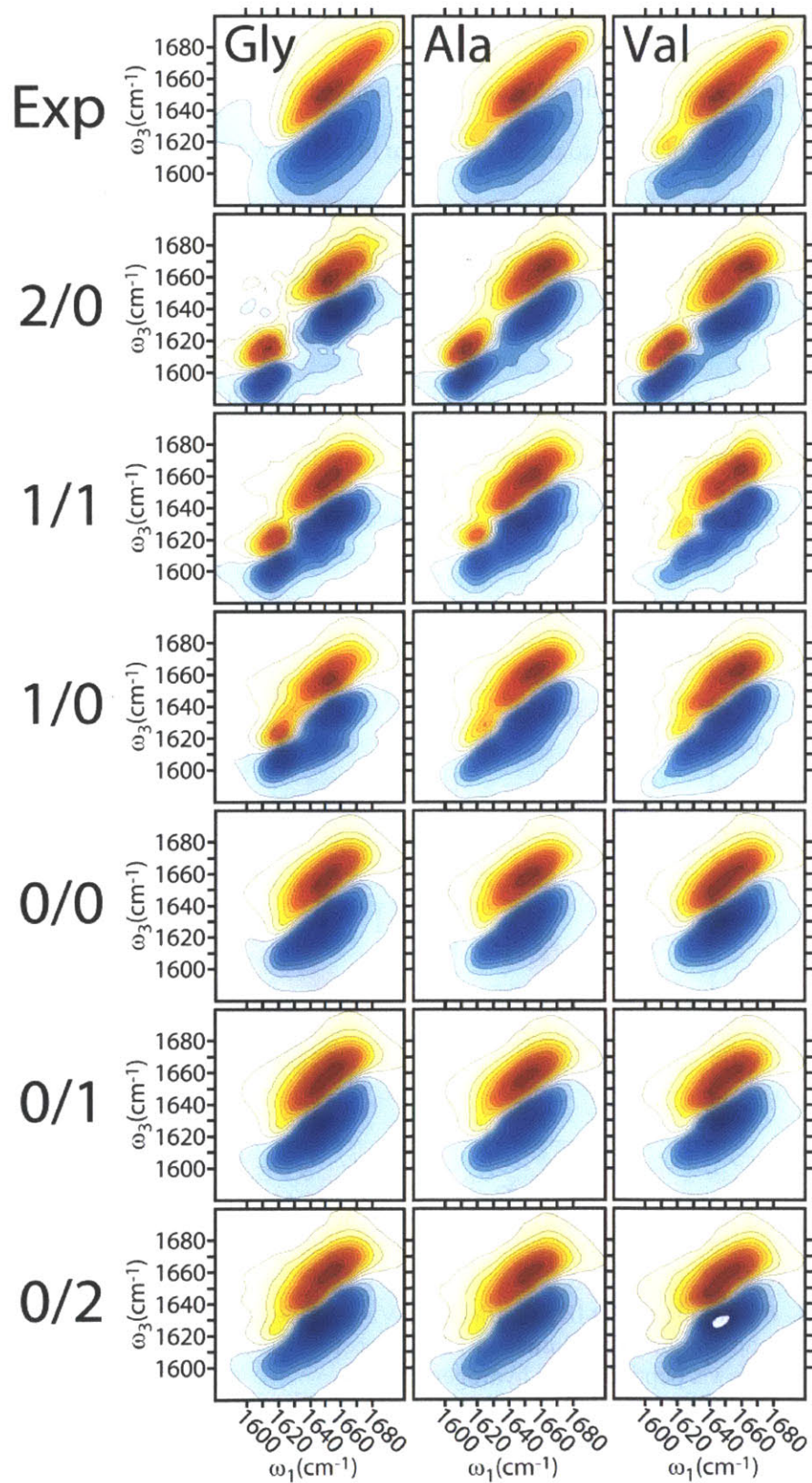


Figure 7. Comparison of the experimental data (Top row) to the six simulated surfaces generated for each peptide.

sheet's parallel mode. The parallel mode was found to be more intense in the experimental data where it is observed as a subtle shoulder on the blue side of the amide I' band (Figure 1). Since the proline red shift is the primary observable used to predict structure in this study a small difference in intensity on the blue side of the spectra does not diminish the value of the model.

To assign the hydrogen bond population of each peptide, 2D IR spectra were simulated for the six hydrogen bond bins (Figure 7). The non-linear intensity dependence in conjunction with the additional frequency dimension of a 2D IR spectrum reveals information on hydrogen bond composition not seen in FTIR spectra. The salient features common to all of the peptides simulated are summarized in Figure 8. Now better resolved, the red-shift of the proline resonance provides a count of the number of

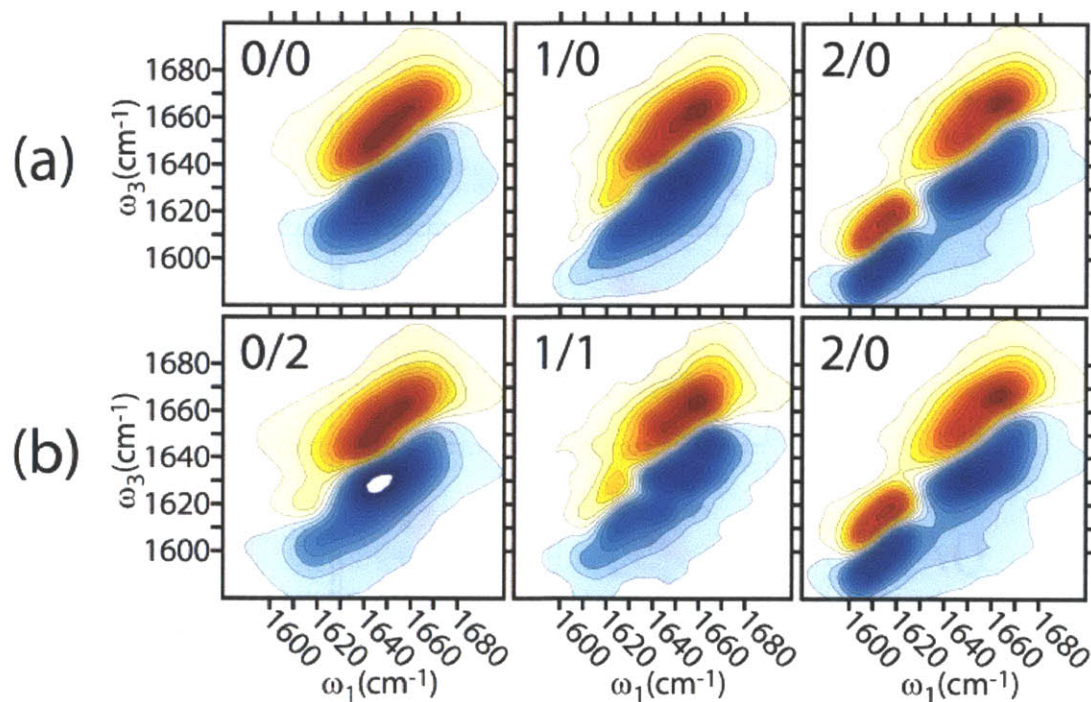


Figure 8. (a) Simulated 2D IR spectra of the Val peptide as a function of peptide-peptide hydrogen bond number. (b) Simulated 2D IR spectra of the Val peptide as a function of peptide-peptide vs. peptide-water hydrogen bonds.

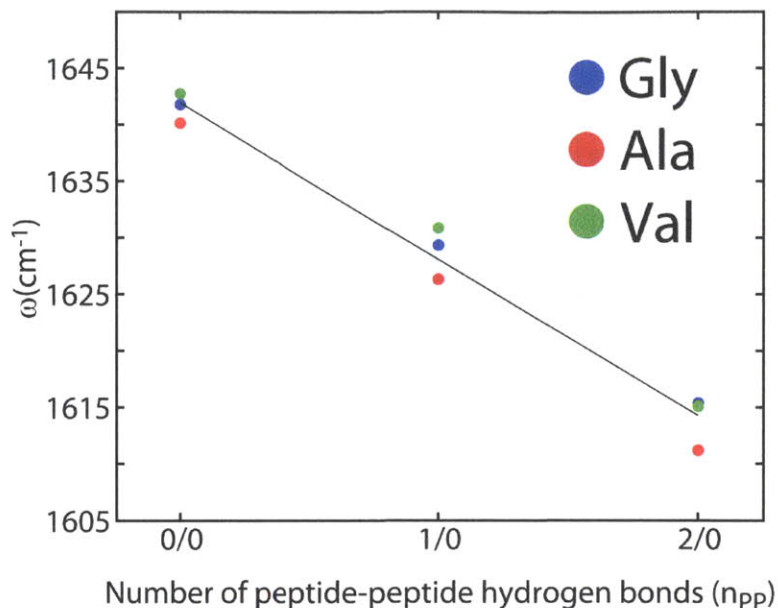


Figure 9. Average frequency of the simulated proline peak maximum for the 0/0, 1/0 and 2/0 bins. The linear fit to this data produced an equation relating the frequency of the proline peak and the number of peptide-to-proline hydrogen bonds: $\omega_p(\text{cm}^{-1}) \approx -13.8 \cdot n_{pp} + 1642$.

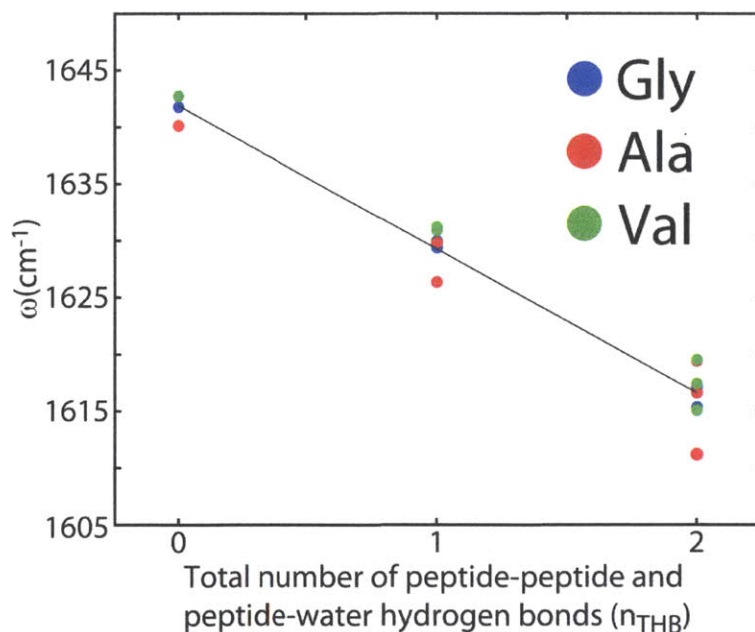


Figure 10. Average frequency of the simulated proline peak maximum for all six bins plotted as a function of the total number of peptide and water hydrogen bonds to proline. The linear fit to this data produced an equation relating the frequency of the proline peak and the total number of hydrogen bonds to proline: $\omega_p(\text{cm}^{-1}) \approx -12.6 \cdot n_{THB} + 1642$.

hydrogen bonds to proline (Figure 8a). Fitting the simulated lowest mode peak frequency as a function of the total number of hydrogen bonds to proline (n_{THB}) for all six states gives the equation $\omega_p(\text{cm}^{-1}) \approx -12.6 \cdot n_{\text{THB}} + 1642$. (Figure 9 and 10)

In addition, spectral modeling of the six bins has shown that for these peptides the proline peak is more intense when proline is hydrogen bonded to a peptide amide unit than when proline is hydrogen bonded to water. This effect is amplified by the non-linear intensity dependence of a 2D IR spectrum, thereby providing a method of distinguishing whether proline's hydrogen bonds are donated by peptide groups or water. Although precise lineshapes remain a challenge to simulate, features such as the

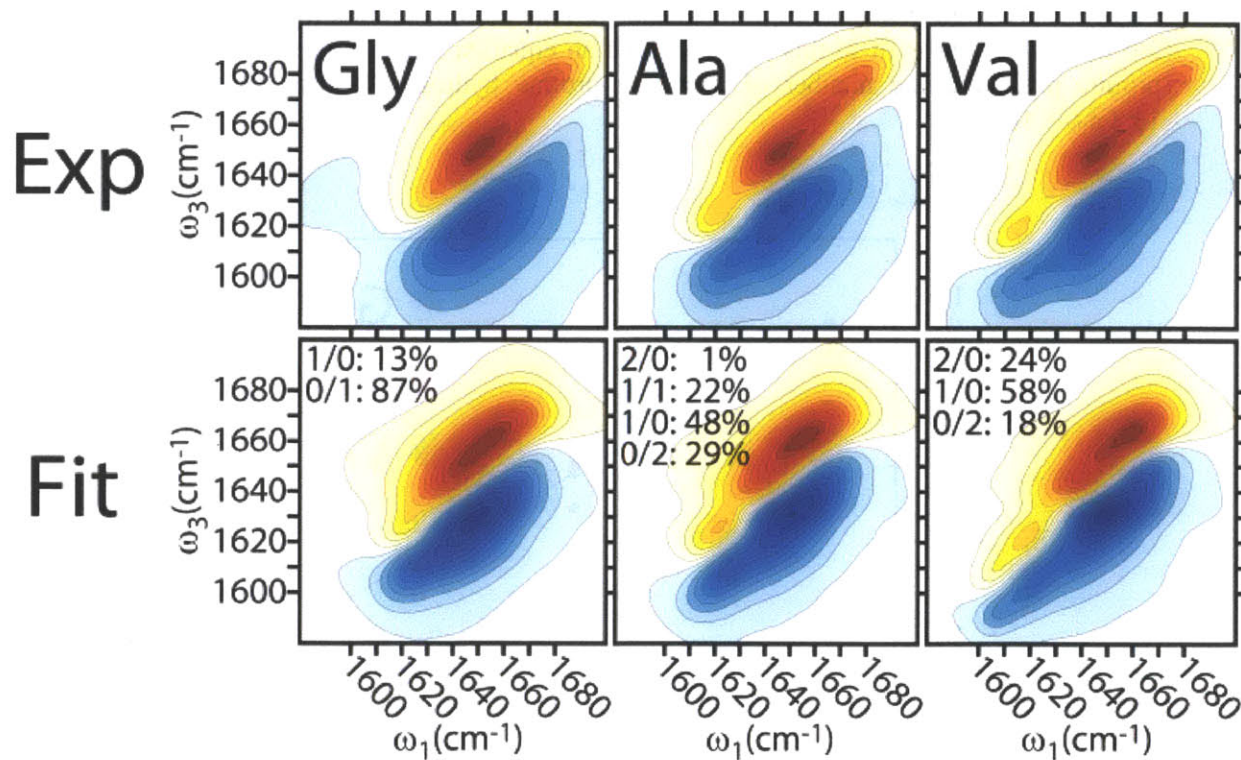


Figure 11. (Top row) Experimental 2DIR data, (Bottom row) - Fitted spectra, generated by fitting the experimental data over the frequency range 1575 to 1650 cm^{-1} with the simulated spectra for all six peptide bins using a genetic algorithm containing a least squares fitness function and no additional constraints. Since the amide I' band maximum is located at $\sim 1645 \text{cm}^{-1}$ this fit range allowed for fitting both the proline peak shift and its intensity relative to the main amide band. Listed to the right of each surface is the population of individual HB bin's predicted from the fitting routine. Contours for all 2D

IR surfaces presented in this paper are plotted from -1.0 to 1.0 in equally spaced steps of 0.08.

relative intensity or frequency splitting of the proline and main amide band are characteristic enough to uniquely distinguish those states with varying peptide-to-proline hydrogen bonded configurations. These conformers can be grouped as irregular $i \rightarrow i+3$ and $i \rightarrow i+4$ turns (2/0 or 2 hydrogen bond (HB) turn), $i \rightarrow i+3$ β -turns or $i \rightarrow i+4$ bulged turns (1/0 and 1/1, or 1HB turn), compact disordered states (0/0 and 0/1), and the 0/2 states which primarily consists of extended conformers.

In order to assign an approximate hydrogen bond composition for each peptide the experimental 2D IR surfaces were fit for frequencies $<1650 \text{ cm}^{-1}$ with the simulated 2D IR spectra for all of the peptide bins using a genetic algorithm containing a least squares fitness function and no additional constraints (Figure 11). The Gly peptide was found to contain approximately 15% 1HB turns, and $<1\%$ 2HB turns. The remainder consists of highly disordered peptide configurations. The Ala peptide was found to contain approximately 70% 1HB turns, $<10\%$ 2HB turns, and the remainder in disordered condensed or extended states. This is consistent with a high population of type II β turns and bulged turns for the folded state of the Ala peptide. Finally, the Val peptide was found to contain approximately 60% 1 HB turns and $\sim 25\%$ 2HB turns, which is consistent with a high population of type II β turns, bulged turns and irregular β turns. Although these populations are imprecise they provide a significant improvement over MD based predictions (Figure 12 and Table 1).

Bin	Population (%)		
	200 ns	27.5 ns	Fit
2/0	5.9	4.9	~ 25
1/1	3.2	1.3	<5
1/0	12.7	11.5	50-60
0/0	19.6	19.1	<5
0/1	47.1	49.9	<5
0/2	11.1	12.9	~ 20

Table 1. Bin populations calculated from a 200 ns MD trajectory, 27.5 ns MD trajectory, and the 2D IR experimental surface fitting for the Val peptide.

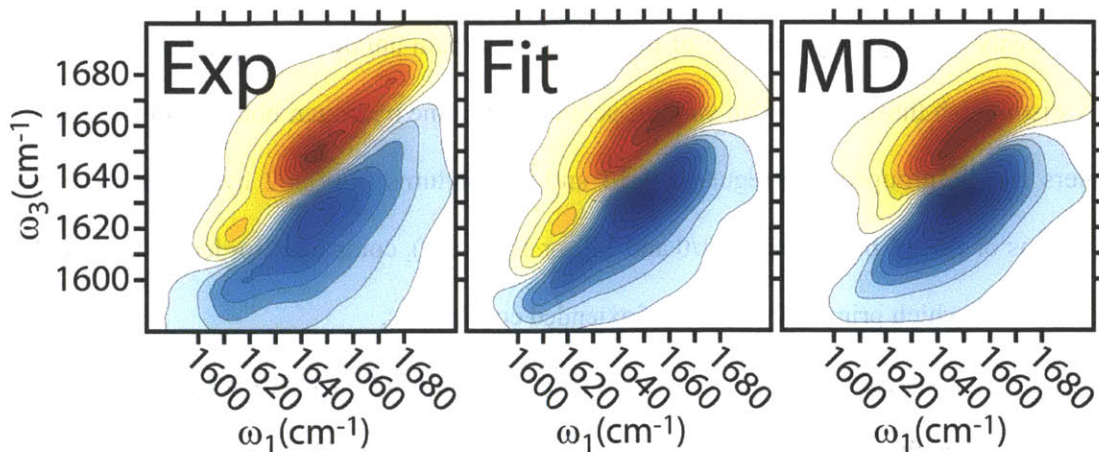


Figure 12. Comparison of the experimental spectrum for the Val peptide to simulated spectra using the bin populations generated for the fitting of the experimental data and populations calculated from a 200 ns MD trajectory.

4.4 Conclusions

The 2D IR data shows a systematic change in turn conformational preference as a function of point mutation. Here as the size of the **X** amino acid's side chain decreases the number of proline hydrogen bonds to the peptide decreases and the number of proline hydrogen bonds to water increases. This indicates that the preference for the **X**-Pro-Gly sequence to form a β -turn increases in the order Gly < Ala < Val. The observed stabilization of the **X**-Pro-Gly turn may arise from a number of factors, including steric clashes between the side chain of the point mutation and proline's pyrrolidine ring and hydrophobic clustering of the **X** and Val side chains.

The Val peptide's irregular 2/0 β -turn was found to result from the formation of dual hydrogen bonds between proline and the N-H moieties of two sequential amino acids following the PG loop. A search of 75587 PDB entries using the ϕ/ψ for the Val peptides VPGV segment in a 2/0 turn, found 158 protein segments containing a 4 amino acid sequence with these ϕ/ψ angles using a search tolerance of 30°. ²⁵ Of these sequences 21% had a proline in the *i* position, 21% had a proline in the *i+1* position, and

87% have a glycine in the $i+2$ position. Clearly, the Gly residue plays a crucial role in accommodating the tight 2/0 turn conformation in this sequence. Interestingly, a similar conformation was previously observed by X-ray diffraction in the cyclic peptide (APGVGV)₂ where the cyclic constraint generates a closed turn.²⁶

To summarize, this chapter has presented three findings which should prove useful in future studies of structure in proline-containing systems. First, the X point mutation provides a robust means of tuning the compactness or stability of turns in engineered proteins and biopolymers. Second, the frequency shift of the lowest frequency amide I' vibration in the FTIR spectrum can be employed as a measure of the number of hydrogen bonds to proline. This finding is particularly interesting because it provides a simple method for interrogating protein structure in proline rich tissue samples such as elastin and collagen matrices that are not readily addressed by other methods. Third, through the use of 2D IR it is possible to assign the hydrogen bond composition of these turns. With further refinement of the modeling methods, it will be possible to quantitatively assign the population of the different conformations in the equilibrium structural ensemble of each peptide from the 2D IR spectrum. This analysis will thereby provide a new approach for characterizing conformational states in intrinsically disordered proteins. In addition, it can be used in the future to benchmark the accuracy of commonly used force fields to describe structural and dynamical properties of small peptides.

4.5 Acknowledgments

The work presented in this chapter was reproduced with limited alterations from the article "Identifying Residual Structure in Intrinsically Disordered Systems: A 2D IR Spectroscopic Study of the GVGXPGVG Peptide. Joshua Lessing, Santanu Roy, Mike Reppert, Marcel Baer, Dominik Marx, Thomas La Cour Jansen, Jasper Knoester, Andrei Tokmakoff; J. Am. Chem. Soc. 2012, 134, 5032-5035." Peptide synthesis, FTIR data collection, 2D IR data collection, experimental data analysis, and fitting of the 2D IR surfaces was performed by J. L., the MD simulations were performed by Santanu Roy

and Marcel Baer, spectral modeling was performed by Santanu Roy and Mike Reppert. The primary investigators for this work were Dominik Marx, Thomas La Cour Jansen, Jasper Knoester, and Andrei Tokmakoff.

4.6 References

- (1) Anton V. Persikov, J. A. M. R., Alan Kirkpatrick, Barbara Brodsky *Biochemistry* **2000**, *39*, 14960.
- (2) William R. Gray, L. B. S., Judith A. Foster *Nature* **1973**, *246*, 461.
- (3) Kathryn J. Coyne, X.-X. Q., J. Herbert Waite *Science* **1997**, *277*, 1830.
- (4) Janelle E. Jenkins, M. S. C., Emily B. Butler, Randolph V. Lewis, Jeffery L. Yarger, Gregory P. Holland *ChemComm* **2010**, *46*, 6714.
- (5) Olin D. Anderson, F. C. G., Ryan E. Yip, Nigel G. Halford, Peter R. Shewry, J. M. Malpica-Romero *Nucleic Acids Res.* **1989**, *17*, 461.
- (6) Thomas la Cour Jansen, J. K. *J. Chem. Phys.* **2006**, *124*, 044502.
- (7) Thomas la Cour Jansen, A. G. D., Tim M. Watson, Jonathan D. Hirst, Jasper Knoester *J. Chem. Phys.* **2006**, *125*, 044312.
- (8) Santanu Roy, J. L., G. Meisl, Ziad Ganim, Andrei Tokmakoff, Jasper Knoester, Thomas L. C. Jansen *J. Chem. Phys.* (accepted).
- (9) Richards, F. M. *Ann. Rev. Biophys. Bioeng.* **1977**, *6*, 151.
- (10) Adam W. Smith, A. T. *J. Chem. Phys.* **2007**, *126*, 045109.
- (11) Adam W. Smith, J. L., Santanu Roy, Ziad Ganim, Chunte Sam Peng, Thomas L. C. Jansen, Jasper Knoester, Andrei Tokmakoff *J. Phys. Chem. B* **2010**, *114*, 10913.
- (12) Munira Khalil, N. D., Andrei Tokmakoff *J. Phys. Chem. A* **2003**, *107*, 5258.

- (13) Peter Hamm, M. L., Robin M. Hochstrasser *J. Phys. Chem. B* **1998**, *102*, 6123.
- (14) Kijeong Kwac, H. L., Minhaeng Cho *J. Chem. Phys.* **2004**, *120*, 1477.
- (15) Roger Rousseau, E. S., Axel Kohlmeyer, Dominik Marx *Biophys. J.* **2004**, *86*, 1393.
- (16) Marcel Baer, E. S., Axel Kohlmeyer, Roger Rousseau, Dominik Marx *J. Phys. Chem. B* **2006**, *110*, 3576.
- (17) Eduard Schreiner, C. N., Björn Ludolph, Revanur Ravindra, Nikolaj Otte, Axel Kohlmeyer, Roger Rousseau, Roland Winter, Dominik Marx *Phys. Rev. Lett.* **2004**, *92*.
- (18) Rachel Glaves, M. B., Eduard Schreiner, Raphael Stoll, Dominik Marx *ChemPhysChem* **2008**, *9*, 2759.
- (19) Berk Hess, C. K., David van der Spoel, Erik Lindahl *J. Chem. Theory Comput.* **2008**, *4*, 435.
- (20) William L. Jorgensen, J. T.-R. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- (21) William L. Jorgensen, J. C., Jeffry D. Madura, Roger W. Impey, Michael L. Klein *J. Chem. Phys.* **1983**, *79*, 926.
- (22) Berk Hess, H. B., Herman J. C. Berendsen, Johannes G. E. M. Fraaije *J. Comput. Chem.* **1997**, *18*, 1463.
- (23) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.
- (24) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (25) Adel Golovin, K. H. *BMC Bioinformatics* **2008**, *9*.
- (26) Karle, I. L.; Urry, D. W. *Biopolymers* **2005**, *77*, 198.

Chapter 5

Isotope-Edited 2D IR Study of the GVGVPGVG Peptide

5.1 Introduction

Intrinsically disordered sequences (IDSs) are amino acid sequences that do not form a stable secondary or tertiary structure in solution under physiological conditions. These IDSs instead are found to exist as a heterogeneous ensemble of conformations that may interconvert on timescales from picoseconds to microseconds. Once considered rare and ineffectual, over the past decade it has become clear that IDSs play a crucial role in biological systems.¹⁻³ Bioinformatics base predictions estimate that IDSs are common in proteins responsible for physiologic and pathologic regulation, recognition, and signaling.^{4,5} Intrinsically disordered regions (IDRs) and intrinsically disordered proteins (IDPs) are predicted to be heavily expressed in numerous diseases including cancer, diabetes, Parkinson's, Alzheimer's and cardiovascular disease.⁴⁻⁷ As a result developing the means to characterize the heterogeneous ensemble of an IDS is a required step for the advancement of human health and disease prevention.

Detailed structural characterization of samples containing a high percentage of IDSs is difficult using methods such as x-ray crystallography, fluorescence, NMR, ESR and SAXS.⁸⁻¹¹ Never the less these techniques have been employed to gain valuable insight into the properties of IDSs. For example, IDSs of ≥ 30 amino acids and 10 to 29 amino acids are observed in $\sim 10\%$ and $\sim 40\%$, respectively, of the x-ray crystal structures in the PDB¹¹ The recent identification of patterns in these sequences, through the use of bioinformatics base methods, has aided in the creation of algorithms for predicting IDSs. In solution NMR experiments, the deviation of chemical shifts and NOE values from there expected random coil values have been used to detect transient secondary structure.¹²⁻¹⁴ Never the less, none of the techniques currently used to study IDSs have a time resolution that is faster than the life time of a transient conformation in a highly disordered sequence. As a result averaged measurements are used for structure determination thereby diminishing the ability to resolve structural heterogeneity in the ensemble. One promising technique for studying IDSs is isotope-edited two-dimensional infrared spectroscopy (2D IR) because of its ability to measure protein secondary structure with picosecond time resolution.¹⁵⁻²¹ Since the conformations of an IDS exchange on a longer time scale than the 5-6 ps required for the 2D IR measurement, it is possible to use 2D IR to distinguish between transient structures in a heterogeneous ensemble.

In this chapter we use 2D IR spectroscopy in conjunction with spectral modeling from molecular dynamics simulations to identify the dominant structures that exist in the equilibrium ensemble of an elastin-like peptide. Elastin is a ubiquitous IDP found in the extracellular matrix of animal connective tissue. This protein contributes to organ morphology, provides the restoring force for bodily tissues after mechanical deformation, and is a signaling protein in the regulation of smooth muscle cell activity during morphogenesis and disease.²²⁻³² Numerous

models have been proposed for the structure of the elastin protein including the phenomenological single phase classical rubber,^{33,34} and liquid drop models,³⁵ as well as the structure based oiled coil,³⁶ β -spiral,³⁷ cis-trans isomerization,³⁸ and three phase models.³⁹ Though since the structure of elastin remains unassigned, the molecular mechanism of elastin's elasticity and signaling are still open to debate. This dearth of information on elastin's molecular structure results from both the fibrous structure of the natural protein and the high mobility of its peptide backbone. The mature elastin fiber is composed of a protein microfibril which is bound to and enveloped by a cross-linked network of the precursor protein tropoelastin where the microfibril is believed to act as a scaffold for the deposition and alignment of tropoelastin.^{40,41} The mature elastin fiber cannot be crystalized and as a result is an unsuitable sample for X-ray diffraction. The fiber is also highly insoluble in water and as a result is an unsuitable sample for solution NMR. In recent year's solid state NMR experiments have been performed providing insight into the structure of dry and semi-hydrated elastin fibers.^{42,43} Never the less a solid state NMR based structure of elastin under more native hydrated conditions has remained elusive due to the increase in mobility of elastins peptide chains upon addition of water. As a result many researchers have studied the water soluble tropoelastin as well as peptides consisting of repeating units of amino acid sequences found in the natural protein which are termed elastin-like peptides (ELP). With the exception of small cyclic ELPs,^{44,45} these samples, like the mature elastin fiber, cannot be crystalized and have fast conformational motions that are difficult to resolve with NMR. Since 2D IR spectroscopy can provide structural information on fast time scales and can interrogate non-crystalline samples it is an ideal tool for studying elastin.

5.2 Background

For humans the native elastin protein is encoded by a single gene on chromosome 7.⁴⁶ This gene codes for several domains, the two most common of which are hydrophobic and cross-linking domains. The hydrophobic domains are alanine, glycine, leucine, isoleucine, proline, and valine rich and are responsible for the elasticity of the protein.^{36,47} The cross-linking domains are either proline or alanine rich with a periodic incorporation of lysine residues for cross-linking to neighboring elastin proteins.^{36,47}

Elastogenesis begins in the rough endoplasmic reticulum of fibroblasts, or smooth muscle cells, which generate the soluble 760 residue ~70 kDa elastin protein precursor tropoelastin.⁴⁸ During synthesis tropoelastin binds to a 67 kDa galactoselectin elastin-binding protein (EBP) which functions as a chaperon by preventing the thermodynamically preferred aggregation of tropoelastin that occurs under physiological conditions.⁴⁹ After secretion from the cell the EBP interacts with galactosugars attached to the microfibril network in the extracellular space triggering the release of tropoelastin. This free tropoelastin aligns itself with the scaffolding of the local microfibril network and is then posttranslationally modified by lysyl oxidase. Lysyl oxidase is a copper-dependent catalyst that converts lysine side chains to allysine residues which are required for cross-linking.⁵⁰ Next, allysine and lysine residues covalently bond to their allysine and lysine counter parts on neighboring tropoelastins to form lysinonorleucine, merodesmosine, desmosine and isodesmosine cross-links.^{51,52} This reaction irreversibly incorporates the tropoelastin monomer into the elastin fiber. This incorporated monomer will subsequently be cross-link to newly deposited tropoelastin proteins, repeating the cycle of elastogenesis which ultimately ends in the generation of a complete elastin fiber.

The elastin fiber has remarkable material properties whose absence causes numerous pathologies. As a result elastin has been the subject of extensive study in the fields of medicine, biology and materials science. Elastin networks are crucial to organ morphology and when they are malformed as is the case in supra-avalvular aortic stenosis,^{24,29,31,32} elastoderma,⁵³ atherosclerosis,^{30,31} marfan syndrome,^{54,55} and cutis laxa,²⁸ it results in pronounced medical complications and in some cases death. In the field of material science the novel properties of elastin and its synthetic mimics have been exploited for use as particles for drug delivery,⁵⁶ transducers,⁵⁷ separatory materials,⁵⁸⁻⁶⁰ and as a scaffold for growing synthetic tissue.⁶¹

At low temperatures the elastin protein exists in an extended state, which upon heating contracts and aggregates to form a coacervate phase.⁶²⁻⁶⁸ Elastin coacervation is a thermodynamically reversible process in which tropoelastin monomers self-assemble to form a semi-ordered 50-60% hydrated phase where the monomers align their lysine rich regions for cross-linking. This counterintuitive structuring upon heating is known as an inverse temperature transition (ITT).⁶⁹ Elastin also forms a coacervate upon addition of salt or raising the pH.⁶³ In the case of human tropoelastin the optimal coacervation conditions are found at 37 °C, 150mM NaCl and pH 7-8 which is similar to the physiologic conditions found in the human body.⁶³ In its folded state above the ITT, elastin becomes a robust elastomeric material that can perform repeated cycles of extension and contraction without rupture. For example elastin in human lung parenchyma, which is under continuous dynamic stress, was found to have a mean life span of 74 years.⁷⁰ This resilience is due in part to its strength, capable of lifting weights that are 1000 times larger than the weight of the dry matrix.⁷¹

5.3 Current models for elastin

Over the years many models have been proposed to explain the viscoelasticity of elastin. In this section several of the more commonly referenced models will be discussed. Currently no one model can be used to self-consistently explain the properties of elastin. None the less these models in the aggregate provide clues as to the actual structure of the peptide.

The classical rubber model, proposed by Hovee and Flory in 1974, is one of the earliest models apply to explain elastin's viscoelasticity. For this work they performed calorimetric force extension studies on an elastin sample immersed in mixtures of water, ethylene glycol, and alcohol concluding that elastin behaves like a standard rubber. For a classical rubber the restoring force results from the configurational entropy gain associated with transitioning from an extended state where the polymer chains are aligned to a contracted state where the polymer chains are randomly oriented. Since this mechanism is based on the ordering of the polymer chains it applies to both dry rubbers and rubbers containing solvent molecules in the polymer matrix. As a result this model fails to explain the elasticity of elastin since in the absence of water elastin becomes brittle.³⁵ This also implies that interactions between water, which acts as a plasticizer, and the peptide are at the heart of the mechanism governing elastin's elasticity.⁷² An alternate entropy driven mechanism could be given in which elastin transitions from a contracted state where the hydration water is disorder to an extended state where a portion of the hydration water forms ordered networks. Although not proposed in the work of Hovee and Flory many models have been created that incorporate the effects of waters configurational entropy and will be discussed later in this section.

In an early attempt to explain water's role in the protein's mechanism of elasticity, the oiled coil and liquid drop models were proposed building on the two-phase model of S. M. Partridge.^{35,36,73,74} The liquid drop model, proposed in 1970 by Weis-Fogh and Anderson, describes the elastin fiber as a two-phase system consisting of a cross-linked network of protein globules with water in the spaces in between (figure 1).³⁵ In its contracted state these globules take on a spherical shape thereby minimizing the water-protein interface. Upon extension these spheres become elongated, generating a larger surface area and in turn bringing buried hydrophobic side-chains that are at the interior of the globules to the surface. Since the solvation

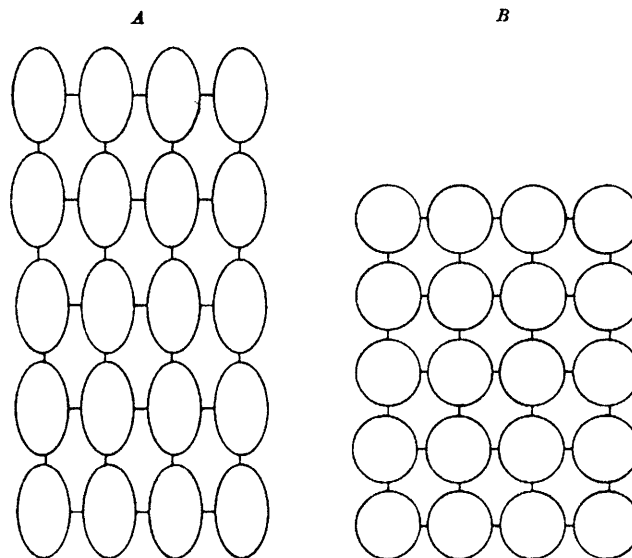


Figure 1. Structure of elastin proposed by Partridge in 1966,⁷⁴ as well as Weis-Fogh and Anderson in 1970.³⁵ They described the elastin fiber as a cross linked network of collapsed peptide globules, represented by circles and ellipses, with water in the spaces in between. Image (A) represents the extended state of the fiber and image (B) represents the contracted relaxed state of the fiber. This figure was reproduced from the work of Weis-Fogh and Anderson.³⁵

of a hydrophobic group requires the generation of an ordered network of water to envelop it this hydrophobic hydration creates the entropic driving force for the hydrophobic collapse of the protein returning it to a contracted state.

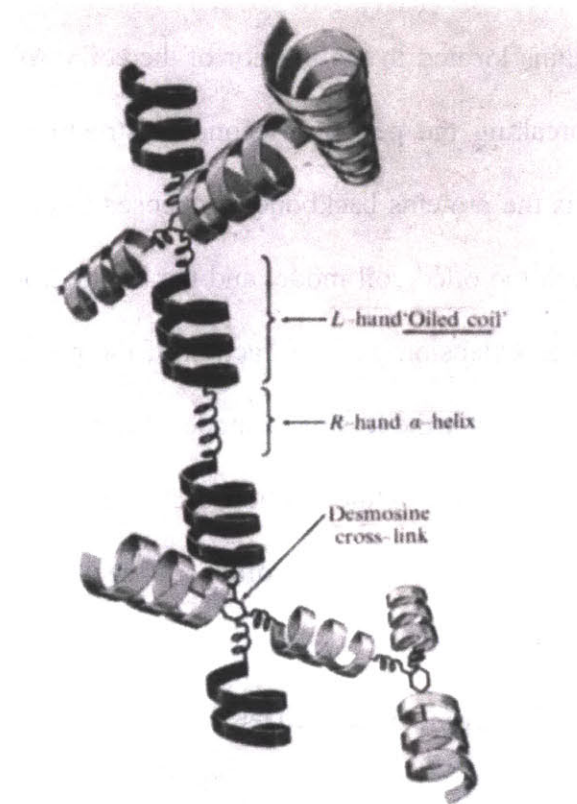


Figure 2. Proposed oiled coil model for elastin taken from the work of Gray, Sandberg, and Foster.³⁶ Here the elastin fiber consists of alternating segments of cross linking α -helical regions and elastic oiled coils.

In contrast Gray, Sandberg, and Foster in 1973 proposed the oiled coil model based on the discovery that sequences like PGVGV are found as repeating monomer units in an elastin sample isolated from copper-deficient pig aortae.³⁶ The oiled coil model like the liquid drop model also describes elastin as a two phase system where a network of cross-linked protein units are suspended in water. The difference between these models is that the liquid drop model

describes these protein units as globular whereas the oiled coil model describes them as fibrillar. The distinction here is that these fibrillar units contain a defined secondary structure (figure 2). According to the oiled coil model the peptide chain of the fibrillar unit forms a coil with broad loops 12-15 Å in diameter with Gly oriented towards the exterior of the coil and Pro, Val and other hydrophobic side chains located in the interior of the coil. As a result the recoil force in this model comes from breaking the peptides secondary structure where the primary site of hydration upon extension is the proteins backbone as opposed to the side-chains as proposed in the liquid drop model. Both the oiled coil model and the liquid drop model adequately explain the thermodynamics of force extension and contraction of the polymer and the role that water plays in generating elasticity. Never the less these models has been criticized based on the belief that the fast conformational motions of the peptide back bone would prevent the globules or fibers from remaining intact.⁷⁵ These arguments often cite the NMR experiments of Torchia and Piez that found that 80% of the backbone carbonyl carbons in a sample of calf ligamentum nuchae had a rotational relaxation time of $\tau_R \sim 40$ ns.⁷⁶

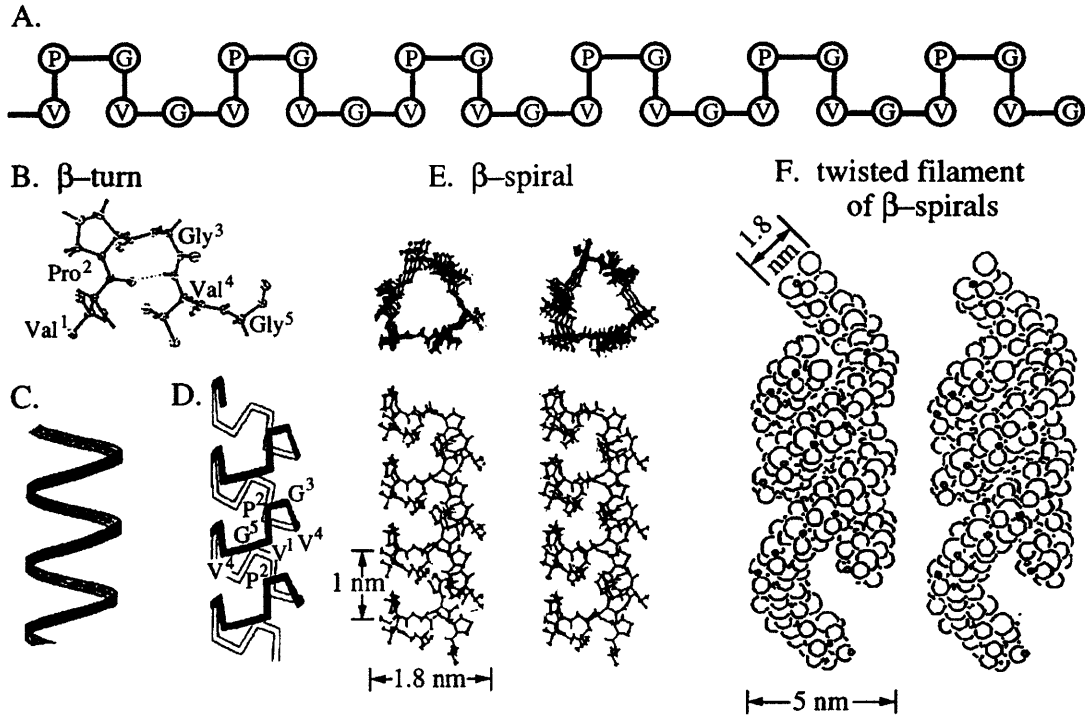


Figure 3. The β -spiral structure of elastin proposed by Dan W. Urry and coworkers. (A) Schematic representation of the poly(GVGVP) sequence demonstrating how the PG unit creates periodic β -turns. (B) PG turn structure from the crystal structure of the cyclo-(GVGVP)₃ peptide. (C and D) Helical representations of the poly(GVGVP) sequence with and without the PG β -turns. (E) Stereo plot of the β -spiral structure. (F) Proposed structure for the association of β -spirals to form twisted filaments drawn with a space-filling united residue representation. This figure was reproduced from the work of Dan W. Urry.²³

An alternative to these models was proposed by Urry which described the structure of the ELP (VPGVG)_n as a β -spiral (figure 3) based in part on the NMR and x-ray crystal structure for cyclo-(VPGVG)₃.^{37,44,77-79} In this model upon heating the ELP transitions from an extended state to a folded helical structure. The proposed helix has a pitch of 10 Å for each successive loop around the helical axis with no hydrogen bonds between successive loops. Each loop consists of three VPGVG units with the PG in a β -turn conformation that is parallel to the helical axis. Unlike the structures proposed for the oiled coil and liquid drop models the β -spiral has the

hydrophobic segments of the peptide chain on its exterior surface and has large cavities on this surface allowing water into the interior of the protein. This model is consistent with the findings from CD and UV resonance Raman studies that upon folding the population of β -turns in the sample increases.^{79,80} This model also provides an explanation for the similarity between the cyclo-(VPGVG)₃ and linear ELP NMR, CD, and UV resonance Raman (UVR) spectra.⁷⁹⁻⁸¹ Never the less since the β -spiral structure lacks a hydrophobic core this model cannot account for water's role in imparting elasticity. In addition the β -spiral model relies on the formation of a structure with long range order which is unlikely given the high conformational entropy of elastin's peptide backbone. Finally MD simulations by Daggett and coworkers have shown that an idealized β -spiral is not stable in water and upon heating collapses to form a compact amorphous state with only short range order.^{75,82,83}

Using CD and UVR measurements the work of Asher and coworkers also showed similarities between the spectrum for a linear peptide, GVG(VPGVG)₂VP, and the cyclo-(VPGVG)₃. Unlike the work of Urry and coworkers though, they did not detect a pronounced change in conformation for the linear peptide upon heating. Instead they found that the linear peptide at 20 °C has a high population of β -strands, and type III β -turns (Note: A type III β -turn contains the torsion angle $\phi_{i+1} = -60^\circ$, $\psi_{i+1} = -30^\circ$, $\phi_{i+2} = -60^\circ$, $\psi_{i+2} = -30^\circ$)^{84,85} with only a small increase in type II β -turns with heating to 60 °C. This small increase in type II β -turns is in contrast to the large increase that would be expected for the formation of a β -spiral for a longer sequence.

Unlike the work of Asher and coworkers, other authors propose more substantial changes in conformation for small ELPs upon moving through the ITT. For example Reiersen et. al. studied the temperature dependent conformational changes in a family of 15 peptides based on

the sequence GVG(VPGVG)_n where n = 1-5 using CD spectroscopy.⁸⁶ They found that all peptides studied had an ITT which upon heating is observed as a loss of random coil structure and a gain of type-II β -turn structure which may be in equilibrium with multiple β -turn conformations. In addition they found that the entropy and enthalpy change for the phase transition was similar for all peptides and as a result proposed that the folding of elastin is non-cooperative and that the VPGVG monomer was the fundamental unit of elasticity. Never the less since this work was performed on peptides ≤ 28 amino acids it does not conclusively disprove the existence of a cooperative phase transition for larger peptides

Winter, Marx, and coworkers also proposed that the small GVGVPGVG (GVGn1) peptide has an ITT. This proposal is based on CD, FTIR, differential scanning calorimetry, pressure perturbation calorimetry, and MD simulations that show two state behavior for the peptide.⁸⁷⁻⁸⁹ According to these studies the peptide is in an unfolded state at low temperatures, 10-30 °C, consisting of disordered structures and γ -loops and a folded state at temperatures above the ITT, 40-60 °C, consisting of β -strands and type II β -turns. From there MD simulations they found that upon folding the peptide forms 2-3 peptide-peptide hydrogen bonds and loses ~20 water molecules that are found to solvate backbone hydrophilic groups at low temperatures. In contrast to the liquid drop model they observe no change in the number of solvation waters encapsulating the side-chains. Thus according to Winter, Marx, and coworkers back-bone hydrogen bond contacts drive the phase transition not the waters of hydrophobic hydration. This conclusion is in direct disagreement with the liquid drop model but to a small extent is consistent with the β -spiral and oiled coil models which rely on secondary structure formation driving the ITT. Never the less since these studies were performed on an eight amino acid peptide it is not

conclusive proof that hydrophobic hydration does not drive the phase transition of longer ELP's and the native protein.

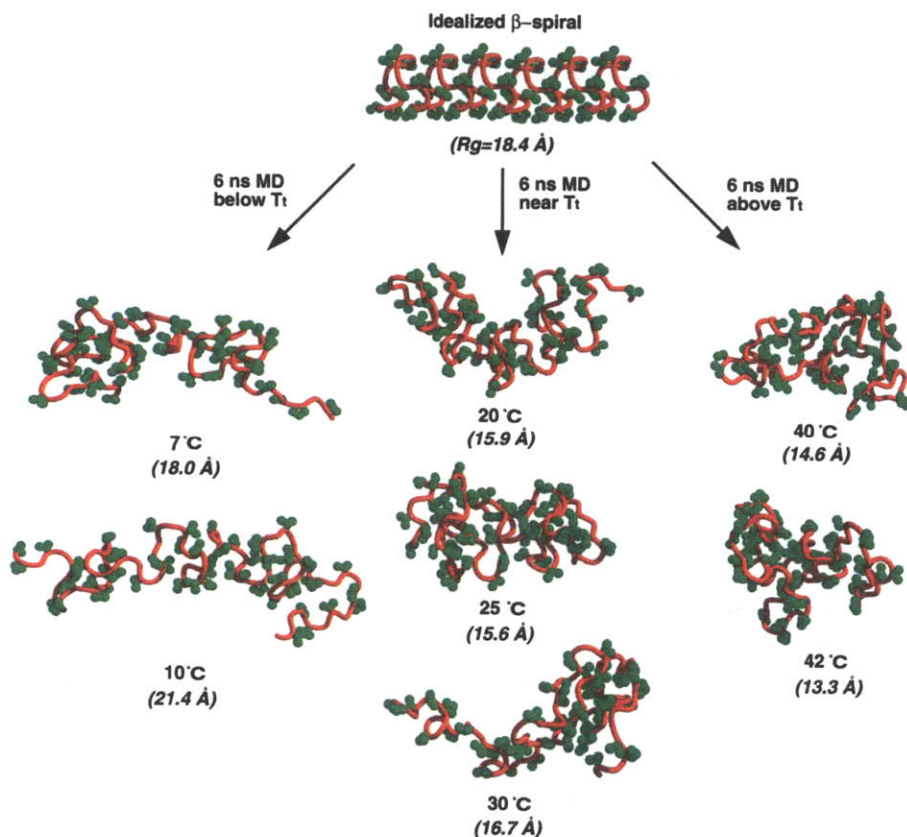


Figure 4. MD structures for the (VPGVG)₁₈ peptide at temperatures from 7 to 42 °C, 6 ns after starting from a β-spiral conformation. The peptides main chain is drawn in red and side chain atoms in green. This image was reproduced from the work of Daggett and coworkers.⁸³

In opposition to the work of Winter, Marx, and coworkers, Daggett and coworkers proposed a hydrophobic collapse model. This proposal was based on both equilibrium and force extension MD simulations for the (VPGVG)₁₈ peptide.^{75,82,83} In this description at low temperatures the peptide exists in an extended state with 36% short range β-spiral structure, existing on a 5 residue length scale, which upon heating decreases to 26% β-spiral structure as the peptide undergoes a hydrophobic collapse (figure 4). This collapse is accompanied by a

reduction in solvent accessible surface area for the hydrophobic side chains as well as an increase in the population of β -strands, β -turns, and the formation of a hydrophobic core. As a result, like the β -spiral, the model of Daggett and coworkers is consistent with the experimental observation of β -turn formation above the ITT. Unlike the β -spiral model, the model of Daggett and coworkers does contain a hydrophobic core and therefore is also consistent with thermodynamics studies on the protein under force extension. In this respect the model of Daggett and coworkers is consistent with the liquid drop model since they both claim that hydrophobic hydration plays a role in the proteins elasticity. But unlike the oiled coil and liquid drop models the globules described by Daggett and coworkers are loose with the protein moving through large amplitude cycles of extension and contraction with the backbone remaining solvated even in the collapsed state. As a result this hydrophobic globule model also does not contradict the finding that elastin has highly mobile chains in its collapsed state. One limitation to the model of Daggett and coworkers is that it is based on simulations initiated from a β -spiral structure. As a result given the short time scale of their simulations, < 100 ns in total, the conformational ensemble may be biased towards a high population of β -spiral like structures.

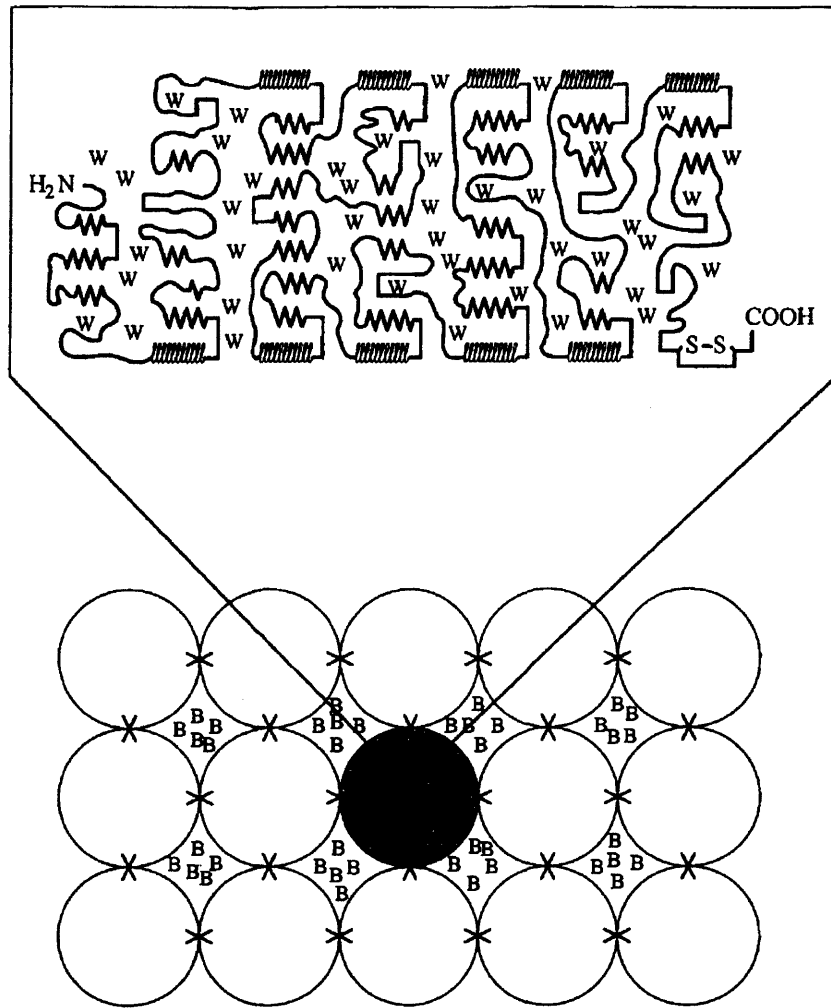


Figure 5. Depiction of the three phase model proposed by Debelle et. al. In the lower portion of the image is a cross linked (crosses) network of globular tropoelastin proteins (circles) with bulk water (B) in the interstitial space. Above is a representation of the structure with in a single tropoelastin globule. Here α -helical cross linking domains are at the surface of the globule with a semi ordered peptide rich in β turns and strands with hydration water molecules solvating these structures. This image is reproduced from the work of Debelle et. al.³⁹

Using a combination of sequence analysis, FTIR, NIR FT-Raman and CD spectroscopy on natural elastin samples, Debelle et. al. proposed a modified version of the liquid drop model of Weis-Fogh and Anderson.^{39,90-92} According to Debelle et. al., water swollen elastin is a three phase system consisting of 1) peptide chains made up of both hydrophobic and cross-linking

amino acid sequences, 2) hydration water which is tightly bound to the surface of the peptide, and 3) bulk water that resides in the interstitial space between the peptide globules (figure 5). By separating the water into two populations this model stresses the role of changes in waters conformational entropy in generating the polymers elasticity. In this description in its relaxed state the peptide chains are predominantly in unordered conformations with the bulk solvent water interacting strongly with hydration water thereby preventing the hydration water from hindering the mobility of the peptide chain. Upon stretching the conformational equilibrium shifts towards ordered chains with possible formation of β -sheet structure. Simultaneously the bulk water is expelled from the interstitial spaces between the globules which reduces the number of interactions between bulk and hydration waters thereby allowing the hydration water to restrain the peptide chains. This combination of structured restraining water and secondary structure formation generates a large drop in entropy for the fiber. Upon removing this elongation force elastin regains this entropy by recoiling according to a classical mechanism of elasticity. This model is appealing because it accounts for the role of both water and peptide conformational entropy. Never the less given the complexity of the sample used in their experiments there results cannot definitively confirm this model.

Finally an alternative to the models presented above was proposed by Valiaev et. al. who observed a temperature independent extensional transition in the single molecule force-extension measurements of SKGPG(VPGXG)₁₈₀ where X can be Gly, Ala, or Val with the substitution ratio 3:2:5. This work showed that there is a transition at 200 to 260 pN in the force extension curves for the ELP which closely resembles extension measurements for poly(l-proline). Based on this observation the authors proposed a cis-trans isomerization mechanism for elastin. Though since the extensional transition in this study is temperature independent it does not

explain the finding that tropoelastin's elasticity increases substantially upon heating past the phase transition.⁸¹ As a result the transition measured by Valiaev et. al. could be separate from the mechanism that imparts elasticity to the protein in vivo. An alternate explanation of the findings of Valiaev et. al. is that the observed transition occurs at extension forces that are beyond what is experienced by the natural fiber during regular physiologic function.

In spite of the disagreements in secondary structure assignments the studies listed above and many other show that linear and cyclic monomer, oligomer, and polymer of ELP sequences display similar spectroscopic features.^{23,37,44,45,79,80,86,87,93-97} As a consequence a thorough assignment of the vibrational spectrum of a model ELP can provide insight into the structure of the family of elastin proteins. The elasticity of elastin is generally assumed to originate from the conserved amino acid sequence PGV, PGGV, PVGG, PGVGV, PGVGVA, and PGFGVGAGV.^{36,98} The most common of these sequences both in the natural fiber and in spectroscopic studies of elastin is PGVGV. As a result the GVGn1 peptide was chosen for study with 2D IR and FTIR.

In chapter 4 it was established, based on an experimental and theoretical study of the proline amide I' vibration, that the GVGn1 peptide contains a high population of closed turn structures. This investigation is continued here with a study of isotope labeled analogs of the GVGn1 peptide. By implementing the labeling strategies outlined for the characterization of TZ2 in chapter 3 it was possible to refine the description of the peptide turn. Next the effects of salt concentration and temperature on turn conformation will be presented. Finally the effect of peptide length on the turn conformation is explored.

5.4 Experimental Overview

The library of ^{13}C O and $^{13}\text{C}^{18}\text{O}$ isotopologues synthesized for this work are shown in figure 6. This library includes the unlabeled or “wild type” (WT) peptide along with 5 single labeled peptides and 3 multiply labeled peptides. The name assigned to each amide site listed from the N to the C terminus is G3[†], V4[†], G5[†], V1, P2, G3, and V4.

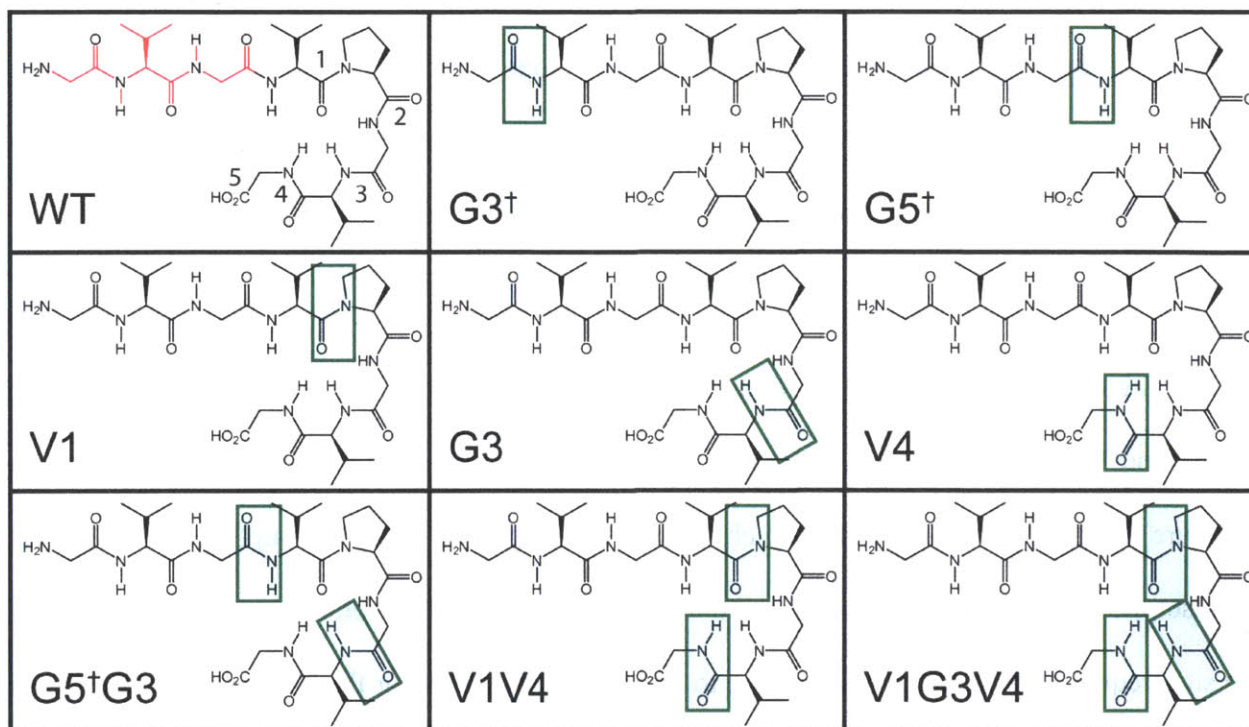


Figure 6. Library of peptides used for assigning GVGn1 structure. Boxes indicate the location of isotope labeled amide sites. All labeled amide sites were labeled with $^{13}\text{C}^{18}\text{O}$ except for the G3 site in V1G3V4 which was labeled with $^{13}\text{C}^{16}\text{O}$ for increased peak resolution. The monomer unit VPGVG is drawn in black in the WT structure with amide sites numbered to demonstrate the naming convention used in this work.

This naming scheme reflects the fact that VPGVG is a monomer unit in the elastin biopolymer. As a result the N terminal sequence G3[†]V4[†]G5[†] is numbered in accordance with their position in the previous monomer unit which is distinguished with a †. It should also be noted that this scheme assigns the name of an amide unit based on the amino acid to its N terminus. For

example an isotope label on the amide unit between G3 and V4 is referred to here as the G3 label.

Five single $^{13}\text{C}^{18}\text{O}$ labeled peptides were synthesized, the turn labels V1, G3 and V4, the mid-strand label G5[†] and the N-terminal label G3[†]. As discussed in chapter 3 isotope labeling decouples the vibrations of an amide site from those of its neighbors providing a local probe of peptides structure. By studying the line-shape of the labeled peak it is possible to gain local information on hydrogen bonding, structural heterogeneity and the rate at which these heterogeneous environments interconvert. In addition evaluating the frequency dependent changes in the main amide I' band and cross peaks generated upon labeling provides clues as to the sites contribution to the secondary structure of the peptide.

The dual $^{13}\text{C}^{18}\text{O}$ labels G5[†]G3 and V1V4 and the triple label V1G3V4, with $^{13}\text{C}^{18}\text{O}$ on the V1 and V4 sites and ^{13}CO on the G3 site, were synthesized for measuring peptide secondary structure in analogy to the TT label in chapter 3. In the case of TZ2 the coupling of the T3 and T10 isotope labels across a semi-ordered unit of β -sheet structure generated a measurable shift in the site frequency of the TT peak relative to the single label peak for the T10 peptide. The same approach was taken here for the multiply labeled peptides by comparing the FTIR and 2DIR spectrum for each multiple label to their corresponding single label analogs. Since the labels in the G5[†]G3, V1V4 and V1G3V4 peptides are located at positions around the turn they are capable of reporting on changes to the turns structure.

These labels are first presented in the section below at 10 °C to provide a description of the average structure of the peptide. Next temperature and salt dependent data will be presented to demonstrate the changes that occur to the structure of the turn as the peptide is driven through

the ITT. Finally data for GVG(VPGVG)_n where n = 1-6 and (GVGVP)₂₅₁ will be presented to determine if the length of the peptide effects the turn conformation.

5.5 Experimental Methods

All peptides presented in this chapter with the exception of the (GVGVP)₂₅₁ were prepared using Fmoc-based solid phase peptide synthesis as described in chapter 2 using Gly-Wang resin. Peptides synthesized using SPPS were purified by reverse phase HPLC using a two-phase buffer gradient: (buffer A) 0.1% TFA in H₂O and (buffer B) 80% acetonitrile, 0.085% TFA in H₂O and were desalted where indicated below. These peptides were subsequently 3x dissolved in an HCl solution and lyophilized to remove TFA from the sample.^{99,100} The recombinantly expressed (GVGVP)₂₅₁ was obtained from Dan W. Urry of Bioelastics Research Ltd. and was used without further purification. Isotope labeled Fmoc amino acids were purchased as 99% enriched 1-¹³C labels from Cambridge Isotope Laboratories. To create the ¹³C¹⁸O labels the carboxyl oxygen's of the 1-¹³C labeled Fmoc amino acids were then labeled with ¹⁸O to ~97% isotope enrichment via acid hydrolysis in H₂¹⁸O purchased from Isoflex USA.

Samples for 2D IR and FTIR experiments were first dissolved in D₂O and lyophilized in order to remove H₂O and exchange acidic protons in the sample for deuterium. The dry deuterated product was then dissolved in pH = 1.0 DCl in D₂O with 0, 75 or 150 mmole of KD₂PO₄ or pH = 1.0 DCl in D₂O with 150 mmole of the mixture K₂DPO₄ and KD₂PO₄ in a 1:1 molar ratio. A sample of the 150 mmols solution of the K₂DPO₄ and KD₂PO₄ mixture was analyzed by Northeast Environmental Laboratory, Inc. (Danvers, MA) which measured the concentration of the potassium, chloride and orthophosphate ions to be K⁺ = 207 mmole, Cl⁻ =

256 mmole, and $\text{PO}_4^{3-} = 130$ mmole. Peptide samples were prepared at concentrations from 15.0 to 50.0 mg/ml with no spectral changes observed over this range. Next samples were placed between two 1mm thick CaF_2 windows with a 50 μm Teflon spacer generating a ~ 25 μL sample volume. The sample was then mounted in a temperature controlled brass sample holder. FTIR data was collected with a Nicolet 380 FTIR spectrometer in a dry air purged environment at a resolution of 2.0 cm^{-1} . 2D IR spectra were collected in a dry air purged environment using 100 fs broad band mid-infrared pulses with a τ_1 evolution time that was scanned to 3.1 ps for rephasing spectra and 2.5 ps for non-rephasing spectra. Data was collected in both *ZZZZ* and *ZZYY* polarization geometries at a variety of waiting times ranging from $\tau_2 = 0.0$ to 1.5 ps as will be indicated below.

5.6 FTIR and 2D IR spectra

The FTIR spectra for the desalted GVGn1 isotopologues at 10 °C in pH = 1.0 DCl in D_2O with 150 mmole of $\text{KD}_2\text{PO}_4\text{:K}_2\text{DPO}_4$ (1:1 by mol) are presented in figure 7. The amide I' spectrum of GVGn1 contains several notable peak features which will be briefly reviewed in this section. The peak at $\omega = 1614$ cm^{-1} , in the FTIR spectra of WT, G3^\dagger , G5^\dagger , G3, V4 and $\text{G5}^\dagger\text{G3}$, is attributed to the amide I' vibration of proline. The peaks at $\omega = 1640$ and 1670 cm^{-1} are assigned to the ν_\perp and ν_\parallel antiparallel β -sheet modes.^{101,102} The peak at $\omega = 1660$ cm^{-1} is assigned to random coil vibrations since GVGn1 is an eight residue peptide with a central proline and is therefore unable to form an α -helix. Finally, the non-zero offset on the blue side of the spectrum is the result of the COOD stretching mode that is centered at ~ 1725 cm^{-1} . Isotope labeled peaks appear on the red side of the spectrum with peak maxima at frequencies ranging from 1554.4 cm^{-1}

¹ for the ¹³C¹⁸O labeled V1 site in the V1 peptide to 1610.3 cm⁻¹ for the ¹³CO labeled G3 site in the V1G3V4 peptide. The combination of the ¹³CO, ¹³C¹⁸O, and proline red shifts of 41, 60, and 27 cm⁻¹ respectively can provide a high degree of site specific resolution in a single spectra as evidenced by the V1G3V4 peptide.

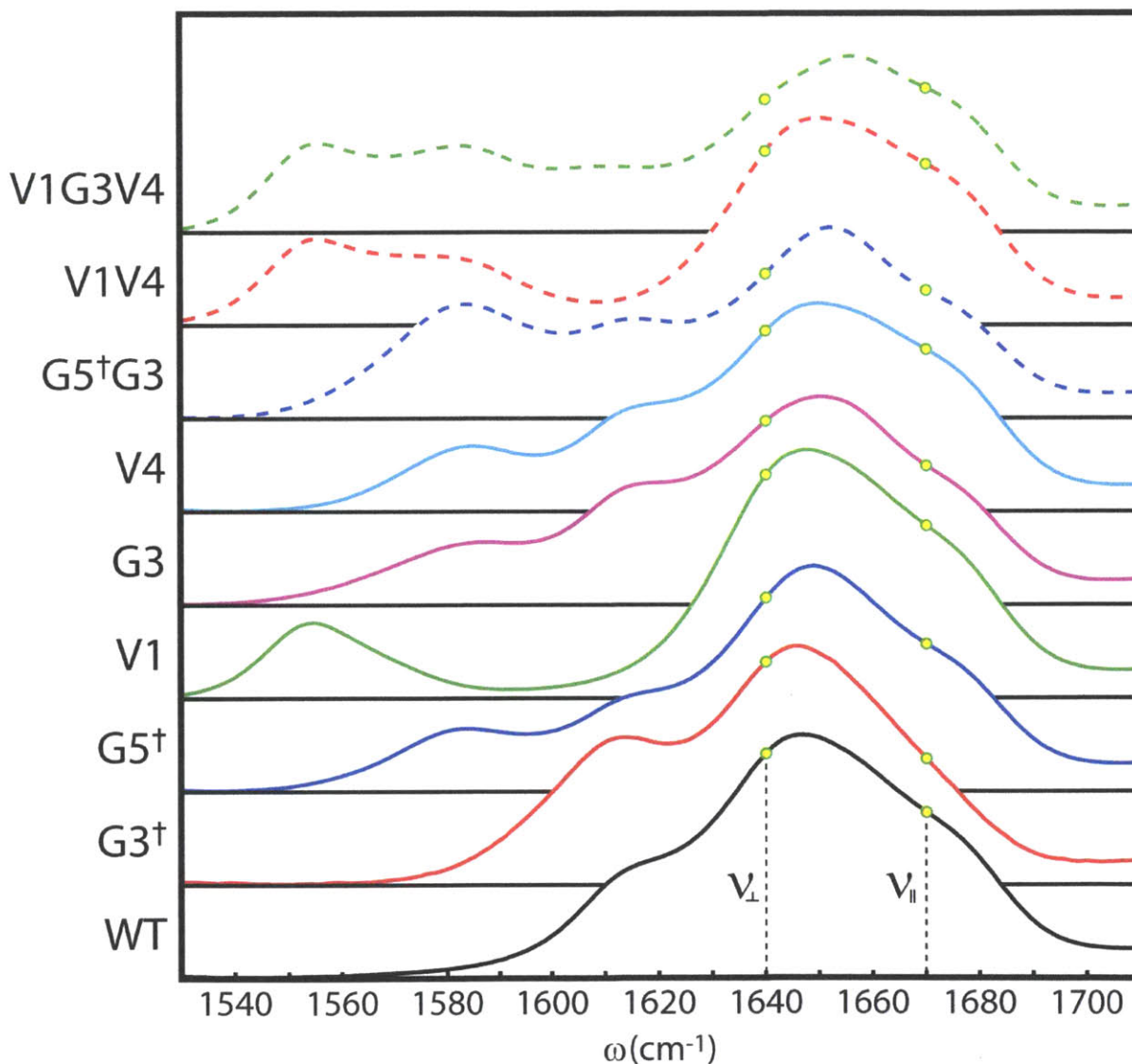


Figure 7. FTIR spectrum for the GVGn1 peptide library collected at 10 °C on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). The green and circle are placed at 1635 and 1675 cm⁻¹ to indicate the location of the β-sheet ν_{\perp} and ν_{\parallel} modes respectively.

The 2D IR spectra for the desalted GVGn1 isotopologues at 10 °C with a waiting time of $\tau_2 = 150$ fs in pH = 1.0 DCl in D₂O with 150 mmole of KD₂PO₄:K₂DPO₄ (1:1) are presented in figure 8. Due to the non-linear scaling of intensity in a 2D IR surface peaks are better resolved

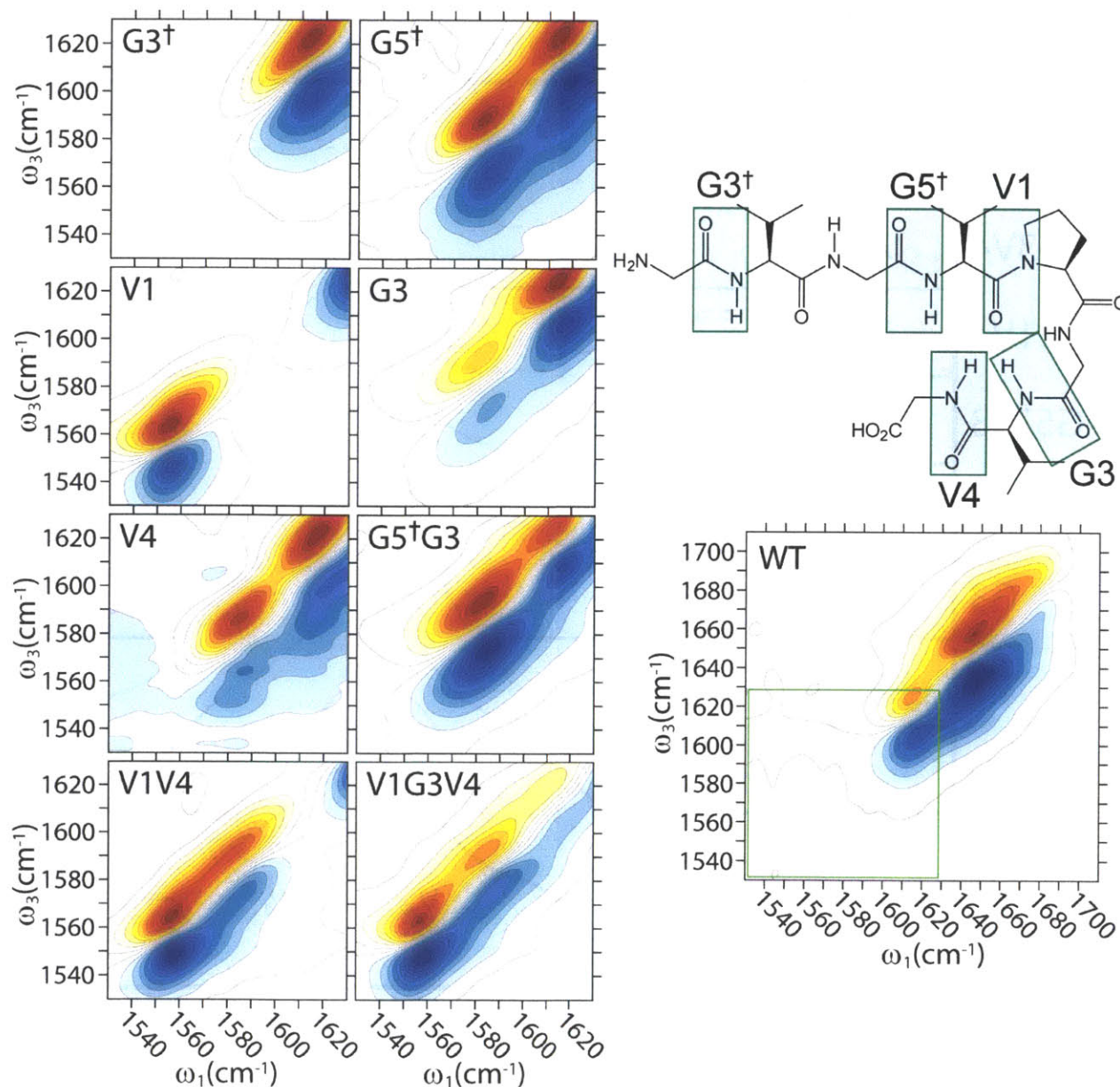


Figure 8. 2DIR spectrum for the library of GVGn1 peptides used for assigning the peptides structure. Spectrum were collected in ZZZY polarization at 10 °C and $\tau_2 = 150$ fs on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). Note that due to the desalting of the peptides these 150 mmol spectrum contain line-

shapes similar to spectrum collected for peptides prepared without a desalting step and dissolved in pH = 1.0 DCl in D₂O with no additional salt added.

than in the FTIR spectrum making the identification of peak features from a 2D IR surface a robust metric for characterization. With the exception of G3[†] the 2D IR surfaces for all of the peptides contain well resolved isotope peaks. As a result these data sets can be easily fit with Gaussian peaks and interpreted using basic line-shape metrics such as center frequency (ω), on-diagonal FWHM (Δ), off-diagonal FWHM (Γ), peak rotation and ellipticity ($E = (\Delta^2 - \Gamma^2)/(\Delta^2 + \Gamma^2)$). Fitting the 2DIR surfaces gave peak frequencies of G5[†]: $\omega = 1582.0 \text{ cm}^{-1}$, V1: $\omega = 1554.4$ and 1572.2 cm^{-1} , G3: $\omega = 1581.0 \text{ cm}^{-1}$, and V4: $\omega = 1584.1 \text{ cm}^{-1}$. In the case of the G3[†] peptide it was not possible to extract the isotope peak frequency from the 2D IR spectrum because its resonance is nearly degenerate with that of unlabeled proline. As a result FTIR difference spectra were used to assign the isotope peak frequency giving an isotope peak of G3[†]: $\omega = 1604.4 \text{ cm}^{-1}$ (figure 9). Values for ω , Δ , Γ , and E are presented for all labeled sites in table 1.

	ω	Δ	Γ	E
G3 [†]	1604.4			
G5 [†]	1582.0	25.5	9.6	0.75
V1 (2/0)	1554.4	19.3	9.3	0.62
V1 (1/0)	1572.2	16.3	10.0	0.46
G3	1581.0	34.6	10.4	0.84
V4	1584.1	24.4	7.6	0.82

Table 1. Center frequency (ω), on diagonal FWHM (Δ), off diagonal FWHM (Γ), and ellipticity (E) for the isotope labeled sites of GVGn1. The values for ω , Δ , and Γ are presented here in cm^{-1} and the value for the ellipticity is given by $E = (\Delta^2 - \Gamma^2)/(\Delta^2 + \Gamma^2)$. Peak metrics for all compounds with the exception of G3[†] were measured by fitting their

2D IR spectra. Since the $G3^\dagger$ isotope peak is partially overlapped in frequency with the proline resonance it was not possible to accurately fit its 2D IR surface. As a result the value for ω was measured from the $[G3^\dagger - WT]$ FTIR difference spectrum presented in figure 9.

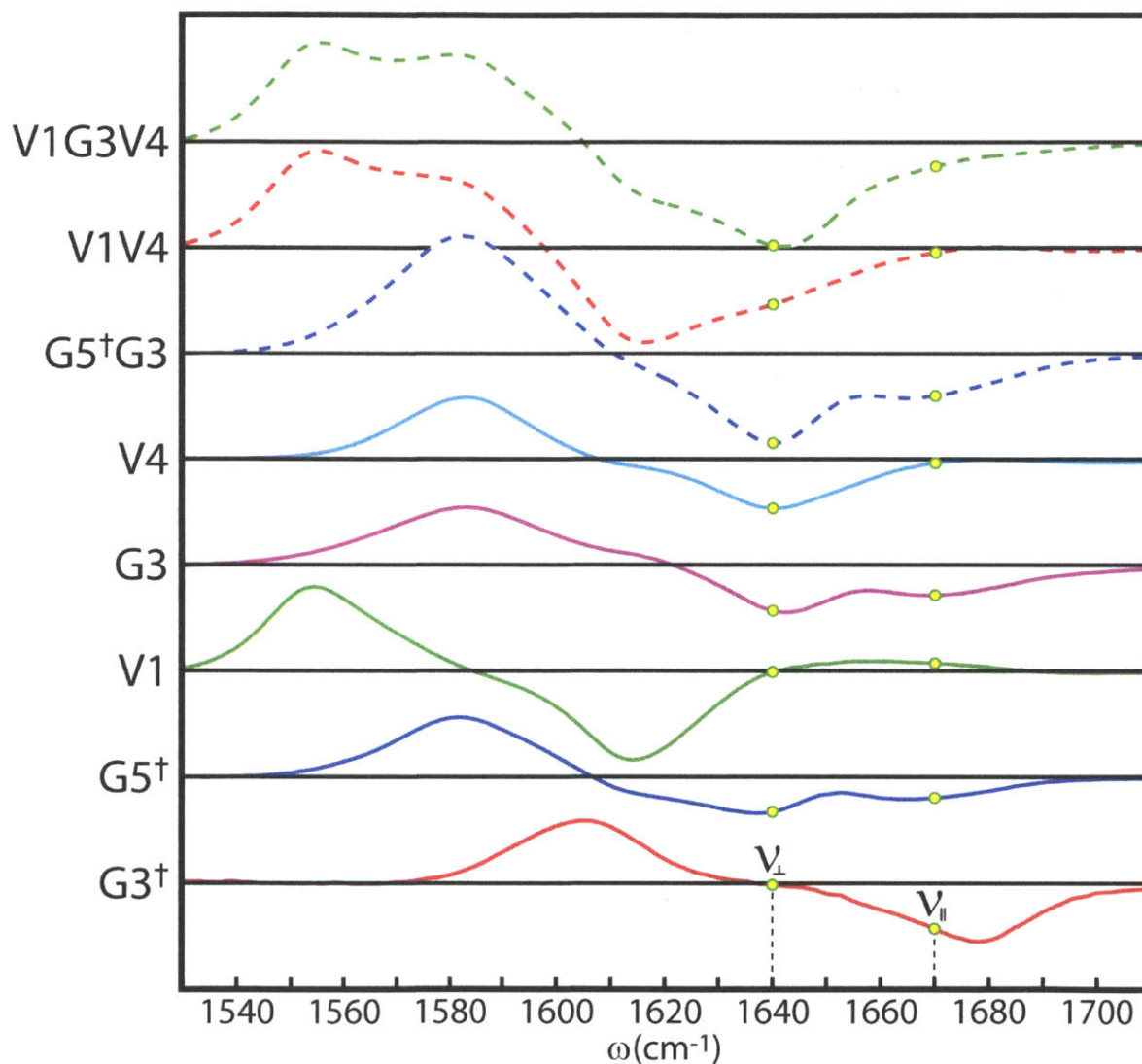


Figure 9: [(Isotope Label) – WT] FTIR difference spectrum for the GVGn1 peptide library collected at 10 °C on desalted peptides in pH = 1.0 DCl in D_2O with 150 mmol of KD_2PO_4 .

5.7.1 Single Site analysis: V1

Peak frequency values provide insight into the local environment surrounding each labeled site. In chapter 4 it was established, based on the peak frequency of the proline resonance, that the GVGn1 peptide contains a high population of 1/0 and 2/0 hydrogen bond turns. This assertion is confirmed here by the presence of an asymmetric V1 peak line-shape (figure 10).

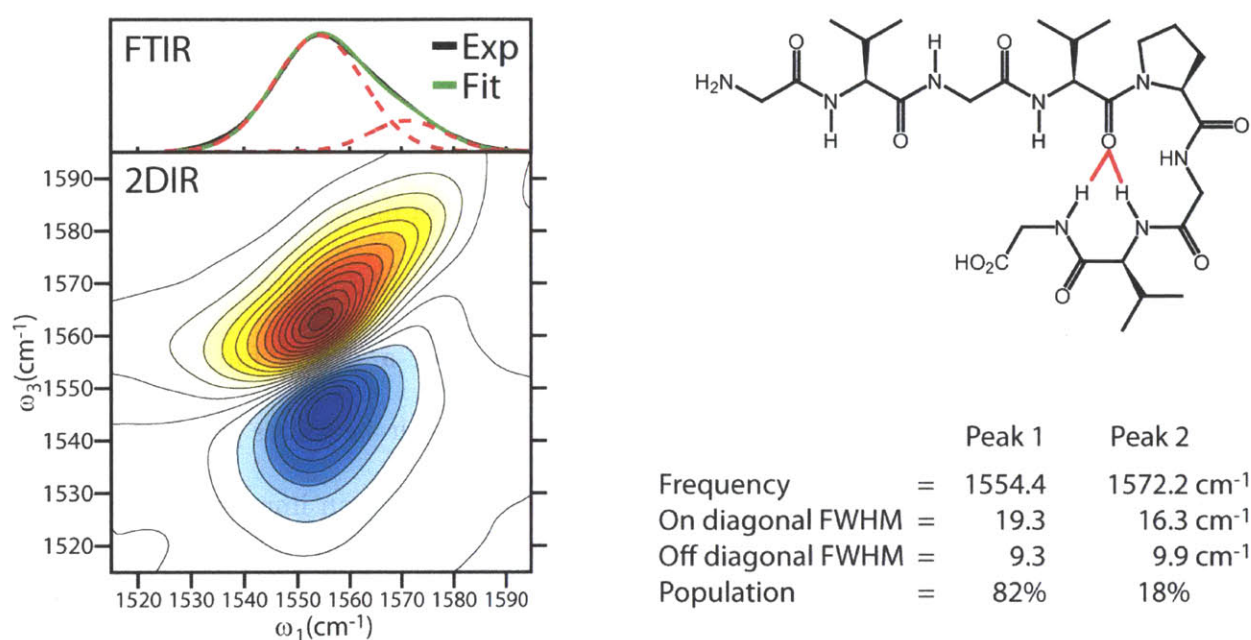


Figure 10. Experimental V1 FTIR and ZZZY 2DIR spectrum collected at 10 °C on a desalted peptide at $\tau_2 = 150$ fs in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄. The FTIR spectrum was fit with 2 Gaussian peaks giving populations for the 2/0 and 1/0 turns of 82% and 18%. In the upper right hand corner of the figure is a schematic representation of the GVGn1 peptide in a 2/0 turn with hydrogen bonds depicted as red lines.

The asymmetric peak results from the over-lap of the 1/0 and 2/0 peaks at 1572.2 and 1555.4 cm⁻¹ respectively. The frequency difference of $\Delta = 17.8$ cm⁻¹ observed here is more similar to the 16cm⁻¹ red shift per hydrogen bond found for the K8 label of TZ2 in chapter 3 then the 13.8 cm⁻¹ red shift per hydrogen bond found for the proline site in chapter 4. Based on the fitting of the

2DIR spectrum the V1 peak was found to have an $\Delta = 19.3 \text{ cm}^{-1}$, $\Gamma = 9.3 \text{ cm}^{-1}$, and $E = 0.62$ for the 2/0 peak and $\Delta = 16.3 \text{ cm}^{-1}$, $\Gamma = 9.9 \text{ cm}^{-1}$, and $E = 0.46$ for the 1/0 peak. We find that the 2/0 structure contains a more heterogeneous distribution of local environments based on the 3 cm^{-1} larger Δ . In addition this data shows that the ensemble of 2/0 turns is likely to exchange less rapidly than the 1/0 ensemble based on the 2/0 peak having a 1.2 cm^{-1} smaller Γ and a 35% increase in ellipticity relative to the 1/0 peak. A fitting to the V1 FTIR isotope peak was used to assign the relative population of the 1/0 and 2/0 turn giving 18% and 82% respectively. These values are not in agreement with the results given in chapter 4 where the populations for the 1/0 and 2/0 turns were reported to be 55-60% and $\sim 25\%$ respectively. Never the less both predictions show a high population of a closed turn. In contrast the MD simulations predict an open turn with an 11.5% and 4.9-5.9% population for the 1/0 and 2/0 turn respectively.¹⁰² As demonstrated in chapter 4 only peptide conformations with a closed turn reproduce the pronounced proline peak shift in the GVGn1 spectrum. As a result the values presented here and in chapter 4, which predict a high population of a closed turn, are an improvement over MD based results. These discrepancies also demonstrate the value of fully isolating peak features through the use of isotope labeling where here the combination of the proline and $^{13}\text{C}^{18}\text{O}$ redshift changes the V1 site frequency by $\Delta\omega = -87 \text{ cm}^{-1}$.

The 2D IR data presented above was collected at a waiting time of $\tau_2 = 150 \text{ fs}$ which is small relative to the time scale of molecular motion. The result is a 2D IR surface that presents a nearly static picture of the conformational ensemble. By studying the isotope labeled peaks as a function of waiting time it is possible to track the configurational changes that occur for this ensemble providing insight into the molecular dynamics of the peptide.¹⁰³ As a function of waiting time these isolated peaks transition from being on diagonally elongated to symmetric as

the structure of the local environment and in turn the sites amide I' transition frequency evolves. Encoded in these waiting time dependent spectral changes is information on local solvent exposure, conformational flexibility, and chemical exchange. As described above the value of Δ for an isotope peak is the result of the frequency for that transition being modulated by the distribution of heterogeneous local environments that exist for that labeled amide site in the ensemble. As a function of waiting time transitions that were initially at a frequency of ω_1 evolve to a new frequency ω_3 due to changes in the local environment, resulting in the generation of two off diagonal peaks at (ω_1, ω_3) and (ω_3, ω_1) . In the case of a rearrangement of the local solvent environment or a transient conformational fluctuation of the peptide these cross peaks generate a symmetric line shape with a rotation of the tilt of the on diagonal peak maximum. In the case of chemical exchange these cross peaks can generate a distinct feature in the off diagonal region of the spectrum.

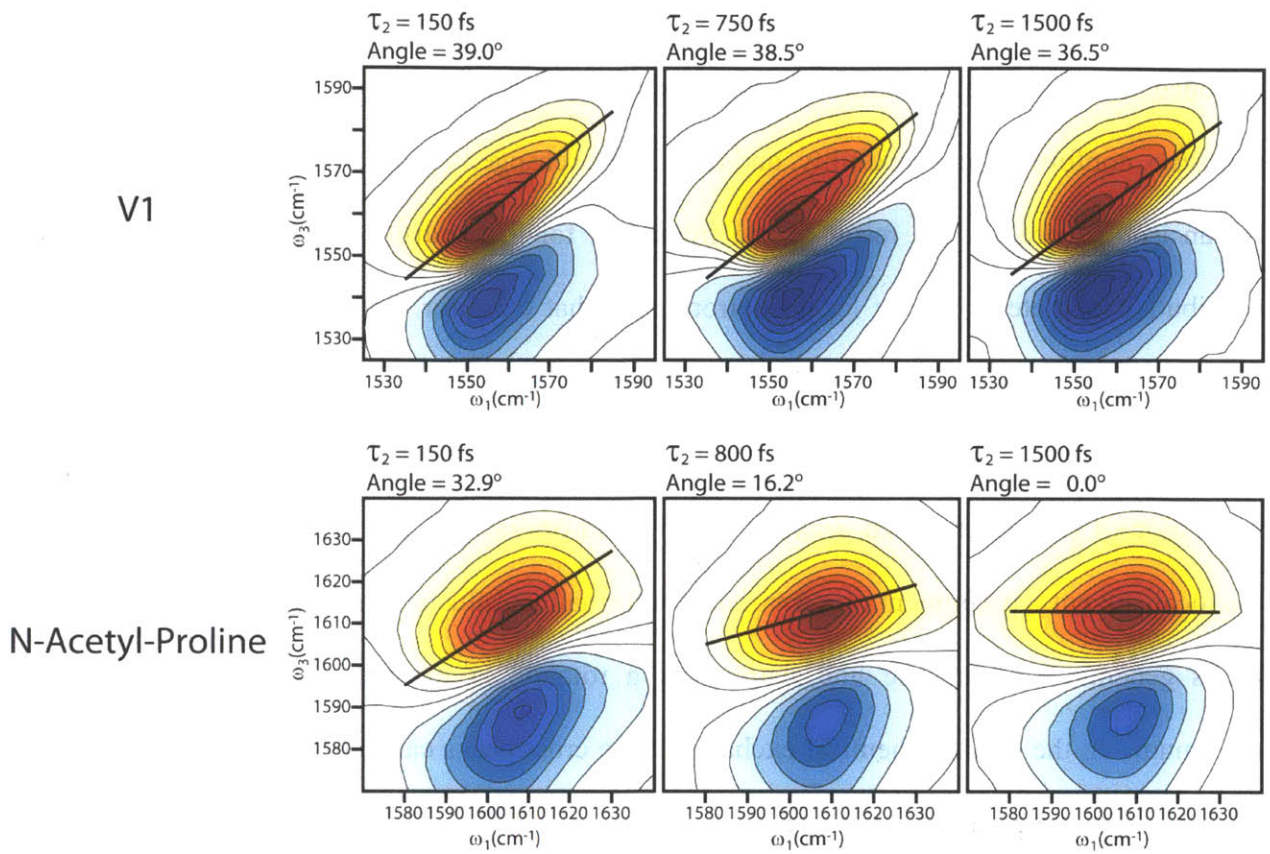


Figure 11: Waiting time 2DIR data for non-desalted V1 and N-Acetyl-Pro-COOD collected in ZZZZ polarization at 10 °C in pH = 1.0 DCl in D₂O with no additional salt added. The black lines in the figure are the result of fitting the fundamental peak maximum.

A comparison of the 2D IR waiting time spectrum for V1 and N-Acetyl-Proline-COOD provides further proof of the coexistence of a 1/0 and 2/0 turn in the equilibrium ensemble of GVGn1 (Figure 11). Figure 11 shows that when the proline amide unit is not incorporated into a peptide, as is the case for N-Acetyl-Proline-COOD, it contains a more symmetric line-shape that becomes homogeneous with waiting time. The free proline peak has $\Delta = 27.6 \text{ cm}^{-1}$ which is 69% and 43% larger than the Δ for the 1/0 and 2/0 peaks respectively. This time scale and line width is consistent with the presence of a water exposed amide unit since water exposure is known to

generate broad line shapes and the observed relaxation time for this peak is similar to the relaxation time of the OH stretch of HOD in D₂O.¹⁰⁴ In contrast the V1 peak does not rotate as a function of waiting time. This lack of rotation combined with the peaks asymmetry is consistent with the existence of two overlapping peaks. In this scenario both peaks do become homogeneous with waiting time but since they are displaced along the diagonal axis it gives the appearance of a single static inhomogeneously broadened peak.

5.7.2 Single Site analysis: G3 and V4

Ab initio calculations have shown that each hydrogen bond to the amide C=O generates a 20 cm⁻¹ red-shift to the site frequency whereas each hydrogen bond to the amide N-H generates a 10 cm⁻¹ red-shift.¹⁰⁵ As a result the observation that both V4 and G3 have similar site energies and peak rotation as a function of waiting time (figure 12) reinforces the proposed structure from chapter 4 in which these sites on average have their N-H pointing in the same direction relative to the peptide turn.

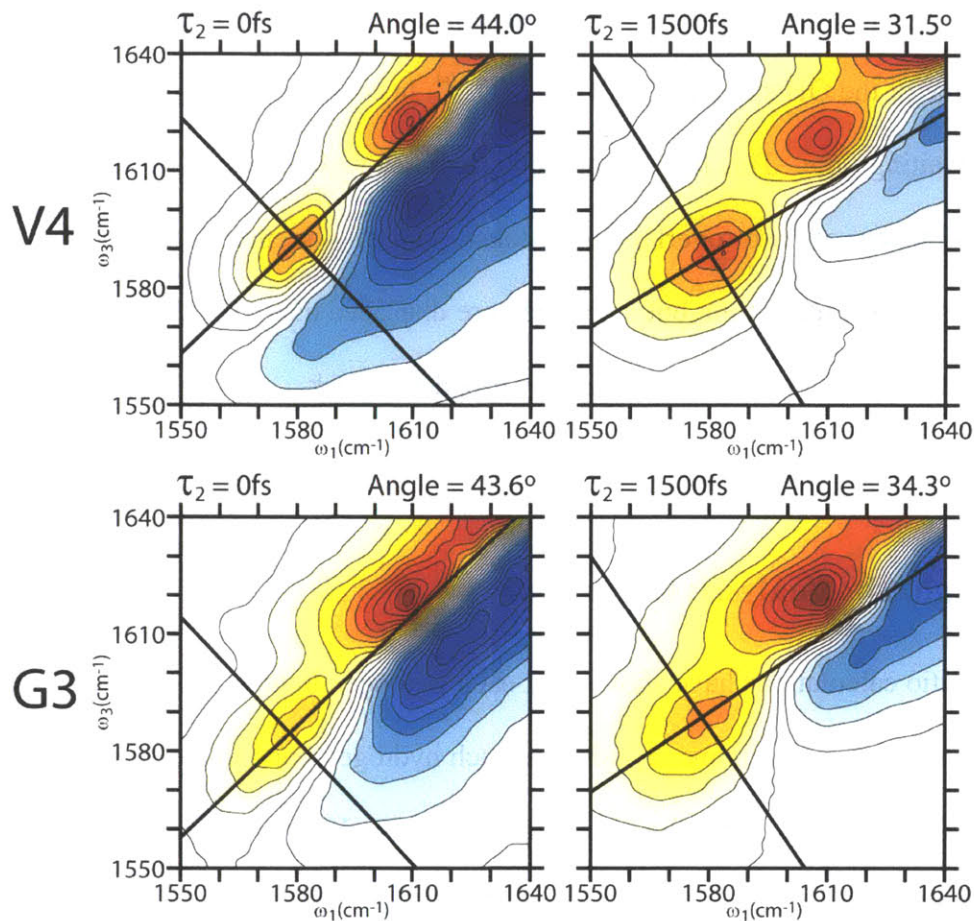


Figure 12. Waiting time 2DIR data for non-desalted V4 and G3 collected in *ZZZZ* polarization at 10 °C in pH = 1.0 DCl in D₂O with no additional salt added. Each figure has two perpendicular lines, one indicating the fit line for the fundamental peak maximum and a perpendicular line in the anti-diagonal direction provided as a guide for the eye.

This result is in contrast to the pattern of site energies and waiting time dynamics expected for a standard hairpin with a type II β turn. For a standard hairpin these values are found to oscillate along the peptide backbone as a result of the alternating orientation for amide units as observed for TZ2 in the folded and frayed state (chapter 3 figure 11b).

Despite these similarities differences do exist between the G3 and V4 sites. First the G3 site is found to have a Δ of 34.6 cm^{-1} which is 42% larger than the value of 24.4 cm^{-1} found for V4.

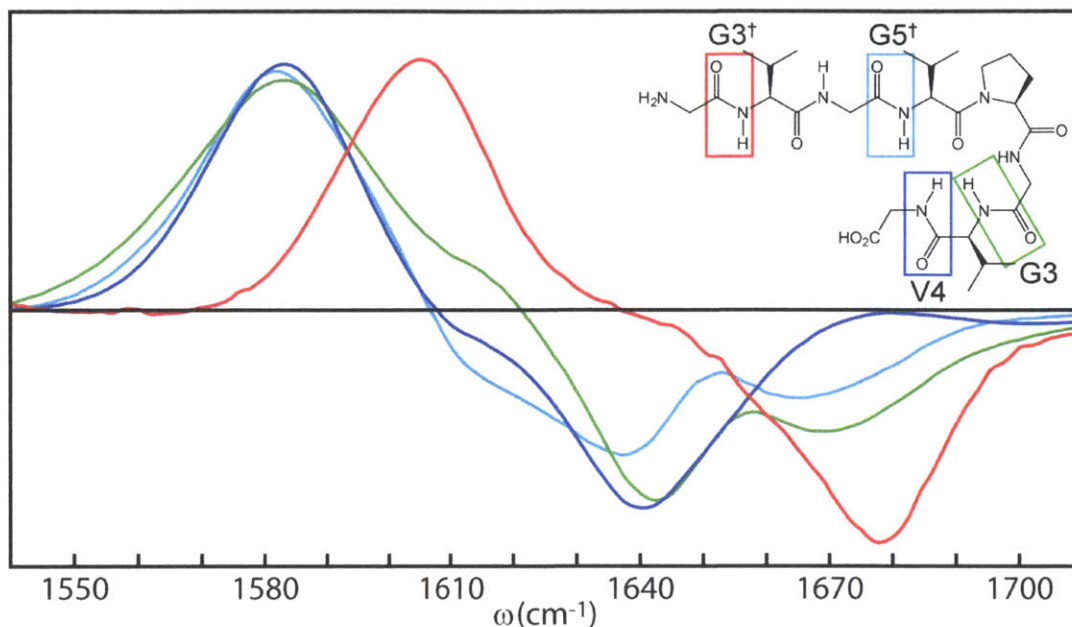


Figure 13. [(Isotope Label) – WT] FTIR difference spectrum for G3[†] (Red), G5[†] (Cyan), G3 (Green), and V4 (Blue), collected at 10 °C on desalted peptides in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄.

Second the G3 and V4 sites display different loss features in the FTIR difference spectrum (Figure 13). In figure 13 we see that for V4, intensity is primarily lost from the ν_{\perp} and random coil region of the spectrum which is consistent with a partially disordered site that participates in a conformation containing a cross strand hydrogen bond. In contrast the G3 label displays loss features across the amide I' spectrum potentially demonstrating a larger diversity of possible solvent conformations for the G3 amide site. Both G3's broader loss spectrum and increased Δ is consistent with a less constrained amide site do to the absence of steric hindrance created by a side-chain moiety. Alternatively the broad loss feature could be the result of a strong coupling

between the G3 site and its neighboring amide units. In the G3 FTIR difference spectrum there is a local maximum at $\omega = 1657.6 \text{ cm}^{-1}$ and two local minimum at $\omega = 1642.2$ and 1669.3 cm^{-1} . These values are similar to those for the random coil, ν_{\perp} , and ν_{\parallel} peaks which have center frequencies at 1660, 1630, and 1670 cm^{-1} . As a result this loss pattern is also consistent with a strong coupling generating a large splitting between ν_{\perp} and ν_{\parallel} with only a small contribution to the random coil region. Finally the FTIR difference spectrum for G3 gains intensity whereas V4 loses intensity in the region from 1610 to 1620 cm^{-1} corresponding to the center of the proline peak. The observed changes could result from a change in coupling with the G3 site gaining and the V4 site losing its coupling to proline upon labeling. The G3 and V4 coupling to the V1 site will be revisited in section 5.8 using 2D IR spectrum for analysis.

5.7.3 Single Site analysis: $G5^{\dagger}$

$G5^{\dagger}$ has a site frequency of 1582.0 cm^{-1} which is similar to the site frequency found for G3 and V4. Based on this finding the $G5^{\dagger}$ site is likely to on average have a solvent exposed C=O and a N-H that is pointed towards the center of the peptide as was predicted for G3 and V4. Though to confirm this assertion 2D IR waiting time measurements must first be carried out in order to compare the rate of nodal rotation for the $G5^{\dagger}$ isotope peak to the rates of G3 and V4 in order to determine if there solvent environments have similar dynamics. In addition the $G5^{\dagger}$ FTIR loss feature is found to be similar to the G3 loss feature in that it contains a depletion in intensity at ν_{\perp} , ν_{\parallel} , and random coil frequencies. In contrast the $G5^{\dagger}$ loss feature is broader and in particular has smaller losses in the ν_{\perp} and ν_{\parallel} regions compared to G3 which reflects a weaker coupling of the $G5^{\dagger}$ site with its neighboring amide units. Finally the $G5^{\dagger}$ site is found to have a

$\Gamma = 9.6 \text{ cm}^{-1}$ giving an ellipticity of $E = 0.75$. This value for E is 10.7% and 8.5% smaller than the values measured for $G3$ and $V4$ respectively. The smaller ellipticity for $G5^\dagger$ may indicate that this site is more solvent exposed and since the $V4$ site is on average hydrogen bonded to $V1$, as opposed to $G5^\dagger$ as would be the case in a standard hairpin, the $G5^\dagger$ site is not likely to participate in peptide-peptide hydrogen bonds with sites in the turn.

5.7.4 Single Site analysis: $G3^\dagger$

Since the GVGn1 peptide is enriched in glycine residues it displays a high degree of conformational flexibility. As a result the N-terminus can bend inwards towards the turn to form peptide-peptide hydrogen bonds. For this reason the $G3^\dagger$ label was synthesized to explore the local environment around the N-terminus (Figure 14).

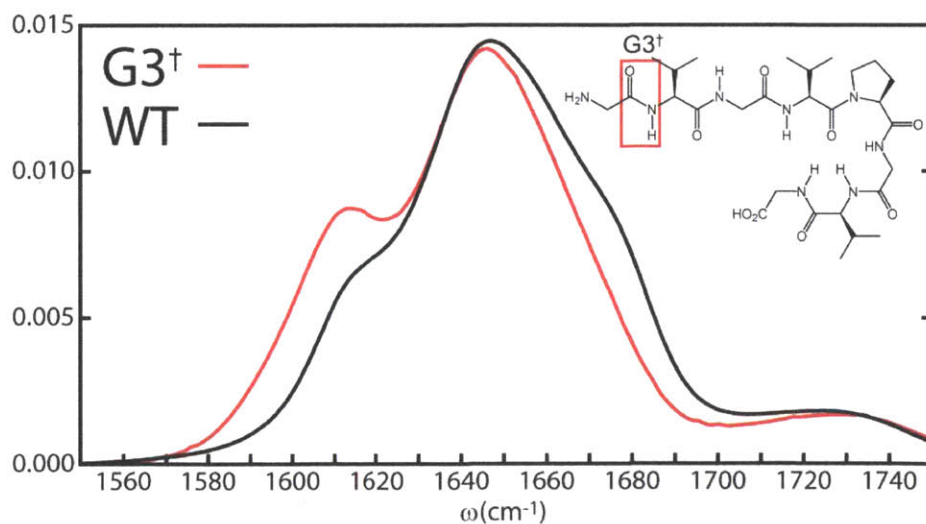


Figure 14. FTIR spectrum for $G3^\dagger$ (Red) and WT (Black) collected at 10 °C on desalted peptides in pH = 1.0 DCl in D_2O with 150 mmol of KD_2PO_4 .

We find that the G3[†] site energy is located at 1604.4 cm⁻¹ which is blue-shifted by 21.4 cm⁻¹ relative to the average site frequency for the G5[†], G3, and V4 sites. This indicates that the G3[†] site has fewer hydrogen bonds or a different orientation of its amide C=O relative to the turn than the G5[†], G3, and V4 sites. In addition unlike the G5[†], G3, and V4 sites the depletion in intensity in the FTIR difference spectrum for G3[†] is concentrated in a frequency region to the blue of the ν_{\parallel} with a small random coil component. Here the absence of a loss feature at the ν_{\perp} frequency indicates that the G3[†] site on average does not participate in peptide-peptide hydrogen bonds. These results for G3[†] are similar to the findings for the S1 label of TZ2 in chapter 3. In chapter 3 it was proposed that the S1 site has a high degree of conformational mobility, which would be expected for a terminal amide group, but simultaneously has an amide C=O that is shielded from the solvent generating a high frequency S1 peak. Given the high site frequency for G3[†] it is likely that it also has an amide C=O that is shielded from the solvent. This assignment could be explained by a conformation in which the N-terminus bends back towards the turn with its amide C=O oriented towards the V1 and V4 hydrophobic side-chains. Though modeling of the G3[†] site energy as a function of peptide conformation will be necessary to verify this assertion.

5.8 Multiple Site Analysis

In this section off diagonal peak structure generated from the formation of excitonic vibrations will be explored to determine the strength of the peptide turn. To quantify the stability of the turn the G5[†]G3, V1V4 and V1G3V4 peptides were synthesized. By spectroscopically isolating these combinations of amide sites it is possible to detect cross strand couplings that are indicative of turn formation.

This approach was utilized in chapter 3 in the form of the TT dual label to measure the stability of the mid-strand peptide contacts. In this example the isotope peak of the TT double label was found to be red-shifted relative to the T10 single label isotope peak due to cross strand coupling between the local vibrations of T3 and T10. Since the $G5^\dagger$ and G3 site energies are nearly degenerate, $\Delta = 1.0 \text{ cm}^{-1}$, these local vibrations can also mix to generate two new peaks whose Δ frequency splitting and cross peak intensities are related to the coupling strength

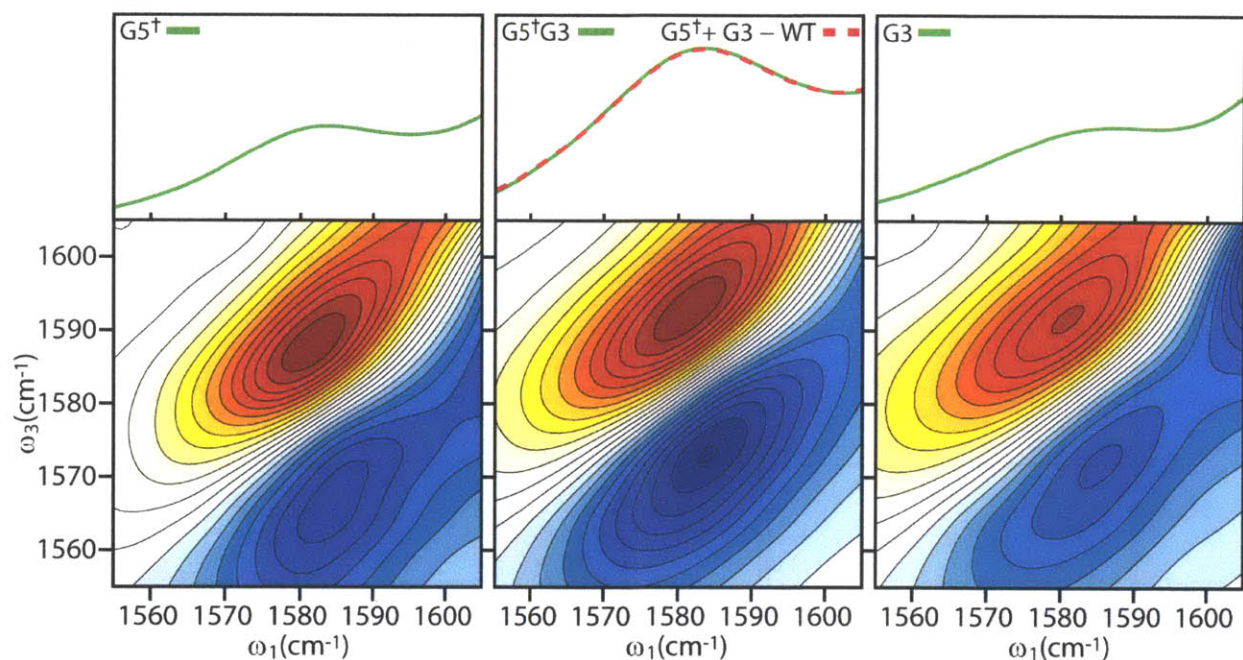


Figure 15. FTIR and 2DIR spectrum for $G5^\dagger$, G3, and $G5^\dagger G3$ collected at 10 °C on desalted peptides in pH = 1.0 DCl in D_2O with 150 mmol of KD_2PO_4 . 2DIR spectrum were collected in ZZYY polarization at $\tau_2 = 150$ fs.

between the sites. For the conditions studied here no indications of cross strand coupling of the $G5^\dagger$ and G3 sites were detected as is demonstrated in figure 15. Visual inspection of the FTIR and 2DIR data for $G5^\dagger$, G3 and $G5^\dagger G3$ peaks indicate that the dual labeled $G5^\dagger G3$ spectrum is the result of a linear combination of the single labeled $G5^\dagger$ and G3 spectrum. In addition the

peak metrics for the $G5^\dagger G3$ peak, $\nu_{\max} = 1581.7 \text{ cm}^{-1}$, $\Delta = 29.7 \text{ cm}^{-1}$, $\Gamma = 10.0 \text{ cm}^{-1}$, and $E = 0.79$ are found to be similar to the average of the values for the $G5^\dagger$ and $G3$ spectrum $\nu_{\max \text{ average}} = 1581.5 \text{ cm}^{-1}$, $\Delta_{\text{average}} = 30.1 \text{ cm}^{-1}$, $\Gamma_{\text{average}} = 10.0 \text{ cm}^{-1}$, and $E_{\text{average}} = 0.79$.

Unlike the $G5^\dagger G3$ spectra the $G3$, $V1V4$, and $V1G3V4$ spectra do contain indications of cross strand coupling. In figure 16 it can be seen that both the $V1V4$ FTIR and 2DIR spectra contain features that are not the result of a linear combination of their corresponding single labeled spectra. In the $V1V4$ FTIR spectra these differences are seen as an increase in peak intensity for the $V1$ $2/0$ peak and an increase in intensity on the blue edge of the $V4$ peak. In the 2D IR spectrum for $V1V4$ and $V1G3V4$, cross peaks are observed between the $V4$ and $V1$

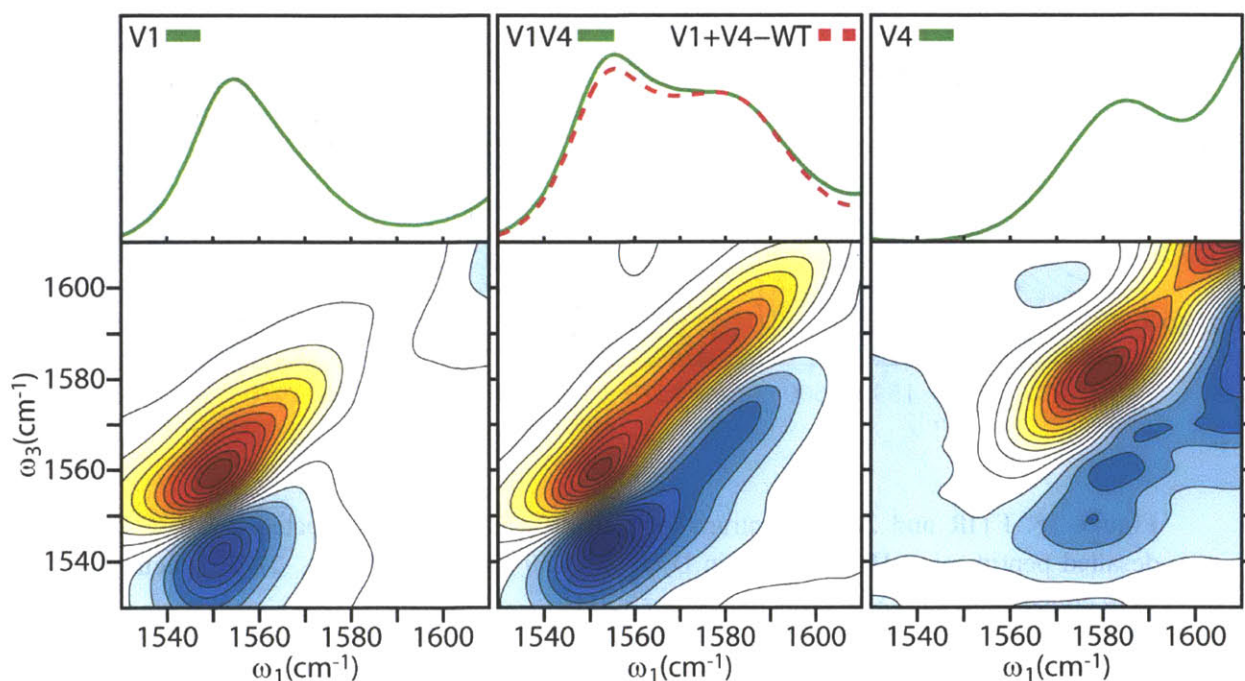


Figure 16. FTIR and 2DIR spectrum for $V1$, $V4$, and $V1V4$ collected at $10 \text{ }^\circ\text{C}$ on desalted peptides in $\text{pH} = 1.0$ DCl in D_2O with 150 mmol of KD_2PO_4 . 2DIR spectrum were collected in ZZYY polarization at $\tau_2 = 150 \text{ fs}$.

isotope peaks (figure 17). These cross-peaks are most easily observed in the lower right corner of the spectrum as a negative ridge along $\omega_3 \sim 1544$ and 1562 cm^{-1} . A cross peak is also observed in the G3 peptide 2D IR spectrum between the G3 isotope peak and the unlabeled proline V1 peak. This cross peak is most easily observed in the upper left corner of the spectrum as a positive ridge along $\omega_3 \sim 1624 \text{ cm}^{-1}$. Since the G3 and V4 sites are separated from the V1 site by 1 and 2 amino acids respectively, these cross peaks are more likely to result from a through space as opposed to through bond coupling. Based on the modeling results for the proline peak in chapter 4 it is likely that these cross peaks are the result of a hydrogen bond between the G3 and V4 amide N-H moiety's and the V1 amide C=O. Simultaneously the G5[†]G3 2D IR spectrum did not indicate the existence of a longer range coupling across the turn. As a result it is concluded that the GVGn1 forms a closed but loose β -turn with V1 to G3 and V1 to V4 hydrogen bond contacts. To further explore these findings in the future it would be prudent to synthesize the ^{13}CO analogs of the V4 and G3 peptides as well as the ^{13}CO labeled G3V4 peptide. By using a ^{13}CO instead of a $^{13}\text{C}^{18}\text{O}$ label the V4 and G3 site frequencies would be nearly degenerate with the unlabeled V1 site. In the case of the proposed ^{13}CO labeled G3 peptide the frequency splitting between the labeled G3 site and the unlabeled proline site would be $\Delta\omega_{\text{max}} = 3.7 \text{ cm}^{-1}$. As a result this peptide should display a dramatically red shifted isotope peak akin to what was observed for the TZ2 TT label if the peptide has a G3 to V1 hydrogen bond.

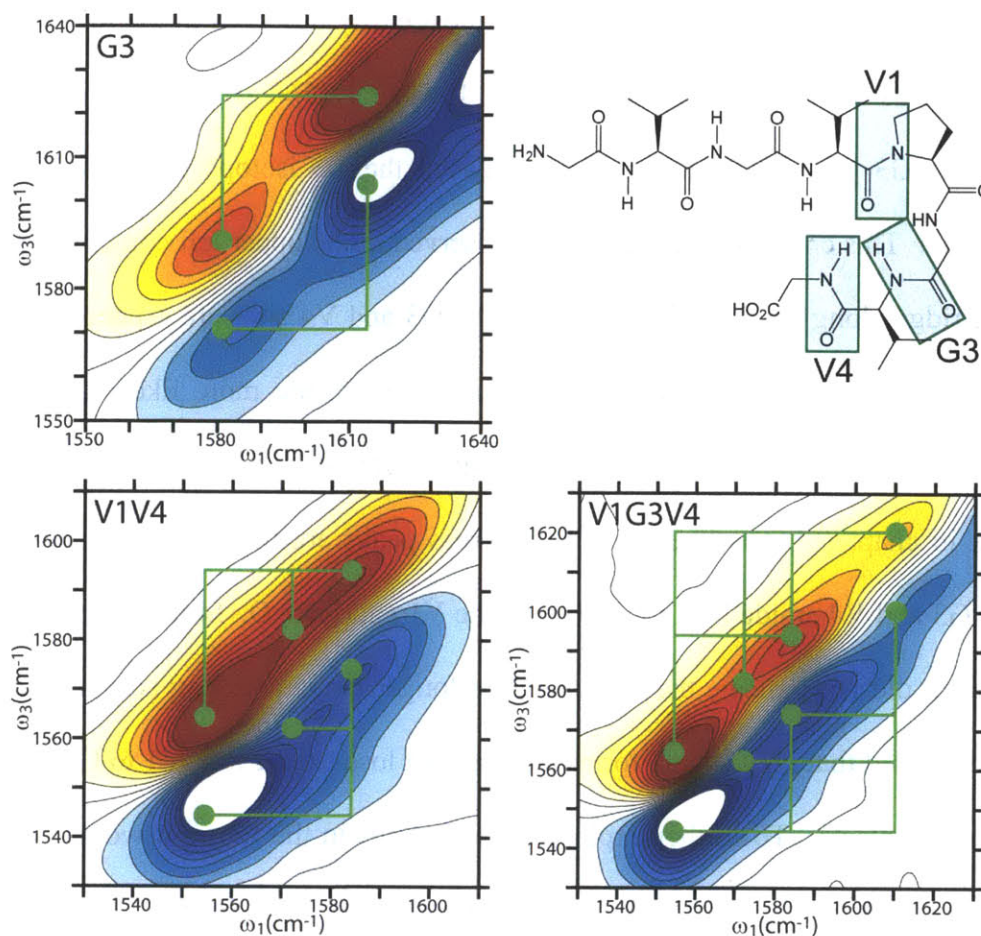


Figure 17. 2DIR spectrum for G3, V1V4, and V1G3V4 collected at 10 °C on desalted peptides in pH = 1.0 DCI in D₂O with 150 mmol of KD₂PO₄. 2DIR spectrum were collected in ZZYY polarization at $\tau_2 = 150$ fs. Green circles indicate the location of isotope peaks and the corresponding green lines are provided to indicate the location of cross peaks. In the G3 spectrum the green dots are placed over the peak maximum of the G3 and unlabeled V1 peaks with $\omega_{\max} = 1581.0$ and 1614.0 cm⁻¹ respectively. In the V1V4 spectrum the green dots are placed over the labeled V1 2/0 and 1/0 peaks with $\omega = 1554.4$ and 1572.2 cm⁻¹ along with a dot over the V4 peak maximum at $\omega_{\max} = 1584.1$ cm⁻¹. The V1G3V4 peptide has green dots over the V1 2/0 and 1/0 peaks with $\omega = 1554.4$ and 1572.2 cm⁻¹ along with a dot over the V4 and G3 peak maxima at $\omega_{\max} = 1584.1$ and 1610.3 cm⁻¹ respectively.

5.9 Temperature and Salt Dependence

As discussed earlier in this chapter tropoelastin and long ELPs undergo a phase transition from a structure that is soluble at low temperature to a 50-60% hydrated coacervate phase above

the ITT.^{23,63,67} Never the less the changes to the amide I' spectrum of the small ELPs studied here are less dramatic (Figure 18).

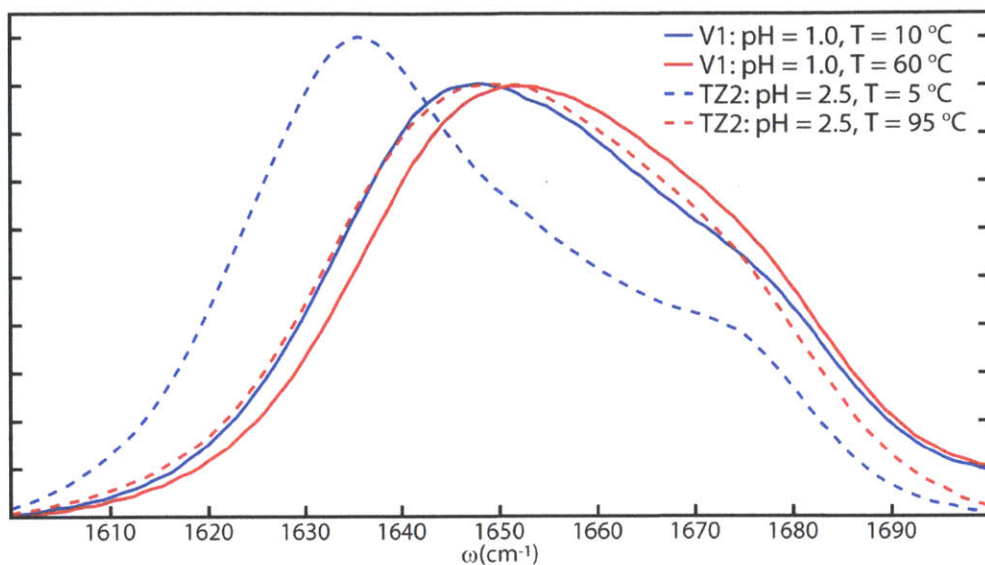


Figure 18: FTIR spectrum for non-desalted TZ2 in pH = 2.5 DCI in D₂O with no salt added and non-desalted V1 in pH = 1.0 DCI in D₂O with no salt.

Figure 18 is a comparison of FTIR spectrum taken at the end points of the TZ2 and V1 melting and “folding” transitions respectively. As discussed in chapter 3, at low temperatures the TZ2 peptide displays a large ν_{\perp} contribution that diminishes upon heating as the cross strand beta-sheet contacts become thermally disordered. In contrast the high temperature spectrum has a small ν_{\perp} contribution and is dominated by ν_{\parallel} and random coil vibrations. Interestingly the high temperature TZ2 spectrum is qualitatively similar to the V1 spectrum for the putatively “unfolded,” 10 °C, and “folded,” 60 °C, states. This comparison demonstrates both that the GVGn1 peptide exists largely in frayed and disordered conformations at all temperatures and that the changes in secondary structure for the GVGn1 peptide across the “folding” transition are small.

Only small changes are observed in the GVGn1 amide I' spectrum upon an increase in temperature and upon addition of salt. In figure 19 temperature dependent FTIR spectra are presented for WT, V1, G3, and G5[†]G3 in a solution containing DCl in D₂O and in a solution containing DCl and KD₂PO₄ in D₂O.

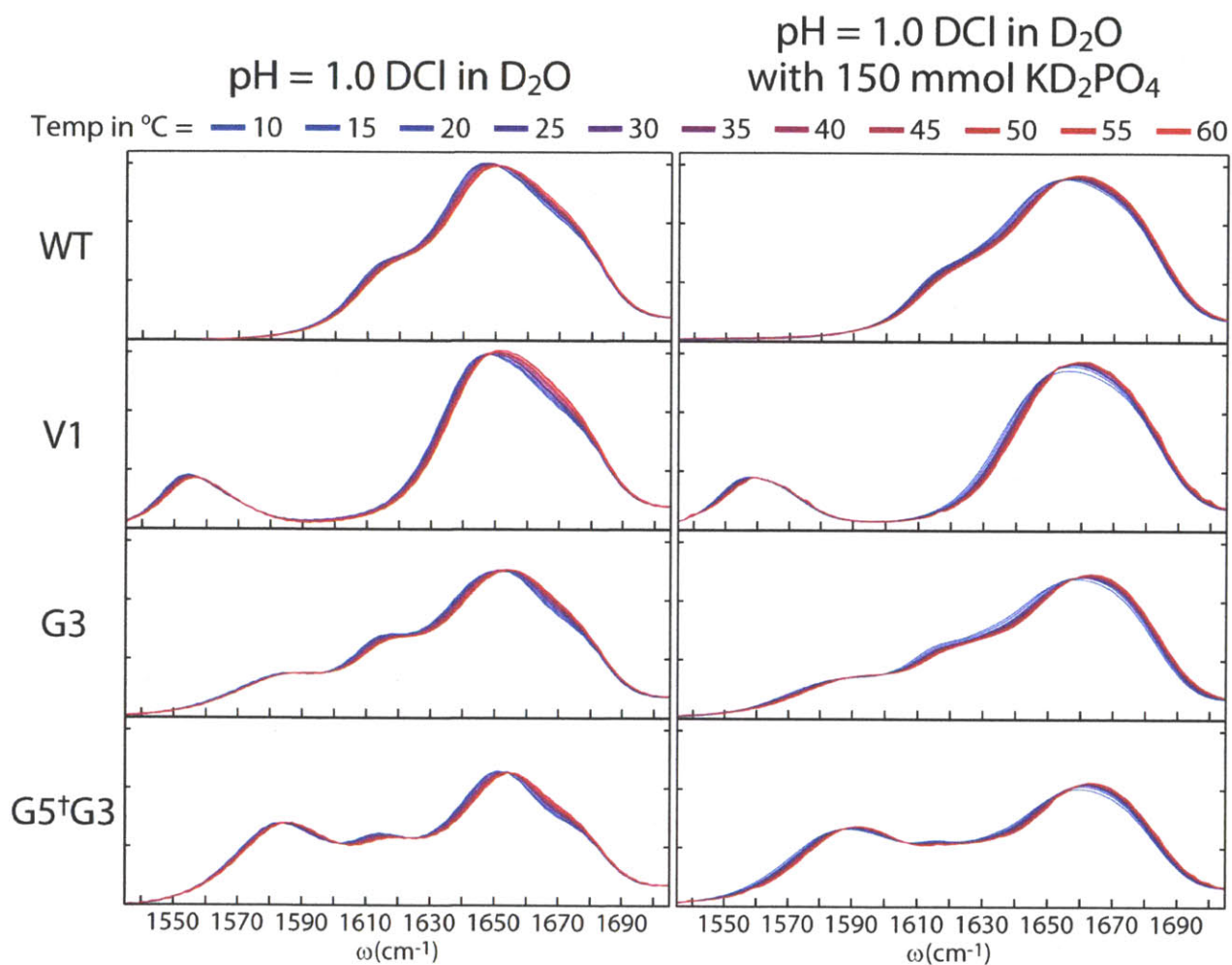


Figure 19. FTIR spectrum for non-desalted WT, V1, G3, and G5[†]G3 collected as a function of temperature from 10-60 °C in 5 °C steps. Spectrum in the left column were prepared in pH = 1.0 DCl in D₂O and in the right column in pH = 1.0 DCl in D₂O with 150mmol of KD₂PO₄ added.

Here KD_2PO_4 and DCl were added to the solution to create the D_4PO_4^+ , K^+ , and Cl^- ions in order to induce protein aggregation.^{106,107} From this data we see that no substantial change is observed for any peptide as a function of temperature regardless of salt concentration. Instead a small blue shift is observed for the entire amide I' line-shape upon heating. Since this blue shift is small and continuous with temperature it is attributed to nonspecific thermally induced disordering of the sample and is not likely to be the result of a structural phase transition. Similarly addition of salt, at constant temperature, generates a blue shift to the amide I' spectrum (Figure 20). Though in contrast to the temperature dependent data, the addition of salt to the sample also generates a subtle change in amide I' line-shape. This change is most pronounced in the region from 1660 to 1690 cm^{-1} where an increase in intensity is observed as a function of salt addition.

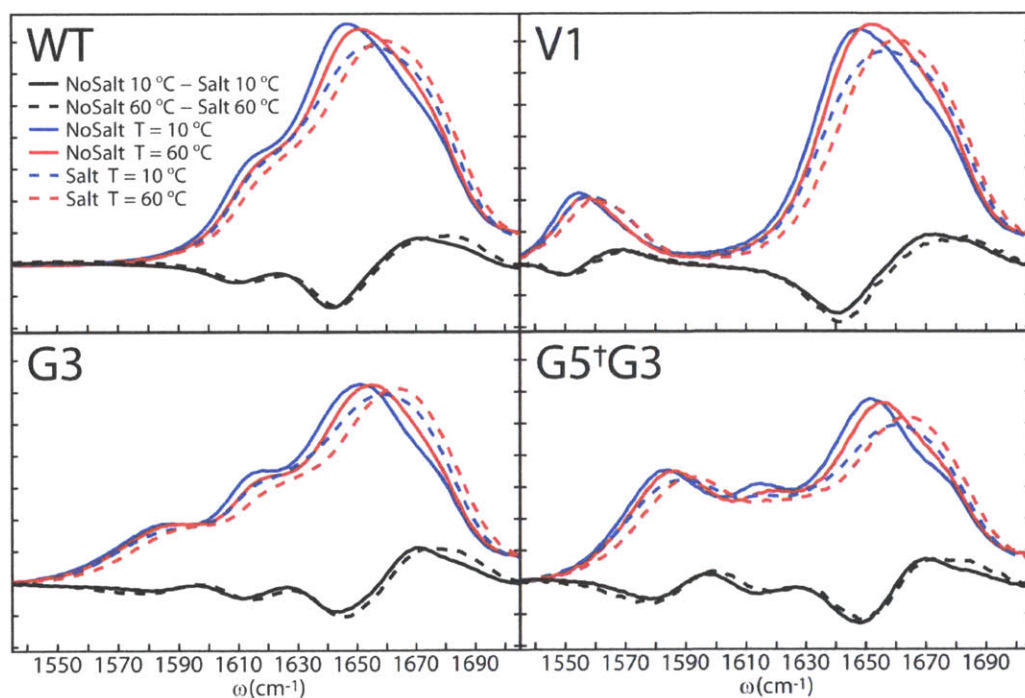


Figure 20. FTIR spectrum for non-desalted WT, V1, G3, and $\text{G5}^\dagger\text{G3}$ collected at $T = 10$ and 60 $^\circ\text{C}$. Data collected in $\text{pH} = 1.0$ DCl in D_2O is displayed with red and blue solid lines and data collected in $\text{pH} = 1.0$ DCl in D_2O with 150mmol of KD_2PO_4 added is displayed with red and blue dashed lines. The FTIR difference spectrum (No salt – Salt) at constant temperature is displayed in black.

In order to quantify the effect of salt addition on turn conformation the V1 isotope peak was fit with 2 Gaussians to determine the degree of conformational change upon heating and salt addition (Figure 21).

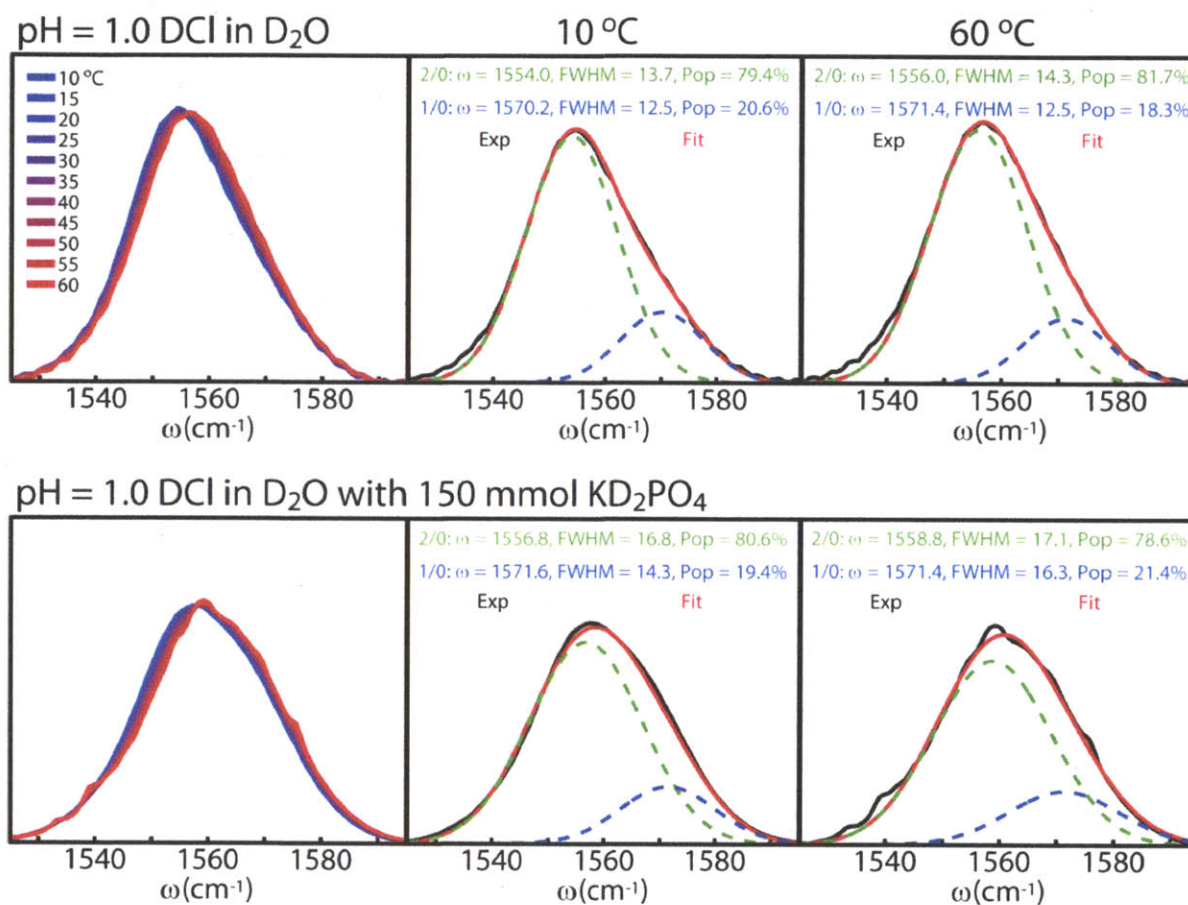


Figure 21. FTIR spectrum collected on the non-desalted V1 peptide. (Left) Temperature dependent FTIR spectrum collected from T = 10-60 °C in 5 °C steps, (Center) Fit to the 10 °C spectrum, and (Right) Fit to the 60 °C spectrum.

For the V1 peptide in pH = 1.0 DCl in D₂O with no additional salt the V1 peak at 10 °C is found to have a 2/0 and 1/0 peak centered at 1554.0 and 1570.2 cm⁻¹, with populations of 79.4% and 20.6%, and FWHM or 13.7 and 12.5 cm⁻¹ respectively. Upon heating to 60 °C the 2/0 and 1/0

peaks shift by $\Delta\omega = +2.0$ and $+1.2 \text{ cm}^{-1}$, with a population change of $\Delta P = +2.3\%$ and -2.3% , and a $\Delta\text{FWHM} = +0.6$ and 0.0 cm^{-1} respectively. Similar temperature dependent results are found for the V1 peptide in $\text{pH} = 1.0$ DCl in D_2O with $150 \text{ mmol KD}_2\text{PO}_4$. Upon heating in the presents of salt the 2/0 and 1/0 peaks blue shift by $\Delta\omega = +2.0$ and -0.2 cm^{-1} , with a population change of $\Delta P = -2.0 \%$ and $+2.0 \%$, and a $\Delta\text{FWHM} = +0.3$ and $+2.0 \text{ cm}^{-1}$ respectively. A small but slightly more substantial change is observed between the salt and no salt data. Upon addition of salt at $10 \text{ }^\circ\text{C}$ the 2/0 and 1/0 peaks change by $\Delta\omega = +2.8$ and $+1.4 \text{ cm}^{-1}$; $\Delta\text{FWHM} = +3.1$ and $+1.8 \text{ cm}^{-1}$; $\Delta P = +1.2\%$, and -1.2% respectively. At $60 \text{ }^\circ\text{C}$ the 2/0 and 1/0 peaks change by $\Delta\omega = 2.8$ and 0.0 cm^{-1} ; $\Delta\text{FWHM} = 2.8$ and 3.8 cm^{-1} ; $\Delta P = -3.1\%$ and $+3.1\%$ respectively. The observation that the addition of salt generates a larger peak shift then an increase in temperature is also observed for the COOD resonance (Figure 22).

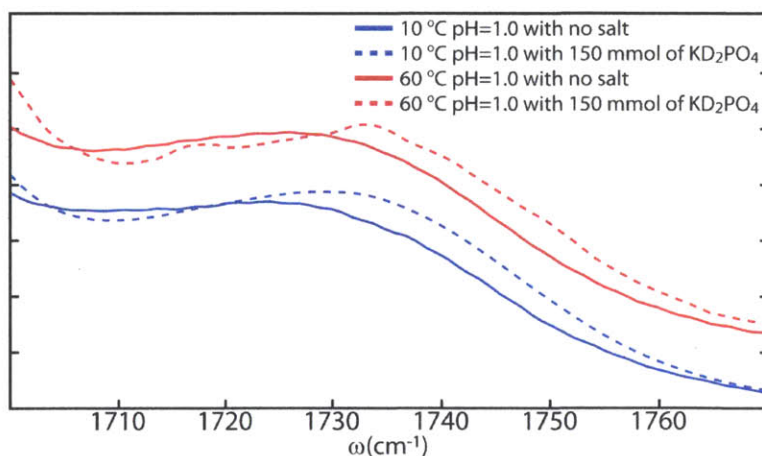


Figure 22. FTIR spectrum collected on the non-desalted V1 peptide plotted for the COOD region.

The COOD resonance is found to blue shift by 2.7 cm^{-1} upon heating from $10 \text{ }^\circ\text{C}$ to $60 \text{ }^\circ\text{C}$ and by 4.0 cm^{-1} when dissolved in the $150 \text{ mmol of KD}_2\text{PO}_4$ solution. Though as was the case for the amide I' resonance these changes are continuous as a function of temperature and salt

concentration and therefore are not likely to be the result of a structural phase transition. From these observations it is clear that the V1 and COOD peaks blue-shift upon heating or addition of salt but the underlying 2/0 and 1/0 turn populations are unaffected to within the error of the measurements performed.

Clearer evidence that increasing the concentration of charged species generates a larger change to the amide I' spectrum than heating is seen in the V1 2DIR spectrum (Figure 23).

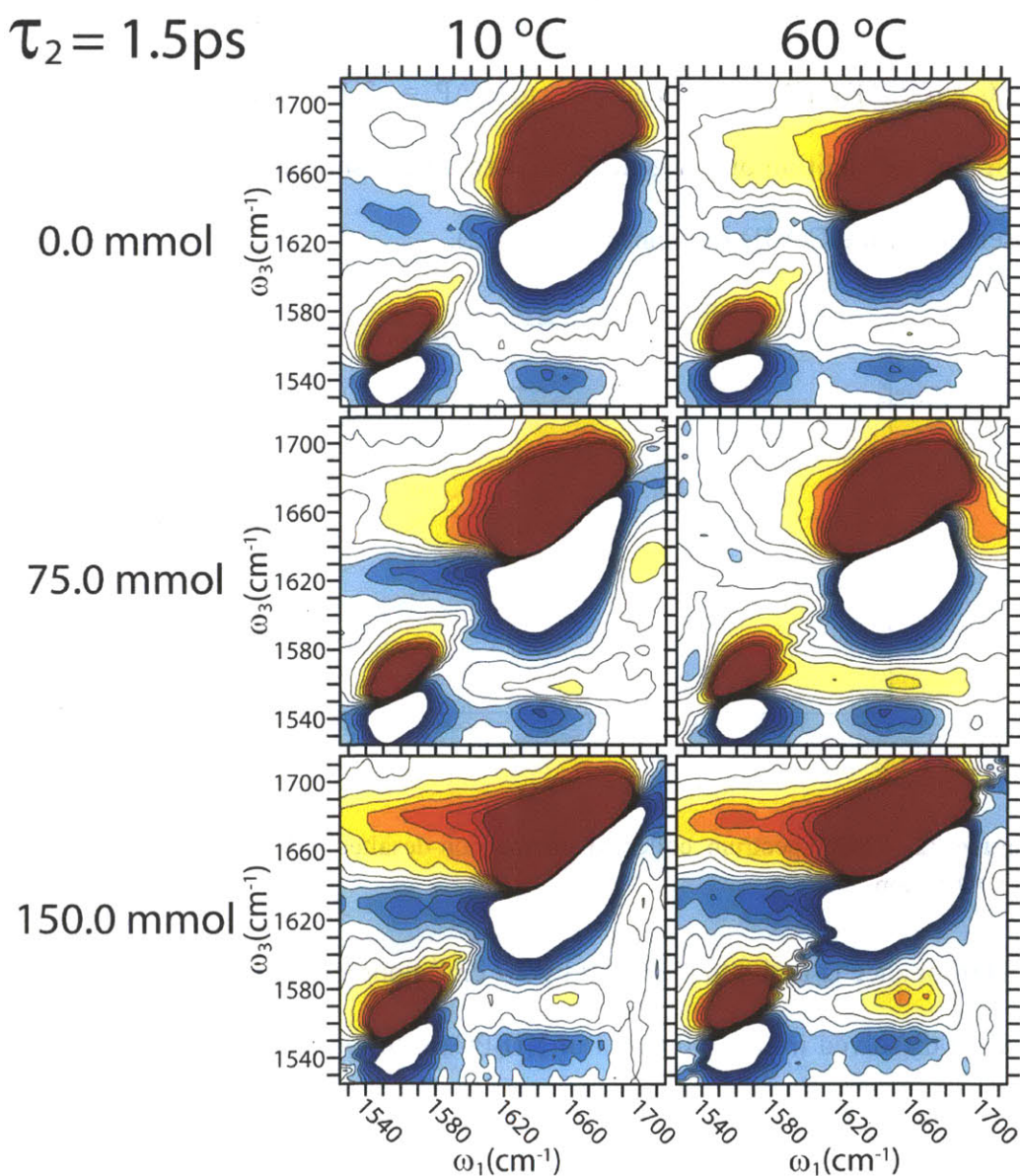


Figure 23. 2DIR spectra of the non-desalted V1 peptide as a function of temperature and salt concentration. All spectra were collected in a ZZZZ polarization geometry with $\tau_2 = 1.5\text{ps}$.

In this figure the V1 $\tau_2 = 1.5\text{ps}$ amide I' 2DIR spectrum is given as both a function of salt concentration and temperature. In contrast to the FTIR spectrum the 2DIR spectrum shows pronounced changes as a function of salt concentration. As stated earlier upon salt addition the amide I' line-shape changes in the region from 1660 to 1690 cm^{-1} . In the 2DIR spectrum this on diagonal line-shape change occurs in conjunction with the growth of a cross peak between the 1660 to 1690 cm^{-1} region and the V1 isotope peak. This cross peak is found to grow asymmetrically as a function of salt addition with a larger contribution on the blue side centered at $\omega_3 = 1678.0\text{ cm}^{-1}$. Since the $G3^\dagger$ FTIR difference spectrum shows a pronounced loss feature centered at 1677.8 cm^{-1} the cross peak could be the result of the N-terminus interacting with the V1 site lending credence to the proposal from section 5.7.4 that the $G3^\dagger$ bends inward towards the turn. Further experiments and modeling will be required to generate a structure frequency correlation for the $G3^\dagger$ site to confirm the existence of the proposed structure. Never the less the presence of a cross peak demonstrates the existence of through space coupling and as a result this data shows that salt addition induces structuring or a hydrophobic collapse.

5.10 Size dependent FTIR

Increasing the size of the peptide is found to increase the red shift of the proline peak. This observation is illustrated in figures 20 and 21 which show the temperature dependent FTIR spectra and single value decomposition (SVD) analysis for $GVG(VPGVG)_{1-6}$ and $(GVGVP)_{251}$.

In figure 24 we see that the FTIR spectrum for all peptides contain a similar line-shape with an increase in the red-shift of the proline resonance as a function of increasing chain

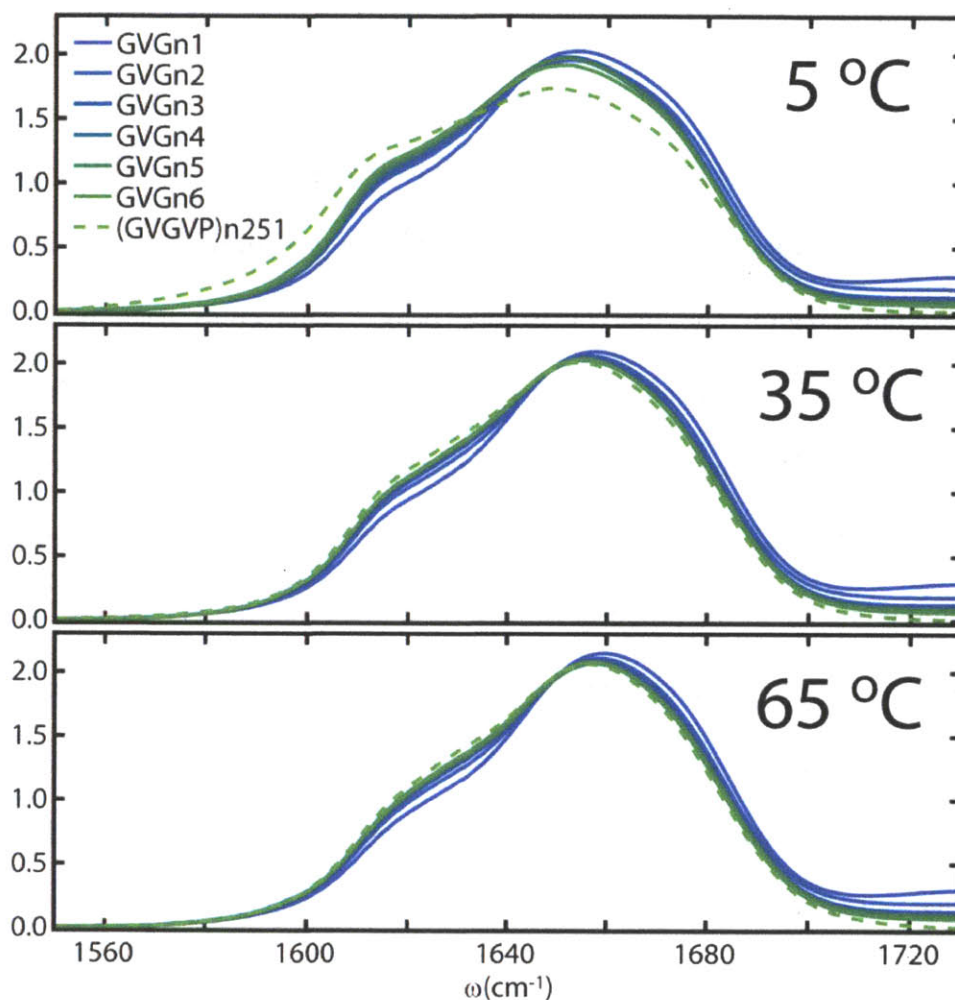


Figure 24. Temperature dependent FTIR spectra for non-desalted GVG(VPGVG) $_n$ where $n = 1-6$ and the (GVGVP) $_{251}$ polymer in pH = 1.0 DCl in D $_2$ O with 150 mmol of KD $_2$ PO $_4$.

length at 5.0 °C. This red-shift diminishes upon heating giving a similar proline peak shift for all peptides at 65 °C. Interestingly the first and second SVD components for all samples retain

similar peak features that differ only in intensity. As a result our understanding of the spectrum of the smaller ELPs may provide the requisite insight to assign the spectrum of larger ELPs.

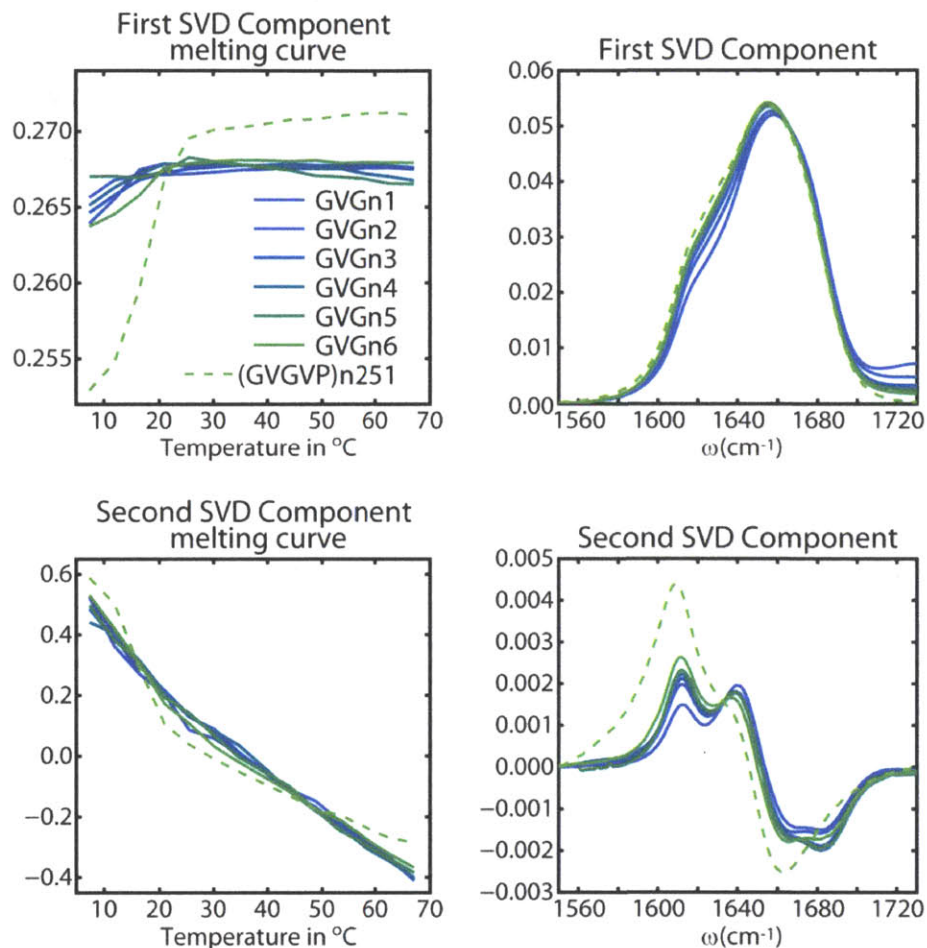


Figure 25. First and Second SVD component spectrum and corresponding melting curves for $\text{GVG}(\text{VPGVG})_n$ where $n = 1-6$ and the $(\text{GVGVP})_{251}$ polymer in $\text{pH} = 1.0$ DCl in D_2O with 150 mmol of KD_2PO_4 .

In figure 25 we see that $(\text{GVGVP})_{251}$ has an abrupt change in its first and second component SVD melting curves centered at 17 °C. In addition in the FTIR data for $(\text{GVGVP})_{251}$ in figure 26 shows that the majority of the temperature dependent red shift of the proline peak occurs over the temperature range 10-20 °C. Since this sample was found to transition from a low temperature state that does not scatter light to one that scatters light above 17 °C the

observed temperature dependent changes in the IR spectrum are believed to be the result of the ITT.

The second component spectrum for $\text{GVG}(\text{VPGVG})_n$, where $n = 1-6$, were found to contain a local maximum at $\omega \sim 1612$ and 1640 cm^{-1} and a local minimum at $\omega \sim 1665$ and 1681 cm^{-1} . The local maxima represent a blue shift or loss of peak intensity from the proline 2/0 peak and ν_{\perp} antiparallel β -sheet modes with temperature and the local minima represent a blue shift or gain of peak intensity for the ν_{\parallel} antiparallel β -sheet mode and an unassigned mode at $\omega \sim 1681 \text{ cm}^{-1}$ with temperature. In order to characterize the conformational changes that occur for the peptide across the phase transition the $(\text{GVGVP})_{251}$ FTIR spectra were compared to the simulated 2/0, 1/0, and 0/0 FTIR spectrum of GVGn1 from chapter 4.

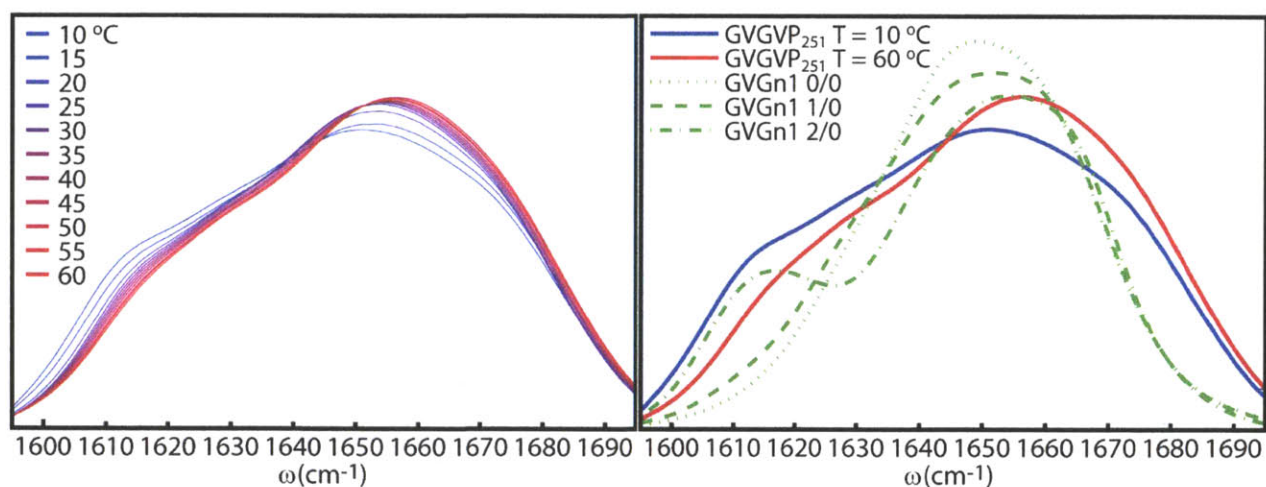


Figure 26. (Left) Temperature dependent FTIR spectrum for $(\text{GVGVP})_{251}$ in $\text{pH} = 1.0$ DCl in D_2O with 150 mmol of KD_2PO_4 . (Right) A comparison between the $(\text{GVGVP})_{251}$ FTIR spectrum with the simulated 2/0, 1/0, and 0/0 spectrum for the GVGn1 peptide.

In figure 26 we see that the proline peak in the FTIR spectrum for $(\text{GVGVP})_{251}$ at $10 \text{ }^\circ\text{C}$, (proline peak maximum at $\omega = 1608.3 \text{ cm}^{-1}$), most closely resembles the simulated GVGn1 2/0 spectrum (proline peak maximum at $\omega = 1616.9 \text{ cm}^{-1}$). Since the proline resonance acts as a local

probe of turn structure the red shifted proline peak for the (GVGVP)₂₅₁ polymer might have the same interpretation as in the case of the GVGn1 peptide. If this assumption is correct then the above temperature dependent FTIR spectrum would suggest a new explanation for elastin's ITT in which the peptide transitions from a low temperature state containing primarily 2/0 turns which upon heating diminishes as the population of 1/0 turns increases. Using the equation from chapter 4 for determining the number of peptide-peptide hydrogen bonds to proline, $\omega_p(\text{cm}^{-1}) \approx -13.8 \cdot n_{pp} + 1642$, gives $n_{pp} = 2.4$. This unlikely prediction may potentially reflect the fact that the red shift of the proline site in (GVGVP)₂₅₁ may originate from more than just the formation of the 2/0 turn.

The differences between the structure of GVGn1 and (GVGVP)₂₅₁ are not limited to the relative population of 2/0 and 1/0 turns. In figure 27 is a comparison of the 2D IR ZZYY spectrum of GVGn1 and (GVGVP)₂₅₁ at T = 10 and 15 °C respectively. In addition to the broader more red shifted proline peak for (GVGVP)₂₅₁, we observe the formation of a cross peak between the ν_{\perp} and ν_{\parallel} antiparallel β -sheet modes. This indicates that the polymer, in its unfolded state, is more structured than the GVGn1 peptide under similar conditions. One potential explanation is that the 1/0 turn in (GVGVP)₂₅₁ may contain a population of type II β -turns with both a V1 to G3 and a G5[†] to V4 cross strand hydrogen bond generating a unit of β -sheet structure. In order to assign the mechanism of the ITT in (GVGVP)₂₅₁ first this β -sheet structure has to be identified as well as determining the ΔG associated with changes in this structure relative to the ΔG associated with changes in the turn conformation, main-chain hydration and side-chain hydration during the ITT.

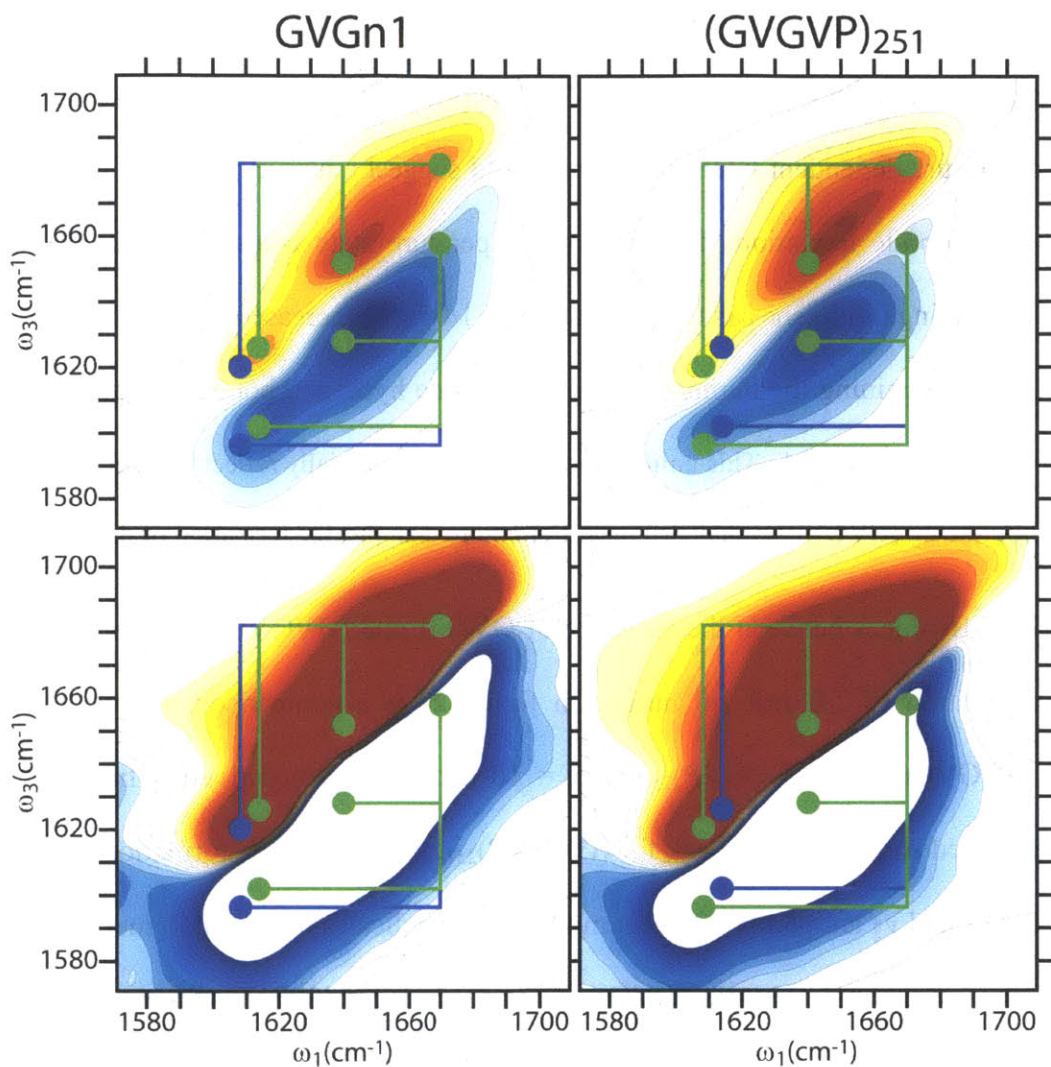


Figure 27. ZZZY 2DIR spectrum for the GVGn1 (Left column) and (GVGVP)₂₅₁ (Right column) peptides. Desalted GVGn1 was collected at $T = 10\text{ }^{\circ}\text{C}$ and $\tau_2 = 150\text{ fs}$ in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄:K₂DPO₄ (1:1). (GVGVP)₂₅₁ was collected at $T = 15\text{ }^{\circ}\text{C}$ and $\tau_2 = 150\text{ fs}$ in pH = 1.0 DCl in D₂O with 150 mmol of KD₂PO₄. Dots are placed to mark the location of the antiparallel β -sheet modes, $\nu_{\perp} = 1640$ and $\nu_{\parallel} = 1670\text{ cm}^{-1}$, as well as the location of the GVGn1 and (GVGVP)₂₅₁ proline peak maximum at $\omega_{\text{max}} = 1614$ and 1608.3 cm^{-1} respectively.

5.11 Conclusion

The GVGn1 peptide on average contains a high population of closed 2/0 and 1/0 turns for all temperatures and salt concentrations investigated. The V1 C=O site is found to be buried in the interior of the turn forming transient hydrogen bonds with the N-H of the G3 and V4 amide sites. The hydrogen bonds are seen in the 2D IR spectrum of G3, V1V4, and V1G3V4 as cross peaks between their corresponding isotope peaks. The transient nature of these cross strand hydrogen bond contacts are reflected in the G5[†]G3 spectrum which shows no indication of coupling between the G5[†] and G3 amide sites. It was shown that the G5[†] site is not likely to contribute a peptide-peptide hydrogen bond to the turn but instead it is partially solvent exposed owing to its location on the periphery of the peptide turn. This solvent exposure was found to decrease for the N-terminus as demonstrated by the G3[†] spectrum which contains a blue shifted isotope peak. From these results it can be concluded that the GVGn1 contains a closed irregular turn with the G3 and V4 N-H oriented towards the V1 C=O and may potentially contain a highly mobile N-terminus that is buried in the hydrophobic core of the peptide turn. This turn structure is durable showing no significant changes with temperature or salt concentration. Based on the work presented in chapter 4 this stability is most likely the result of steric clashes between the V1 side-chain with the proline pyrrolidine ring. It was shown that the addition of salt generates the largest changes to the amide I' spectrum and a possible structuring of the N-terminus through the formation of a β -strand or interaction with the V1 site in the turn. It was also found that the addition of salt generates the most pronounced changes to the amine amide I' band and the off diagonal cross peak region while not generating a significant shift in isotope peak frequencies. As a result the driving force for the phase transition for VPGVG sequences is likely due to a

hydrophobic effect involving the aggregation of Val side-chains to form a hydrophobic core and as a result the phase transition does not project significantly onto the amide I' spectrum.

5.12 Acknowledgments

The author would like to thank Carlos Biaz, Kevin Jones, Chunte Sam Peng and Michael Reppert for assistance with 2D IR experiments. The author would also like to thank Jongjin Kim and Sam Chunte Peng for aid with the synthesis of the isotope labeled peptides and the size dependent peptides GVGn2 through GVGn6, respectively. Finally the author would like to thank Dan W. Urry of Bioelastics Research for providing the (GVGVP)₂₅₁ sample.

5.13 References

- (1) Satoshi Fukuchi, K. H., Keiichi Homma, Takashi Gojobori, Ken Nishikawa *BMC Struct. Biol.* **2011**, *11*.
- (2) J. J. Ward, J. S. S., L. J. McGuffin, B. F. Buxton, D. T. Jones *J. Mol. Bio.* **2004**, *337*, 635.
- (3) Vladimir N. Uversky, A. K. D. *Biochim. Biophys. Acta* **2010**, *1804*, 1231.
- (4) Vladimir N. Uversky, C. J. O., A. Keith Dunker *Annu. Rev. Biophys.* **2008**, *37*, 215.
- (5) Lilia M. Iakoucheva, C. J. B., J. David Lawson, Zoran Obradović, A. Keith Dunker *J. Mol. Bio.* **2002**, *323*, 573.
- (6) H. Jane Dyson, P. E. W. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.

- (7) Fink, A. L. *Curr. Opin. Struct. Biol.* **2005**, *15*, 35.
- (8) Eliezer, D. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23.
- (9) Tanja Mittag, J. D. F.-K. *Curr. Opin. Struct. Biol.* **2007**, *17*, 3.
- (10) Chen H., R. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 516.
- (11) Tanguy Le Gall, P. R. R., Marc S. Cortese, Vladimir N. Uversky, A. Keith Dunker *J. Biomol. Struct. Dyn.* **2007**, *24*, 325.
- (12) J. A. Marsh, V. K. S., Z. Jia, J. D. Forman-Kay *Protein Sci.* **2006**, *15*, 2795.
- (13) Hyun Lee, K. H. M., Ranjith Muhandiram, Kyu-Hwan Park, Jae-Eun Suk, Do-Hyung Kim, Jun Chang, Young Chul Sung, Kwan Yong Choi, Kyou-Hoon Han *J. Biol. Chem.* **2000**, *275*, 29426.
- (14) R. Mohana-Borges, N. K. G., G. J. Kroon, H. J. Dyson, P. E. Wright *J. Mol. Biol.* **2004**, *340*, 1131.
- (15) Adam W. Smith, J. L., Santanu Roy, Ziad Ganim, Chunte Sam Peng, Thomas L. C. Jansen, Jasper Knoester, Andrei Tokmakoff *J. Phys. Chem. B* **2010**, *114*, 10913.
- (16) Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. *Acc. Chem. Res.* **2008**, *41*, 432.
- (17) Thomas la Cour Jansen, J. K. *Acc. Chem. Res.* **2009**, *42*, 1405.
- (18) Yung Sam Kim, R. M. H. *J. Phys. Chem. B* **2009**, *113*, 8231.
- (19) Kolano, C.; Helbing, J.; Kozinski, M.; Sander, W.; Hamm, P. *Nature* **2006**, *444*, 469.
- (20) Sang-Hee Shim, R. G., Yun L. Ling, David B. Strasfeld, Daniel P. Raleigh, Martin T. Zanni *Proc. Nat. Acad. Sci. U.S.A.* **2009**, *106*, 6614.

- (21) Carlos R. Baiz, C. S. P., Mike E. Reppert, Kevin C. Jones, Andrei Tokmakoff
Analyst **2012**, *137*, 1793.
- (22) Bernadette Vrhovski, A. S. W. *Eur. J. Biochem.* **1998**, *258*, 1.
- (23) Urry, D. W. *J. Phys. Chem. B* **1997**, *101*, 11007.
- (24) Dean Y. Li, A. E. T., Beth B. Boak, Donald L. Atkinson, Gregory J. Ensing,
Colleen A. Morris, Mark T. Keating *Hum. Mol. Gen.* **1997**, *6*, 1021.
- (25) Daniel P. Wendel, D. G. T., Kurt H. Albertine, Mark T. Keating, Dean Y. Li *Am.
J. Respir. Cell Mol. Biol.* **2000**, *23*, 320.
- (26) Robert R. Mercer, J. D. C. *J. Appl. Physiol.* **1990**, *69*, 756.
- (27) Hans Oxlund, J. M., A. Viidik *J. Biomech.* **1988**, *21*, 213.
- (28) Qirui Hu, A. S., Carla Sens, Jiwon Choi, Zoltan Szabo, Barry C. Starcher, Russell
H. Knutsen, J. Michael Shipley, Elaine C. Davis, Robert P. Mecham, Zsolt Urban *Matrix Biol.*
2010, *29*, 621.
- (29) Satyajit K. Karnik, J. D. W., Lise Sorensen, Benjamin S. Brooke, Lisa D. Urness,
Dean Y. Li *Matrix Biol.* **2003**, *22*, 409.
- (30) Satyajit K. Karnik, B. S. B., Antonio Bayes-Genis, Lise Sorensen, Joshua D.
Wythe, Robert S. Schwartz, Mark T. Keating, Dean Y. Li *Development* **2003**, *130*, 411.
- (31) Benjamin S. Brooke, A. B.-G., Dean Y. Li *Trends Cardiovasc. Med.* **2003**, *13*,
176.
- (32) Dean Y. Li, B. B., Elaine C. Davis, Robert P. Mecham, Lise K. Sorensenk, Beth
B. Boak,; Ernst Eichwald, M. T. K. *Nature* **1998**, *393*, 276.
- (33) K. L. Dorrington, N. G. M. *Biopolymers* **1977**, *16*, 1201.
- (34) C. A. J. Hoeve, P. J. F. *Biopolymers* **1974**, *13*, 677.

- (35) Torkel Weis-Fogh, S. O. A. *Nature* **1970**, 227, 718.
- (36) William R. Gray, L. B. S., Judith A. Foster *Nature* **1973**, 246, 461.
- (37) C. M. Venkatachalam, D. W. U. *Macromolecules* **1981**, 14, 1225.
- (38) Valiaev, A.; Lim, D. W.; Oas, T. G.; Chilkoti, A.; Zauscher, S. *J. Am. Chem. Soc.* **2007**, 129, 6491.
- (39) Laurent Debelle, A. J. P. A. *Biochimie* **1999**, 81, 981.
- (40) Low, F. N. *Anat. Rec.* **1962**, 142, 131.
- (41) Patricia Brown-Augsburger, T. B., Joel Rosenbloom, Robert P. Mecham *Biochem J.* **1996**, 318, 149.
- (42) Perry, A.; Stypa, M. P.; Tenn, B. K.; Kumashiro, K. K. *Biophys. J.* **2002**, 82, 1086.
- (43) Pometun, M. S.; Chekmenev, E. Y.; Wittebort, R. J. *J. Biol. Chem.* **2004**, 279, 7982.
- (44) Cook, W. J.; Einspahr, H.; Trapane, T. L.; Urry, D. W.; Bugg, C. E. *J. Am. Chem. Soc.* **1980**, 102, 5502.
- (45) Karle, I. L.; Urry, D. W. *Biopolymers* **2005**, 77, 198.
- (46) Michael J. Fazio, M.-G. M., Edith Passage, Mon-Li Chu, Donald Black, Ellen Solomon, Jeffrey M. Davidson, Jouni Uitto *Am. J. Hum. Genet.* **1991**, 48, 696.
- (47) Zena Indik, H. Y., Norma Ornstein-Goldstein, Paul Sheppard, Noel Anderson, Joan C. Rosenbloom, Leena Peltonen, Joel Rosenbloom *Proc. Natl. Acad. Sci. U.S.A.* **1987**, 84, 5680.
- (48) L. Debelle, A. M. T. *Int. J. Biochem. Cell. Biol.* **1999**, 31, 261.
- (49) Aleksander Hinek, M. R. *J. Cell Biol.* **1994**, 126, 563.

- (50) H. A. Lucero, H. M. K. *Cell. Mol. Life Sci.* **2006**, *63*, 2304.
- (51) Gillian Francis, R. J., John Thomas *Biochem. J.* **1973**, *136*, 45.
- (52) Patricia Brown-Augsburger, C. T., Thomas Broekelmann, Carolyn Sloan, Robert P. Mecham *J. Biol. Chem.* **1995**, *270*, 17778.
- (53) Richard L. Kornberg, S. S. H., Aarne I. Oikarinen, Lois Y. Matsuoka, Jouni Uitto *N. Engl. J. Med.* **1985**, *312*, 771.
- (54) Gigante, A. M. D. C., C. M.D.; Greco, F. M.D. *J. Ped. Ortho.* **1999**, *19*, 283.
- (55) P. A. Abraham, A. J. P., William H. Carnes, Jouni Uitto *J. Clin. Invest.* **1982**, *70*, 1245.
- (56) Elizabeth R. Wright, V. P. C. *Adv. Drug Delivery Rev.* **2002**, *54*, 1057.
- (57) Sarah R. MacEwan, A. C. *Pept. Sci.* **2010**, *94*, 60.
- (58) Partridge, S. M. *Nature* **1967**, *213*, 1123.
- (59) Mahmoud Reza Banki, L. F., David W. Wood *Nat. Methods* **2005**, *2*, 659.
- (60) Wan-Yi Wu, C. M., Filomena Califano, Reza Banki, David W. Wood *Nat. Protoc.* **2006**, *1*, 2257.
- (61) Eugene D. Boland, J. A. M., Kristin J. Pawlowski, David G. Simpson, Gary E. Wnek, Gary L. Bowlin *Front. Biosci.* **2004**, *9*, 1422.
- (62) Dino Volpin, I. P.-R., Dan W. Urry, Lorenzo Gotte *J. Biol. Chem.* **1976**, *251*, 6871.
- (63) Bernadette Vrhovski, S. J., Anthony S. Weiss *Eur. J. Biochem.* **1997**, *250*, 92.
- (64) Tetsuji Yamaoka, T. T., Yuuki Seto, Tomoko Tada, Shigeru Kunugi, David A. Tirrell *Biomacromolecules* **2003**, *4*, 1680.
- (65) Dan W. Urry, B. S. *Nature* **1969**, *222*, 795.

- (66) M. Manno, A. E., V. Martorana, P. L. San Biagio, D. Bulone, M. B. Palma-Vittorelli, D. T. McPherson, J. Xu, T. M. Parker, D. W. Urry *Biopolymers* **2001**, *59*, 51.
- (67) Keeley, F. W.; Bellingham, C. M.; Woodhouse, K. A. *Phil. Trans. R. Soc. Lond. B* **2002**, *357*, 185.
- (68) Judith T. Cirulis, F. W. K. *Biochemistry* **2010**, *49*, 5726.
- (69) Urry, D. W. *J. Protein Chem.* **1988**, *7*, 1.
- (70) Steven D. Shapiro, S. K. E., Michael A. Province, John A. Pierce, Edward J. Campbell *J. Clin. Invest.* **1991**, *87*, 1828.
- (71) Urry, D. W. *Angew. Chem. Int. Ed.* **1993**, *32*, 819.
- (72) Dan W. Urry, S. P., Jie Xu, David T. McPherson *J. Am. Chem. Soc.* **1997**, *119*, 1161.
- (73) Partridge, S. M. *Chemistry and Molecular Biology of the Intercellular Matrix*; Academic Press Inc: New York, 1970; Vol. 1.
- (74) Partridge, S. M. *The physiology and biochemistry of muscle as a food*; Univ. Wisconsin Press, Madison, 1966.
- (75) Bin Li, V. D. *J. Muscle Res. Cell Motil.* **2002**, *23*, 561.
- (76) D. A. Torchia, K. A. P. *J. Mol. Biol.* **1973**, *76*, 419.
- (77) Urry, D. W. *Adv. Exp. Med. Biol.* **1974**, *43*, 211.
- (78) Urry, D. W.; Long, M. M.; Gross, E. *Crit. Rev. Biochem. Mol. Biol.* **1976**, *4*, 1.
- (79) Dan W. Urry, T. L. T., Hiroshi Sugano, Kari U. Prasad *J. Am. Chem. Soc.* **1981**, *103*, 2080.
- (80) Zeeshan Ahmed, J. P. S., Sanford A. Asher *Biopolymers* **2008**, *91*, 52.

- (81) Urry, D. W.; Shaw, R. G.; Prasad, K. U. *Biochem. Biophys. Res. Commun.* **1985**, *130*, 50.
- (82) Li, B.; Alonso, D. O. V.; Bennion, B. J.; Daggett, V. *J. Am. Chem. Soc.* **2001**, *123*, 11991.
- (83) Li, B.; Alonso, D. O. V.; Daggett, V. *J. Mol. Biol.* **2001**, *305*, 581.
- (84) Lewis, P. N.; Momany, F. A.; Scheraga, H. A. *Biochim. Biophys. Acta* **1973**, *303*, 211.
- (85) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425.
- (86) Herald Reiersen, A. R. C., Anthony R. Rees *J. Mol. Bio.* **1998**, *283*, 255.
- (87) Nicolini, C.; Ravindra, R.; Ludolph, B.; Winter, R. *Biophysical Journal* **2004**, *86*, 1385.
- (88) Eduard Schreiner, C. N., Björn Ludolph, Revanur Ravindra, Nikolaj Otte, Axel Kohlmeier, Roger Rousseau, Roland Winter, Dominik Marx *Phys. Rev. Lett.* **2004**, *92*.
- (89) Rousseau, R.; Schreiner, E.; Kohimeyer, A.; Marx, D. *Biophys. J.* **2004**, *2004*, 1393.
- (90) Debelle, L.; Wei, S. M.; Jacob, M. P.; Hornebeck, W.; Alix, A. J. P. *Eur. Biophys. J.* **1992**, *21*, 321.
- (91) Debelle, L.; Alix, A. J. P.; Jacob, M.-P.; Huvenne, J.-P.; Berjot, M.; Sombret, B.; Legrand, P. *J. Bio. Chem.* **1995**, *270*, 26099.
- (92) Debelle, L.; Alix, A. J. P.; Wei, S. M.; Jacob, M.-P.; Huvenne, J.-P.; Berjot, M.; Legrand, P. *Eur. J. Biochem.* **1998**, *258*, 533.
- (93) Peter C. Groß, W. P., Michael Zeppezauer *Verlag der Zeitschrift für Naturforschung* **2003**, *58*, 873.

- (94) Dana Kurková, J. K., Pavel Schmidt, Jiří Dybal, José Carlos Rodriguez-Cabello, Matilde Alonso *Biomacromolecules* **2003**, *4*, 589.
- (95) Pavel Schmidt, J. D., José Carlos Rodriguez-Cabello, Virginia Reboto *Biomacromolecules* **2005**, *6*, 697.
- (96) Younhee Cho, L. B. S., Satoshi Iimura, Yanjie Zhang, Jaibir Kherb, Ashutosh Chilkoti, J. Martin Scholtz, Paul S. Cremer *J. Am. Chem. Soc.* **2009**, *131*, 15188.
- (97) Vesna Serrano, W. L., Stefan Franzen *Biophys J.* **2007**, *93*, 2429.
- (98) Sandberg, L. B.; Leslie, J. G.; Leach, C. T.; Alvarez, V. L.; Torres, A. R.; Smith, D. W. *Pathologie-biologie* **1985**, *33*, 266.
- (99) Valery V. Andrushchenko, H. J. V., Elmar J. Prenner *J. Pept. Sci.* **2007**, *13*, 37.
- (100) Stéphane Roux, E. Z., Bernard Rousseau, Maïté Paternostre, Jean-Christophe Cintrat, Nicolas Fay *J. Pept. Sci.* **2008**, *14*, 354.
- (101) Adam W. Smith, A. T. *J. Chem. Phys.* **2007**, *126*, 045109.
- (102) Lessing, J.; Roy, S.; Reppert, M.; Baer, M.; Marx, D.; Jansen, T. L. C.; Knoester, J.; Tokmakoff, A. *Journal of the American Chemical Society* **2012**, *134*, 5032.
- (103) Roberts, S. T.; Loparo, J. J.; Tokmakoff, A. *J. Chem. Phys.* **2006**, *125*, 084502.
- (104) Sean T. Roberts, K. R., Andrei Tokmakoff *Acc. Chem. Res.* **2009**, *42*, 1239.
- (105) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491.
- (106) Yanjie Zhang, P. S. C. *Curr. Opin. Chem. Biol.* **2006**, *10*, 658.
- (107) Younhee Cho, Y. Z., Trine Christensen, Laura B. Sagle, Ashutosh Chilkoti, Paul S. Cremer *J. Phys. Chem. B* **2008**, *112*, 13765.

Appendix

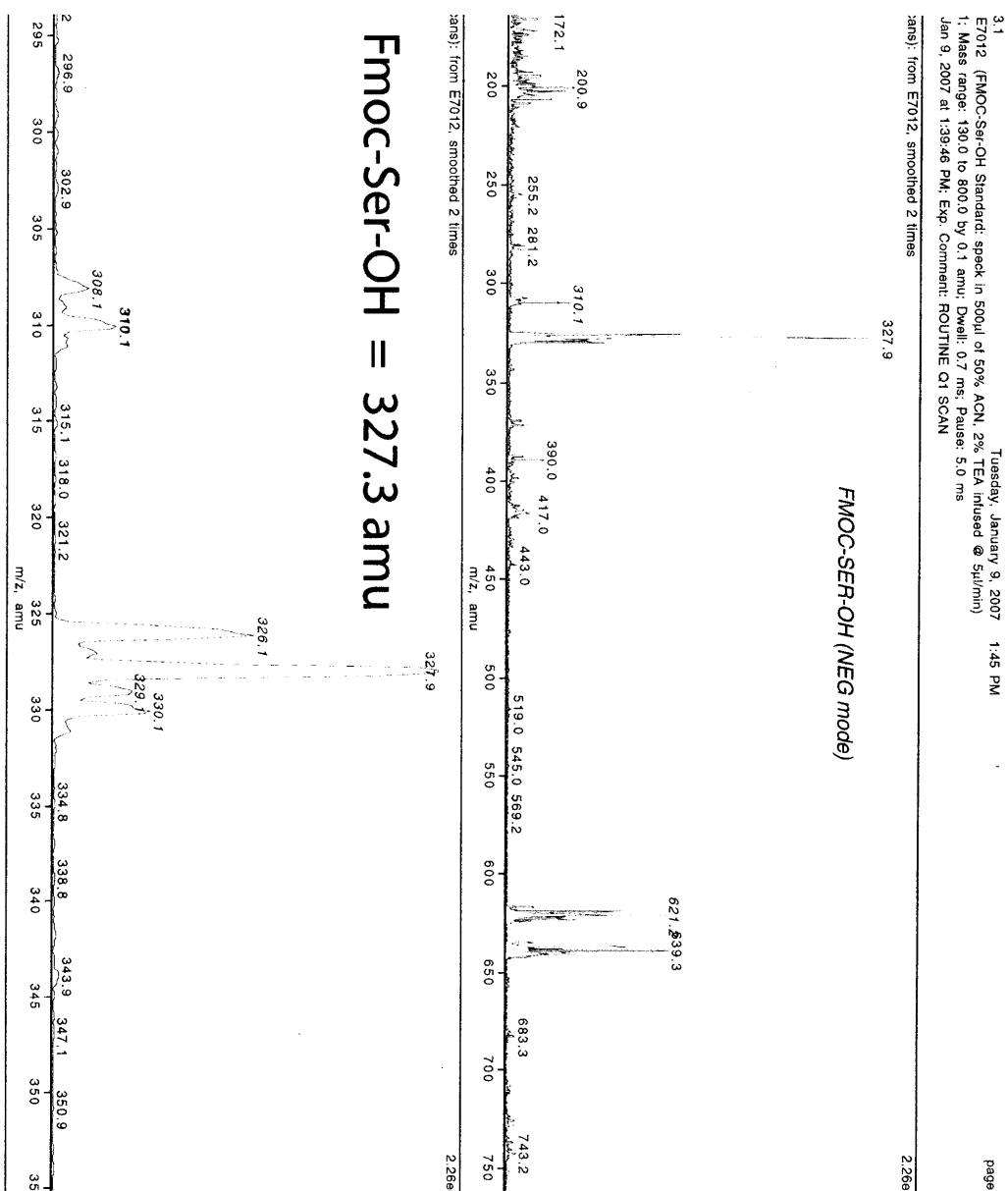


Figure 1. Electrospray mass spectra in negative ion mode of Fmoc-Ser-OH after one round of ^{18}O isotope labeling. The molecular weight of neutral unlabeled Fmoc-Ser-OH is 327.3 amu. Here the Fmoc-Ser-OH natural abundance peak is at 326.1 m/z and the isotope labeled peaks are at 327.9, 329.1, and 330.1 m/z.

1: E7019 (FMOC-SER-OSI: spinoff) in 50% ACN, 2% TEA
1: Mass range: 150.0 to 900.0 by 0.1 amu; Dwell: 0.7 ms; Pause: 5.0 ms
1: Jan 18, 2007 at 5:04:05 PM; Exp. Comment: ROUTINE Q1 SCAN

scans): from E7019, subtracted (scan 0 from E7013-7019NEG blank), smoothed 2 times 444.0 4,0096 cps

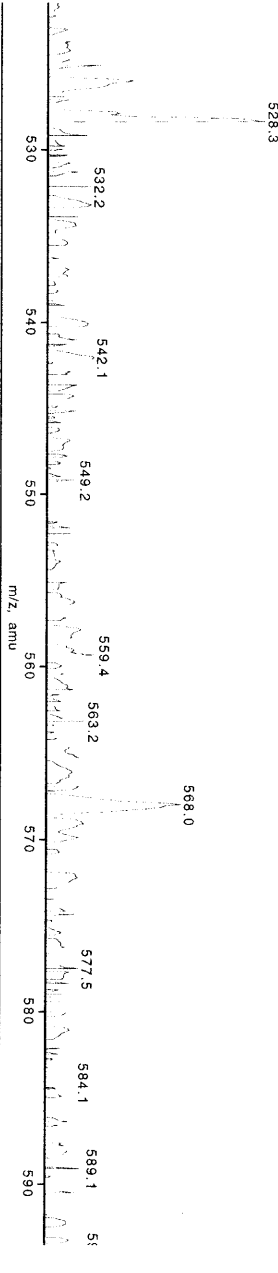
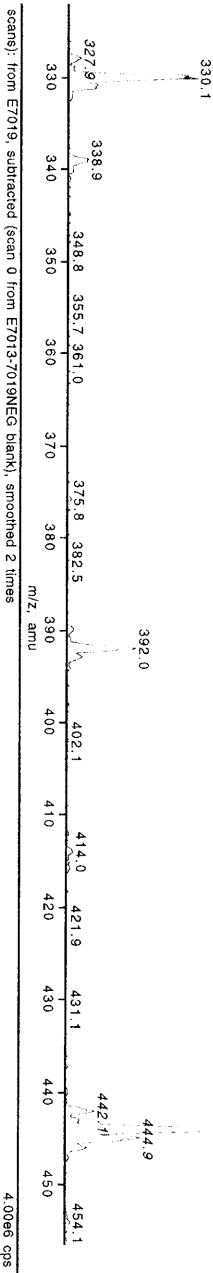
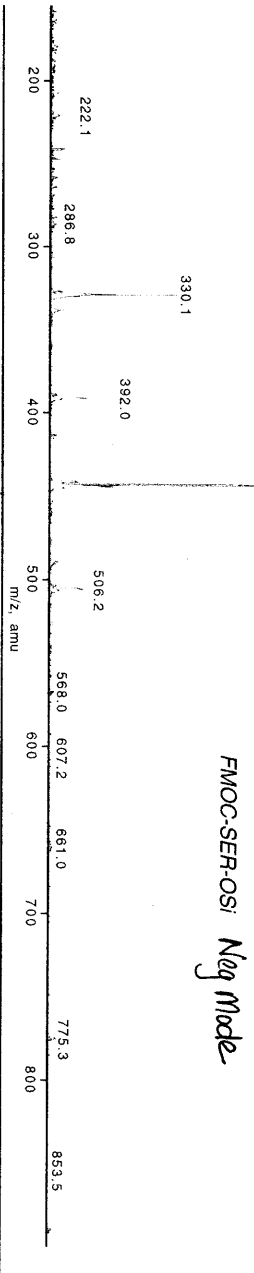


Figure 2. Elec
Fmoc-Ser-OSi(
OSi(CH₃)₂(C(CH₃)₃)
444.9 m/z.

TZ2-UL = NH₂-SWTWENGKWTWK-CONH₂

1:0052312V412014 2012.05.23 16:28:01N009-1001.D

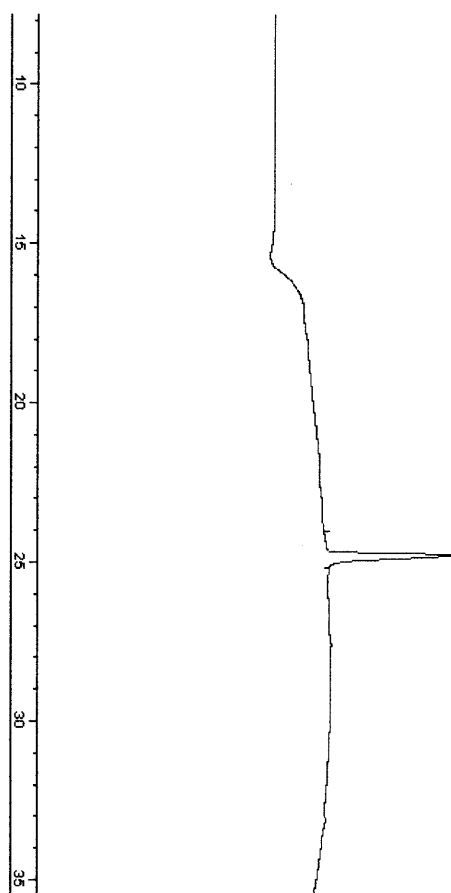
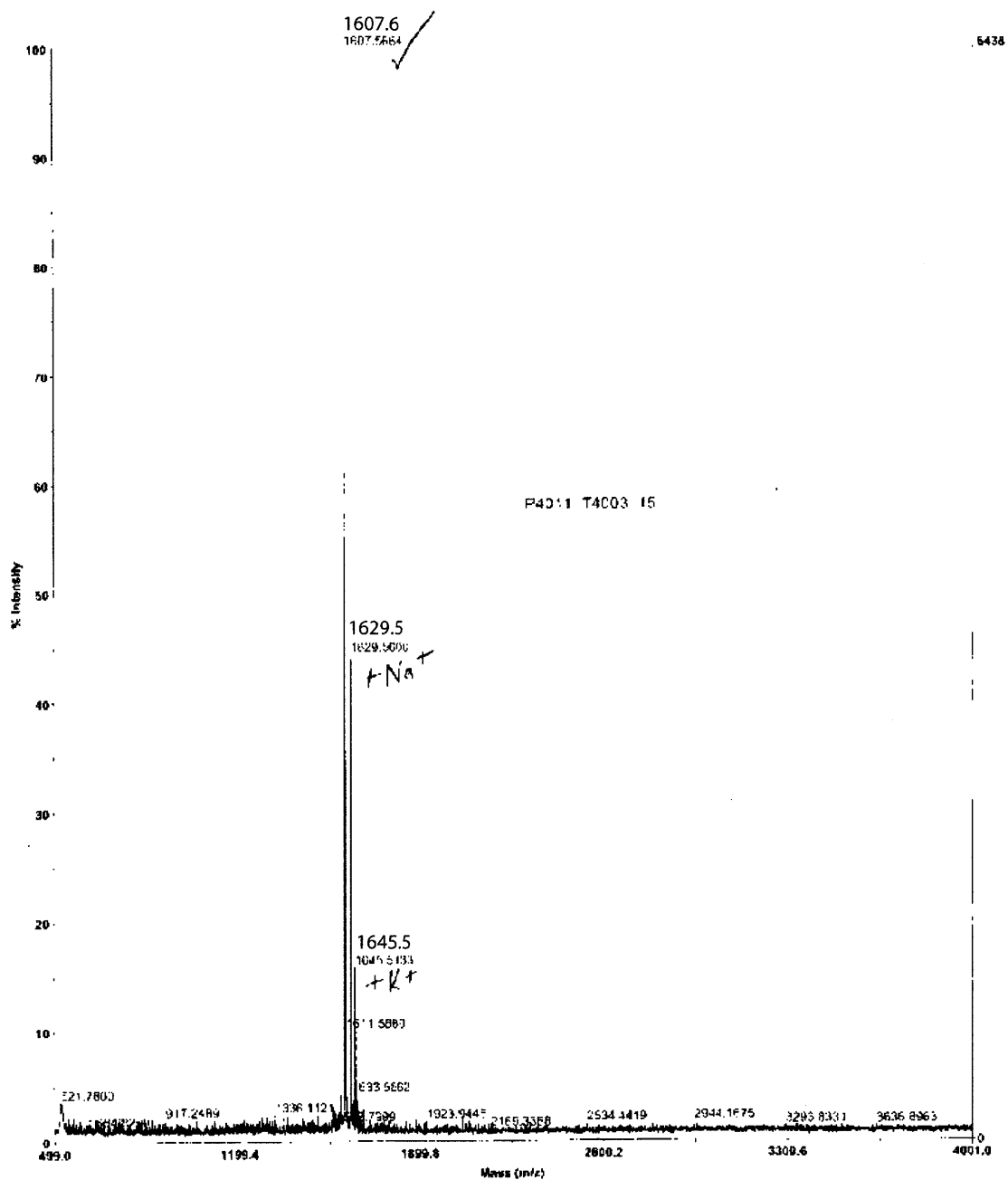


Figure 3. Analytical HPLC of the unlabeled TZ2 peptide (TZ2-UL) using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

Voyager Spec #1=>BC=>RSM10000=>NF0.7[BP = 1607.6, 6439]



D:\...P4011_T4003_15.dat
Acquired: 13:54:00, February 05, 2004

Figure 4. MALDI of the unlabeled (UL) TZ2 peptide. The molecular weight of neutral TZ2-UL is 1607.8 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 1607.6, 1629.5, and 1645.5 m/z respectively.

S1 = NH₂-S(C18O)WTWENGKWTWK-C

n (062312)612014 2012-05-23 16:26:01(10:1101.D)

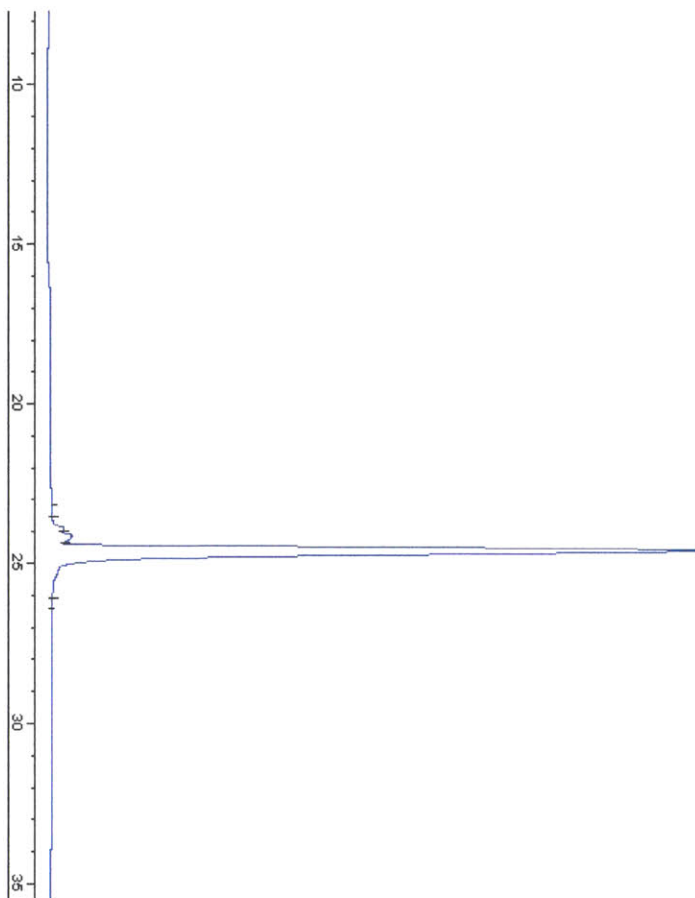


Figure 5. Analytical HPLC of TZ2-S1 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

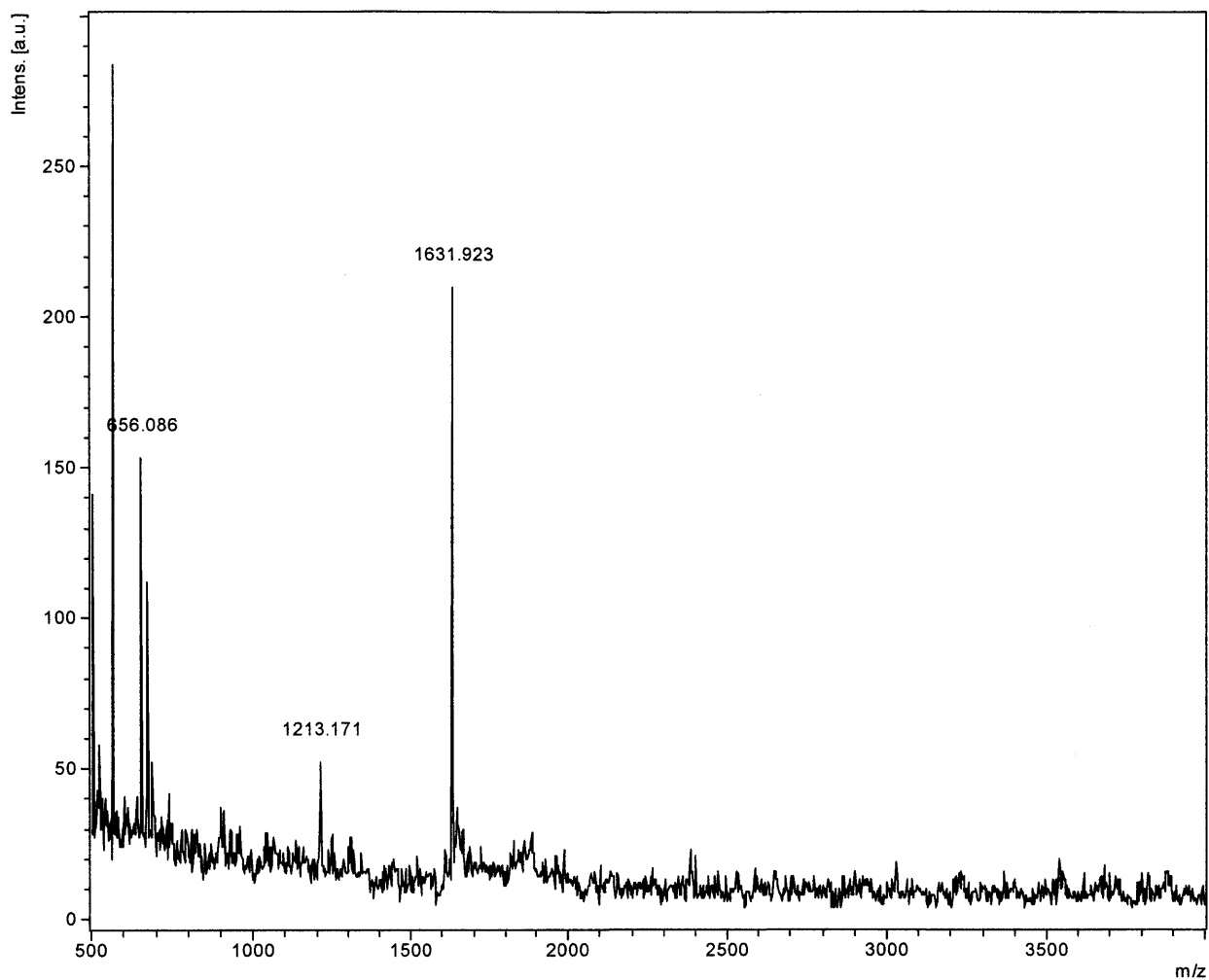


Figure 6. MALDI of the TZ2-S1 peptide. The molecular weight of neutral TZ2-UL is 1607.8 amu. Here the Na^+ peaks appear at 1631.9 m/z.

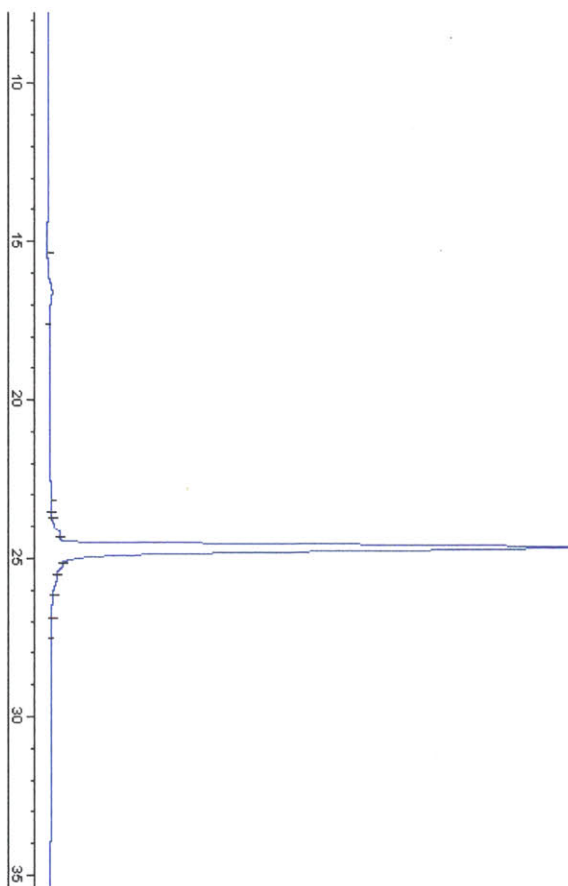


Figure 7. Analytical HPLC of TZ2-K8 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

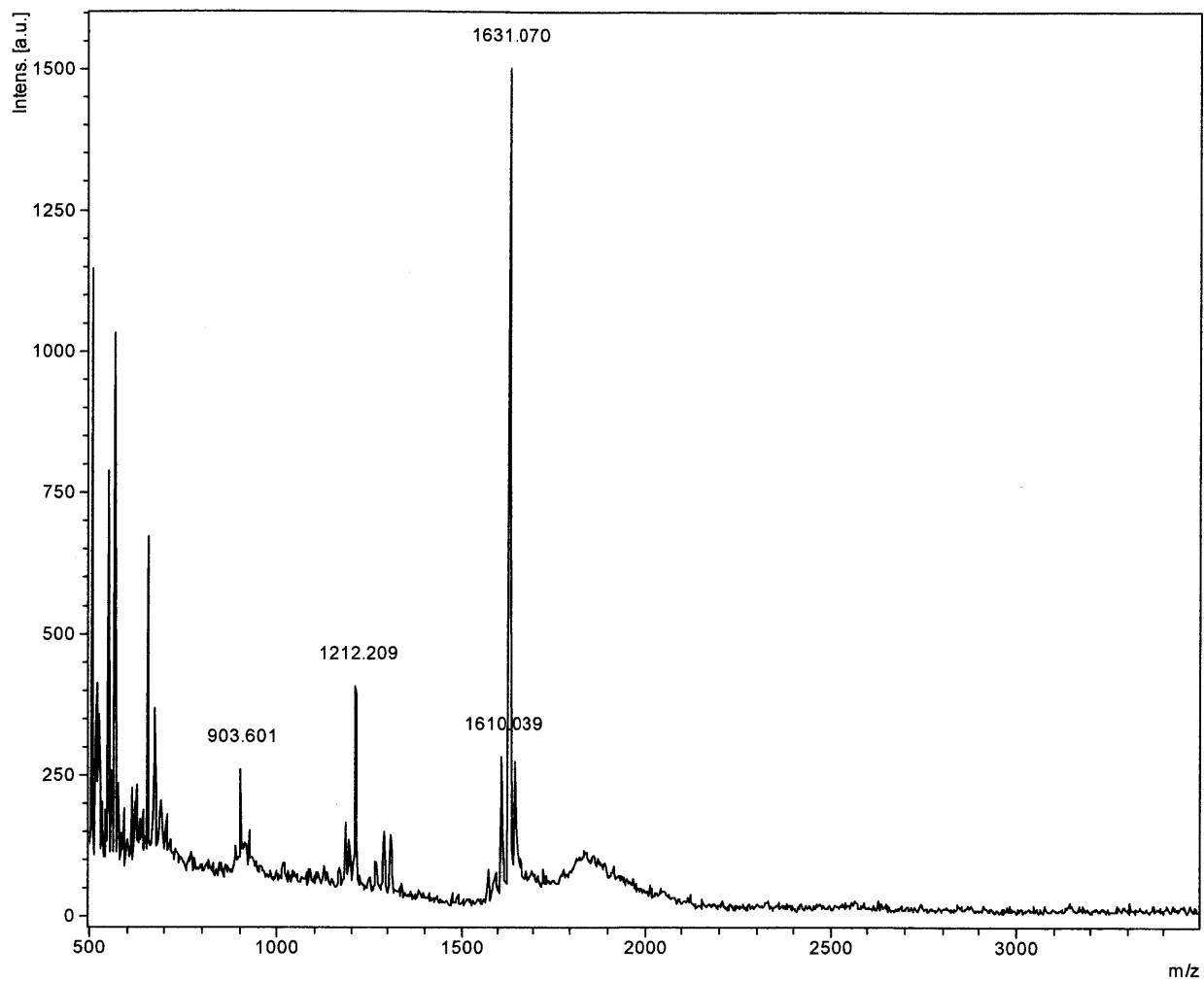


Figure 8. MALDI of the TZ2-K8 peptide. The molecular weight of neutral TZ2-UL is 1607.8 amu. Here the Na^+ peaks appear at 1631.1 m/z.

T10 = NH₂-SWTWENGKWT(C₁₈O)WK-I

71 (05231216-12014 2012-05-23 16:26:01007-0801.D)

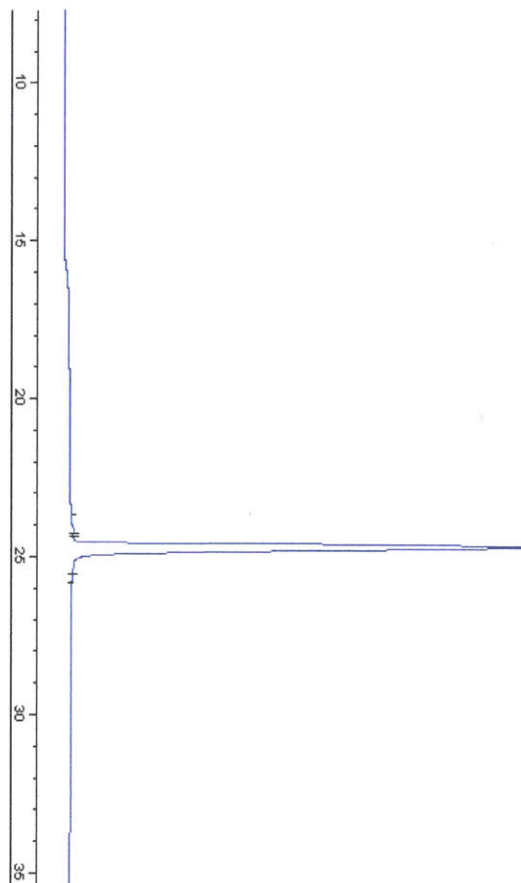


Figure 9. Analytical HPLC of T10 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

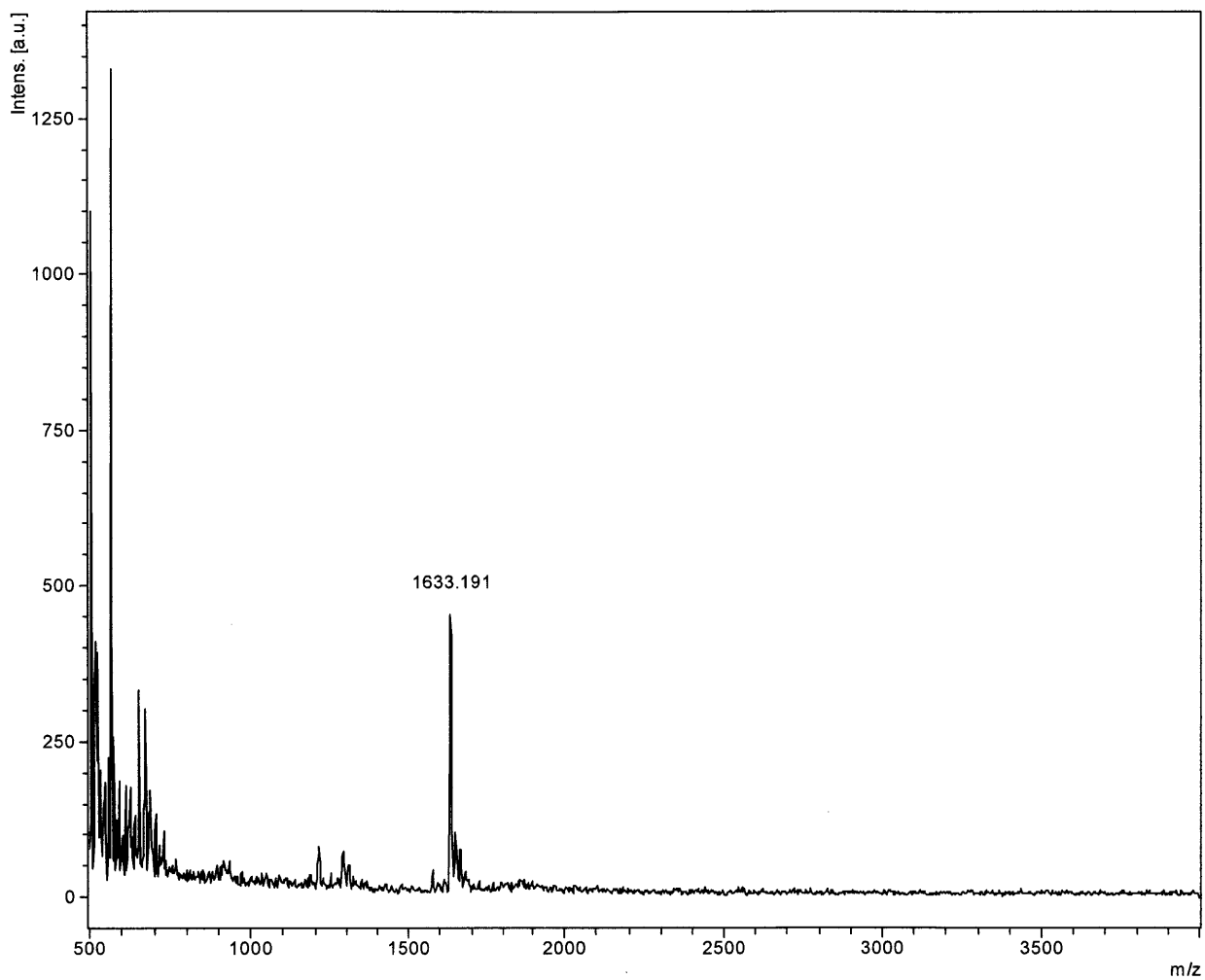


Figure 10. MALDI of the TZ2-T10 peptide. The molecular weight of neutral TZ2-UL is 1607.8 amu. Here the Na⁺ peaks appear at 1633.2 m/z.

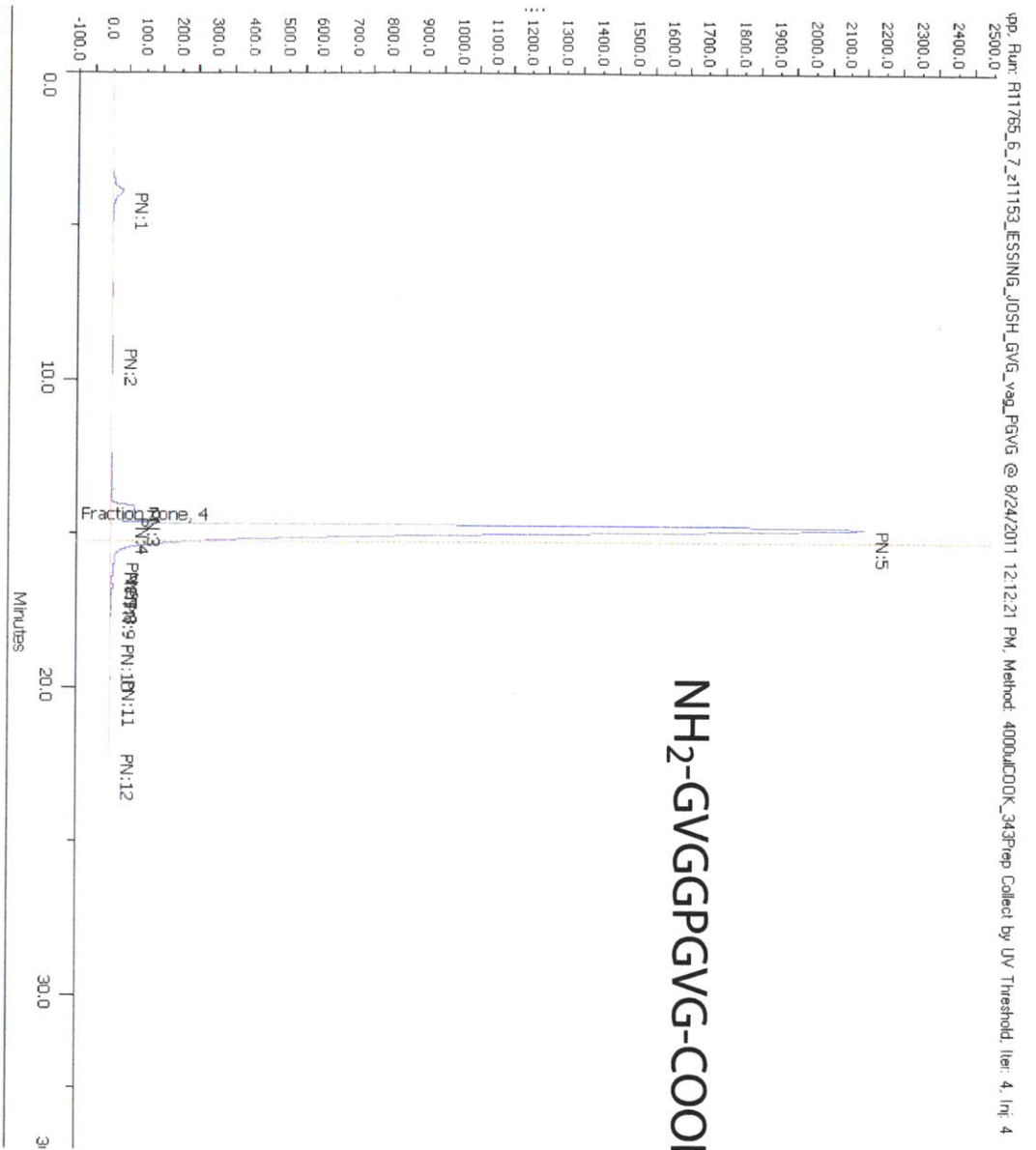
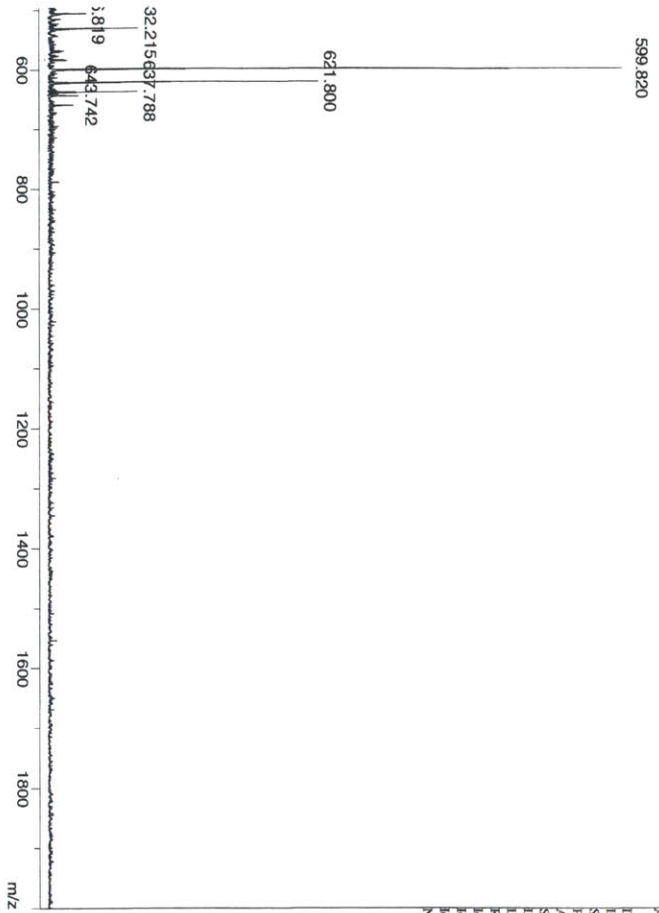


Figure 11. Analytical HPLC of GVGGPGVG using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

NH₂-GVGGPVG-COOH



Acquisition method name: D:\Esthinky_Methods\RP_PepM
 Date of acquisition: 2011-08-26T14:10:07.807-04:00
 Instrument type: msnucleo
 Serial instrument number: 256969.00047
 PFT: delay in [ns]: 160 ns
 Acquisition operation mode: Reflector
 Sample name (file name prefix): Z11153_3_VO_E11V1
 Laser repetition rate in Hz: 60 Hz
 Linear detector voltage: 2.698 kV
 Reflector detector voltage: 1.618 kV
 Ion source voltage: 19 kV
 Ion source: lens voltage: 15.5 kV
 Ion source: lens voltage: 9.5 kV
 Number of shots: 200



nics flexAnalysis

printed: 8/26/2011 2:10:28 PM

OLYMERS\2011\August\082611\Z11153_3_VO_E11V1



Figure 12. MA of the neutral 599.8, 621.8 a

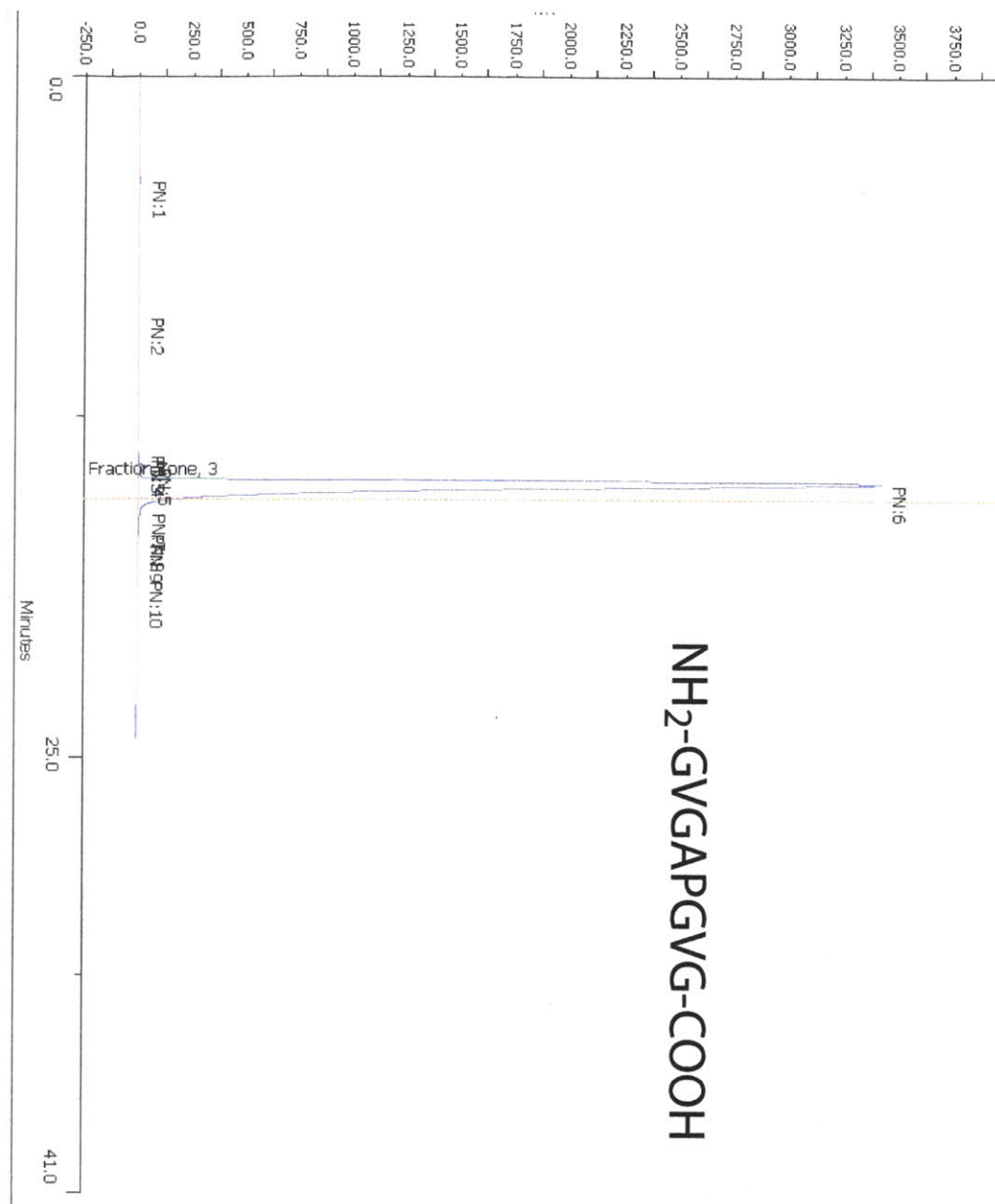
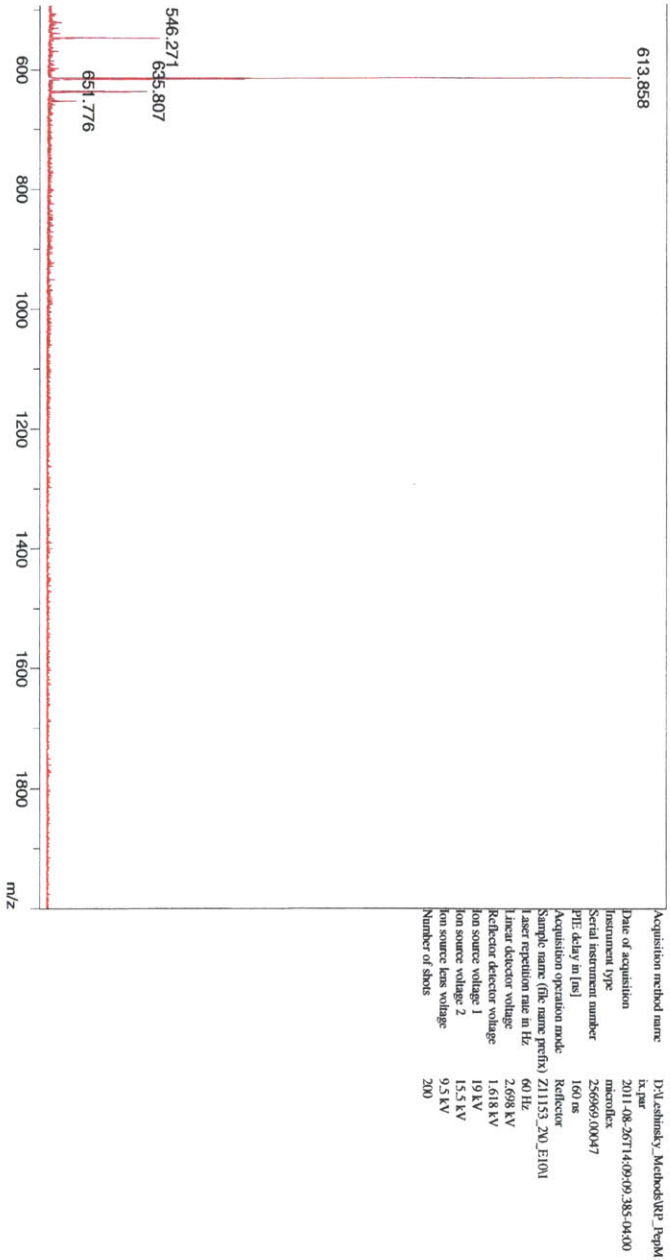


Figure 13. An phase buffer 0.043% TFA in

NH₂-GVGAPGVG-COOH



nics flexAnalysis

printed: 8/26/2011 2:09:24 PM

OLYMERS\2011\August\082611\Z11153_200_E101



Figure 14. MA of the neutral 613.9, 635.8 a

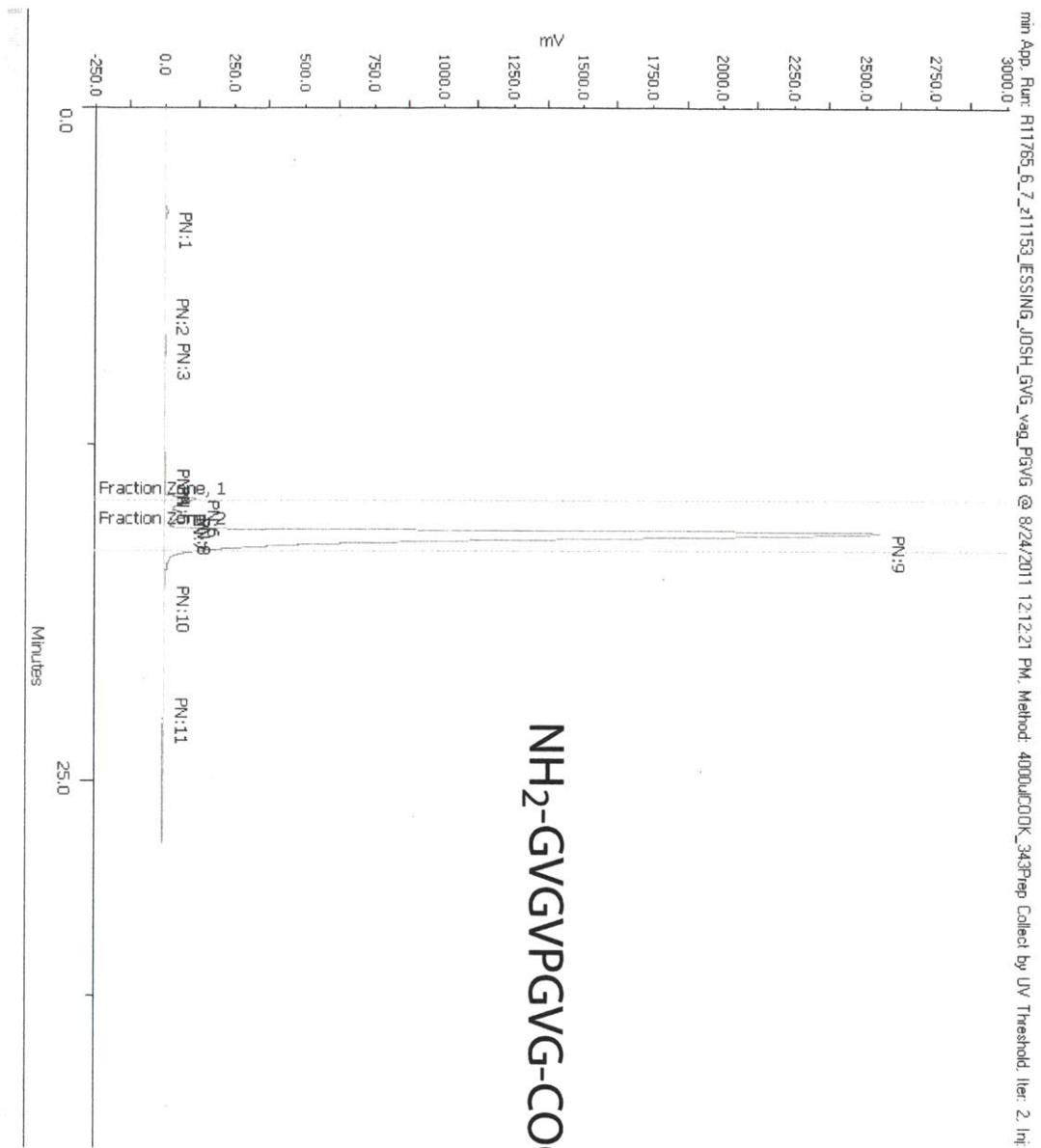
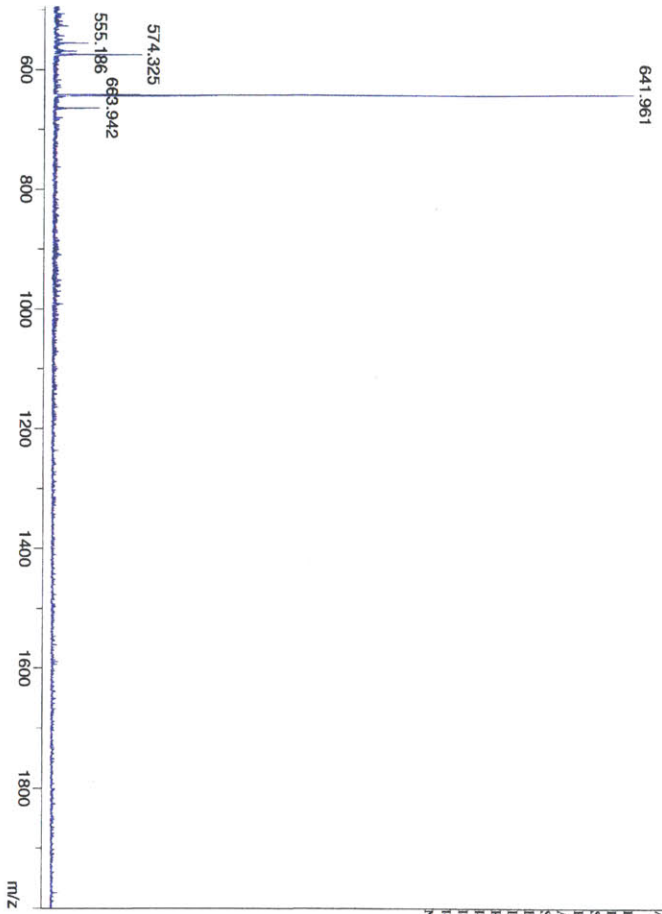


Figure 15. Analytical HPLC of GVGVPVG (Val or GVGN1) using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

NH₂-GVGVPGVG-COOH



Acquisition method name D:\zshinsky_Methods\RP_1p.p
Date of acquisition 2011-08-26T14:08:38.745-04:00
Instrument type microflex
Serial instrument number 25696910047
PTE delay in [ns] 160 ns
Acquisition operation mode Reflector
Sample name (file name prefix) Z11153_1\0_E9V2
Laser repetition rate in Hz 60 Hz
Linear detector voltage: 2.698 kV
Reflector detector voltage: 1.618 kV
Ion source voltage: 1 19 kV
Ion source voltage: 2 15.5 kV
Ion source lens voltage 9.5 kV
Number of shots 200



nics flexAnalysis

printed: 8/26/2011 2:08:53 PM

OLYMERS\2011\August\082611\Z11153_1\0_E9V2



Figure 16. M
molecular wei
Na⁺ peaks app

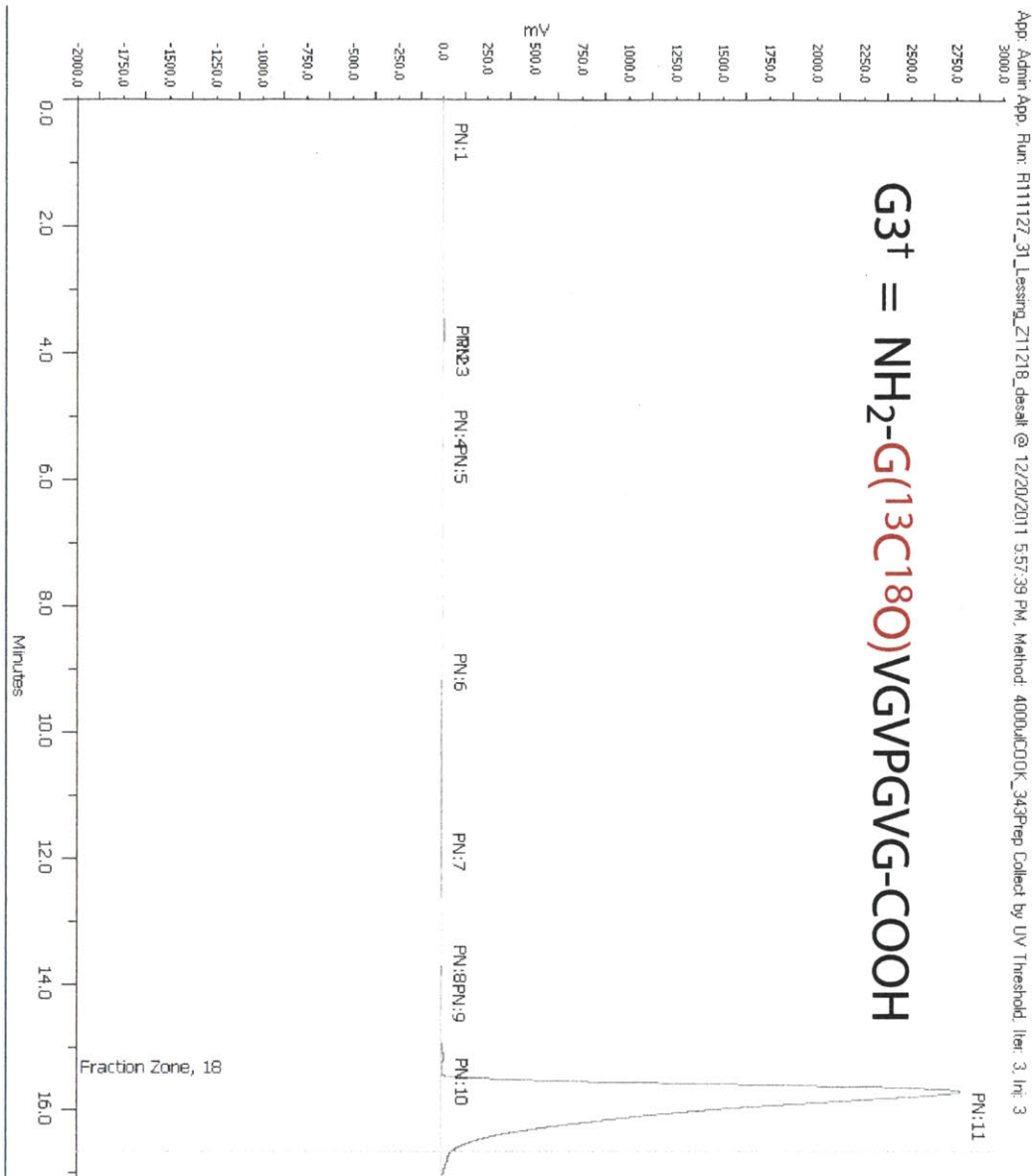
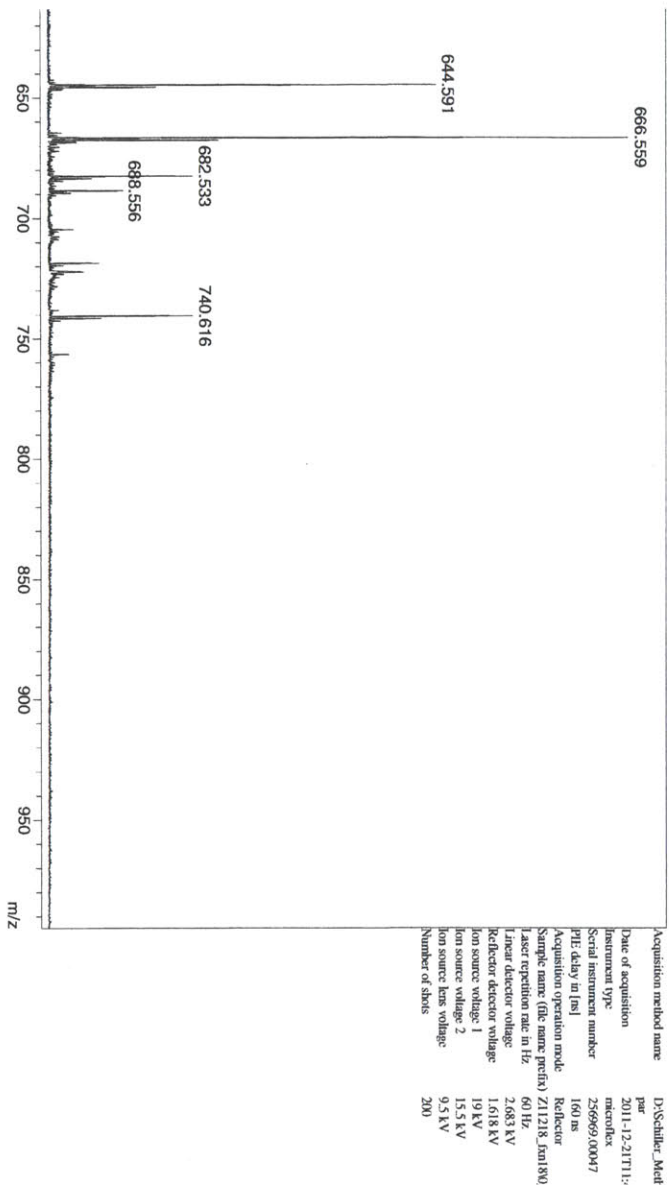


Figure 17. Analytical HPLC of G3⁺ using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.



Ionics flexAnalysis
 147
 printed: 12/21/2011 11:53:11
 JPOLYMERS2011\December2011\2211\Z11218_1m180_D9M



Figure 18. MALDI of the isotope labeled G3⁺ peptide. The molecular weight of the neutral unlabeled GVGN1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 644.6, 666.6, and 682.5 m/z respectively.

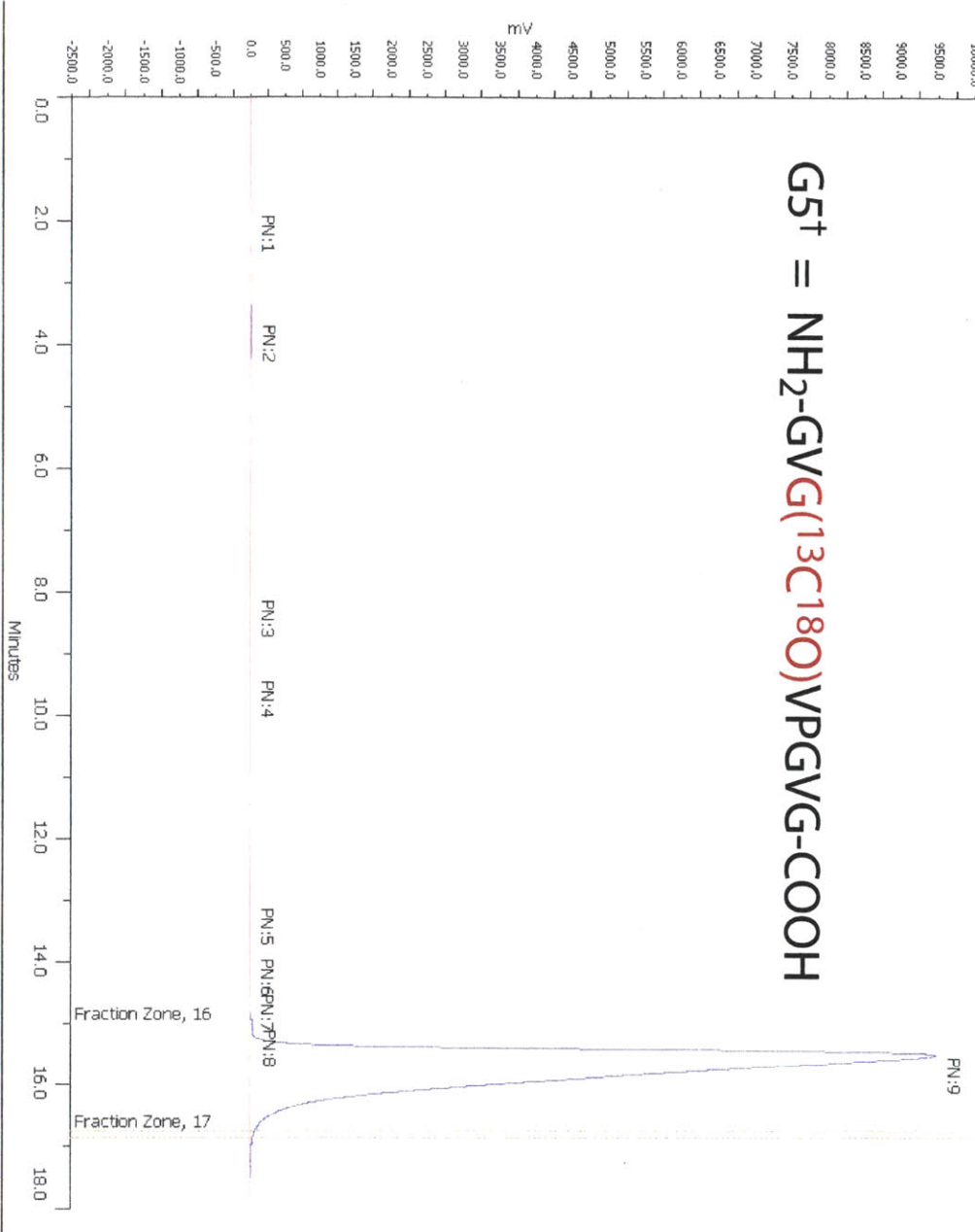
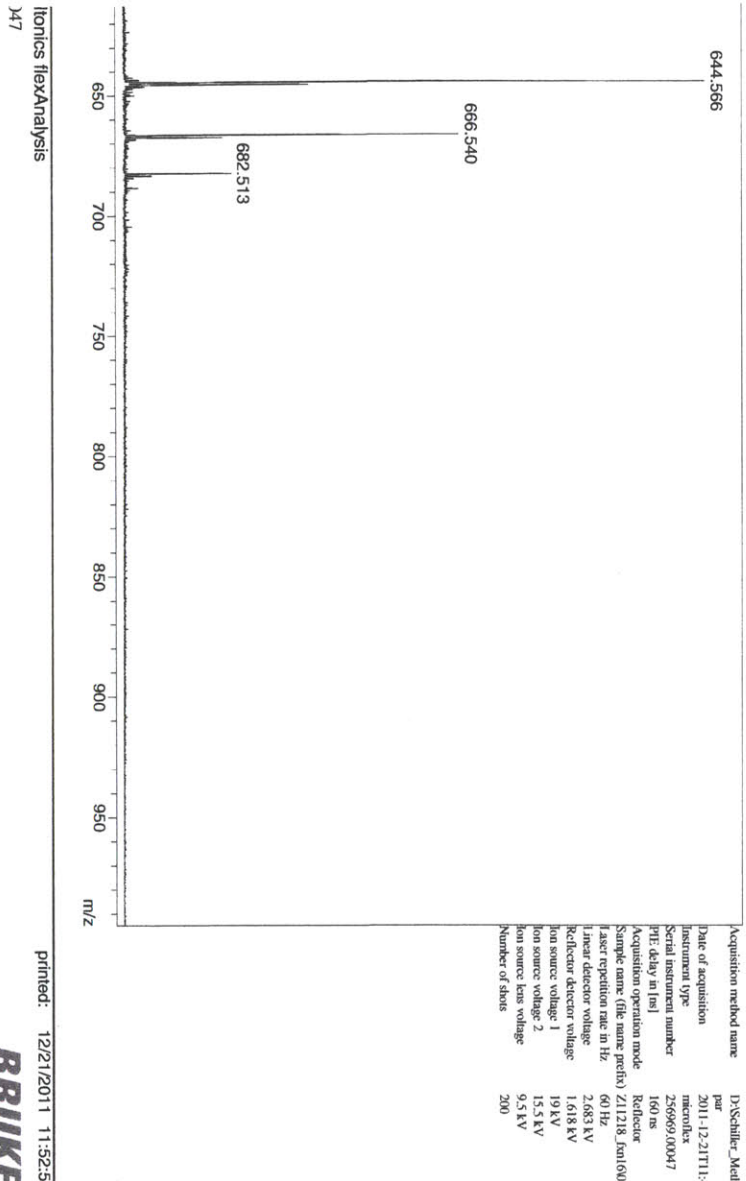


Figure 19. An gradient: (buf H₂O).



OPOLYMERS\2011\December\2011\122111\Z11218_Fxn160_D81



Figure 20. MALDI of the isotope labeled $G5^+$ peptide. The molecular weight of the neutral unlabeled $GVGn1$ peptide is 640.7 amu. Here the main ion, Na^+ and K^+ peaks appear at 644.6, 666.5, and 682.5 m/z respectively.

V1 = NH₂-GVGV(13C18O)PGVG-COOH

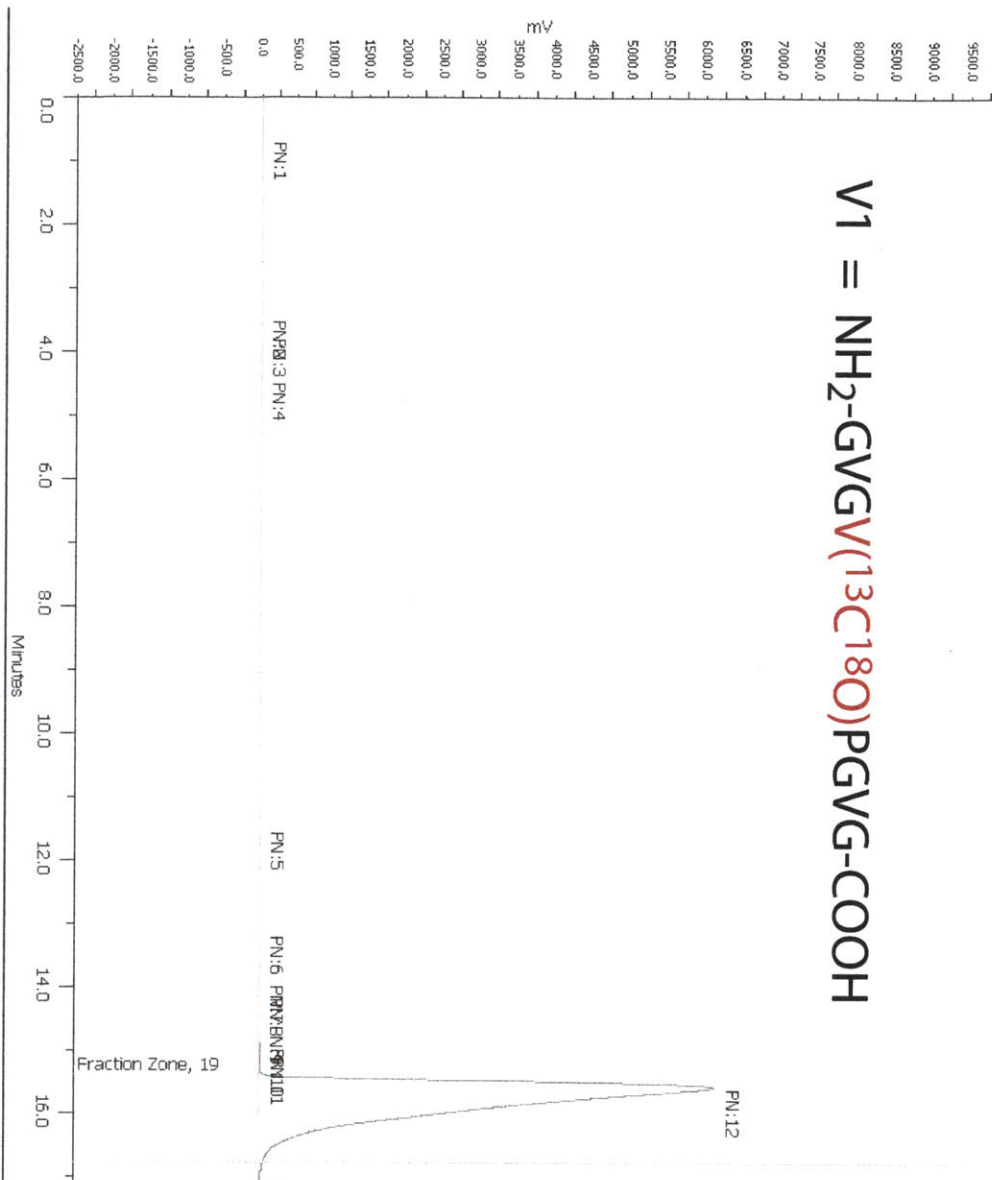
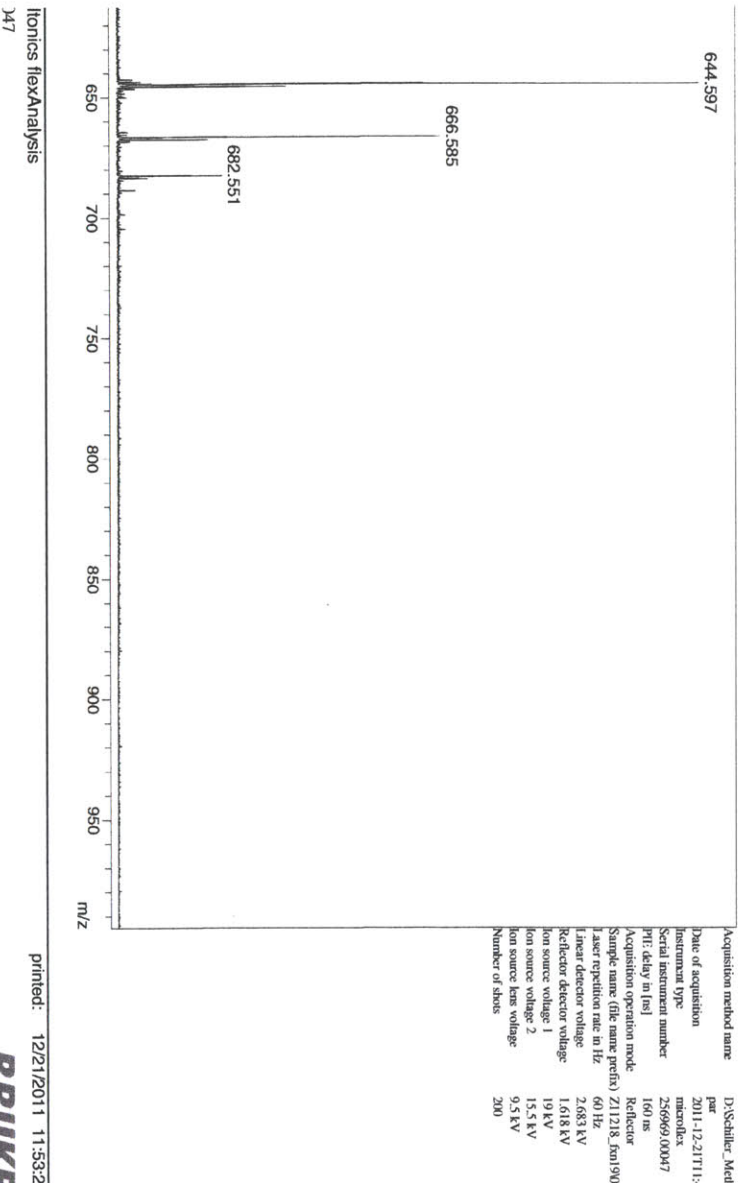


Figure 21. Analytical HPLC of V1 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

V1 = NH₂-GVGV(¹³C¹⁸O)PGVG-COOH



OPOLYMERS\2011\December\2011\2211\Z11218_fm1190_D101



Figure 22. MALDI of the isotope labeled V1 peptide. The molecular weight of the neutral unlabeled GVGN1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 644.6, 666.6, and 682.6 m/z respectively.

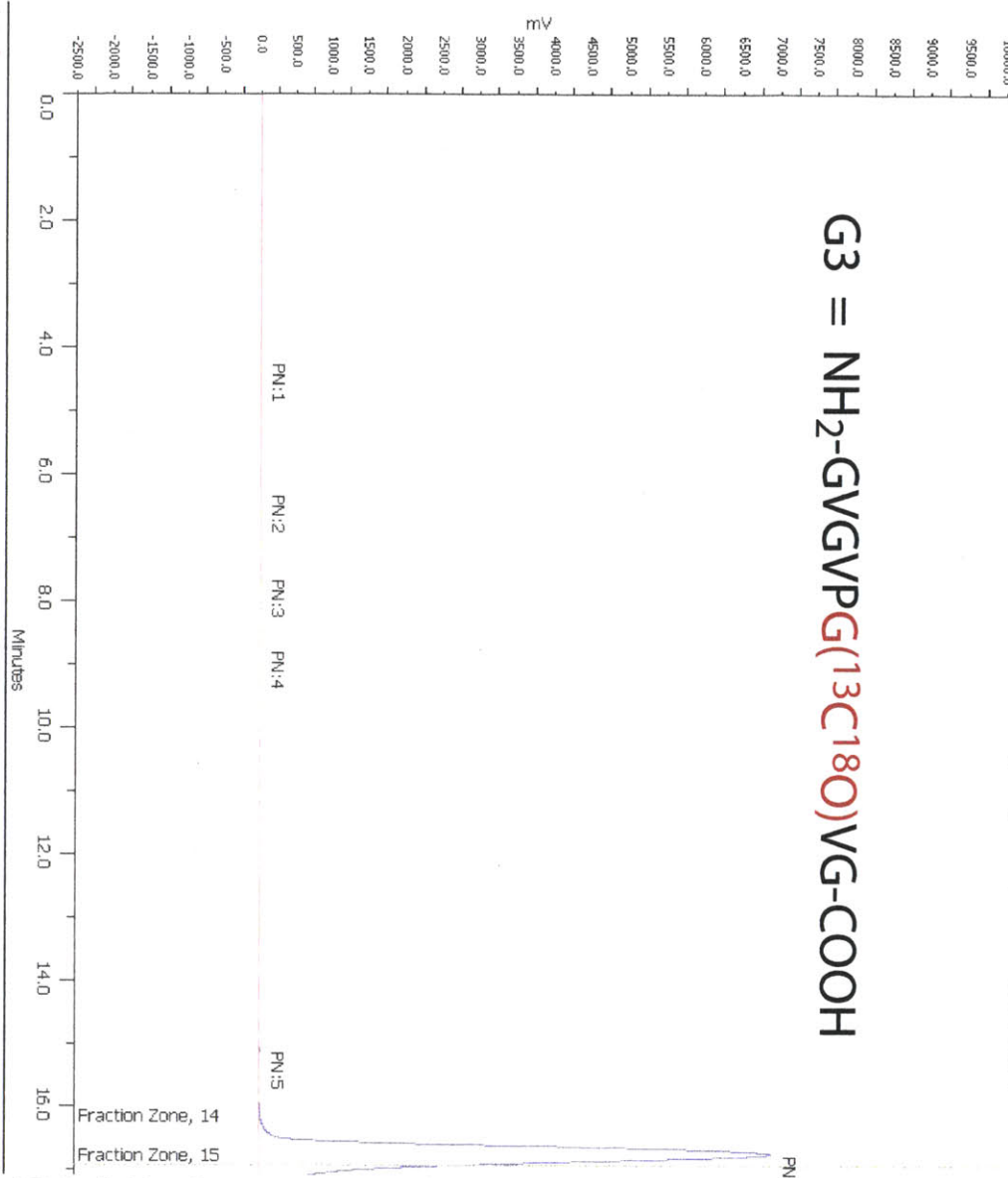
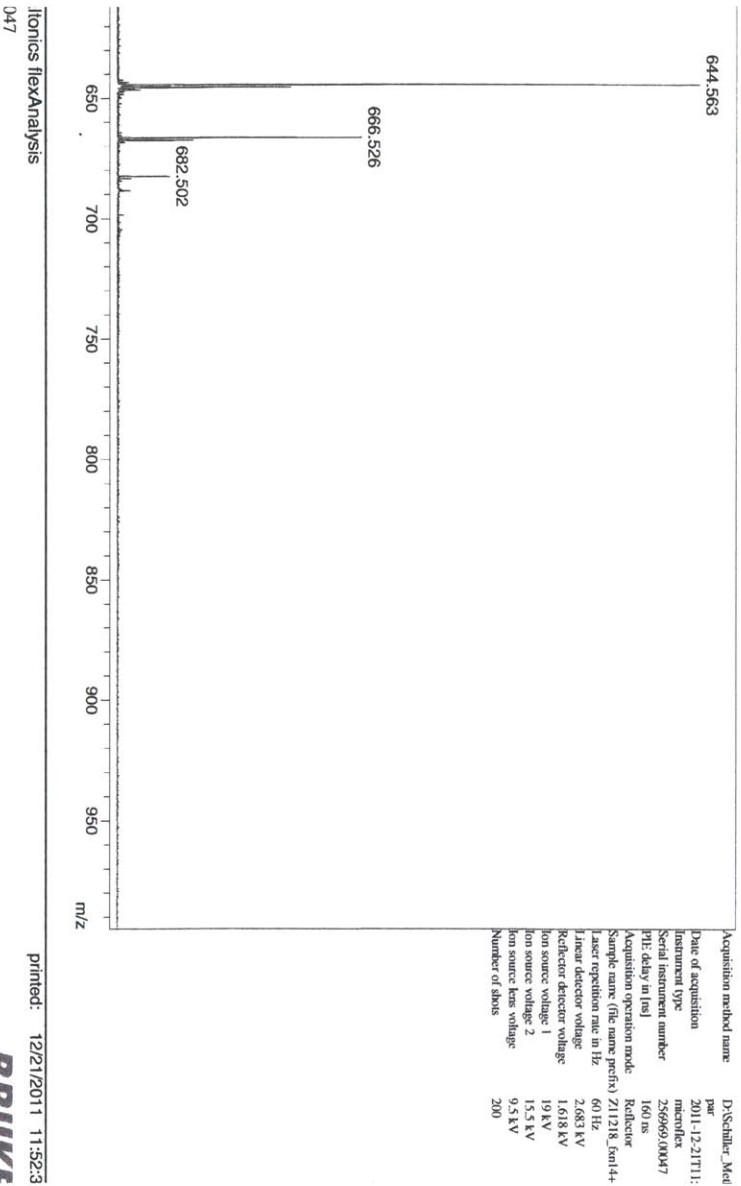


Figure 23. Analytical HPLC of G3 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

G3 = NH₂-GVGVPG(13C18O)VG-COOH



OPOL_YMERS2011\December2011\Z11218_fan44+150_D71



Figure 24. MALDI of the isotope labeled G3 peptide. The molecular weight of the neutral unlabeled GVGn1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 644.6, 666.5, and 682.5 m/z respectively.

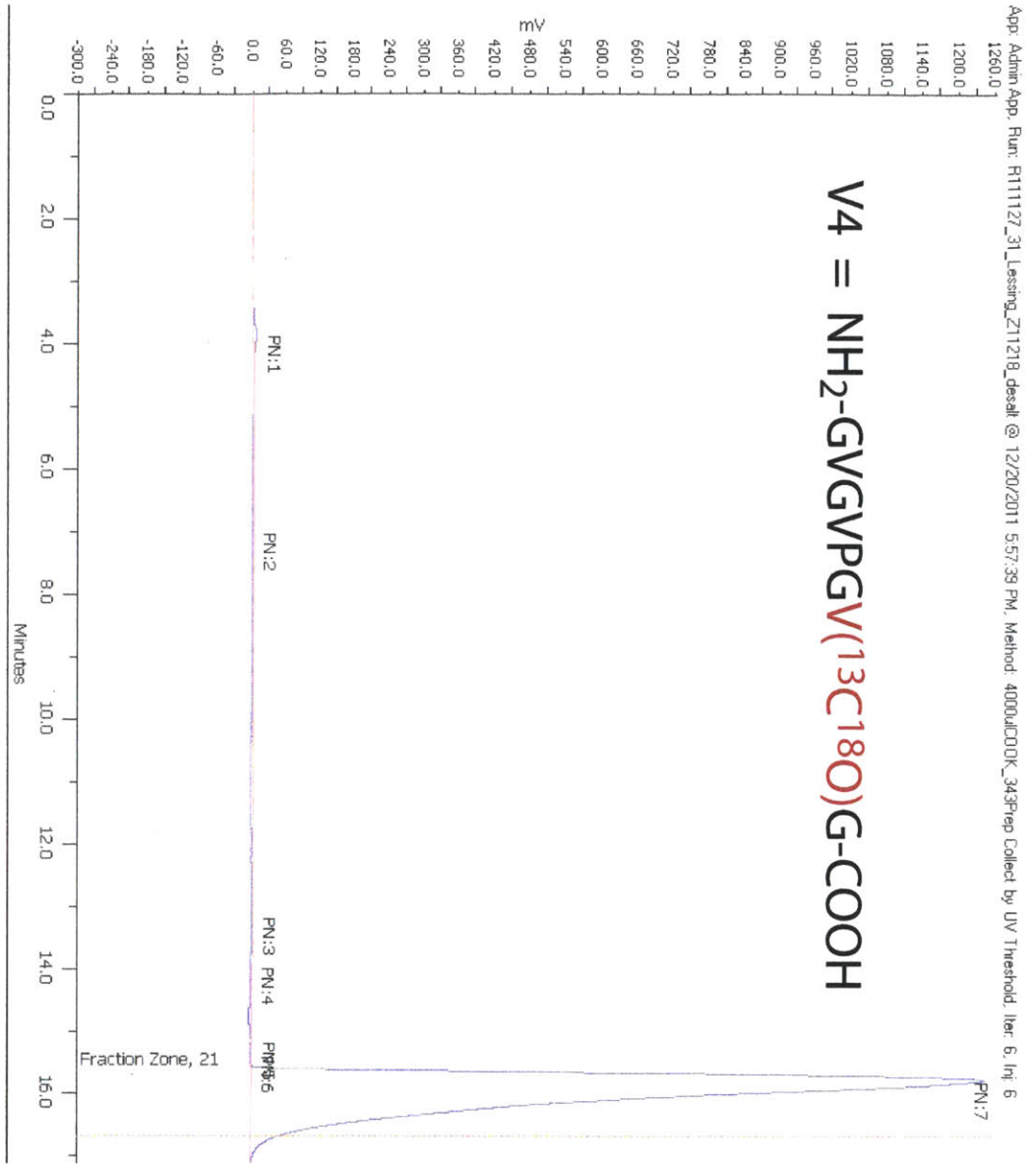


Figure 25. Analytical HPLC of V4 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

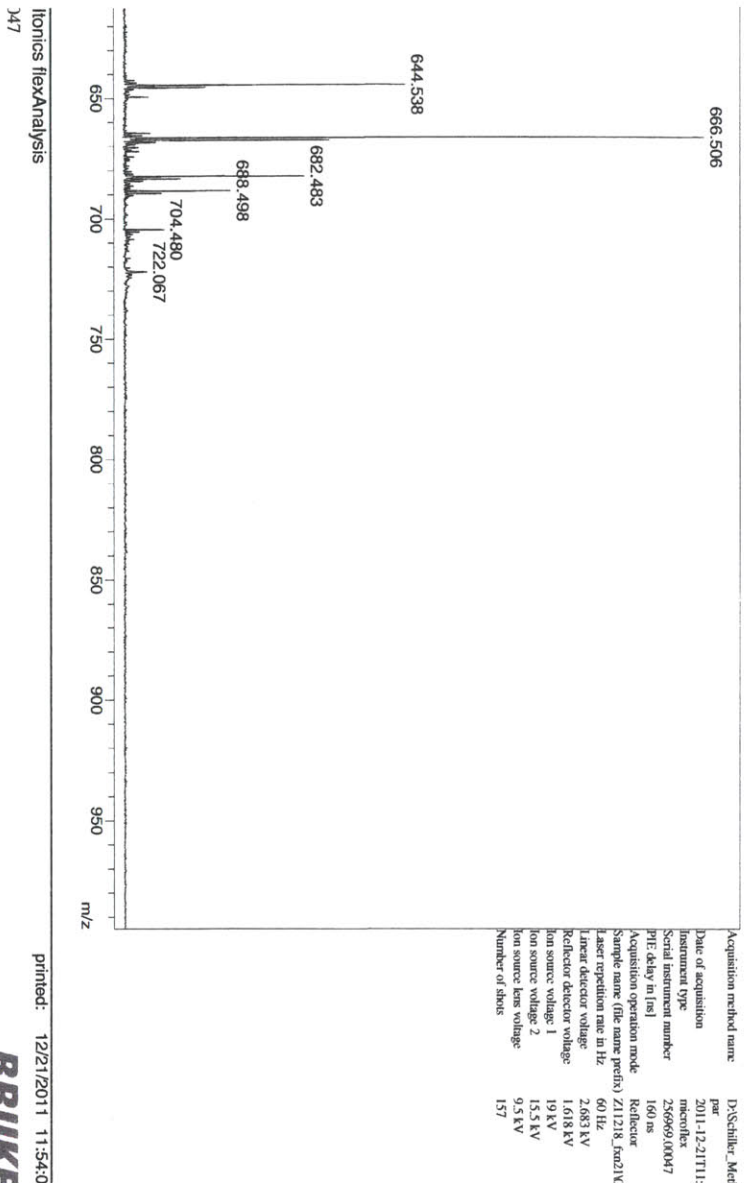


Figure 26. MALDI of the isotope labeled V4 peptide. The molecular weight of the neutral unlabeled GVGn1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 644.5, 666.5, and 682.5 m/z respectively.

G5[†]G3 = NH₂-GVG(13C18O)VPG(13C18O)VG-COOH

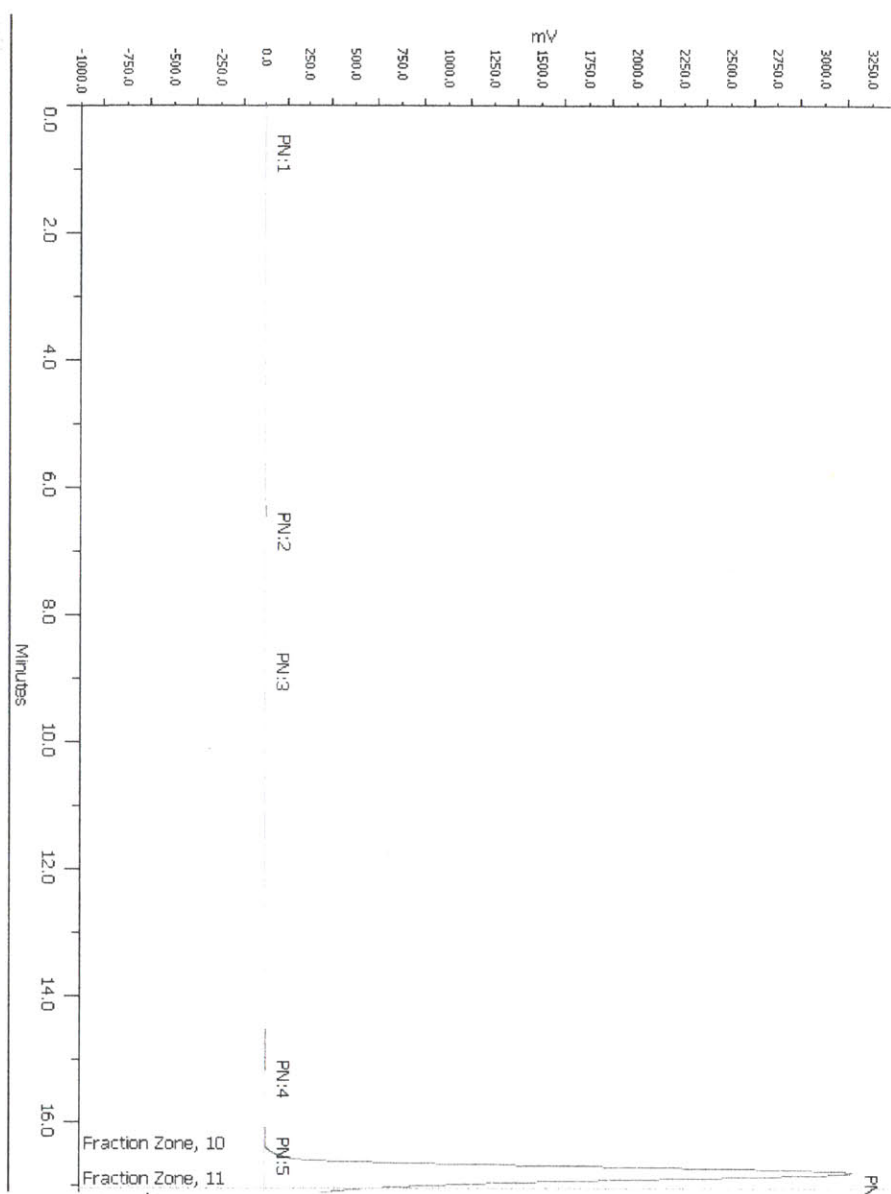
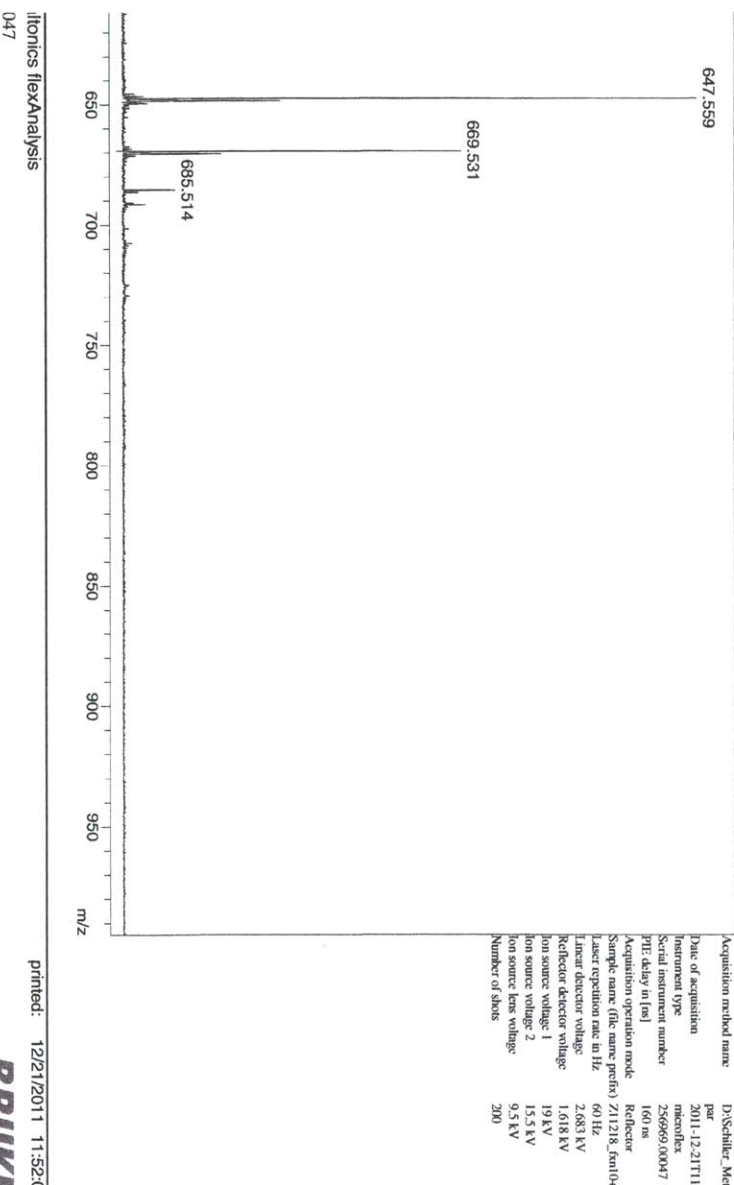


Figure 27. Analytical HPLC of G5[†]G3 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.



OPOLYMERS2011\December2011\Z11218_fm10+1\0_D51

itronics flexAnalysis
047



Figure 28. MALDI of the isotope labeled $G5^+G3$ peptide. The molecular weight of the neutral unlabeled $GVGn1$ peptide is 640.7 amu. Here the main ion, Na^+ and K^+ peaks appear at 647.6, 669.5, and 685.5 m/z respectively.

V1V4 = NH₂-GVGV(13C18O)PGV(13C18O)G-COOH

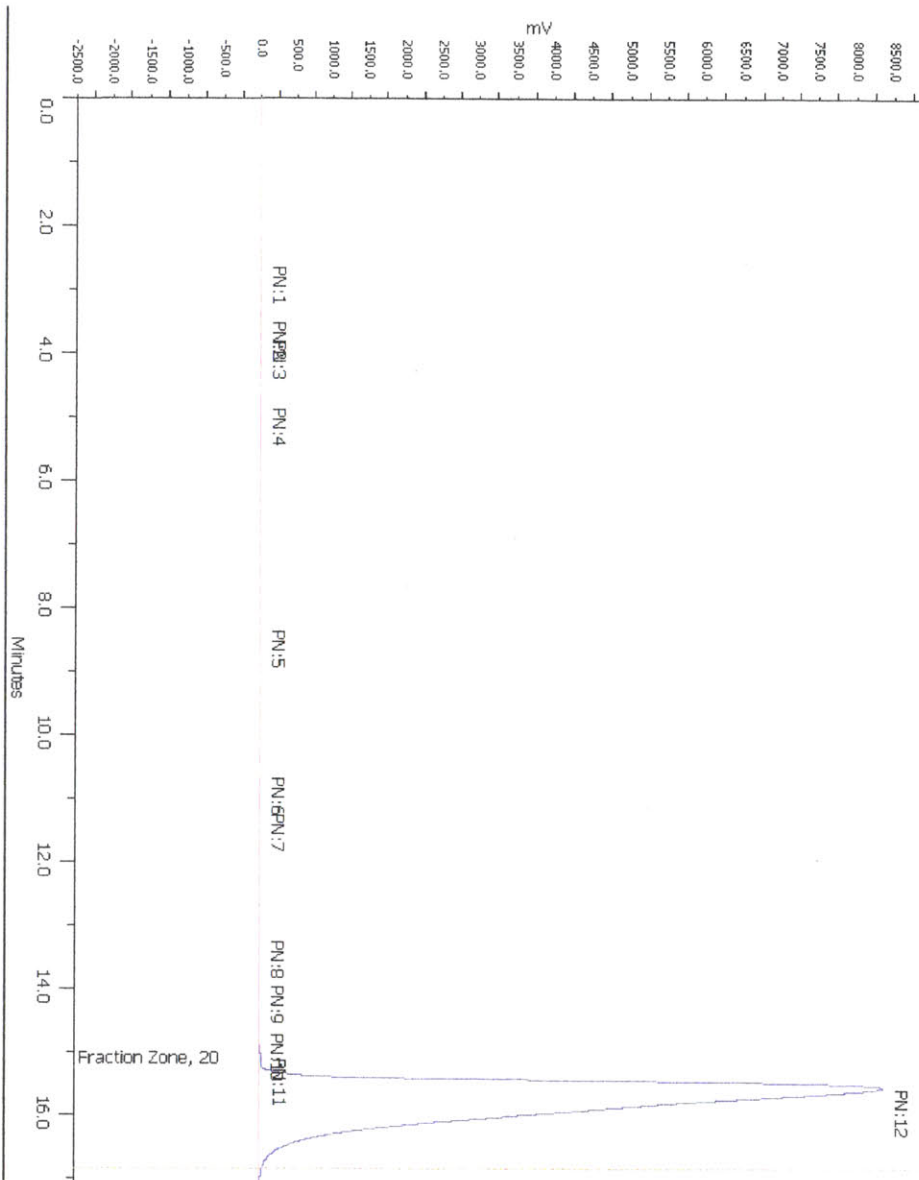


Figure 29. Analytical HPLC of V1V4 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

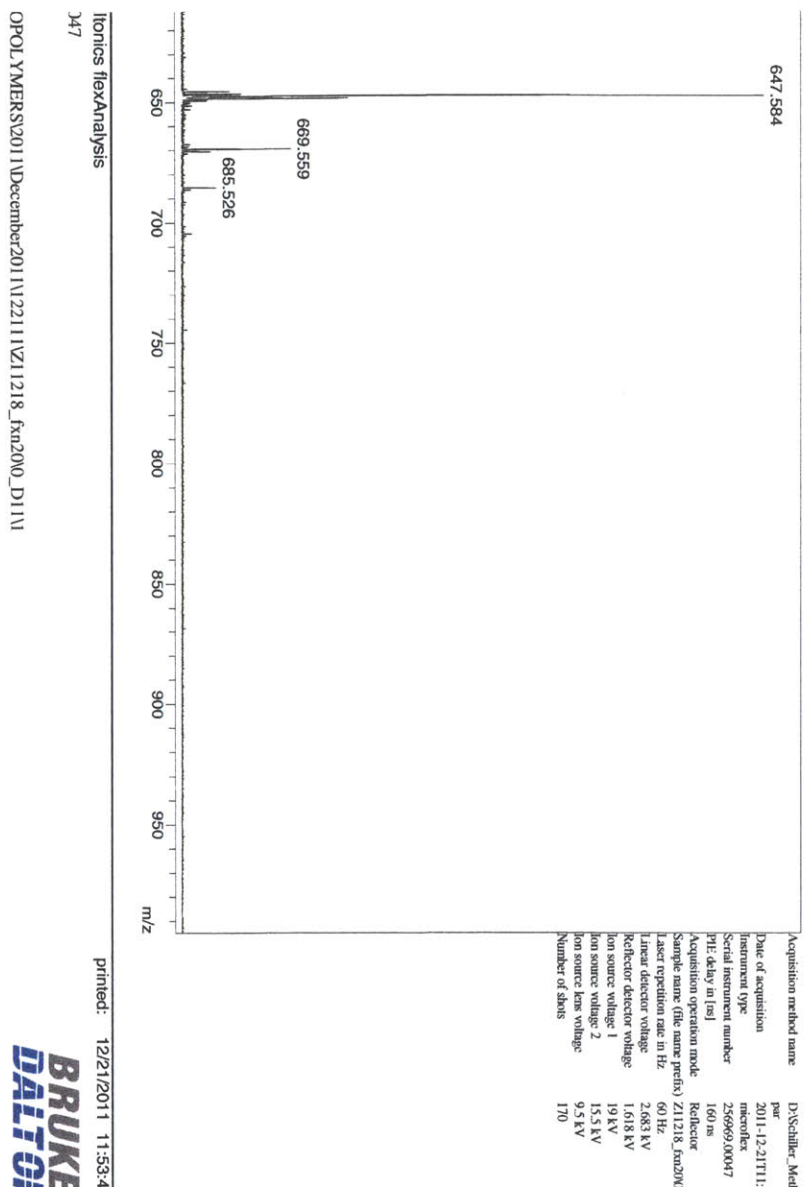


Figure 30. MALDI of the isotope labeled V1V4 peptide. The molecular weight of the neutral unlabeled GVGN1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 647.6, 669.6, and 685.5 m/z respectively.

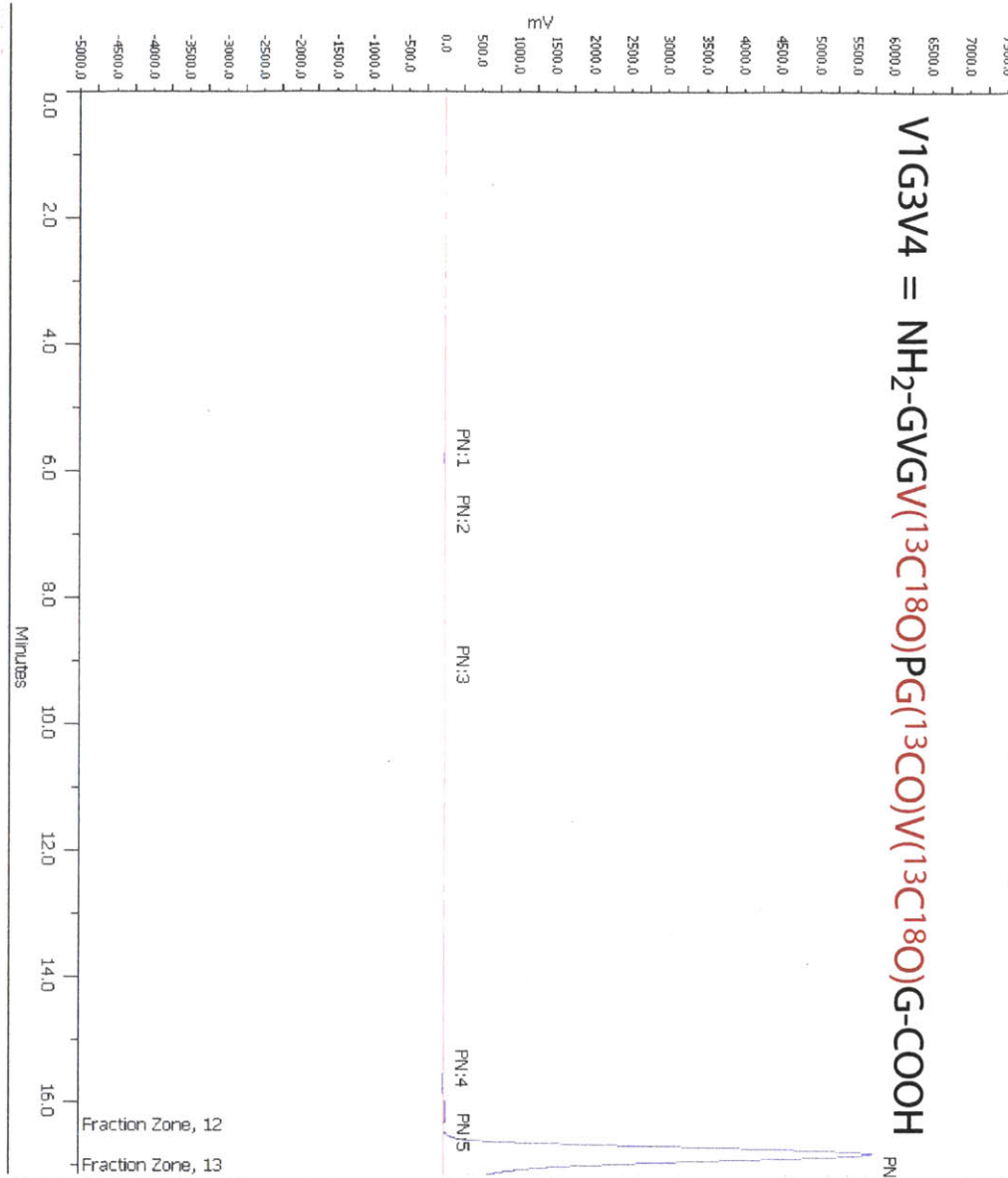


Figure 31. Analytical HPLC of V1G3V4 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

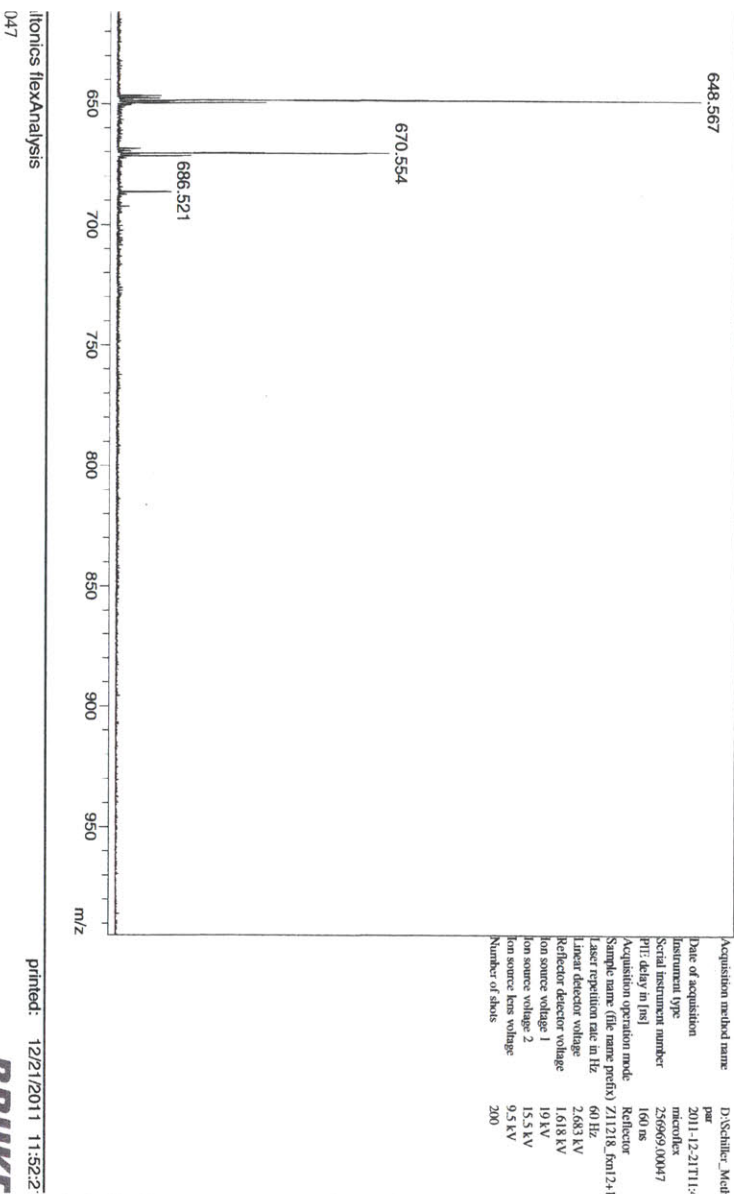


Figure 32. MALDI of the isotope labeled V1G3V4 peptide. The molecular weight of the neutral unlabeled GVGn1 peptide is 640.7 amu. Here the main ion, Na⁺ and K⁺ peaks appear at 648.6, 670.6, and 686.5 m/z respectively.

GVGn2 = NH₂-GVG(VPGVG)₂-COOH

(062312012014 2012-05-23 16:26-01002-0201.D)

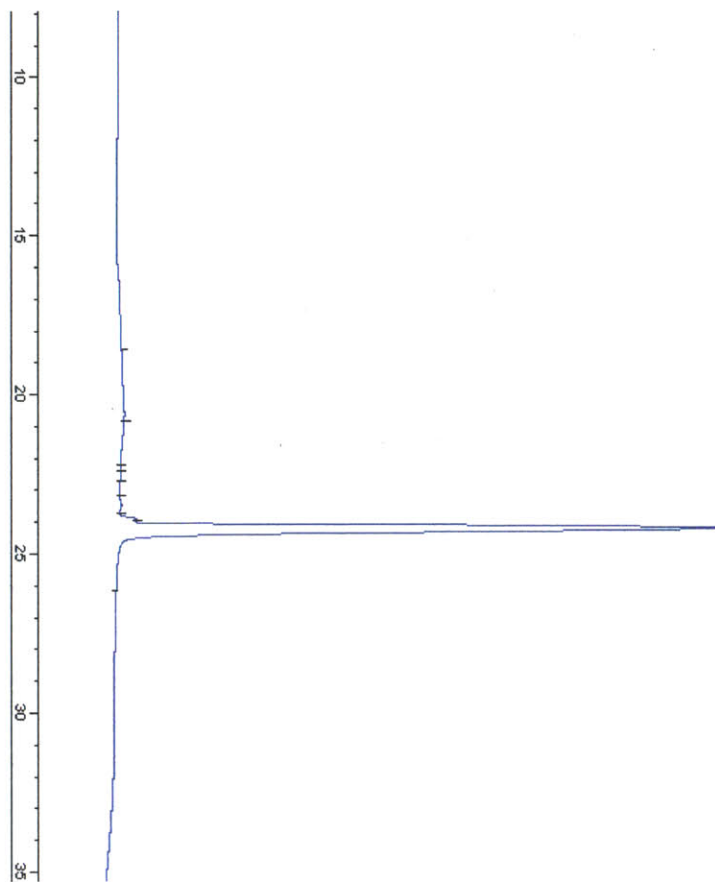


Figure 33. Analytical HPLC of GVGn2 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

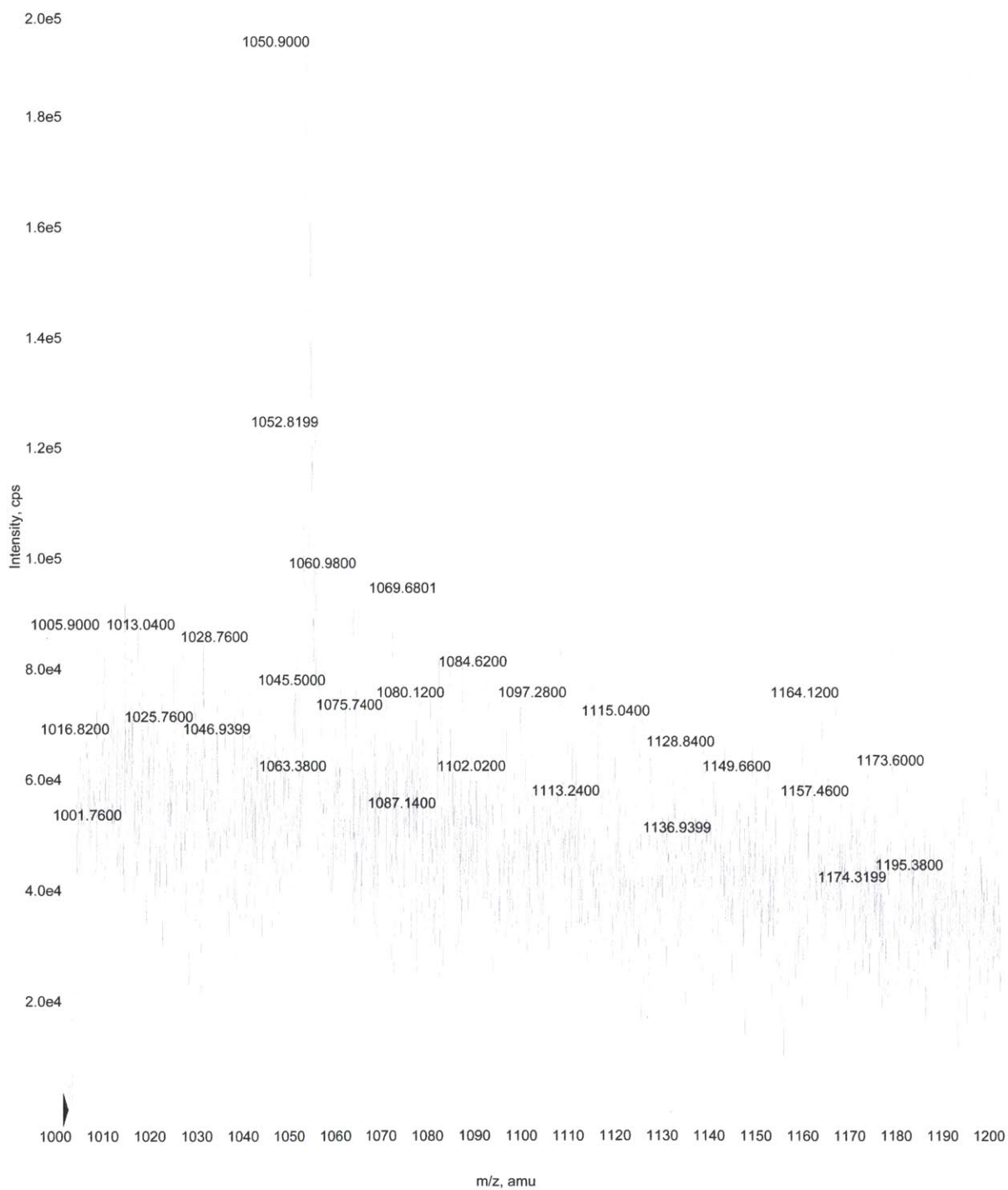


Figure 34. Electrospray mass spectra of the GVGn2 peptide. The molecular weight of the neutral unlabeled GVGn2 peptide is 1050.2 amu. Here the main ion peaks appear at 1050.9 m/z.

GVGn3 = NH₂-GVG(VPGVG)₃-COOH

!052312\612014 2012-05-23 16:28:01\003-0401.D

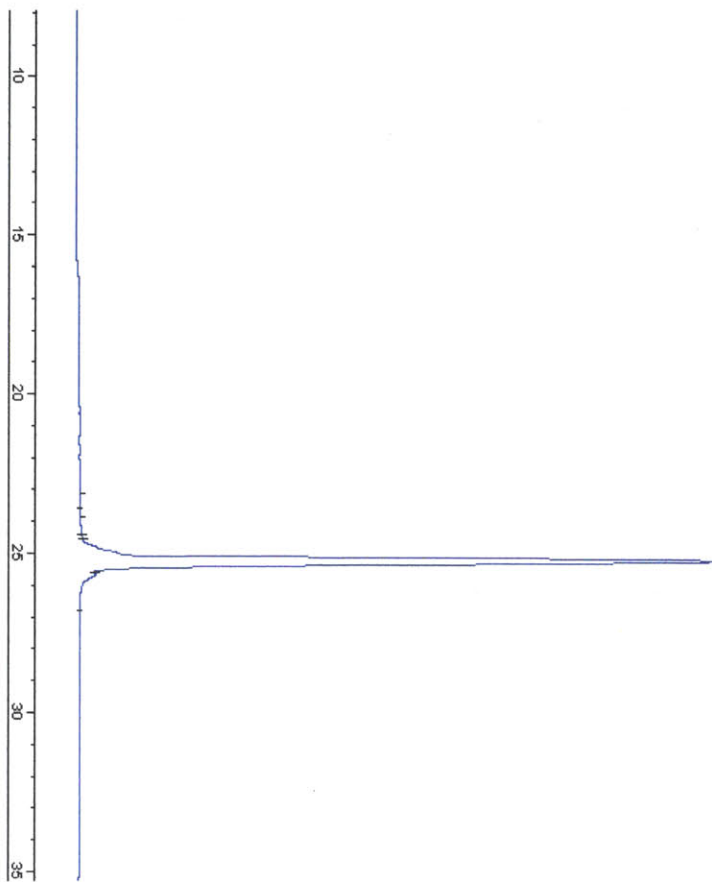


Figure 35. Analytical HPLC of GVGn3 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

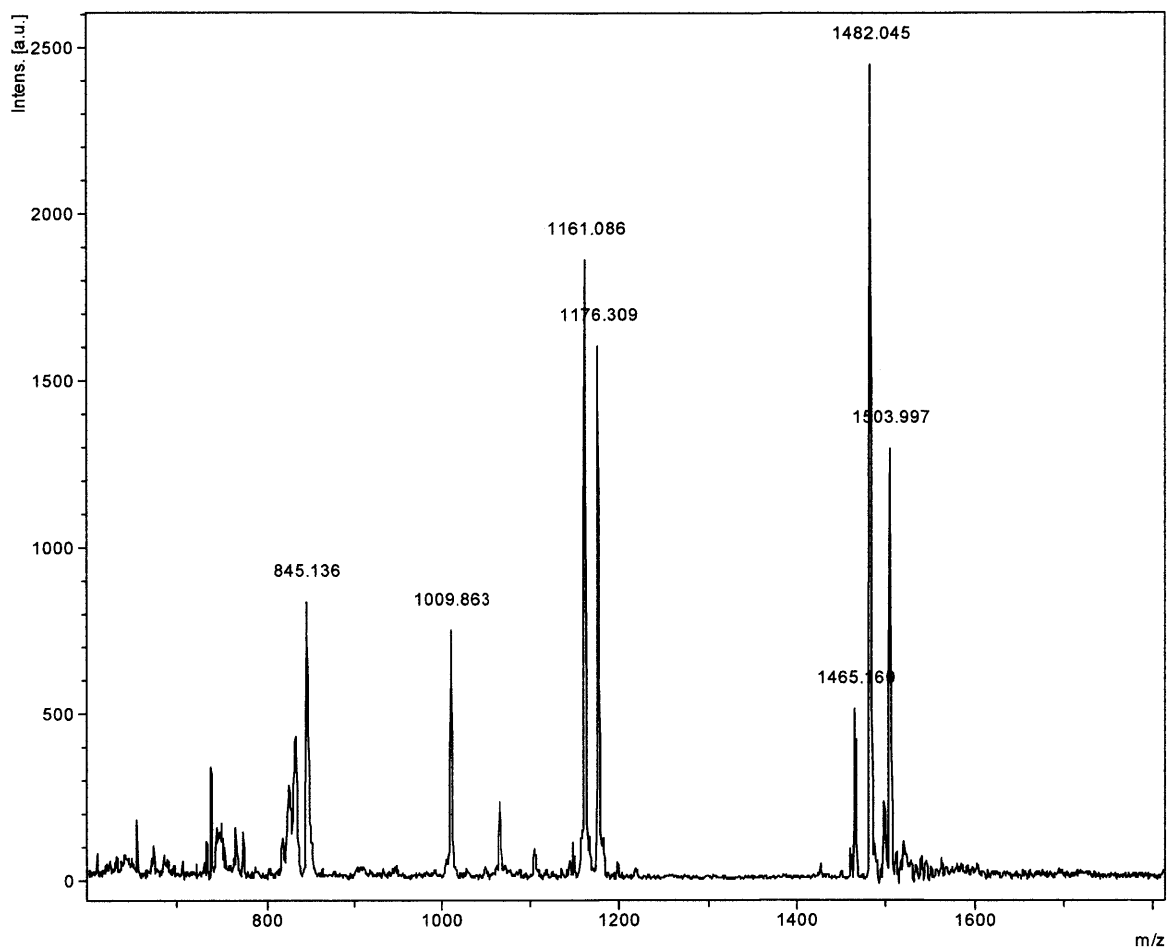


Figure 36. MALDI spectrum of the GVGn3 peptide. The molecular weight of the neutral unlabeled GVGn3 peptide is 1459.7 amu. Here the Na^+ peaks appear at 1482.0 m/z.

GVGn4 = NH₂-GVG(VPGVG)₄-COOH

(05/23/12; 16:20:14; 2012-05-23; 16-25; 011004-050(1.D))

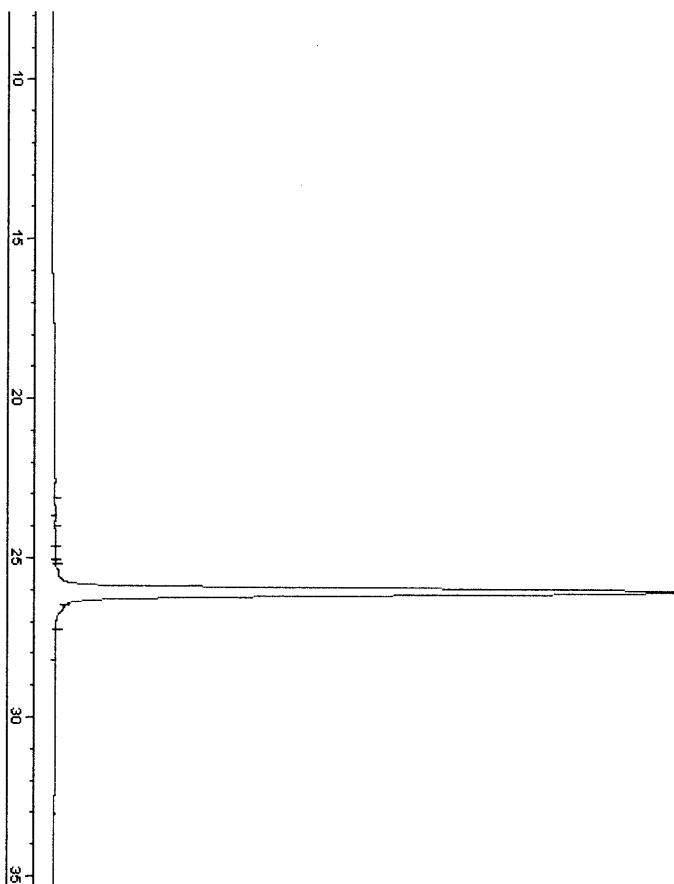


Figure 37. Analytical HPLC of GVGn4 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

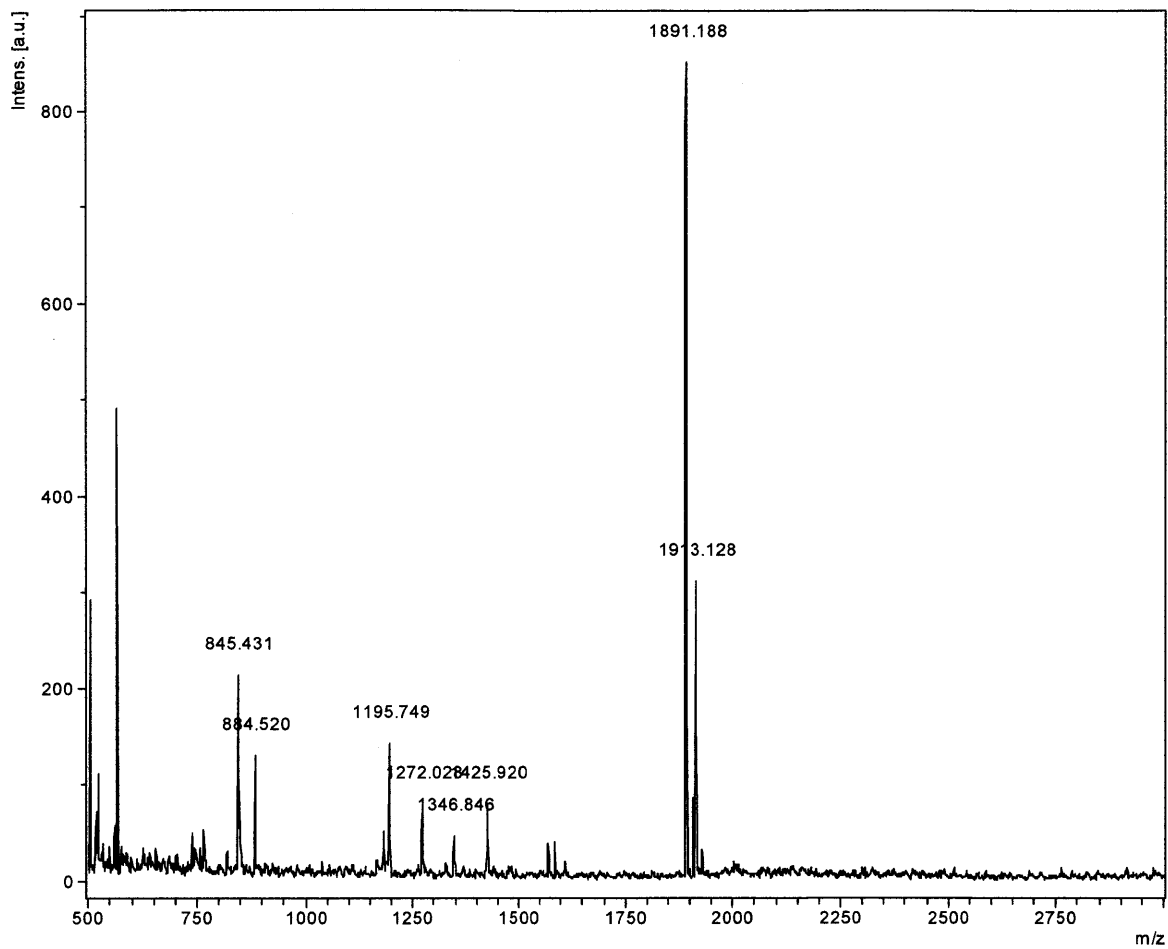


Figure 38. MALDI spectrum of the GVGn4 peptide. The molecular weight of the neutral unlabeled GVGn4 peptide is 1869.2 amu. Here the Na^+ peaks appear at 1891.2 m/z.

5VGN5 = NH₂-GVG(VPGVG)₅-COOH

1:05/23/2014 20:12:05/23-18-28-011005-0801.D

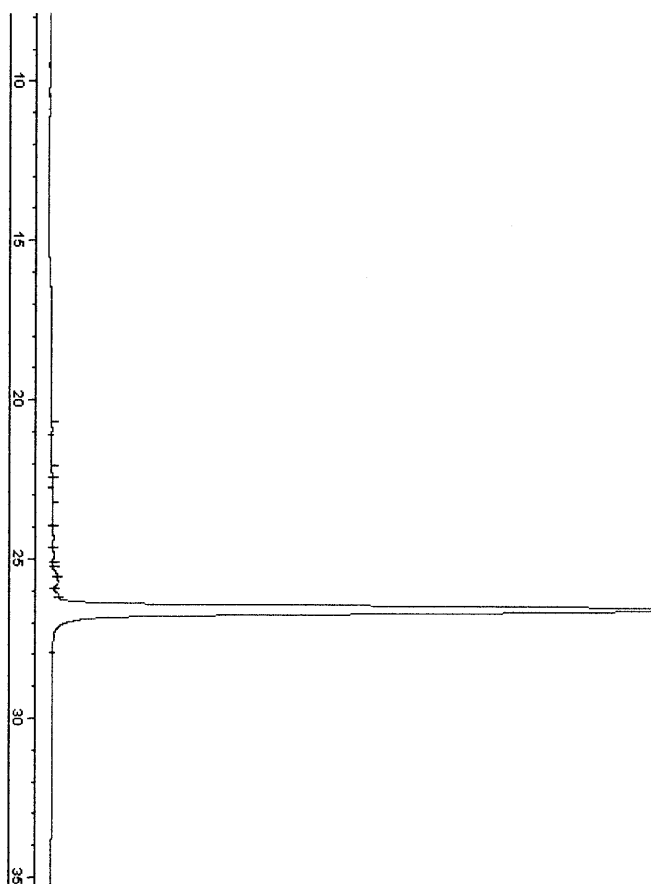


Figure 39. Analytical HPLC of GVGn5 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

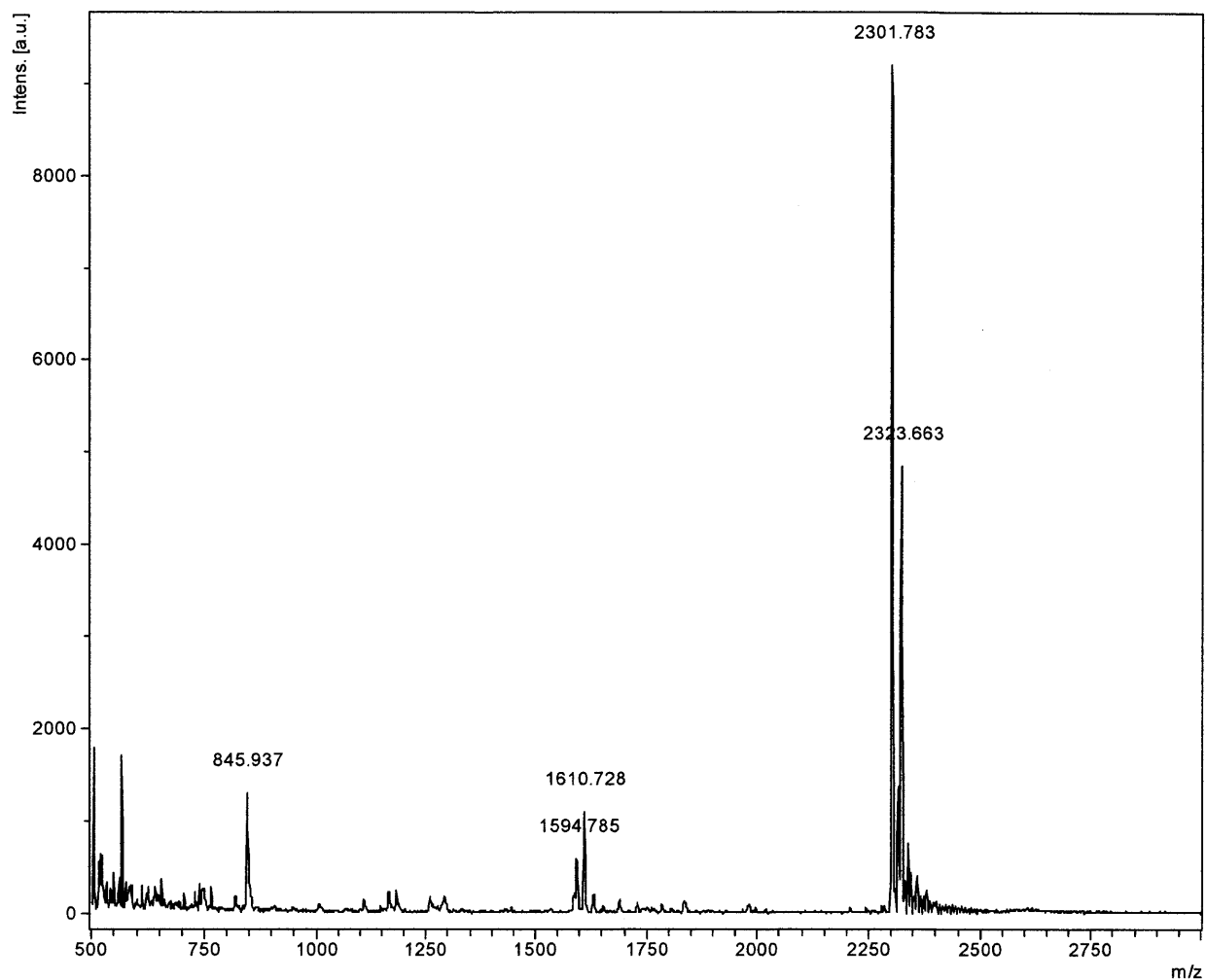


Figure 40. MALDI spectrum of the GVGn5 peptide. The molecular weight of the neutral unlabeled GVGn5 peptide is 2278.7 amu. Here the Na^+ peaks appear at 2301.8 m/z.



(052312)(6/20/14 2012-05-23 16:28:01)006-0701.D

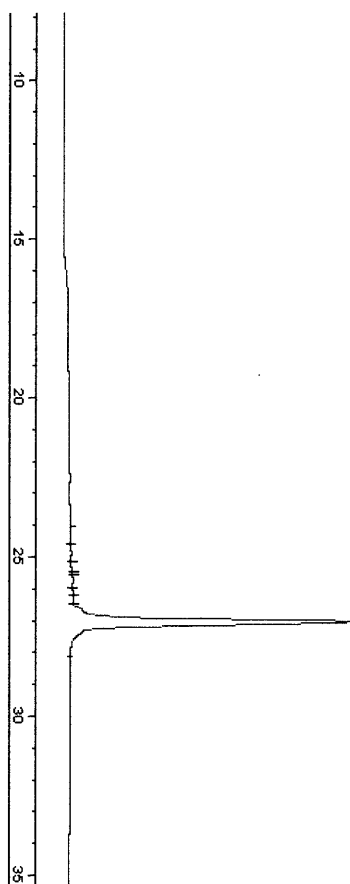


Figure 41. Analytical HPLC of GVGn6 using a C18 HPLC column with a two-phase buffer gradient: (buffer A) 0.05% TFA in H₂O and (buffer B) 80% acetonitrile, 0.043% TFA in H₂O.

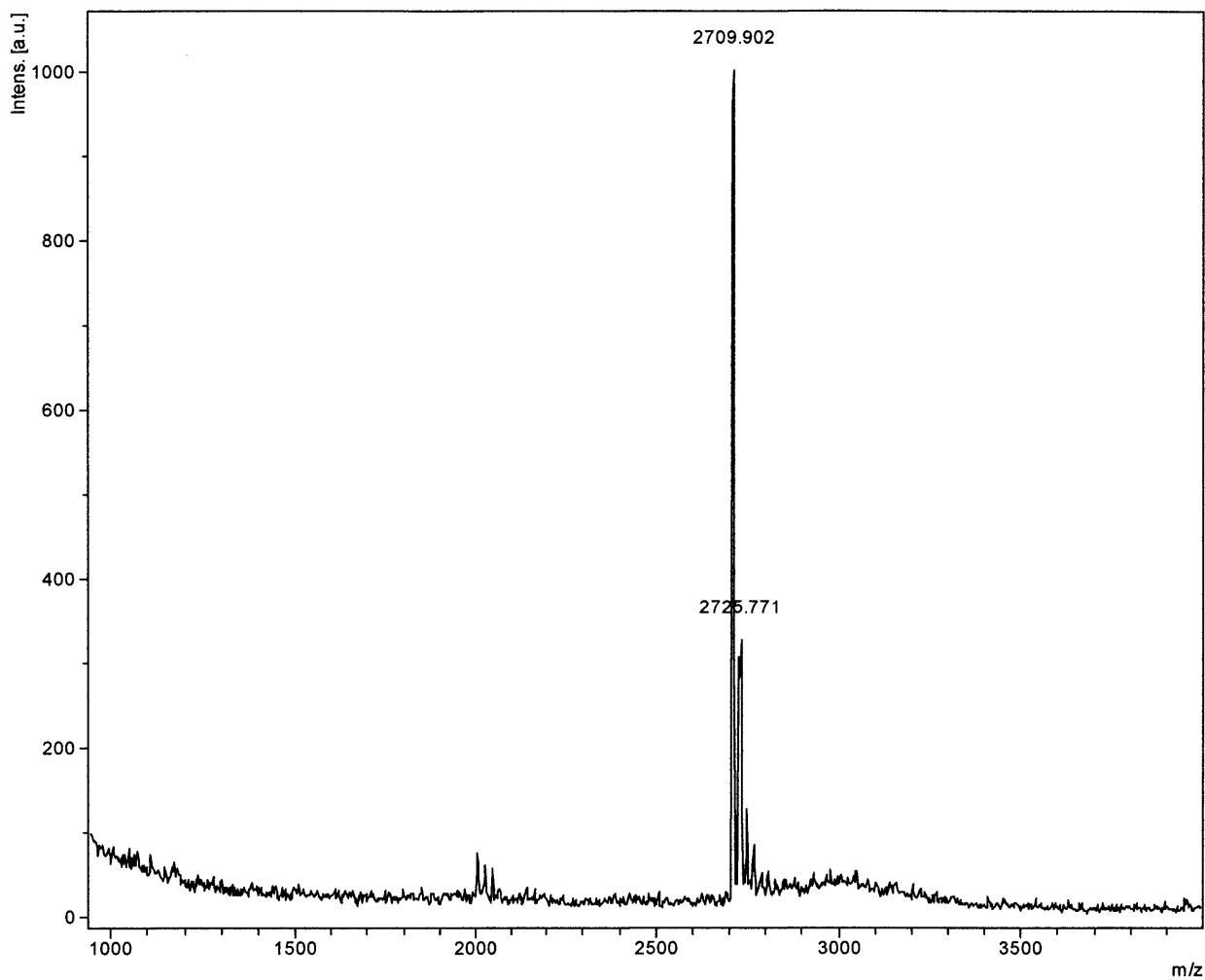


Figure 42. MALDI spectrum of the GVGn6 peptide. The molecular weight of the neutral unlabeled GVGn6 peptide is 2688.2 amu. Here the Na^+ peaks appear at 2709.9 m/z.