INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Probing Peptide Binding to the Second PDZ Domain of hPTP1E using FT-IR and MS

Robert Papp

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science
Concordia, University
Montreal, Quebec, Canada

February 2002

ABSTRACT

Probing Peptide Binding to the Second PDZ Domain of hPTP1E using FT-IR and MS

Robert Papp

Three peptides were evaluated for binding to the second PDZ domain of human protein tyrosine phosphatase 1E using FT-IR and mass spectrometry. This PDZ is known to bind the C-terminal motif X-Ser/Thr-Val. MS demonstrated the formation of a 1:1 complex between PDZ and ENEQVSAV but not with peptides ENEQVCAV and KDDEVYYV due to substitution of Ser in –2 position with Cys or Tyr. Furthermore, the H/D exchange rate of the PDZ-ENEQVSAV, monitored by MS, was reduced compared to free PDZ, suggesting a stable protein-ligand complex had formed.

Thermal denaturation of free PDZ and in combinations with three peptides was investigated using FT-IR and CD spectroscopy. The T_m of free PDZ in D_2O was found to be 47 ± 1°C when monitoring the loss of β-sheets and the appearance of aggregation by FT-IR. Combinations of PDZ with KDDEVYYV, ENEQVCAV and ENEQVSAV revealed T_m values of 48± 1, 49 ± 1 and 57 ± 1°C, respectively. The increased T_m observed with PDZ-ENEQVSAV was attributed to the formation of a stable complex. The thermal denaturation spectra were evaluated by 2D correlation analysis to determine the sequence of unfolding events. β-sheets and turns unfold prior to protein aggregation in free PDZ while in the PDZ-ENEQVSAV complex β-sheets unfold prior to aggregation but turns unfold last, indicating that residual turn-like structures persisted in the denatured state. Evidence from CD spectroscopy suggested that helical structures remained intact during heating to 75°C for PDZ and the PDZ-ENEQVSAV complex.

ACKNOWLEDGEMENTS

At this time, I would like to thank my supervisor Dr. Ann English for her dedication to her students. Her guidance and encouragement was greatly appreciated during my graduate studies. I would also like to thank Dr. Irena Ekiel for the opportunity to work with this project and for many interesting discussions about proteins. I wish to acknowledge Dr. Joanne Turnbull for her help, enthusiasm and efforts as part of my committee. To my labmates Andrea Romeo, Dr. Angelo Filosa, John Wright, Tyrone Sheppard, Christina Esposito and Hasnain Jaffer, I want to thank you for all of your help over the years and I wish you success in the future.

I would also like to thank Dr. Josie Visentini and Dr. Elizabeth Kwong at Merck Frosst Canada for the use of the lab equipment and their understanding while I was completing this work.

This thesis is dedicated to my wife Rania, for just being you, and to my family.

Table of Contents

Page

List of Figures.....viii

List of Tables.....xi

Abbreviations.....xii

1. Studying Protein Interactions.....1

1.1 Protein Structure.....1

1.2 Protein Interactions.....3

1.3 Hydrogen/Deuterium Exchange.....4

1.4 Probes for Protein Structure and Binding.....7

1.4.1 CD Spectroscopy.....7

1.4.2 FT-IR Spectroscopy and 2D Analysis.....9

1.4.3 Mass Spectrometry.....13

1.5 The PDZ Domain.....15

1.6 Statement of Purpose.....20

1.7 Thesis Organization.....21

2. Selectivity of Peptide Binding to the Second PDZ Domain of PTP1E using FT-IR Spectroscopy.....23

Abstract.....23

2.1 Introduction.....24

2.2 Experimental Procedures.....26

2.2.1 Materials.....26

2.2.2 Methods.....26

2.3 Results.....29

Page

2.4 Disscusion.....60

2.5 Conclusion.....65

3. Using Mass Spectrometry as a Probe.....66

Abstract.....66

3.1 Introduction.....67

3.2 Experimental Procedures.....69

3.2.1 Materials.....69

3.2.2 Methods.....69

3.3 Results.....71

3.4 Discussion.....84

3.5 Conclusions.....89

4. General Conclusions and Future Work.....90

4.1 General Conclusions.....90

4.2 Future Work.....92

4.3 References.....94

Appendix A Primary amino acid sequence of the second PDZ domain of hPTP1E...100

Appendix B Estimation of binding constants from melting temperatures.....101

List of Figures

Page

1.1 (a) Base catalyzed H/D exchange (b) Acid catalyzed H/D exchange.....4

1.2 Schematic of H/ D exchange in proteins.....5

1.3 Schematic representation of IR active vibrational modes in peptide bonds.....9

1.4 Mechanism of electrospray ionization.....14

1.5 Rasmol representation of the second PDZ domain from hPTP1E.....17

1.6 Schematic representation of the second PDZ of PTP1E binding the C-terminus of a peptide (INSLV).....19

2.1 (a) Second derivative spectrum of 1mM PDZ in D_2O. (b) Deconvolved spectrum of 1mM PDZ in D_2O.....29

2.2 (a) Second derivative spectrum of 1mM PDZ with 1.2mM ENEQVSAV peptide in D_2O. (b) Deconvolved spectrum of 1mM PDZ with 1.2mM ENEQVSAV peptide in D_2O.....30

2.3 Curve fitting of the deconvolved FT-IR spectrum of the PDZ Domain in D_2O ..32

2.4 Curve fitting of the deconvolved FT-IR spectrum of the PDZ-ENEQVSAV complex in D_2O 32

2.5 Stacked FT-IR spectra of the free PDZ protein (heating cycle).....34

2.6 Stacked FT-IR spectra of the free PDZ protein (cooling cycle).....35

2.7 Stacked FT-IR spectra of the PDZ-ENEQVSAV complex (heating cycle).....36

2.8 Stacked FT-IR spectra of the PDZ-ENEQVSAV complex (cooling cycle).....37

2.9 PDZ (a) Integrated intensity of the aggregation band at 1618cm^{-1} versus temperature (b) Integrated intensity of the β-sheet band at 1639cm^{-1} versus temperature 39

2.10 PDZ-ENEQVSAV (a) Integrated intensity of the aggregation band at 1618cm^{-1} versus temperature (b) Integrated intensity of the β-sheet band at 1639cm^{-1} versus temperature 40

2.11 2.5mM ENEQVCAV in sodium phosphate buffer pH 6.9, 2.5mM DTT and 2.5mM DTPA.....43

List of Figures

Continued

2.12 Deconvolved FT-IR spectrum of 4mM PDZ-ENEQVCAV in sodium phosphate buffer (pH 6.9) containing 4mM DTT and 4mM DTPA.....43

2.13 Deconvolved FT-IR spectrum of 1mM human met-hemoglobin in phosphate buffer (pH 6.9) containing 2mM DTT and 2mM DTPA.....44

2.14 (a) 2.4 mM tyrosine ethyl ester in D_2O pD 7.3 (b) 2.4mM KDDEVYYV peptide in D_2O pD 7.3.....45

2.15 FT-IR Spectrum of 2.0mM PDZ with 2.4mM tyrosine peptide in D_2O pD 7.3...46

2.16 CD spectra of 1mM PDZ in 50mM phosphate buffer (pH 6.9) at increasing Temperatures.....47

2.17 CD spectra of 1mM PDZ with 1.2mM ENEQVSAV in 50mM phosphate buffer (pH 6.9) at increasing temperatures.....48

2.18 Synchronous map of the denaturation of the PDZ domain over the temperature range of 27-59°C.....51

2.19 Asynchronous map of the denaturation of the PDZ domain over the temperature range of 27-59°C 53

2.20 Synchronous map of the denaturation of the PDZ domain over the temperature range of 65-89°C.....54

2.21 Asynchronous map of the denaturation of the PDZ domain over the temperature range of 65-89°C.....55

2.22 Synchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 37-69°C.....56

2.23 Asynchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 37-69°C.....57

2.24 Synchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 65-89°C.....59

2.25 Asynchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 65-89°C.....60

3.1 ESI mass spectrum of ENEQVSAV peptide in 40:60 ACN/0.1% TFA.....72

List of Figures

Continued

3.2 ESI mass spectrum of ENEQVCAV peptide in 40:60 ACN/0.1% TFA.....72

3.3 ESI mass spectrum of KDDEVYYV peptide in 40:60 ACN/0.1% TFA.....73

3.4 ESI mass spectrum of 55μM PDZ in 50mM phosphate buffer (pH 6.9).....74

3.5 High- range ESI mass spectrum of 55μM PDZ domain in 50mM ammonium acetate (pH 6.9)..... 75

3.6 High-range ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide In 50mM ammonium acetate (pH 6.9).....76

3.7 ESI mass spectrum of 55μM PDZ in 50mM ammonium acetate (pH 6.9).. 77

3.8 ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide in 50mM ammonium acetate (pH 6.9).....78

3.9 ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide in 50mM ammonium acetate (pH 6.9)..... 79

3.10 ESI mass spectrum of 50μM PDZ +100μM KDDEVYYV peptide in 50mM ammonium acetate (pH 6.9).....80

3.11 ESI mass spectrum of 50μM PDZ +100μM ENEQVCAV peptide in 50mM ammonium acetate (pH 6.9).....80

3.12 Plot of observed mass vs D_2O exposure time for PDZ (50μM) and PDZ-ENEQVSAV (50μM + 100μM) at 4°C..... 83

B1 van't Hoff plot of β-sheet unfolding for combinations of PDZ with peptides....101

List of Tables

Page

2.1 Curve fitting results of the free PDZ and ENEQVSAV-PDZ complex.....33

2.2 Summary of T_m values obtained by FT-IR spectroscopy for the PDZ domain and combinations with three peptides.....41

2.3 Comparison of the tyrosine band position for six different tyrosine containing samples.....46

2.4 Correlation table summarizing the 2D observations for the PDZ domain at 27 -59°C.....53

2.5 Correlation table summarizing the 2D observations for the PDZ-ENEQVSAV complex at 37-69°C.....58

3.1 Effect of the electrospray ionization parameters for the free PDZ domain and related complexes with peptides.....81

B1 Estimates of binding affinity towards PDZ of three peptides based on T_m measurements and van't Hoff plots.....101

Abbreviations

ACN	acetonitrile
amu	atomic mass units
CD	circular dichroism
DTPA	diethylenetriamine Pentaacetic Acid
DTT	1,4-dithiothreitol
ESI	electrospray ionization
FT-IR	Fourier-transform infrared
ENEQVCAV	Glu-Asn-Glu-Gln-Val-Cys-Ala-Val
ENEQVSAV	Glu-Asn-Glu-Gln-Val-Ser-Ala-Val
ΔG	Gibbs free energy
H/D	hydrogen/deuterium
HPLC	high-performance liquid chromatography
hPTP1E	human protein tyrosine phosphatase 1E
ΔH	van't Hoff enthalpy
K_B	binding constant
K_D	dissociation constant
KDDEVYYV	Lys-Asp-Asp-Glu-Val-Tyr-Tyr-Val
MS	mass spectrometry
m/z	mass-to-charge
NMR	nuclear magnetic resonance
ΔS	entropy
T_m	melting temperature
TFA	trifluoroacetic acid
UV/vis	ultraviolet / visible

CHAPTER 1 - STUDYING PROTEIN INTERACTIONS

1.1 Protein Structure

A protein is made up of a collection of many individual amino acids termed the primary structure or amino acid sequence. In addition, proteins possess the ability to form secondary structures, such as α-helices and β-sheets and tertiary structures, which can result in functional biomolecules such as enzymes, receptors and binding domains. The primary sequence encodes the folding pattern of a protein and ultimately determines its native structure. Folding from a long-chain polypeptide into a three-dimensional protein is not fully understood nor is it considered to be random event (Reinstadler *et al.*, 1996). Folding/unfolding of proteins has been the focus of numerous studies using different techniques such as NMR (Balbach *et al.*, 1995), fluorescence (Dodge and Scheraga, 1996) and FT-IR spectroscopy (Filosa *et al.*, 2001). The primary structure also determines properties such as the iso-electric point, molecular weight and solution charge-state.

The three-dimensional structures observed in proteins, termed the tertiary structure, can be disrupted or denatured from the native state, by the application of heat, chemical denaturants such as urea or guanidine hydrochloride, organic solvents and pH extremes. The biological function of a protein is dependent on the folded solution structure which determines its chemical behavior. Denatured or aggregated proteins retain little or no biological function compared to the native form (Creighton, 1993). Protein function can include but is not limited to catalytic activity, cell structure, chemical transport and binding (Leningher *et al.*, 1993). The denatured state of proteins

does not always involve a complete loss of structure (Shortle 1996) and it is not necessarily a reversible process (van Stokkum *et al.*, 1995).

In many cases, the denaturation of proteins is represented by a two-state model consisting of a native and denatured state. Chemical denaturants preferentially interact with the unfolded state, causing denaturation (Shortle 1996), (Creighton 1993). Organic solvents break up hydrophobic interactions that stabilize the core of the protein while pH extremes alter amino acid charges causing electrostatic repulsion (Leninger *et al.*, 1993). The application of heat increases molecular motions which favor the unfolded state due the increasing entropy. The unfolded state is favored since it becomes a lower free energy state due to the increasing $T\Delta S$ term in Gibbs free energy equation, $\Delta G=\Delta H - T\Delta S$ (Shortle 1996). Since, the hydrophobic core becomes exposed to the solvent some proteins will aggregate through hydrophobic interactions.

Studying the denaturation of proteins is useful for comparing unfolding mechanisms and structural stability. The enthalpy of denaturation can also be used to determine binding constants (Pace and McGrath, 1980); however, the reversibility of unfolding must be considered prior to applying van't Hoff analysis since it assumes a two-state equilibrium unfolding (van Stokkum *et al.*, 1995). Many proteins such as ribonuclease A, cytochrome c, staphylococcal nuclease retain some secondary structures in the denatured state (Shortle, 1996). Proteins may aggregate at high temperatures or gradually unfold all secondary structures (Holzbaur *et al.*, 1990). The denaturation process can occur in multiple stages if the protein contains subunits of different stability (van Stokkum *et al.*, 1995).

1.2 Protein Interactions

The interactions of proteins with other biomolecules is critical to proper cell functioning. An important characteristic of proteins is their ability to specifically bind other proteins, peptides and small molecules which can result in the formation of covalent or non-covalent interactions. Covalent protein interactions usually involve the formation of disulfide bonds between two thiol groups. Non-covalent protein interactions include hydrogen bonding, van der Waals interactions and salt bridges (Dill, 1990). Protein binding is dependent on interactions that occur at the individual amino acid residue level. Therefore specificity is determined by the unique amino acid sequence encoded in the protein.

Ligand binding can induce changes to the secondary or tertiary structure of a protein. For example, Ca^{2+} binding to calmodulin causes conformational re-arrangements which can be detected by spectroscopic methods such as UV/vis, circular dichroism and FT-IR (Trewhella *et al.*, 1989). FT-IR has been used to study peptide binding to calmodulin using ^{13}C labeled peptides (Zhang *et al.*, 1994). The heavier nucleus of ^{13}C causes a significant red shift in the amide I signal which allows for simultaneous monitoring of the protein and target peptide structures. Ligand binding may, however, result in very little conformational changes in the protein (Doyle *et al.*, 1996), therefore to obtain structural details regarding the binding complex, it is often necessary to use higher resolution techniques such as NMR or X-ray crystallography. Detailed analysis by NMR and X-ray is generally employed once the binding has been detected using more simple and rapid methods. These can include Biacore or fluorescence polarization (Sheng and Sala, 2001). Alternatively, binding may be detected by indirect methods that

probe protein interactions. These include H/D exchange kinetics (Ekiel *et al.*, 1998), changes in T_m (Pace and McGrath, 1980) or measuring the enthalpy of binding with calorimetric methods (Wiseman *et al.*, 1989).

1.3 Hydrogen/Deuterium Exchange

Hydrogen can exchange with deuterium from deuterated solvents such as D_2O. The exchange can occur when the hydrogen atom participates in an acid-base equilibrium. Base-catalyzed exchange occurs at pH's above 2.5, where the hydrogen atom is removed as a proton (H^+) from an acidic moiety, leaving a conjugate base. It is then quickly replaced by a deuteron (D^+) from the solvent D_2O (Creighton, 1993). Acid-catalyzed exchange occurs when the proton acceptor gains a deuteron from the solvent D_2O and then releases a proton. These interactions, shown in Figure 1.1, are therefore conditional on the solution pH.

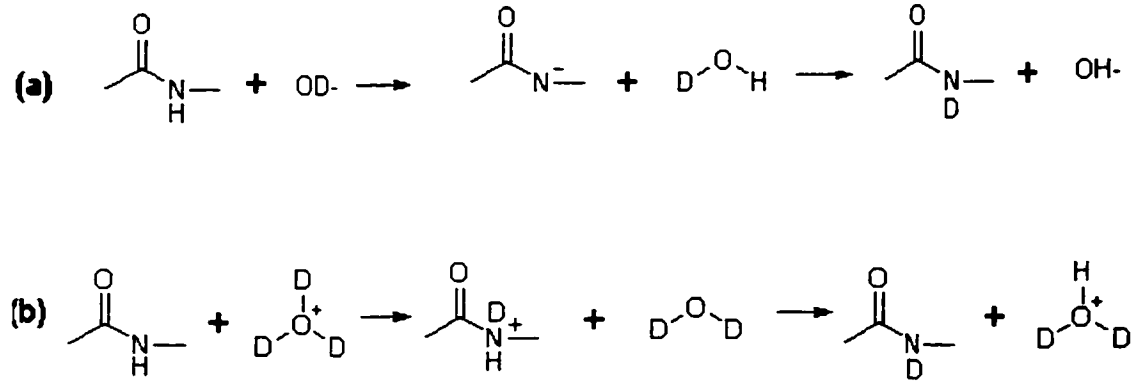

Figure 1.1 (a) Base-catalyzed H/D exchange; (b) Acid-catalyzed H/D exchange

Proteins can exchange hydrogen atoms bound to nitrogen atoms, carboxylic acids, thiols, alcohols and phenols depending on the solution pH (Creighton, 1993). Because

of the abundance of amide linkages in proteins and their involvement in secondary structure, the exchange rate of backbone amide hydrogen atoms is most useful for probing protein structure. The rate of exchange is controlled by temperature and pH. For amide linkages the rate constant of exchange decreases 3-fold with every 10°C reduction in temperature and 10-fold with each unit decrease in pH, reaching a minimum around pH 2.5 where the acid- and base-catalyzed exchange rates are equal (Smith *et al.*, 1997).

The exchange rate of amide-bound hydrogen atoms depends on their degree of exposure to the deuterated solvent. Buried, exchangeable hydrogen atoms exchange more slowly than surface-exposed hydrogen atoms (Wang and Tang, 1996) (Zhang and Smith, 1993). The electrostatic effect of nearby amino acid side chains must also be considered since they can influence the observed pK_a of exchangeable hydrogen atoms. A depiction of the mechanism of H/D exchange in proteins is shown in Figure 1.2.

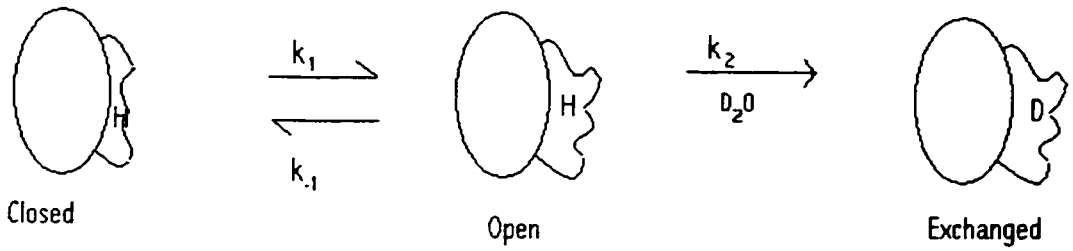

Figure 1.2 Schematic of H/D exchange in proteins.

Two models can be used to characterize the kinetics of H/D exchange for a protein. EX1 kinetics occur when the rate determining step is the rate of transient protein unfolding,

therefore $k_{obs}=k_1$. EX2 kinetics occur when opening and closing are faster than the H/D exchange rate, therefore $k_{obs}= k_2 (k_1/k_{-1})$ (Zhang and Smith, 1996), (Loh *et al.*, 1993).

Deuterium is useful as an isotopic label for proteins. The incorporation of deuterium can be monitored using the ratio of the amide II and amide II' bands detected by FT-IR since the N-H stretch and N-D stretch frequencies are separated by about $100cm^{-1}$ (Jackson and Mantsch, 1995) (Section 1.4.2). Adding a heavier mass to a harmonic oscillator (the N-H bond) lowers its vibrational frequency (Bare *et al.*, 1975). H/D exchange can also be monitored using mass spectrometry, which can detect unit increases in molecular mass during the exchange process (Section 1.4.3). H/D exchange rates can be used to probe a protein under various conditions such as substrate binding (Anderegg and Wagner, 1995) or pH variations (Wang and Tang, 1996)

Mass spectrometry has further evolved to combine proteolytic digestion of proteins with various stages of deuterium exchange (Smith *et al.*, 1997). The mass of the cleaved peptides compared to their theoretical mass in H_2O provides the degree of deuterium incorporation. Equation 1 gives the calculation of percent-deuterium exchange:

(1) $$\%D = (m_1-m_0) / (m_{100}-m_0) \times 100$$

where m_1 is the measured mass of the peptide, m_0 is the mass of the peptide in ^1H solvent and m_{100} is the mass of fully exchanged peptide (Wang F. *et al.*, 1998). Addition of the proteolytic cleavage step allows for the localization of peptide fragments within the protein that exchange slowly or quickly.

1.4 Probes for Protein Structure and Binding

The solution structure of proteins and their interactions can be probed by various instrumental techniques, each with unique benefits and shortcomings. These advantages and disadvantages can be ascribed to the chemical property being measured, whether it is molecular vibrations, magnetic moments, electronic transitions or a physical property such as relative mass. In addition to probing the solution structure of protein, it is important to monitor changes in proteins that occur during interactions with other molecules. This information can provide insight into understanding protein/ligand binding mechanisms.

FT-IR, CD and MS are established techniques for studying proteins (Mizra *et al.*, 1993) and provide complementary information on the solution conformation of a protein. For example, combinations of spectroscopic techniques such as CD and fluorescence with mass spectrometry have been used to study calcium binding proteins (Veenstra *et al.*, 1998).

1.4.1 CD Spectroscopy

CD measures the difference in absorption of left-handed and right-handed circularly polarized light in a sample solution. CD detects optically active molecules and therefore is well suited to protein studies as it is sensitive to the 3D conformation of the amide backbone which ultimately depends on the secondary structure. Aromatic amino acids such as Trp, Phe and Tyr, and disulfide bridges also exhibit CD absorption in the near-UV 250-350nm range (Creighton, 1993).

The far-UV CD (180-250nm) absorption originates from the peptide bond chromophore, found abundantly in proteins. The 3D structure of proteins and polypeptides determines the bond angle and rotation along the individual peptide bonds, which ultimately controls the magnitude and direction of electronic transitions. The two important electronic transitions that result in protein backbone CD signals are $\pi \rightarrow \pi^*$ (from the 3π orbitals of C,N,O) which lies near 190nm, and $\eta \rightarrow \pi^*$ (from a lone pair of O) the lower energy transition at 220nm (Gratzer and Cowburn, 1969) (Woody, 1996). The CD spectral assignments for proteins are based on the observed spectra of homo-polypeptides such as poly-L-lysine which can form an α-helix, a random coil or a β-strand depending on the solution pH (Creighton, 1993). Because CD is sensitive to peptide bond angles, the ordered geometric arrangement of the α-helix results in CD absorptions that are distinct and well characterized compared to other secondary structures such as β-sheets (Sarver and Krueger, 1991) (van Stokkum *et al.*, 1990). The CD spectrum of a right-handed helix typically shows a large negative band at 222nm ($\eta \rightarrow \pi^*$ transition), a smaller negative band at 208nm and a large positive band at 190nm (Woody, 1996). The 208 and 222nm abs can be used to estimate the % helix composition of proteins according to the method of Greenfield *et al.*, (1967). Anti-parallel β-sheets show a small negative band at 215nm ($\eta \rightarrow \pi^*$ transition) and a larger positive band near 198nm ($\pi \rightarrow \pi^*$); however, both assignments depend on the number of strands contained in the β-sheet (Woody 1996). Unordered structures are characterized by a large negative absorption near 200nm (Greenfield and Fasman, 1969). Type I turns resemble helical CD spectra while type II turns resemble β-sheet spectra that are red-shifted by about 5-10nm (Woody, 1996).

The following assumptions are made in order to estimate the secondary structure composition. Firstly, the CD signal of poly-L-lysine approximates secondary structure elements in proteins, secondly, the random coil of poly-L-lysine is a good approximation of unordered protein regions and lastly it is assumed that the 208 - 240 nm spectral region consists only of amide bond absorptions (Greenfield and Fasman, 1969).

CD spectroscopy can be used to probe protein structure during denaturation by monitoring specific wavelengths assigned to structural features. For example, the α-helical absorption which produces two characteristic negative peaks at 208 and 222nm, can be used to monitor protein unfolding (Heyn and Weischet, 1975). CD spectroscopy has also been applied to study protein ligand interactions such as binding of tolmetin to human serum albumin which is demonstrated by a change in the CD signal at 290 and 340nm (Matsuyama *et al.*, 1987).

1.4.2 FT-IR Spectroscopy and 2D Analysis

IR detects bond vibrations that result in a change in the dipole moment of the molecule (Braiman and Rothschild, 1988). For peptide bonds, these vibrations include bending and stretching modes as shown in Figure 1.3 (Bandekar, 1992):

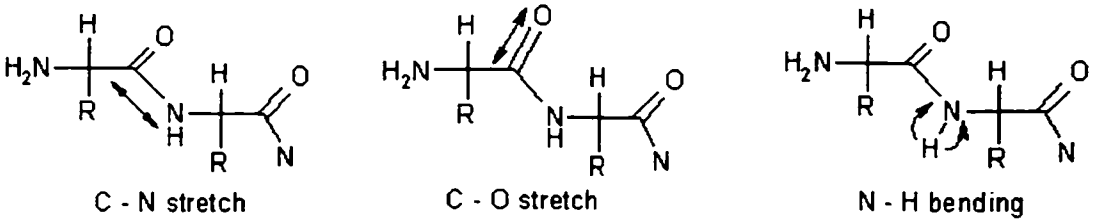

Figure 1.3. Schematic representation of IR active vibrational modes in peptide bonds

FT-IR spectroscopy can distinguish functional structures in proteins because of differences in spatial geometry and hydrogen bonding patterns in the amide backbone. Interactions between neighboring vibrational transition dipoles causes splitting of the C-O stretching modes into distinct FT-IR absorption bands (Tadesse *et al.*, 1991). In addition, the strength and orientation of hydrogen bonding within the peptide backbone will influence the electron density on carbonyl groups and cause frequency shifts (Jackson and Mantsch, 1995). Since the protein backbone contains numerous amide bonds in various secondary structure elements, the FT-IR peaks observed are a sum of the amide absorptions of each structural element. The absorptions that are particularly well resolved include those of β-sheets, β-turns and α-helices. The α-helical absorption can overlap with peaks arising from unordered structures; however, this varies from protein to protein (Jackson and Mantsch, 1995). There are some general assumptions that must be made in order to use FT-IR to estimate protein secondary structure. Firstly, the molar absorptivity of a residue is considered to be the same regardless of its location in a secondary structure element; secondly, side chain absorbance does not interfere with backbone absorption and lastly, bands observed for individual structural elements do not overlap (Goormaghtigh *et al.*, 1990).

There are three distinct frequency regions in protein IR spectra known as amide I (1700-1600cm^{-1}), amide II (1600-1500cm^{-1}) and amide III (1320-1200cm^{-1}) (Dong *et al.*, 1992), the first two being the strongest and most used for protein characterization. The amide I mode consists mainly of C=O stretching (80%) with a smaller contribution from C-N stretching (Jackson and Mantsch, 1995). Bands in this region show very little shift in peak frequencies in water versus deuterium oxide as a solvent, since only a small part

of the signal involves N-H bending. β-Sheet (intramolecular hydrogen bonding) peaks occur at 1620-1640cm^{-1} and may contain a high frequency component at 1670-1680cm^{-1} (van Stokkum *et al.*, 1995). α–Helices show a peak centered at 1645-1655cm^{-1} (Jackson and Mantsch, 1995). These amide bonds are held in a helical formation due to hydrogen bonding between an N-H group and a C=O group that are separated by four bonds. Peaks from β-turns are found at 1665-1675cm^{-1} due to the rigid nature of the turn. The 3_{10} helix is found around 1665cm^{-1} due to the tightly constrained position of the amide nitrogen resulting in higher frequency C-N vibrations. Finally, protein aggregation can be detected by the appearance of a low frequency peak at around 1610-1620cm^{-1} since this involves strong intermolecular hydrogen bonding which forms with the least steric restraint (Jackson and Mantsch, 1995). A high frequency component is often found at 1680cm^{-1}.

Amide II modes consist of N-H bending (60%) and C-N stretching vibrations (40%) (Jackson and Mantsch, 1995) and are useful for monitoring hydrogen/deuterium exchange since peaks in the amide II region will red-shift as much as 100cm^{-1} in D_2O. The replacement of hydrogen with a heavier deuterium nucleus in a chemical bond results in a lower vibrational frequency. Amide III is a combination of many contributing vibrational motions and therefore is not well defined for protein analysis (Jackson and Mantsch, 1995).

Numerous interpretation tools have been implemented for the analysis of FT-IR spectra. These include application of a second derivative function to the spectra, Fourier-self deconvolution and 2D analysis. It is important to note that these techniques enhance the interpretation of spectra but do not increase the resolution since this is determined by

instrument parameters during raw data collection (Goormaghtigh *et al.*, 1990). The second derivative of the original spectra provides sharp, well-defined peaks that are opposite in sign but conserve the peak positions. Fourier self-deconvolution is band-narrowing technique. Application of an enhancement factor produces taller, narrower peaks that appear at the same peak position and maintain the integrated intensity, which is a useful feature for curve fitting FT-IR spectra (Goormaghtigh *et al.*, 1990).

A Generalized 2D analysis of FT-IR spectra has also been developed (Noda, 1993). Mathematical treatment of the data allows for the determination of in-phase (synchronous) and out-of-phase (asynchronous) changes, caused by an applied disturbance. The technique can be used with spectroscopic methods that are sensitive to the applied disturbance. The combination of the resultant synchronous and asynchronous contour plots resolves overlapping FT-IR bands, reveals the sequence of spectral changes, and permits the assignment of structurally related FT-IR bands (Wang Y. *et al.*, 1998).

Synchronous correlation maps consist of auto-peaks that form along the diagonal. These are due to spectral bands that change with the applied disturbance (Ozaki *et al.*, 1996). A cross-peak in the synchronous map relates two bands that are similarly affected by the applied disturbance. A positive cross-peak indicates two bands that both increase or decrease in intensity while a negative cross-peak indicates one band increases and the other decreases (Ozaki *et al.*, 1996). The asynchronous map contains only cross-peaks for bands that show dissimilar variations to a disturbance. Often, these represent structures found in different molecular environments (Noda, 1993). Rules that have been

described in detail by Noda (1993) can be used to establish a sequence of events during the application of a disturbance.

1.4.3 Mass Spectrometry

Mass spectrometry measures the relative mass to charge ratios (m/z) of ionized molecules. A critical step in mass spectrometry is the ionization of molecules since only molecular ions can be detected. The formation of gaseous molecular ions from solution requires desolvation of the sample combined with the application of energy causing a gain or loss of electrons or protons. Often the high energy applied in the ionization source causes fragmentation of the molecular ion prior to detection. Large biomolecules such as proteins and carbohydrates are prone to fragmentation during ionization and thus require a 'gentle' ionization technique to maintain the native solution structure (Huang *et al.*, 1993). Electrospray ionization (ESI) is routinely used for analysis of biomolecules (Loo, 1995). ESI forms charged droplets by the application of a high electric potential, typically 2.5-5 kV, to a sample solution infused through a fine capillary. In positive mode, this electric potential causes a net positive charge to form a Taylor cone at the sample tip from which droplets with a net positive charge are ejected (Kebarle, 2000). The droplet is sprayed and evaporated by the action of a nebulizing gas such as nitrogen. Evaporation may also be assisted by the application of heat. During evaporation of the droplet the positive ions begin to experience coulombic repulsion as they are forced closer together. Eventually, this repulsion exceeds the surface tension of the droplet, the Rayleigh limit, and positive molecular ions are ejected from the droplet (Kebarle, 2000). Figure 1.4 shows the mechanism of the electrospray ionization process.

It is important to note that the molecular ions must already exist in the sample solution and the ESI interface is used to desolvate and focus positive ions to the mass analyzer. Therefore, solution composition is critical when using ESI. Volatile buffers and acids are preferred since they are easily removed during the desolvation process. A positive molecular ion can be formed by the addition of a cation such as a proton, but can often include sodium, potassium or ammonium ions, which can cause erroneous mass determinations if not included in the mass calculation.

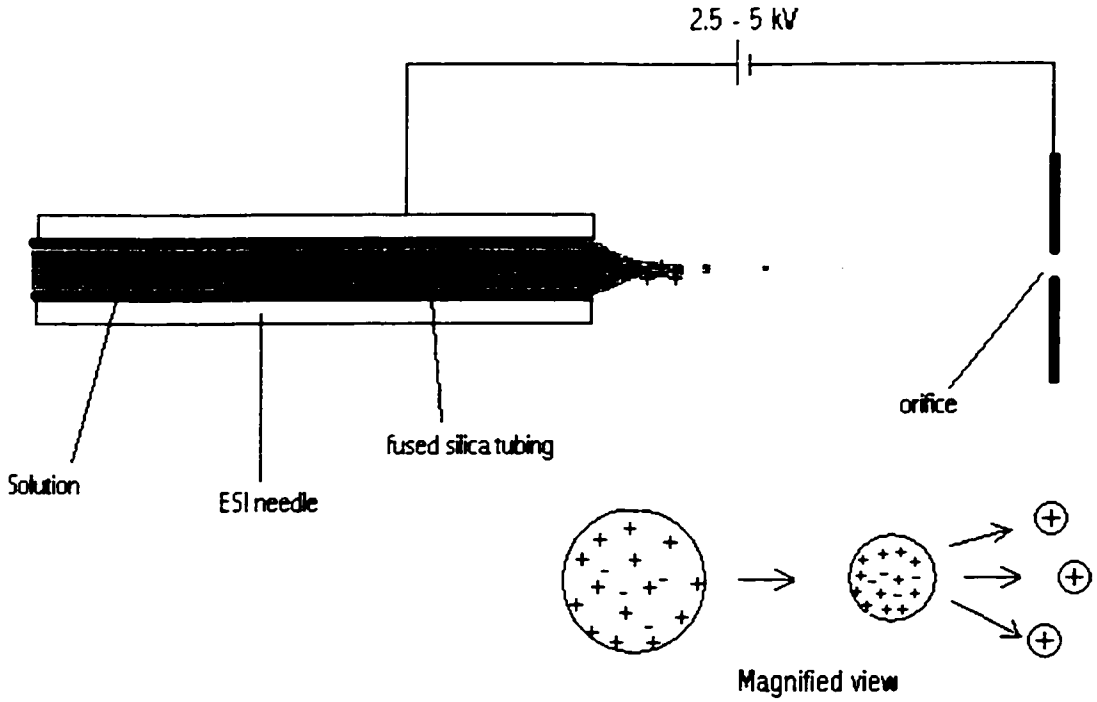

Figure 1.4 Mechanism of electrospray ionization

Polypeptides and proteins containing multiple basic residues such as arginine, histidine and lysine, can be protonated at many sites which in turn give rise to multiple peaks in the mass spectrum. Such peaks can be used to monitor protein interactions and protein conformation. Protein unfolding, either through the application of heat or acidity, can result in the exposure of additional basic residues to the solvent thereby increasing the number of charge states which are detected by the mass spectrometer (Wang and Tang, 1996), (Mizra *et al.*, 1993).

Protein-ligand binding can be monitored by directly observing the mass of the intact protein-ligand complex (Baca and Kent, 1992), (Ganguly *et al.*, 1992). However, non-covalent interactions, such as those that occur during ligand binding can be difficult to observe in the mass spectrum and can be dependent on instrument parameters (Anderegg and Wanger, 1995), (Huang *et al.*, 1993). The ionization process involves the application of heat and a strong electric potential to the protein-ligand complex in solution. The applied energy can cause weak non-covalent interactions to dissociate prior to reaching the mass analyzer. The instrument parameters can influence this effect; for example increasing the ionization voltage reduces the signal observed from a protein-ligand complex (Andergg and Wagner, 1995), (Huang *et al.*, 1993). Therefore, absence of a protein-ligand peak in the mass spectrum does not rule out the presence of a non-covalent complex in solution.

1.5 The PDZ Domain

Signaling cascades involve multiple interactions between different proteins. The interaction between a ligand and receptor can be studied *in vitro*; however, it is difficult

to relate *in vitro* binding to the organization and reactivity found *in vivo*. Many important cell signaling pathways involve non-specific, ubiquitous chemical messengers such as Ca^{2+}, glutamate or NO, which suggests that specificity is controlled by accurate localization of signaling proteins. Although it is difficult to map and understand the complicated architecture of the cell, many new theories of assembly have been proposed from studies on scaffolding and adapter proteins (Craven and Bredt, 1998).

Studying the interactions that occur among functionally related proteins should eventually lead to the understanding of signaling mechanisms in cells. Interactions among functionally related proteins have been detected by techniques such as co-immunopreciption of interacting proteins (Cao *et al.*, 1999). From these studies, many new insights into cell architecture have been reported, and it has been concluded that receptors and ion-channels are not randomly scattered across the cell membrane but must be properly positioned within a signal cascade (Doyle *et al.*, 1996). Proper positioning is accomplished in part by binding domains that bring together the proteins required to carry out a cell function.

Cell Signaling in many instances involves a surface receptor receiving a chemical message from a ligand and transmitting it through a second messenger system. These messages can affect cellular activities such as gene expression, cytoskeletal organization, protein trafficking and cell cycle progression (Ponting *et al.*, 1997). There are many binding domains that have been identified and linked to cell signaling such as SH2, SH3 (src homology), PH (pleckstrin homology), PTB (phosphotyrosine), WW and many others (Oschkinat, 1999).

The name PDZ domain comes from the proteins in which it was first observed. These include post synaptic density protein (PSD-95), Discs-large protein and the Zonula occludens ZO-1, a tight junction protein (Hirao *et al.*, 1998). To date there is a long list of proteins containing one or more PDZ domains (Ponting *et al.*, 1997), (Sheng and Sala, 2001). The PDZ domain is approximately 100 residues in length and has a structural motif containing multiple β-sheets, usually five or six, as well as two or three α–helices (Doyle *et al.*, 1996). This domain forms a compact globular structure with a diameter of 25-30 Å (Ponting *et al.*, 1997). As shown in Figure 1.5, the N and C termini are in close proximity which permits the domain to be 'transferred' as a unit to another protein. An example is the PDZ domain of syntrophin, which is inserted within one of its PH domains (Ponting *et al.*, 1997). Syntrophin binds nNOS through PDZ interactions and has implications in muscular dystrophy (Wang *et al.*, 2000).

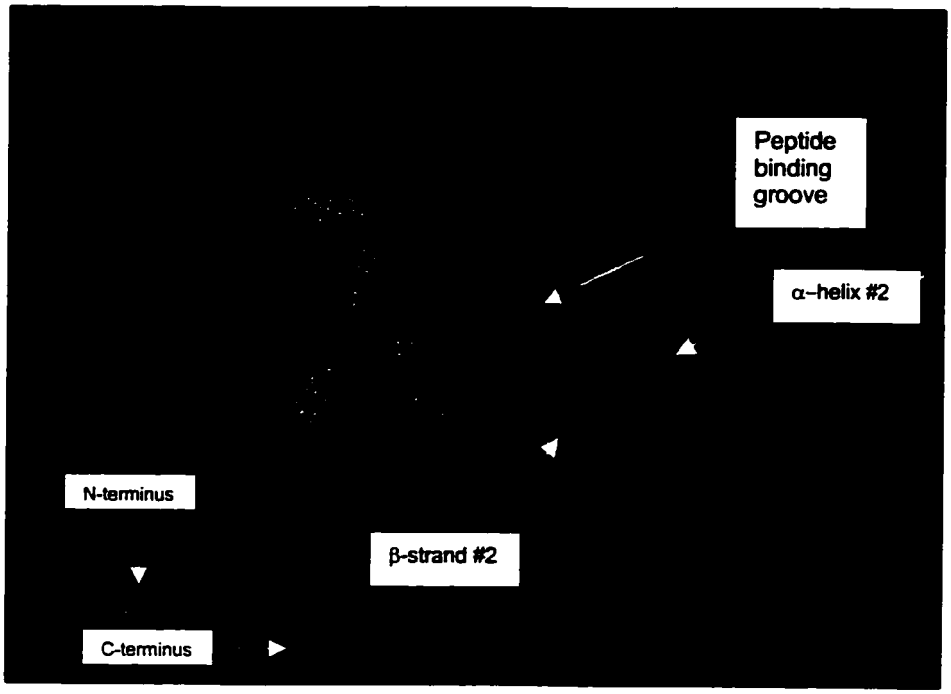

Figure 1.5 Rasmol representation of the second PDZ domain of hPTP1E

A binding groove exists between α–helix #2 and β-strand #2 structures in PDZ domains (Figure 1.5). The C-terminus of the target protein binds in this groove and forms hydrogen bonds with the amide backbone of the adjacent β-sheet. There are many proteins that contain multiple PDZ domains and can therefore bind multiple targets for clustering and scaffolding (Sheng and Sala, 2001). When PDZ domains are found in proteins with catalytic domains, such as a phosphatase, kinase or a synthase, it is believed that they are involved in the control of a signaling pathway (Ponting *et al.*, 1997). The K_D values of PDZ-peptide interactions typically range between 10^{-8} to 10^{-6} M (Sheng and Sala, 2001).

As the study of PDZ domains progresses, many targets proteins have been observed, and classes of PDZ domains have been established based on the properties of the binding target. PDZ domains in Class I bind to the C-terminus of proteins or peptides that end in the following sequence: Ser/Thr-X-hydrophobic-COO^-. The hydrophobic position is usually occupied by Val/Ile. Affinity is based on hydrogen bonding between the hydroxy group of Ser/Thr at the -2 position and an imidazole nitrogen of His in the β-sheet. The hydrophobic functional group of the C-terminal Val/Ile falls into a hydrophobic pocket lined with Phe, Leu, Ile and Leu (Doyle *et al.*, 1996). The terminal carboxylate is cradled in a Gly-Leu-Gly-Phe loop, which stabilizes the negative charge through interactions with amide nitrogens along with an ordered water molecule bound to an Arg (Doyle *et al.*, 1996). Figure 1.6 shows a schematic diagram of the PDZ-peptide interaction and the hydrogen bonds that form (I. Ekiel, personal communication). Note that the sequence of the carboxylate binding loop is Ser-Leu-Gly-Ile (Kozlov *et al.*, 2000). Examples of Class I PDZ domains include the second PDZ of PTP1E that binds

the C-terminus of Fas, and PSD 95 that binds the C-terminus of NMDA receptors and Shaker type K^+ channels (Saras *et al.*, 1997).

PDZ domains in Class II show specificity towards C-terminal sequences of the type, hydrophobic-X-hydrophobic-COO^-. For Class II domains, a hydrophobic residue on the β-sheet replaces His found in Class I (Tochio *et al.*, 1999) and the preferred residues at the -2 position of the incoming C-terminal ligand are Phe, Tyr, Val and Ile (Sheng and Sala, 2001). Examples of class II PDZ domains are found in p55, human LIN-2, Tiam-1 and AF-6 (Sonyang *et al.*, 1997).

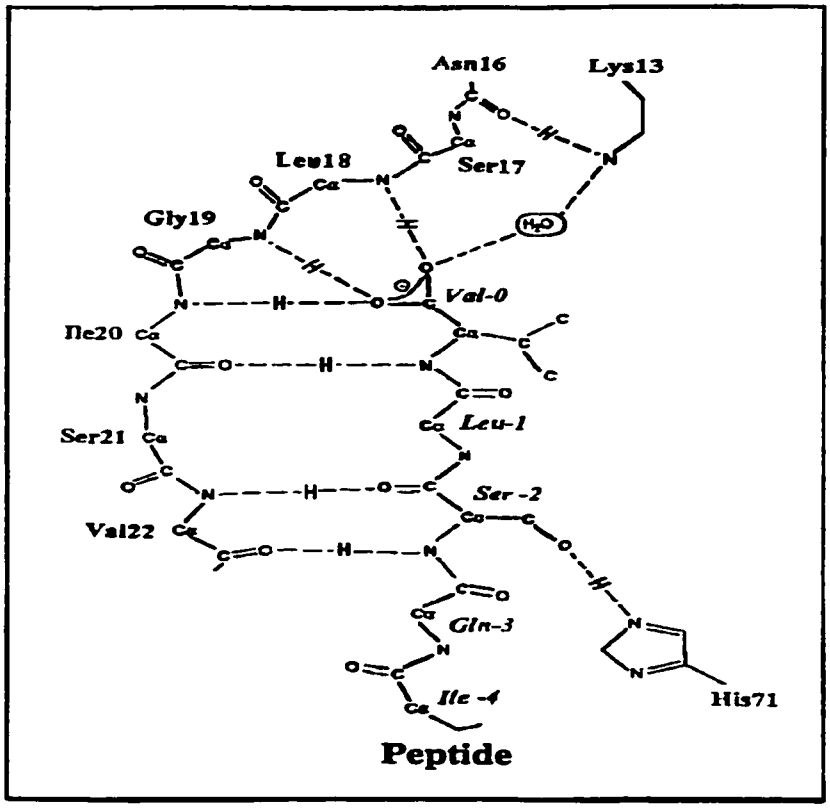

Figure 1.6 Schematic representation of the second PDZ of PTP1E binding the C-terminus of a peptide (INSLV)

Class III PDZ domains bind the C-terminus of peptides containing Asp-X-Val-COO^-. This class selects negatively charged residues at the -2 position from the C-terminus. The negatively charged Asp residue forms a hydrogen bond with the hydroxyl of Tyr and is also influenced by contribution from a nearby positively charged Arg (Tochio *et al.*, 1999). The Val is selected by the same hydrophobic pocket as found in Class I domains. An example of a Class III PDZ domain is found in the peptide binding site of nNOS that binds the melatonin receptor C-terminus (Tochio *et al.*, 1999).

PDZ domains are also capable of forming dimers (Hillier *et al.*, 1999). A thirty residue extension (100-130) of the PDZ domain of nNOS forms two additional small β-strands. Within residues 100-130, there is formation of a tight β-hairpin finger exposing the sequence Glu-Thr-Thr-Phe, which has been shown to bind to a Class I PDZ domain in PSD 95 (Christopherson *et al.*, 1999) or syntrophin (Oschkinat, 1999). The sharp turn is required to fit into the carboxylate binding loop which otherwise expects a C-terminal carboxylate group. This interaction permits dimerization between two PDZ domains (Tochio *et al.*, 2000).

1.6 Statement of Purpose

The binding between the isolated recombinant second PDZ domain of hPTP1E and a target peptide from the C-terminal sequence of the Fas receptor was studied using FT-IR, CD and mass spectrometry. The goal was to gain additional information on the stability and dynamics of peptide-PDZ interactions through the application of instrumental techniques that are complementary to NMR and X-ray crystallography. Details in to the mechanism of PDZ binding were investigated using three different

peptides in order to demonstrate selectivity towards Ser versus Cys or Tyr in the -2 position.

FT-IR, CD and MS were selected because they can rapidly analysis small amounts of protein (0.1mg) with minimal sample preparations. In addition, the protein and peptide interactions are studied in solution which more closely approximates the native environment than the solid state. Application of these methods can potentially provide rapid screening of protein-ligand binding, which could be useful for studying protein-drug interactions. There are currently no previous reports of using these techniques for studying peptide interactions with PDZ domains.

Previous studies have reported PDZ/peptide binding using X-ray crystallography (Doyle *et al.*, 1996) and NMR (Kozlov *et al.*, 2000). X-ray crystallography offers high resolution but it requires a protein crystal, which is often difficult and time consuming to prepare. In addition, the static picture obtained from solid-phase protein crystals may not represent the conformation of the native protein in solution (Surewicz *et al.*, 1993). X-ray crystallography is an invaluable tool for high-resolution structural determinations but is less suited to probing protein conformations under different conditions.

NMR is another high-resolution technique which is invaluable for structural characterization. It also requires lengthy data analysis but it does offer the advantage of determining protein structure in solution. NMR measures spin relaxation rates of atoms such as H, C and F, which occurs on the time-scale of micro-seconds while FT-IR can detect vibrational relaxations that occur on the pico-second timescale (Ozaki and Noda, 1996). Therefore, bond vibrations and tautermerism that are not resolved by NMR are resolved using faster techniques such as FT-IR. Sensitivity limits of NMR require larger

amount of proteins (1mg) compared to FT-IR or mass spectrometry (Smith *et al.*, 1997) which may be an issue if the protein quantities are limited.

1.7 Thesis Organization

This thesis is divided into four interrelated chapters. The first chapter introduces general protein structure, the PDZ domain and the instrumental methods used for this project. The next two chapters are divided according to the instrumental techniques employed. Chapter 2 describes the investigation of peptide binding to the PDZ domain using the spectroscopic methods, FT-IR and CD. Temperature was used to induce denaturation, and 2D analysis of the spectral data was applied to determine the sequence of unfolding events. The study of peptide binding to the PDZ domain by mass spectrometry is discussed in Chapter 3. H/D exchange, monitored by MS, was also used to compare the stability of the free and peptide-bound PDZ domain. Finally, general conclusions and suggestions for further research are contained Chapter 4.

CHAPTER 2 - SELECTIVITY OF PEPTIDE BINDING TO SECOND PDZ OF hPTP1E USING FT-IR SPECTROSCOPY

Abstract

The interactions between the recombinant PDZ domain of human protein tyrosine phosphatase 1E and three different peptides were monitored using FT-IR spectroscopy. Spectra were recorded as a function of temperature to detect differences in the thermal denaturation of the free protein and its complexes. The observed T_m for the PDZ domain was 47°C while that of the PDZ-ENEQVSAV complex was 57°C, consistent with the formation of a stable complex. Combinations with two other peptides, ENEQVCAV and KDDEVYYV, showed only minor increases in the observed T_m. Curve fitting results for the FT-IR spectra indicate a shift in the relative intensity of two β-sheet bands at 1639cm^{-1} (decrease) and 1628cm^{-1} (increase) for the PDZ-ENEQVSAV complex which suggests addition of the peptide to the β-sheet region. These results are consistent with previous X-ray work (Doyle *et al.*, 1996) which showed that the peptide binds as a β-strand in an anti-parallel orientation. 2D correlation analysis of the FT-IR data was employed to determine the sequence of events that occurs during unfolding and revealed a similar pattern between the native and peptide-bound form of the PDZ protein. However, it was observed that with the PDZ-ENEQVSAV complex, β-turns are disrupted after protein aggregation. Evidence from CD spectroscopy suggested that helical structures remained intact during heating to 75°C for PDZ and the PDZ-ENEQVSAV complex.

2.1 Introduction

The PDZ domain is responsible for trafficking and scaffolding protein-protein interactions in cells (Ponting *et al.*, 1997). This occurs through the recognition of specific amino acid sequences found in the C-terminus of target proteins. Examples of PDZ-mediated interactions include linking of nitric oxide synthase to NMDA receptors (Christopherson *et al.*, 1999), the association of Fas receptors with FAP1 (Sato *et al.*, 1995) and aggregation of proteins by INAD in the fruit fly visual system (Tsunoda *et al.*, 1997). X-ray and NMR analysis report structural homology between different PDZ domains which usually contain six β–strands and two α–helices (Doyle *et al.*, 1996, Ekiel *et al.*, 1998). The C-terminal residues from the incoming ligand form hydrogen bonds to a β-strand in an anti-parallel orientation and additional hydrogen bonding occurs with amino acid side-chains from an adjacent α–helix. Hydrogen bonding occurs between Ser/Thr in the –2 position on the target with His-71 on α–helix#2 of the second of PDZ domain of hPTP1E and determines its specificity (Ekiel *et al.*, 1998). The negative charge on the incoming C-terminal carboxylate group is stabilized through interactions with amide nitrogens found in a region of the PDZ domain termed the carboxylate binding loop which contains the conserved sequence, Gly-Leu-Gly-Phe (Doyle *et al.*, 1996, Saras *et al.*, 1996) or Ser-Leu-Gly-Ile for the second PDZ of hPTP1E (Kozlov *et al.*, 2000). Recent reports (Hillier *et al.*, 1999) indicate hetero-dimerization between the PDZ domains of neuronal nitric oxide synthase and syntrophin through the recognition of internal amino-acid sequences. The internal sequence Glu-Thr-Thr-Phe is found within a tight β-hairpin finger of nNOS and fits into the PDZ peptide-binding groove of syntrophin.

FT-IR is used here to probe for possible structural changes that occur during peptide binding and to monitor the thermal denaturation process of the PDZ-peptide complex as compared to the free protein. FT-IR is increasingly used to study proteins as it provides the following advantages: only small amounts of sample are required, the protein is studied in its native solution structure, and the molecular vibrations that are detected by FT-IR occur on a faster time-scale than NMR (Mantsch *et al.*, 1993). In addition, estimates of protein secondary structure composition using FT-IR show comparable results to X-ray crystallograpy (Trewhella *et al.*, 1989), (Filosa *et al.*, 2001)

The melting temperature of a protein is the temperature where unfolding from a native to a denatured state occurs, and has been used to indirectly monitor ligand-binding (Pace and McGrath, 1980). The dominant amide I′ band at 1639cm^{-1} and the appearance of aggregation at 1618cm^{-1} were both used to monitor the thermal denaturation process and to determine the T_m. Also, any differences between free PDZ and the PDZ-peptide combinations that would suggest a binding interaction, such as an increase in T_m or an altered unfolding pattern, were probed using FT-IR.

The FT-IR amide I′ region of the PDZ domain was compared to that of the PDZ-ENEQVSAV complex. Curve fitting the spectra indicates that the intensity of two β-sheet frequencies (1628 and 1639cm^{-1}) were different for the PDZ-ENEQVSAV complex as compared to the free PDZ domain. The 1628cm^{-1} band showed increased intensity as compared to the 1639cm^{-1} band suggesting addition of a short β-type structure within the β-sheet component (Fabian *et al.*, 1993). Thermal denaturation monitored by FT-IR showed a 10°C increase in T_m with the addition of the ENEQVSAV peptide. Increases in T_m with the alternate peptides KDDEVYYV and ENEQVCAV were 1 and 2°C,

respectively. The denaturation process was further evaluated using a generalized 2D-correlation analysis (Noda, 1993) in order to assign secondary structures for the observed spectral bands and to describe the sequence of unfolding events. The 2D results indicate that β-sheets and β-turns are disrupted prior to protein aggregation in free PDZ but in the PDZ-ENEQVSAV complex β-turns are disrupted after protein aggregation.

The use of software tools such as curve fitting and 2D correlation enhance the analysis and interpretation of FT-IR spectra. Through these enhancements it is possible to detect subtle interactions such as binding of a short peptide to a protein.

2.2 Experimental Procedures

2.2.1 Materials

The recombinant isolated second PDZ of hPTP1E as well as the peptides ENEQVSAV, ENEQVCAV and KDDEVYYV were obtained from Dr. Irena Ekiel (Biotechnology Research Insitute, Montreal, Canada). The peptide KDDEVYYV was obtained in crude form and required purification by HPLC on an HP1100 (Hewlett Packard) with a reversed-phase C18 column 150x4.6mm (Waters). The PDZ domain and the two other peptides had been purified as described previously (Ekiel *et al.*, 1998). A buffer solution (pH 6.9) was prepared using sodium phosphate (Fisher). 99.9% D_2O was obtained from Aldrich. Tyrosine ethyl ester (HCl salt) was purchased from Sigma.

2.2.2 Methods

FT-IR spectra were obtained using a Nicolet Magna 550 FT-IR spectrophotometer equipped with a KBr-DTGS detector and purged with dry air. Samples were mounted in

a FT-IR cell between two 13 x 2mm CaF_2 windows separated by a teflon spacer. For amide I spectra, a 50μm teflon spacer was used and for monitoring the SH stretch region a 250μm teflon spacer was employed (Wilmad, New Jersey). Data handling was performed using Omnic 3.1 Software (Nicolet). CD spectra were collected on a Jasco J-710 CD spectrophotometer using a thermostated FT-IR cell with a 50μm teflon spacer. A 5L/min flow-rate of nitrogen was used to eliminate the far-UV interference from oxygen. The short optical pathlength permitted scanning protein concentrations of 1mM in 50mM phosphate buffer without saturating the CD detector. The photomultiplier voltage did not exceed 700V during any of the data collections.

A Hewlett Packard model 8453 UV/VIS spectrophotometer was used for measuring protein concentrations according to the method of Gill and von Hippel (1989). Thermal spectra were collected using a thermostated cell (Harrick), a temperature controller (Omega) and TempSet software (Dwight Analytical Solutions, Ontario). Temp Set software was used to coordinate FT-IR spectra collection with each temperature set-point. Data handling of the thermal spectra was performed using Omnic Software and Ramopn/Gplot. Curve fitting was performed using Grams Peaksolv. 2D analysis was carried out using software written by Dr. Yan Wang (Wang Y. *et al.*, 1998).

Samples of PDZ domain were exchanged three times with 50mM sodium phosphate pH 6.9 using a Millipore Centricon-3 concentrator and a Beckman JA-20 centrifuge. The protein solution was concentrated to 1mM and lyophilized. For FT-IR spectra in deuterated solvent, the lyophilized protein samples of the PDZ domain were diluted with D_2O to a final concentration of 1mM protein and 50mM phosphate pD 7.3. Prior to collecting spectra, the protein sample was allowed to exchange with D_2O for at

least 24 h at 4°C. The buffer spectra were subtracted from PDZ spectra. PDZ/peptide complexes were prepared by reconstituting the 1mM PDZ with a solution containing the 1.2 mM target peptide. A sample of 2mM PDZ with 2.4mM KDDEVYYV was also prepared to observe the Tyr side chain absorption at 1516cm^{-1}. A spectrum of peptide in deuterated phosphate buffer was subtracted from the PDZ/peptide spectra. A vapor spectrum was then subtracted from the resultant spectra according the method of Dong *et al.*, (1992).

The amide I′ (1600-1700cm^{-1}) FT-IR spectra were obtained by averaging 256 scans from 4000-400cm^{-1} using a resolution of 2. Amide I' FT-IR spectra were deconvolved using a half-width-at-half-height of 15 and enhancement factor of 1.5. To monitor the SH stretch region (2500-2600cm^{-1}), a longer pathlength was required and the number of scans was increased to 512, since the ν(SH) stretch is weak. These samples (at concentrations of 2 - 4mM) were prepared in 50mM phosphate buffer in H_2O at pH 6.9 with equimolar DTT and DTPA added to ensure that the cysteine sulfhydryl was reduced. FT-IR spectra of the SH stretching region and the Tyr band, were deconvolved using a half-width-at-half-height of 14 and enhancement factor of 1.9.

Thermal denaturation plots were obtained by collecting the FT-IR spectra at each temperature set-point. The initial temperature was 25°C and it was raised 2°C every 15min up to 89°C. The cooling cycle was also monitored from 89°C to 25°C decreasing 2°C every 15min. Scan time was 9min 18s after each 15min equilibration period. To eliminate the need for background subtraction of each of stacked spectra, these samples were prepared in D_2O.

Thermally-induced protein denaturation was monitored by CD spectroscopy using the thermostated FT-IR cell. The temperature was manually set to the specified temperature and the sample was equilibrated for 15min prior to scanning. Five accumulations were obtained across the 185-300nm spectral range with a step resolution of 0.2nm. The bandwidth was set to 1.0nm and the response was 0.25s. CD spectra of PDZ and PDZ-ENEQVSAV were collected over a temperature range of 25 - 75 °C.

2.3 Results

The amide I′ (1600-1700cm^{-1}) FT-IR spectra of the PDZ protein and ENEQVSAV-PDZ in D_2O are shown in Figures 2.1 and 2.2, respectively. D_2O was used to reduce unwanted interference from liquid water and its vapor, since H_2O absorbs strongly at 1650cm^{-1} while the absorption of D_2O is red shifted to 1200cm^{-1}. Nevertheless, a buffer background and vapor spectrum was subtracted from each spectrum according to method of Dong *et al.*, (1992).

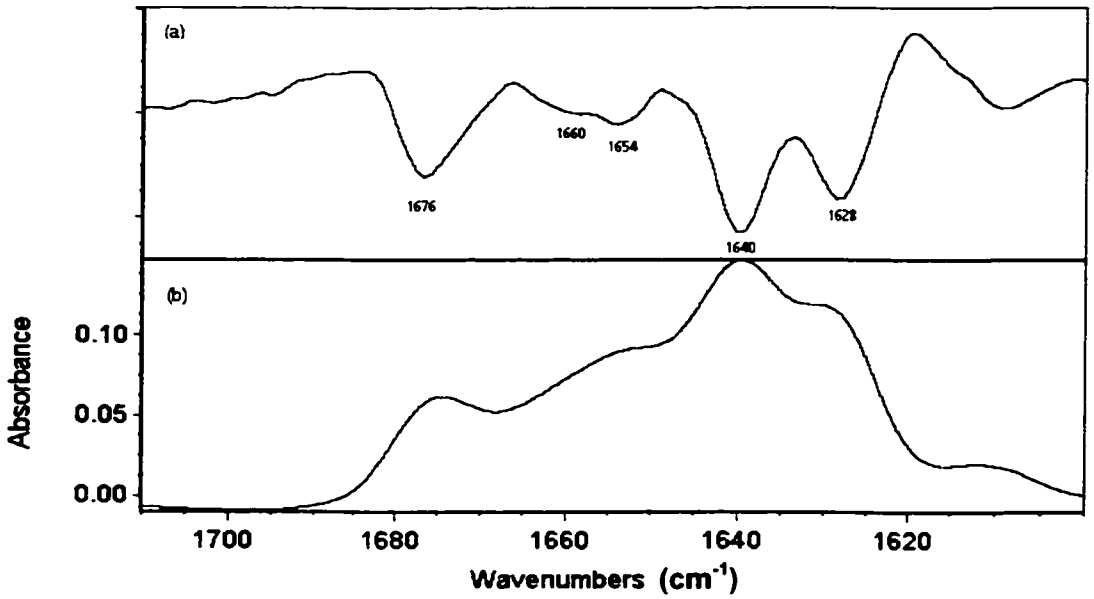

Figure 2.1 (a) Second derivative spectrum of 1mM PDZ in D_2O; (b) Deconvolved spectrum of 1mM PDZ in D_2O (FSD bandwidth set to 15 with an enhancement factor of 1.5)

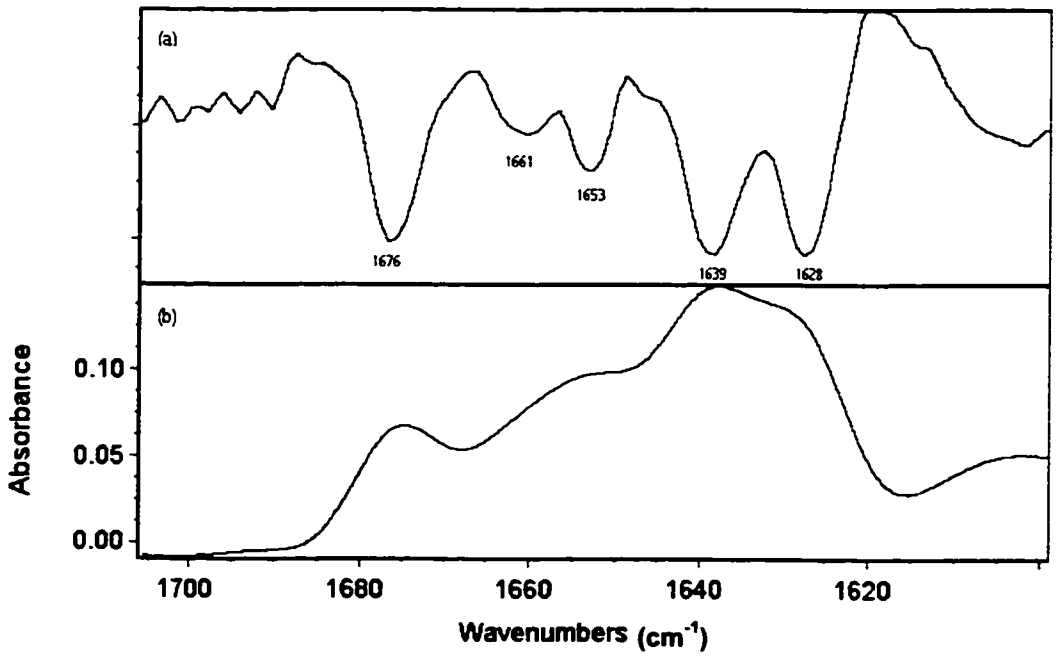

Figure 2.2 (a) Second derivative spectrum of 1mM PDZ with 1.2mM ENEQVSAV peptide in D_2O; (b) Deconvolved spectrum of 1mM PDZ with 1.2mM ENEQVSAV peptide in D_2O (FSD bandwidth set to 15 with an enhancement factor of 1.5)

Apart from minor differences in positions of the peak maxima, the spectra of the free and peptide-bound PDZ domain share the similar spectral features. Both are dominated by bands at 1628 and 1639/1640 cm^{-1}, which are characteristic of β-sheets and 1676cm^{-1} which is characteristic of turn-like structures (Zhang *et al.*, 1994). The lower frequency β-sheet band at 1628cm^{-1} can be attributed to a solvent exposed or tightly hydrogen bonded β-strand (Casal *et al.*, 1988), (Jackson and Mantsch, 1992), (Arrondo *et al.*, 1988) or a short, distorted β-structure added to the outer region of a β-sheet. Two additional peaks are observed in the region of 1650-1661cm^{-1}. The lower frequency component at 1653/1654cm^{-1} represents unordered regions while the higher frequency band at 1660/1661cm^{-1} is assigned to α–helices based on literature assignments (Jackson

and Mantsch, 1995). Side chain absorption appearing between 1606-1611cm^{-1} may be assigned to Arg (Zhang *et al.*, 1994) and Gln (Venyaminov *et al.*, 1990).

The FT-IR spectra for the PDZ and PDZ-ENEQVSAV complex in D_2O were curve fitted in order to monitor any secondary structural changes. Curve fitting was performed by selecting the 6 dominant peaks found in the second derivative spectrum of the amide I′ region 1600-1700cm^{-1}. The deconvolved spectra were fitted using Gaussian peaks. Figures 2.4 and 2.5 contain the curve fitted FT-IR spectrum of the free PDZ domain and the PDZ-ENEQVSAV complex, respectively.

A summary of the curve fitting results for the protein spectra are shown in Table 2.1. The FT-IR spectra of the ENEQVSAV-PDZ complex shows an increase in the 1628cm^{-1} band at the expense of the 1639cm^{-1} band, which has been interpreted as tighter hydrogen bonding within a β-strand or a β-strand with increased interactions with the solvent (Arrondo *et al.*, 1988), (Jackson and Mantsch, 1992) and (Casal *et al.*, 1988). Fabian *et al.* (1993) described the lower frequency β-sheet band as a distorted β-structure formed by very short peptide strands at the outer ends of β-sheets.

The interaction of the peptide with the PDZ domain occurs through anti-parallel hydrogen bonding to a β-strand located within the peptide-binding groove (Kozlov *et al.*, 2000). Therefore, it is likely that the FT-IR results detect the addition of the peptide strand which manifests itself as a relative increase in the lower frequency β-sheet region. There is also some increase in turns (1675cm^{-1}) upon binding the ENEQVSAV peptide. The side chain contributions from Arg and Gln residues were also found to decrease in intensity and red-shift from 1611cm^{-1} to 1606cm^{-1} upon binding the ENEQVSAV peptide; however this difference was within the experimental error of ~ 2%. The

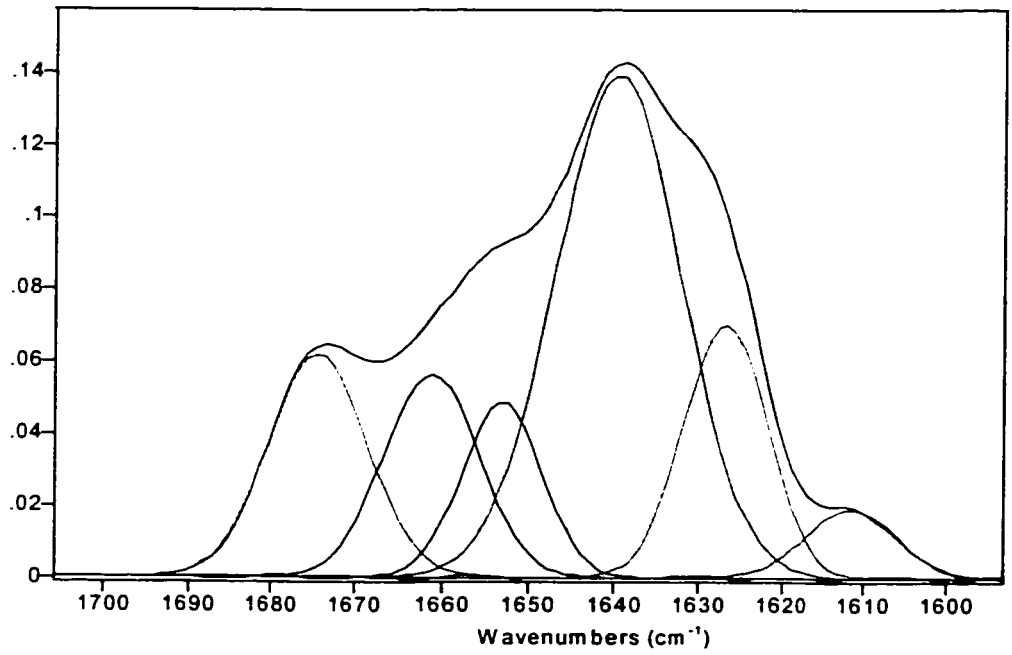

Figure 2.4. Curve fitting of the deconvolved FT-IR spectrum of PDZ in D_2O (FSD bandwidth set to 15 with an enhancement factor of 1.5)

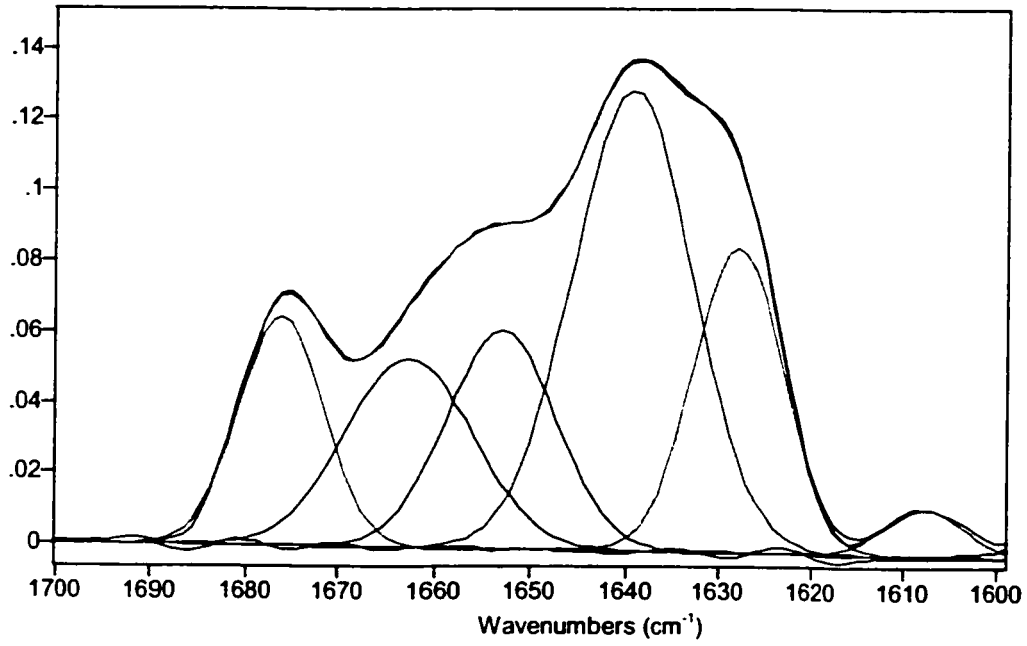

Figure 2.5. Curve fitting of the deconvolved FT-IR spectrum of the PDZ-ENEQVSAV complex in D_2O (FSD bandwidth set to 15 with an enhancement factor of 1.5)

percent relative composition of bands assigned to helices and unordered structures remain essentially unchanged with the addition of the ENEQVSAV peptide.

Table 2.1 Curve fitting results of the free PDZ and ENEQVSAV-PDZ complex.

FT-IR Spectral band	1676 cm^{-1} Turns	1660/61cm^{-1} α–helix	1653/54cm^{-1} Unordered	1639/40cm^{-1} β-sheet	1628cm^{-1} β-sheet	1606-1611cm^{-1} side-chains (Arg and Gln)
PDZ + ENEQVSAV	15.3%	13.1%	14.7%	33.5%	19.4%	3.9%
PDZ	13.1%	12.5%	14.3%	38.5%	15.7%	5.8%
Difference	-2.2%	-0.5%	-0.4%	5.0%	-3.7%	1.9%

Thermal denaturation results

Thermal stress was selected to induce denaturation since it can be precisely controlled and does not interfere with the protein spectrum as would chemical denaturants such as guanidine hydrochloride (Esposito, 2001). The temperature can be increased stepwise and subsequently decreased to verify reversible folding/unfolding. The stacked FT-IR spectra of the heating and cooling of PDZ are shown in Figures 2.5 and 2.6, respectively. Similarly, the heating and cooling spectra for the PDZ-ENEQVSAV complex are shown in Figures 2.7 and 2.8, respectively. From the stacked spectra, the integrated intensity of the β-sheet band at 1639cm^{-1} versus temperature was plotted in Figure 2.9 and was used to obtain the T_m. Likewise, the integrated intensity of the aggregation band centered at 1618cm^{-1} versus temperature was plotted in Figure 2.10.

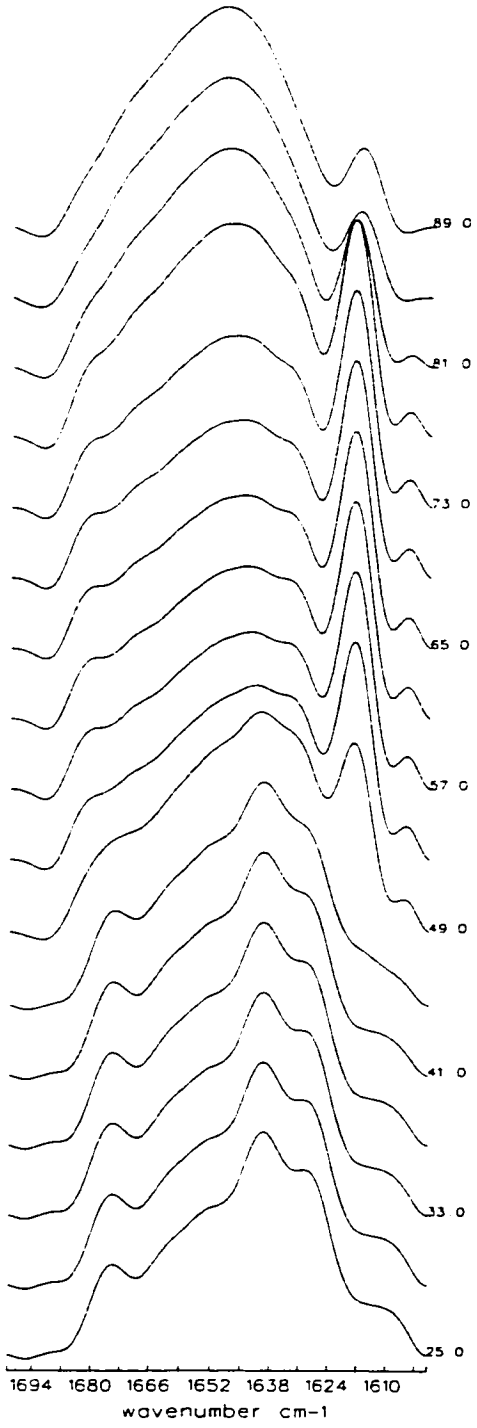

Figure 2.5 Stacked deconvolved FT-IR spectra of the free PDZ protein (heating cycle). FSD bandwidth set to 15 with an enhancement factor of 1.5

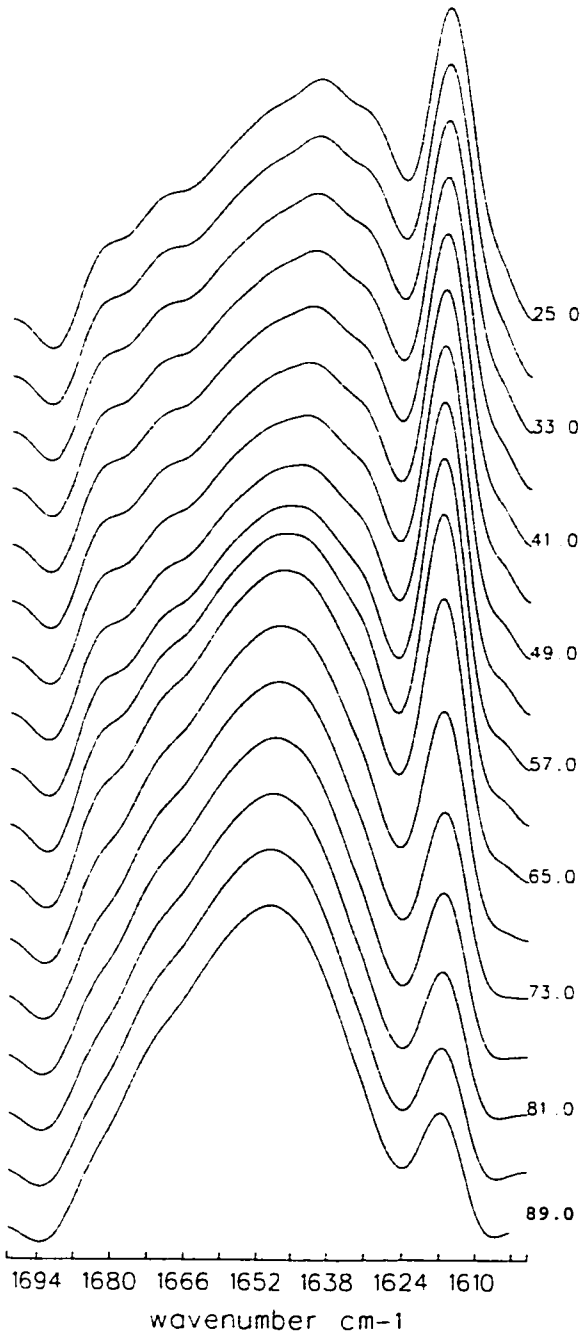

Figure 2.6 Stacked FT-IR spectra of the free PDZ protein (cooling cycle). FSD bandwidth set to 15 with an enhancement factor of 1.5

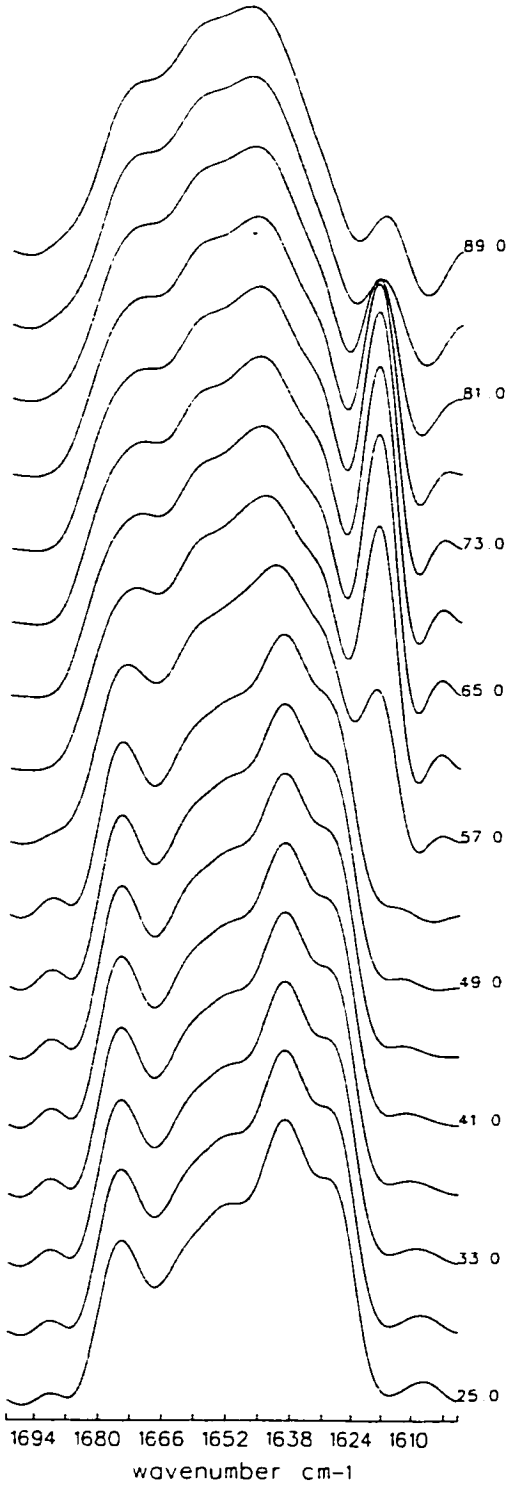

Figure 2.7 Stacked FT-IR spectra of the PDZ-ENEQVSAV complex (heating cycle). FSD bandwidth set to 15 with an enhancement factor of 1.5

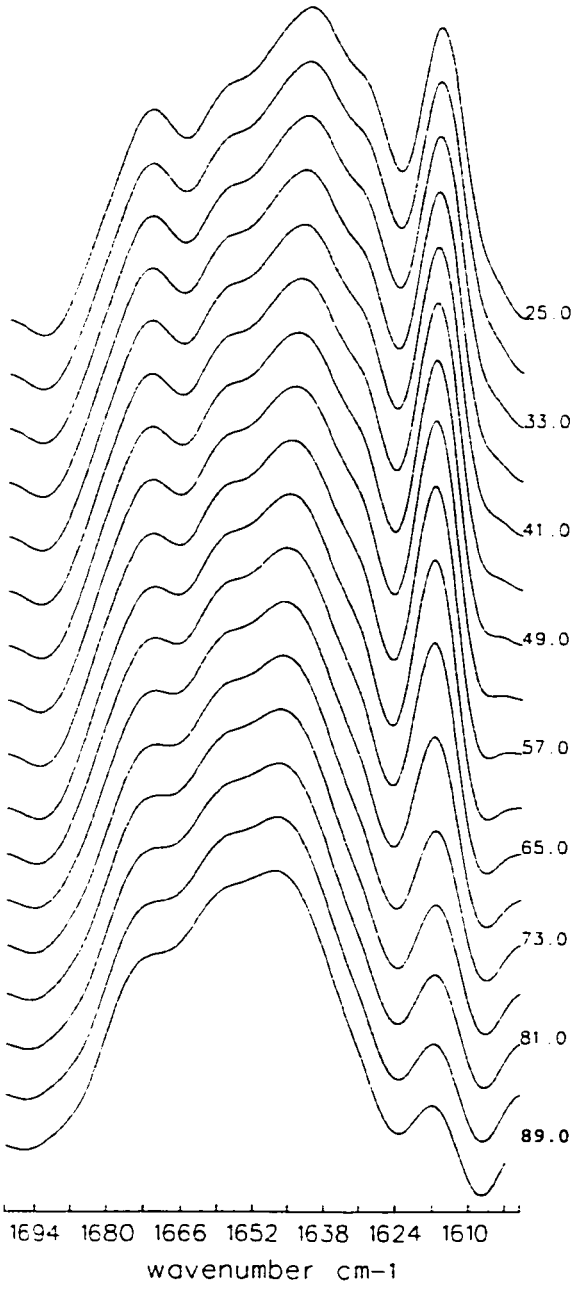

Figure 2.8 Stacked FT-IR spectra of the PDZ-ENEQVSAV complex (cooling cycle). FSD bandwidth set to 15 with an enhancement factor of 1.5

The stacked PDZ spectra in Figure 2.5 show the following features. The amide I′ region remains unchanged from 25°C until denaturation begins at 47°C. The onset of denaturation is indicated by the increase in the band at 1618cm^{-1}, a common feature of denatured proteins (Holzbaur *et al.*, 1990) (van Stokkum *et al.*, 1995), and loss of 1628, 1640 and 1676cm^{-1}. The formation of intermolecular hydrogen bonding between unfolded protein strands results in a lower frequency signal than intramolecular hydrogen bonds found in more constrained β-sheet structures, which showed peaks at 1628 and 1640cm^{-1} (Jackson and Mantsch, 1995). Increasing the temperature above the 47°C transition temperature shows some residual β-structures at 1632cm^{-1} which persists up to temperatures of 79°C. The spectrum of the fully denatured free PDZ domain at 81°C and above show a large featureless peak centered at approximately 1648cm^{-1} which is generally assigned to unordered structures (Jackson and Mantsch, 1995). Upon cooling back to 25°C, some amide I features of the native protein return (Figure 2.6). The bands at 1628, 1640 and 1676cm^{-1} regain part of their original intensity suggesting that the denatured protein can refold residual β-sheet and turn structures. The peak maxima red shifts from 1648 to 1640cm^{-1} indicating that some unordered structures refold to β-sheets. The band at 1618cm^{-1} remains upon cooling indicating that PDZ is irreversibly aggregated.

The stacked spectra of PDZ-ENEQVSAV complex (heating cycle) in Figure 2.7 are similar to those of the free PDZ. The aggregation band appears at 1618cm^{-1}. The spectrum of the denatured complex is also centered at 1648cm^{-1} but pronounced peak shoulders remain at 1661cm^{-1} and 1676cm^{-1} suggesting some residual helical and turn structures persist in the denatured state of the complex. The cooling cycle (Figure 2.8)

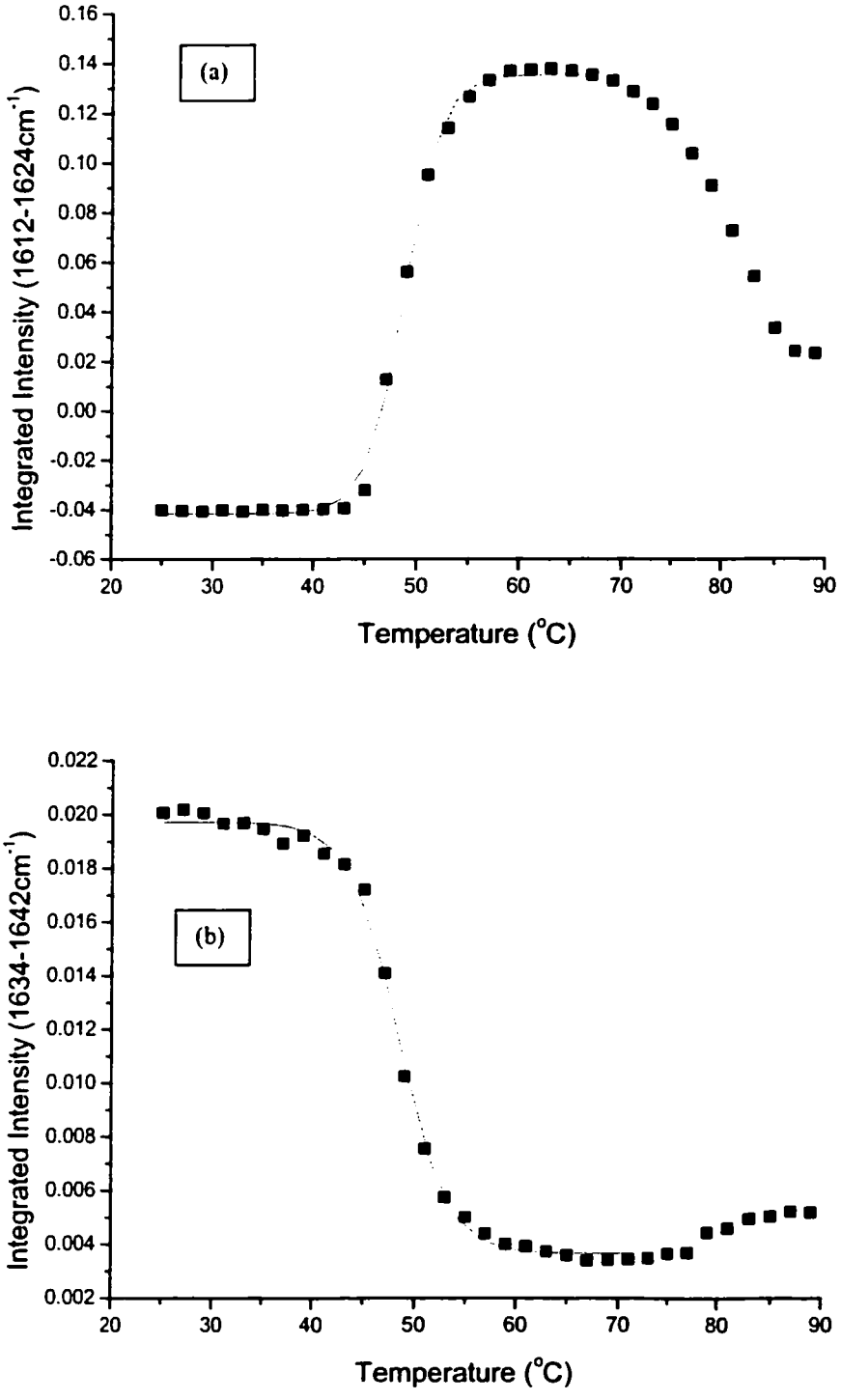

Figure 2.9 PDZ (a) Integrated intensity of the FT-IR aggregation band at 1618cm^{-1} versus temperature; (b) Integrated intensity of the FT-IR β-sheet band at 1639cm^{-1} versus temperature

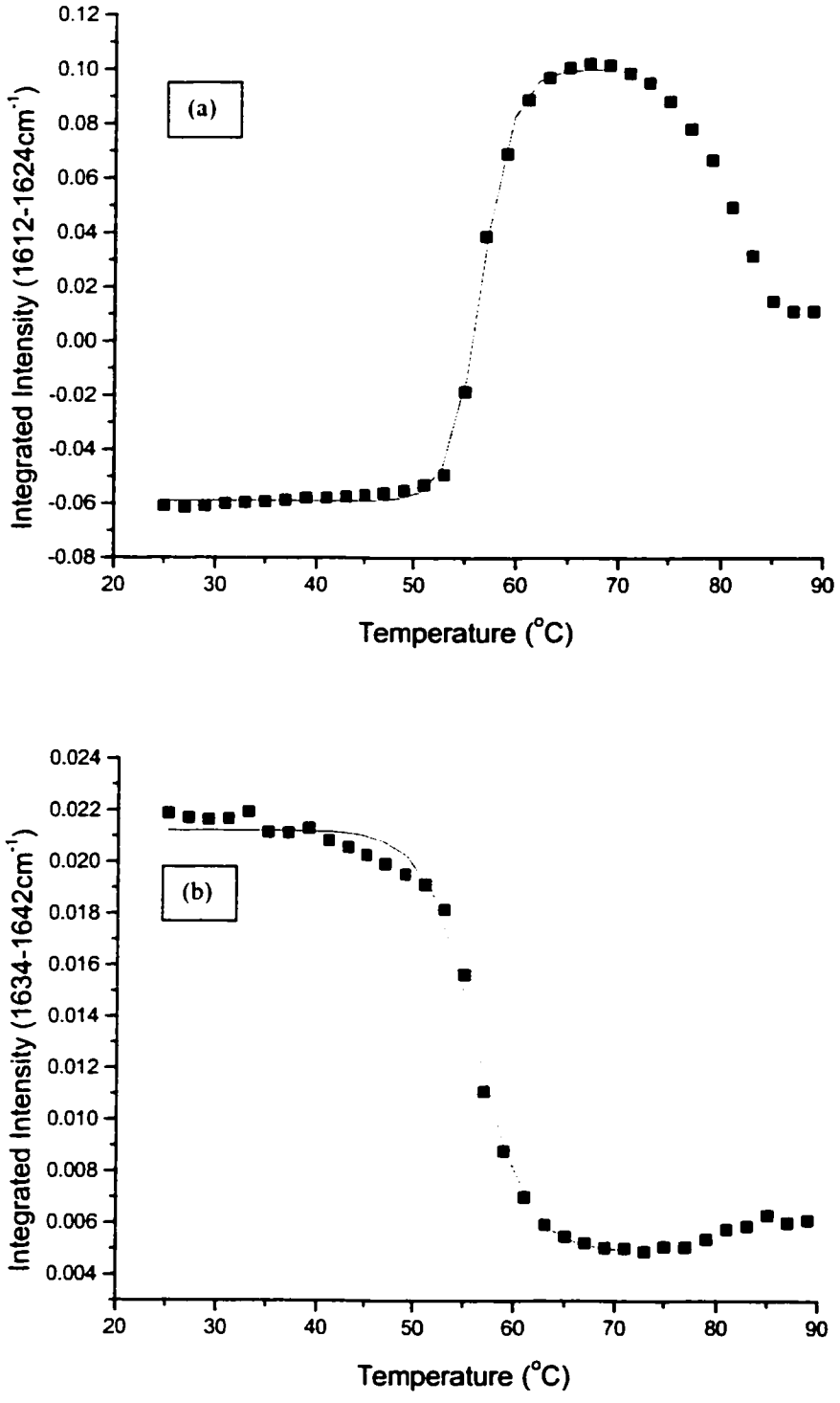

Figure 2.10 PDZ-ENEQVSAV (a) Integrated intensity of the FT-IR aggregation band at 1618cm^{-1} versus temperature; (b) Integrated intensity of the FT-IR β-sheet band at 1639cm^{-1} versus temperature

also shows irreversible aggregation but some amide I peaks at 1628, 1639 and 1675cm^{-1} regain intensity indicating that some residual β-sheet and turn structures have refolded. At temperatures above 70°C the intensity of the aggregation band at 1618cm^{-1} begins to decrease and may indicate that the aggregate strands are disrupted at high temperatures and the protein re-dissolves. This was also observed with the free PDZ. The cooling cycle shows that between 71 - 25°C, the intensity of the 1618cm^{-1} band remains constant. Since, this was observed with free and peptide bound PDZ, the reversibility of unfolding does not appear to be affected by ligand binding.

Stability of the protein sample was monitored by the appearance of aggregation at 1618cm^{-1} and by the loss of the major β-sheet band at 1639cm^{-1}. The T_m value was calculated using the average of these two methods and was found to be the same (within 1°C) regardless of which peak was selected to monitor denaturation. In addition, the T_m was obtained for the combinations of the PDZ domain with ENEQVCAV and KDDEVYYV. Table 2 summarizes the observed T_m values for PDZ protein and combinations with three different peptides.

Table 2.2 Summary of T_m values obtained by FT-IR spectroscopy for the PDZ domain and combinations with three different peptides

Sample	PDZ	PDZ-ENEQVSAV	PDZ-ENEQVCAV	PDZ-KDDEVYYV
Observed T_m(°C)	47 ± 1	57 ± 1	49 ± 1	48 ± 1

The binding of the ENEQVCAV peptide to PDZ domain was investigated by probing probe for hydrogen binding between His-71 in the PDZ domain and the Cys

residue on the peptide. It was hoped that signals in the SH stretching region (≈2500-2600cm^{-1}) would be observed and report on hydrogen bonding with His-71. The SH stretch region is blue shifted far away from interfering bands from water and amide absorptions. The second PDZ of hPTP1E does not contain any Cys residues which could interfere with monitoring the SH stretching region of the peptide spectra. IR bands assigned to thiols are very sensitive to the local environment of the S-H bond (Bare *et al.*, 1975). If upon binding, the thiol group remains in a heterogeneous solvent exposed environment, random polar interactions with the solvent would cause band broadening to occur, and the ν(SH) band would be undetected in the FT-IR spectrum (Bare *et al.*, 1975) (Moh *et al.*, 1987). Figure 2.11 shows the SH stretching region for the free ENEQVCAV peptide, and as predicted, no peaks are observed due to random interactions of the free thiol and the solvent. Figure 2.12 shows the FT-IR spectrum of the PDZ-ENEQVCAV combination. No peaks were detected in the SH stretching region in the presence of the PDZ domain (Figure 2.12). Binding of the ENEQVCAV peptide was therefore not detected through monitoring of the SH stretching region. For comparison purposes, the SH stretching region of human hemoglobin, using similar concentrations and spectral conditions, is shown in Figure 2.13. Human hemoglobin contains 6 Cys (three Cys on two duplicate chains) that appear as three well-resolved peaks. These peaks demonstrate the sensitivity of the SH stretch position and intensity as a probe of the cysteine environment. The Cys residues in human hemoglobin are α-104 (at 2552.6 cm^{-1}), β-112 (2566.3 cm^{-1}) and β-93 (2591.5 at cm^{-1}) (Moh *et al.*, 1987).

A third peptide, KDDEVYYV, was selected which contains a Tyr in the –2 position. The Tyr ring frequency was used to probe insertion of the peptide into

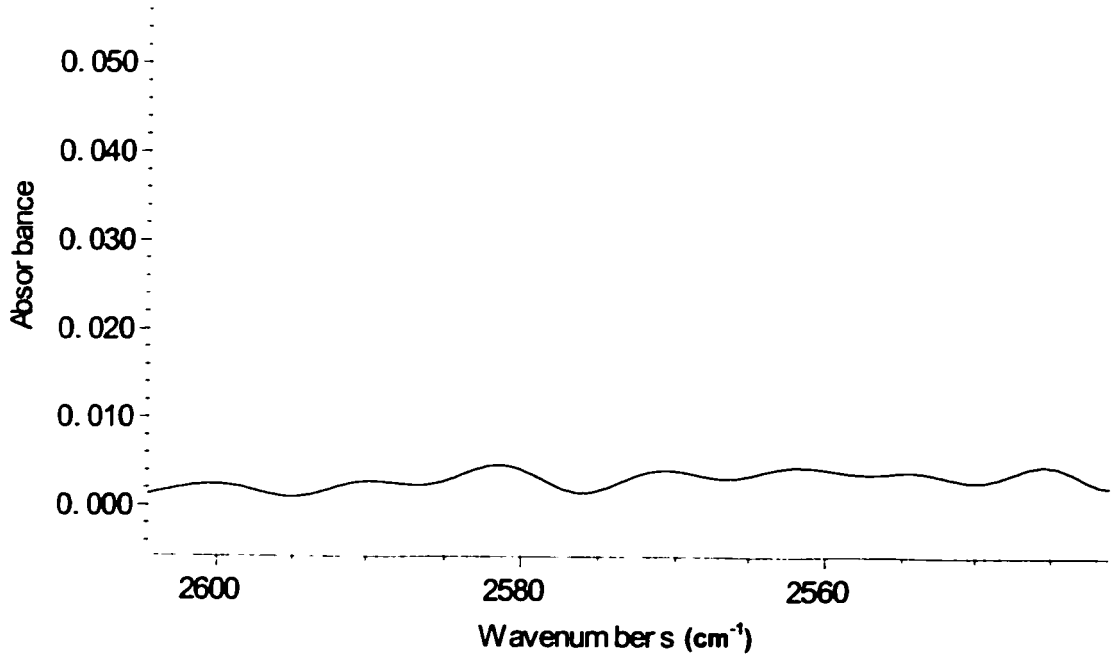

Figure 2.11 2.5mM ENEQVCAV in sodium phosphate buffer pH 6.9, 2.5mM DTT and 2.5mM DTPA (FSD bandwidth set to 14 with an enhancement factor of 1.9)

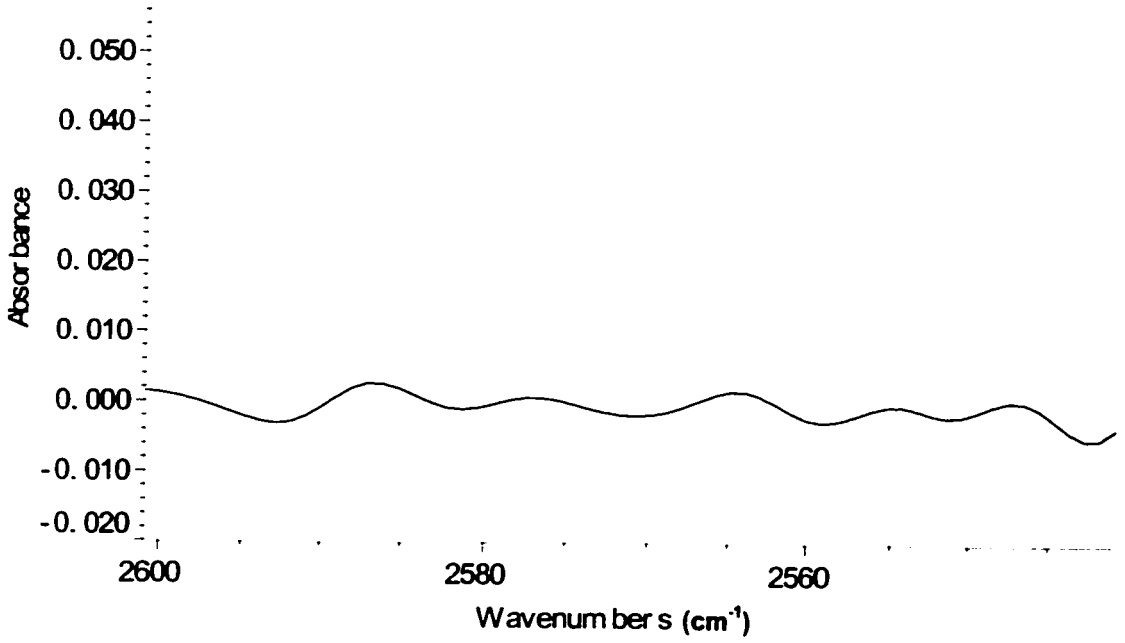

Figure 2.12 Deconvolved FT-IR spectrum of 4mM PDZ-ENEQVCAV in sodium phosphate buffer (pH 6.9) containing 4mM DTT and 4mM DTPA (FSD bandwidth set to 14 with an enhancement factor of 1.9)

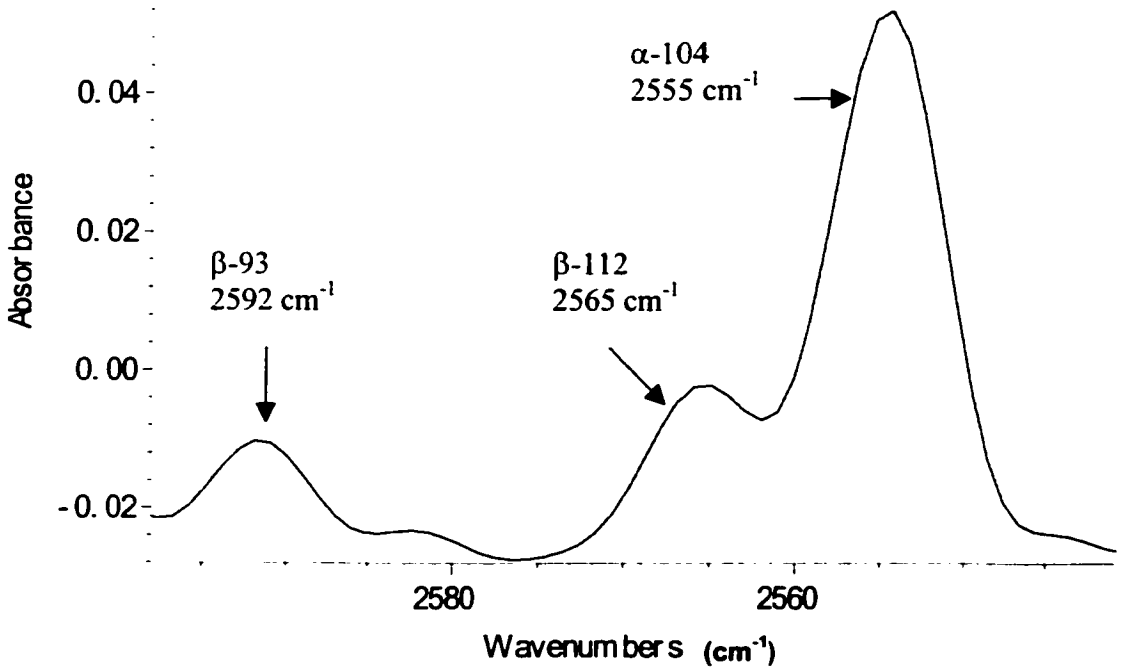

Figure 2.13 Deconvolved FT-IR spectrum of 1mM human met-hemoglobin in phosphate buffer (pH 6.9) containing 2mM DTT and 2mM DTPA (FSD bandwidth set to 14 with an enhancement factor of 1.9)

the peptide-binding groove. The C-C stretching band from the aromatic ring of Tyr appears as a well-resolved peak at 1515cm^{-1}, which has been shown to change peak maxima as the local environment is perturbed (Reinstadler *et al.*, 1996) and has been described as a sensitive marker of protein conformation (Fabian *et al.* 1994). The second PDZ of PTP1E contains a single Tyr residue (Tyr-36) located on β-sheet#3 (Ekiel *et al.*, 1998). Figure 2.14b shows the FT-IR spectrum of the KDDEVYYV peptide at room temperature in D_2O (pD 7.3), and Figure 2.14a the spectrum of tyrosine ethyl ester to demonstrate the contribution of the side-chain absorbance. The KDDEVYYV spectrum shows the Tyr band at 1516.2cm^{-1}. Tyrosine ethyl ester HCl shows the Tyr band at 1516.7cm^{-1}. A smaller peak that appears in both samples at 1614cm^{-1} has been previously described as an unassigned ring vibration of Tyr (Ramelow *et al.*, 1998). The

Tyr band appears at 1515.6cm^{-1} in the PDZ-KDDEVYYV mixture in Figure 2.15. The interpretation of this red shift is difficult due to interference from Tyr-36 in free PDZ, which showed a peak at 1514.7cm^{-1}. Table 2.3 summarizes the peak position of the Tyr band in PDZ and PDZ-peptide combinations. The Tyr band does not change position in PDZ or any peptide combinations except with KDDEVYYV. However, this is a likely result of overlapping peptide (1516.2cm^{-1}) and free PDZ (1514.7cm^{-1}) signals. Unfortunately, it was not possible to resolve the multiple Tyr residues since the frequency shifts are not as significant as with Cys residues (Figure 2.13).

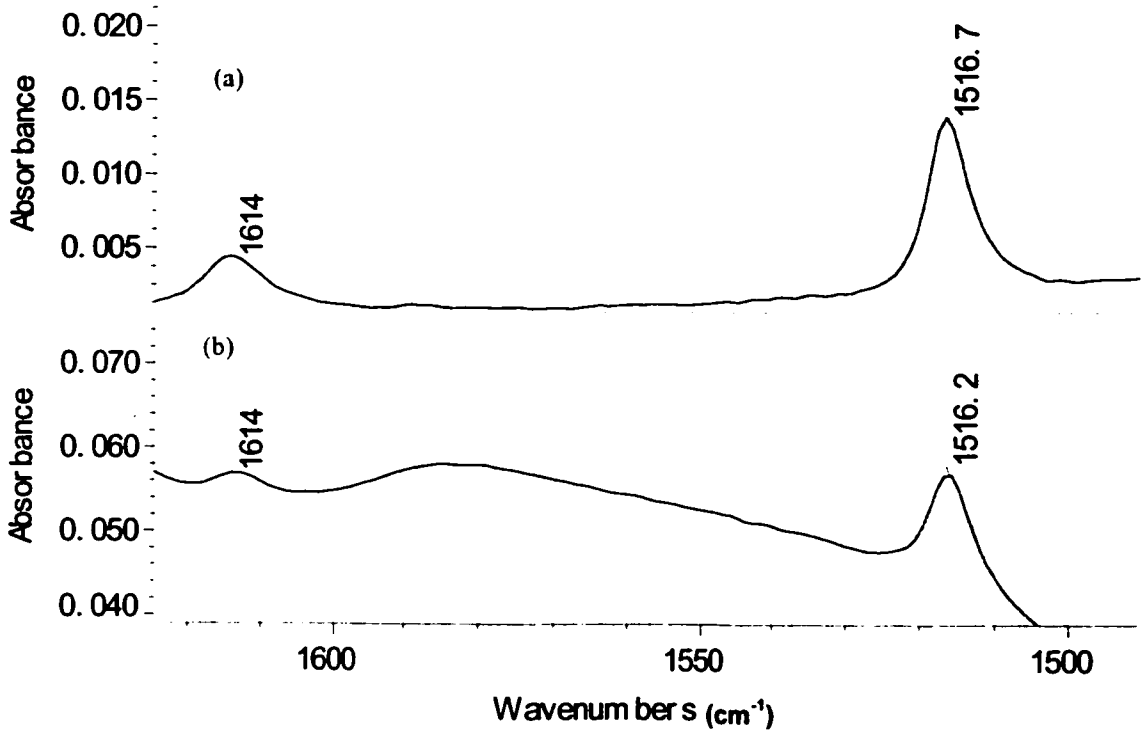

Figure 2.14 (a) FT-IR spectrum of 2.4 mM tyrosine ethyl ester in D_2O (pD 7.3); (b) 2.4mM KDDEVYYV peptide in D_2O (pD 7.3)

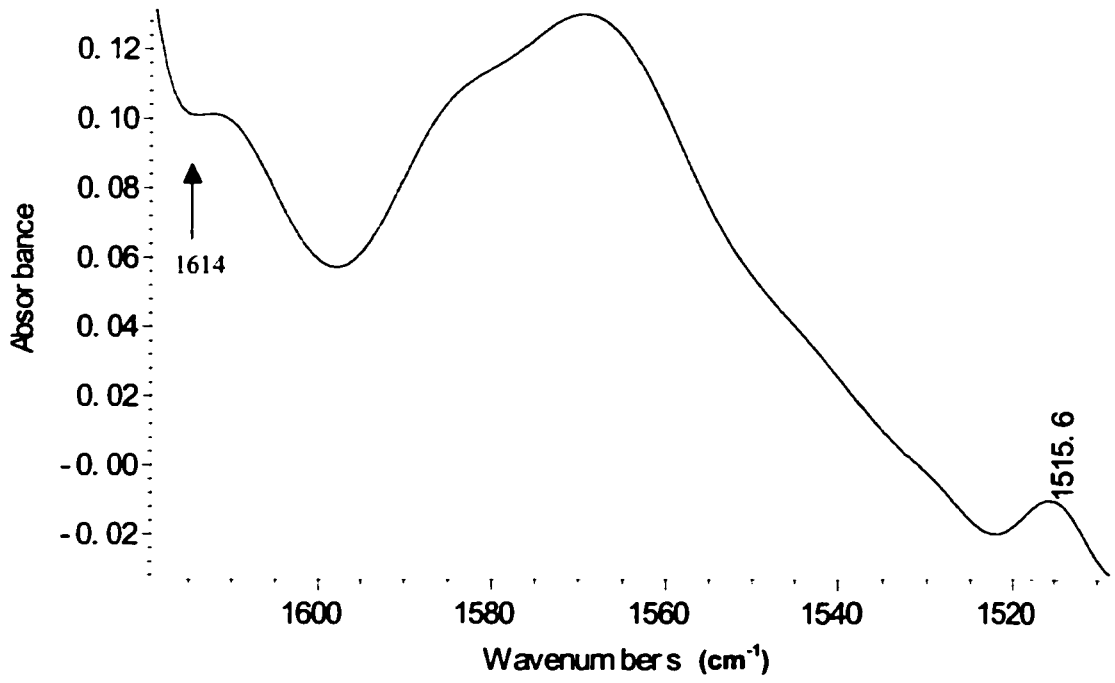

Figure 2.15 FT-IR Spectrum of 2.0mM PDZ with 2.4mM tyrosine peptide in D_2O pD 7.3 (FSD bandwidth set to 15 with an enhancement factor of 1.5)

Table 2.3 Comparison of the Tyr band position for six different Tyr containing samples

	Tyr ethyl ester	KDDEVYYV peptide	Free PDZ	PDZ + KDDEVYYV	PDZ + ENEQVSAV	PDZ + ENEQVCAV
Frequency of Tyr band (cm^{-1})	1516.7	1516.2	1514.7	1515.6	1514.3	1514.6

CD Spectroscopy

Thermal denaturation was also monitored using CD spectroscopy since it is more specific towards the detection of helical structures in proteins as compared to FT-IR spectroscopy (Surewicz *et al.*, 1993). The CD spectra of the free PDZ and PDZ-

ENEQVSAV complex at increasing temperatures are presented in Figures 2.16 and 2.17, respectively. In order to simplify the interpretation of the unfolding transition, only a selected number of spectra are displayed. The CD spectrum of the ENEQVSAV peptide is also presented in Figure 2.17 and shows very little absorbance compared to PDZ.

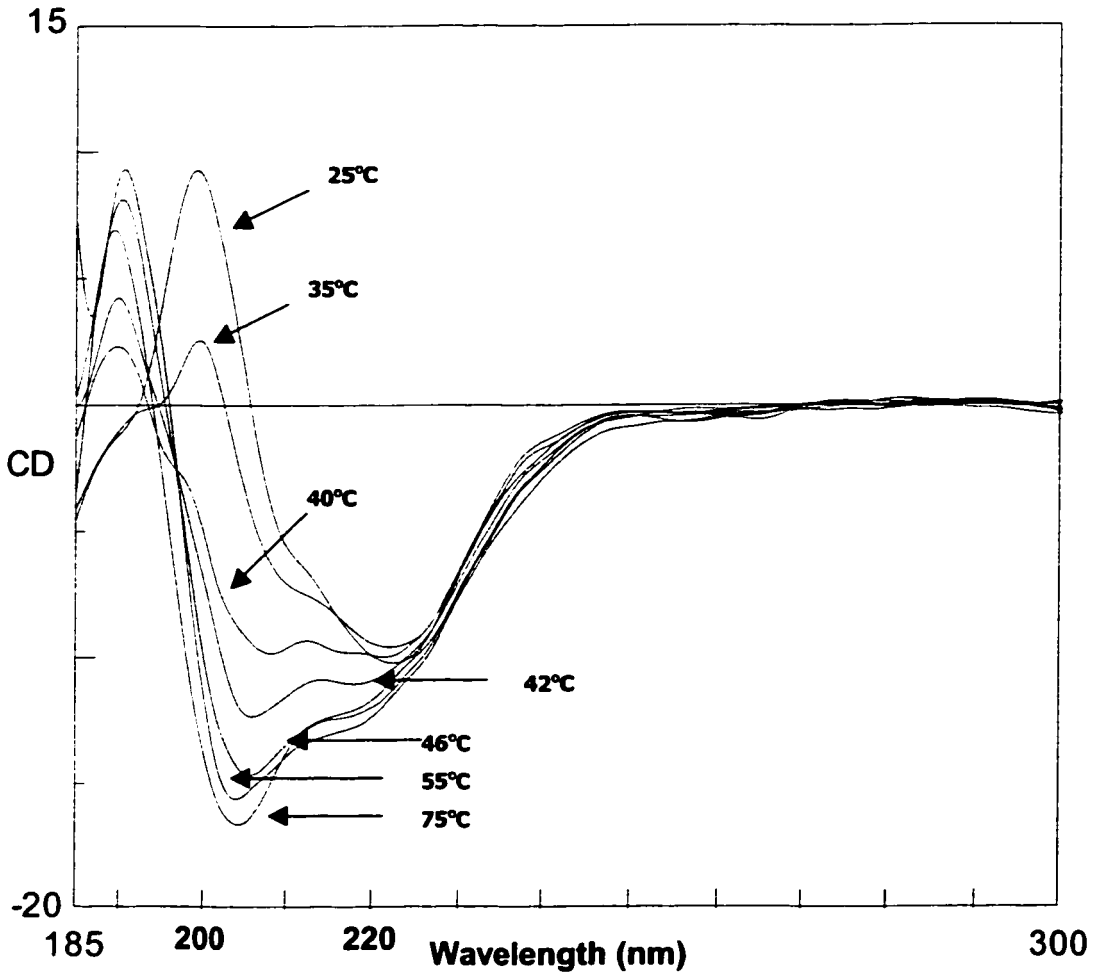

Figure 2.16 CD spectra of 1mM PDZ in 50mM phosphate buffer (pH 6.9) at increasing temperatures

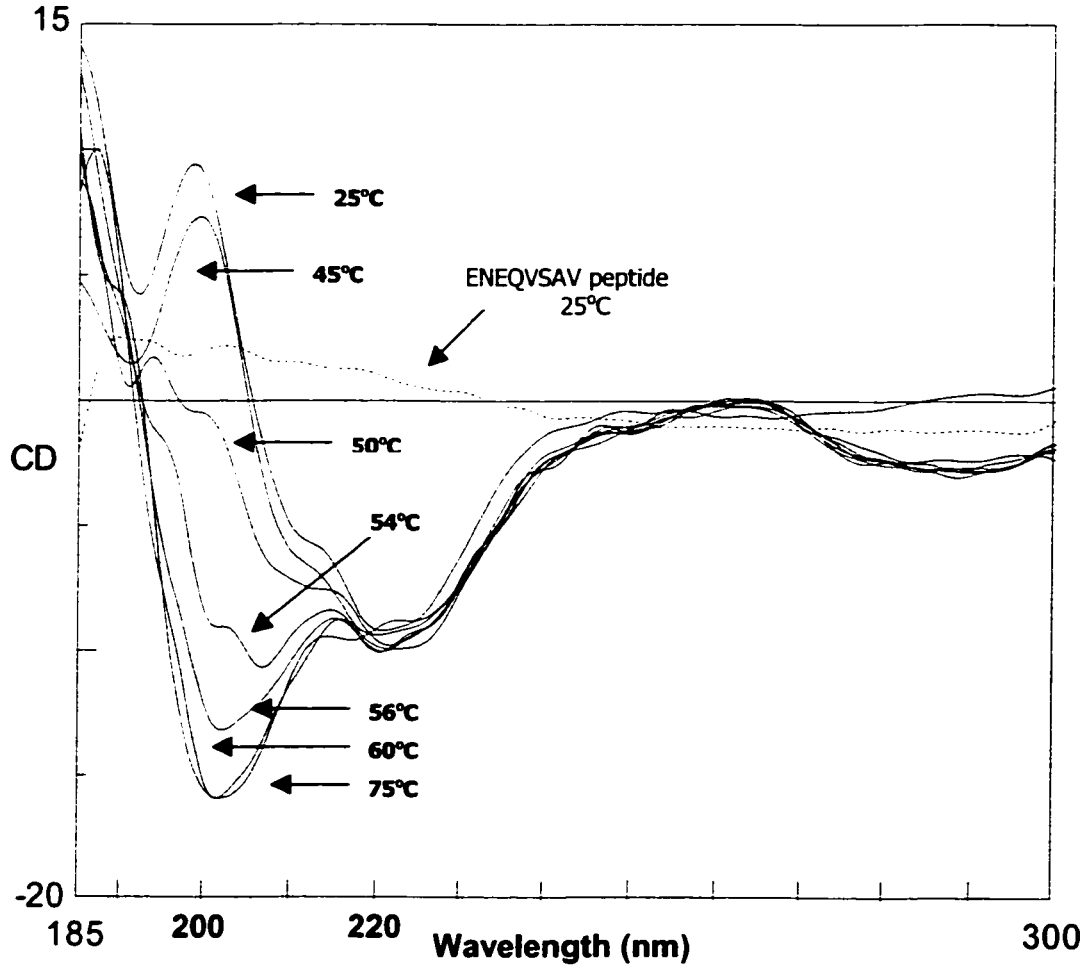

Figure 2.17 CD spectra of 1mM PDZ with 1.2mM ENEQVSAV in 50mM phosphate buffer (pH 6.9) at increasing temperatures

The CD thermal plots of the free and peptide-bound PDZ show similar features, based on the assignments described in Section 1.4.1. The native CD spectrum (25°C) is dominated by a large positive band at 200nm (β-sheets) and shows the double negative helical absorption pattern at 208 and 222nm. The β-sheet band of the PDZ-ENEQVSAV combination appears larger and broader than free the PDZ which may be due to the binding of the peptide as an anti-parallel strand. Even though PDZ consists mainly of β-

sheets (six in total) and only two α–helices, it is common that helical bands dominate the CD spectra (Woody, 1996). As the temperature increased, the positive absorption at 199nm, which is indicative of β-sheets, gradually became an intense negative signal centered at 200nm indicating unfolding to unordered structures. The intensity and position of the helical absorption at 222nm remained essentially unchanged over the temperature range observed indicating that helical structures persist beyond 75°C. By following the 200nm spectral band, it was observed that unfolding occurs at higher temperatures in the PDZ-ENEQVSAV complex as compared to free PDZ. Not enough data points were collected to obtain an accurate measurement of T_m, but CD spectra support the conclusions derived from the FT-IR data summarized in Table 2.2. The CD data shows that the stabilizing effect of the peptide can be localized to the β-sheet region of the protein, which is its expected binding site.

2D Analysis of thermal denaturation

A generalized two-dimensional analysis of spectral data was developed by Noda, (1993) and has been successfully applied to the analysis of numerous types of spectral data. 2D analysis involves monitoring in-phase and out-of-phase changes that occur due to the application of a perturbation, which can include temperature, pH or a denaturant. The benefit of 2D analysis is the resolution of over-lapping spectral bands, facilitated band assignments and information regarding the sequence of events (Ozaki and Noda, 1996).

The mathematical treatment of the spectra involves two equations that are used to create a synchronous (in-phase) contour plot and an asynchronous (out-of-phase) contour

plot. The synchronous plot describes the spectral bands that are changing in-phase as a result of the perturbation. The change can be inversely related, such as the decrease in the intensity of α–helix bands as random-coil bands increase. The synchronous contour plot therefore contains two types of peaks; auto-peaks, which appear along the diagonal and which indicate that a peak is in-phase with itself (always positive) and cross-peaks which correlate in-phase relationships between two different peaks in the perturbed spectra (Noda, 1993).

The asynchronous plot describes the spectral bands that are changing out-of-phase as a result of the perturbation. The asynchronous contour plot therefore contains only cross-peaks which indicate an out-of-phase relation between two different peaks in the perturbed spectra. These can include peaks arising from chemically unrelated structures (Ozaki and Noda, 1996). The sign of the peaks in both types of contour plots can be used to determine the sequence of events and determines if a peak changes intensity before or after another. The use of a correlation table is helpful in summarizing the peak relationships found in both types of contour plots.

The synchronous 2D contour map for the PDZ domain at temperatures of 27-59°C is shown in Figure 2.18. The band assignments are based on the 1D stacked spectra (Figure 2.5). Note that all 2D observations refer to the region of the contour plot above the diagonal. The synchronous map shows an auto-peak along the diagonal at 1618cm^{-1} and a broad auto-peak over the β-sheet range 1624-1640cm^{-1}. The presence of these auto-peaks indicates that these are the principal FT-IR bands that change with increasing temperatures as seen in the 1D spectra (Figure 2.5). Cross-peaks in the synchronous map appear for the following combinations of peaks. A positive cross-peak appears at

1618/1684 cm^{-1} which implies that both aggregation bands increase in intensity with increasing temperatures. A negative cross-peak at 1618/1675cm^{-1} indicates that as the aggregation band increases the turn-like structures decrease. A broad negative cross-peak is observed between the 1618cm^{-1} and the β-sheet range (1624-1640cm^{-1}) indicating that as aggregation increases, the β-sheets are decreasing in intensity. A similar negative cross-peak is observed with 1684cm^{-1} and both β-sheet bands at 1628 and 1639cm^{-1} which also indicates that as aggregation increases, β-sheets are decreasing. A small positive cross-peak is observed at 1628/1675cm^{-1} which suggests that the intensity of the lower frequency β-sheet band and turn-like structures are changing in the same direction.

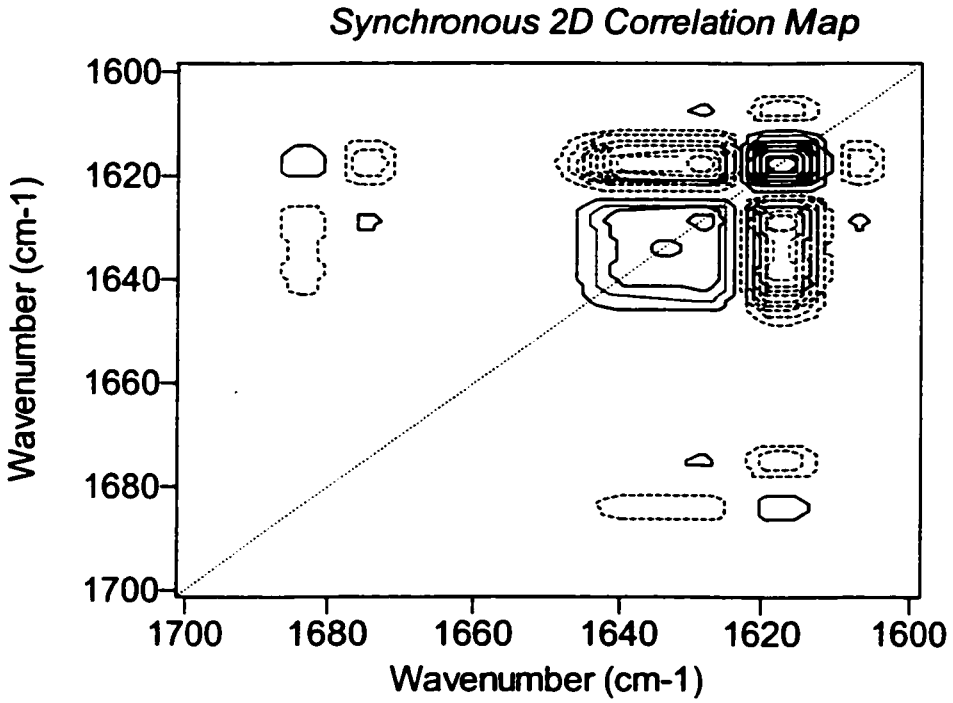

Figure 2.18 Synchronous map of the denaturation of free PDZ over the temperature range of 27-59°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

Based on observations in the 1D spectra (Figure 2.5), both β-sheet and turn bands are therefore decreasing as temperature increases.

The asynchronous 2D contour map for the PDZ domain at temperatures of 27-59°C is shown in Figure 2.19. Cross-peaks in the asynchronous map appear for multiple combinations of peaks. Each combination indicates that there is an out-of-phase relationship between two peaks and intensity changes occur, either at lower or higher temperatures. A broad, negative cross-peak occurs at $1618cm^{-1}$ and the β-sheet range $1624\text{-}1640cm^{-1}$. The observation of a negative cross-peak in the same region in the synchronous plot indicates that the β-sheets ($1624\text{-}1640cm^{-1}$) decrease before aggregation appears ($1618cm^{-1}$). A negative cross-peak occurs at $1618cm^{-1}$ and $1675cm^{-1}$. The observation of a negative cross-peak in the same region in the synchronous plot, indicates that the β-turns ($1675cm^{-1}$) decrease before aggregation appears ($1618cm^{-1}$). Interestingly, a positive cross-peak appears at 1628/1684 cm^{-1}, since this is a negative peak in the synchronous plot, this indicates that this β-sheet band ($1628cm^{-1}$) decreases after the high frequency component of aggregation ($1684cm^{-1}$) increases. However, since the band at $1618cm^{-1}$ is the principle indicator of aggregation, as compared to $1684cm^{-1}$, β-sheets are assumed to unfold prior to aggregation. Other peaks are observed in the asynchronous contour plot; however, since they are not observed in the synchronous plot, it is not possible to assign any temporal relationships (Noda, 1990). Table 2.4 summarizes the sequence of events that occurs during the unfolding of free PDZ.

2D correlation analysis was also used to monitor secondary structure changes that occur after the protein unfolds. Figure 2.20 contains the synchronous map for free PDZ

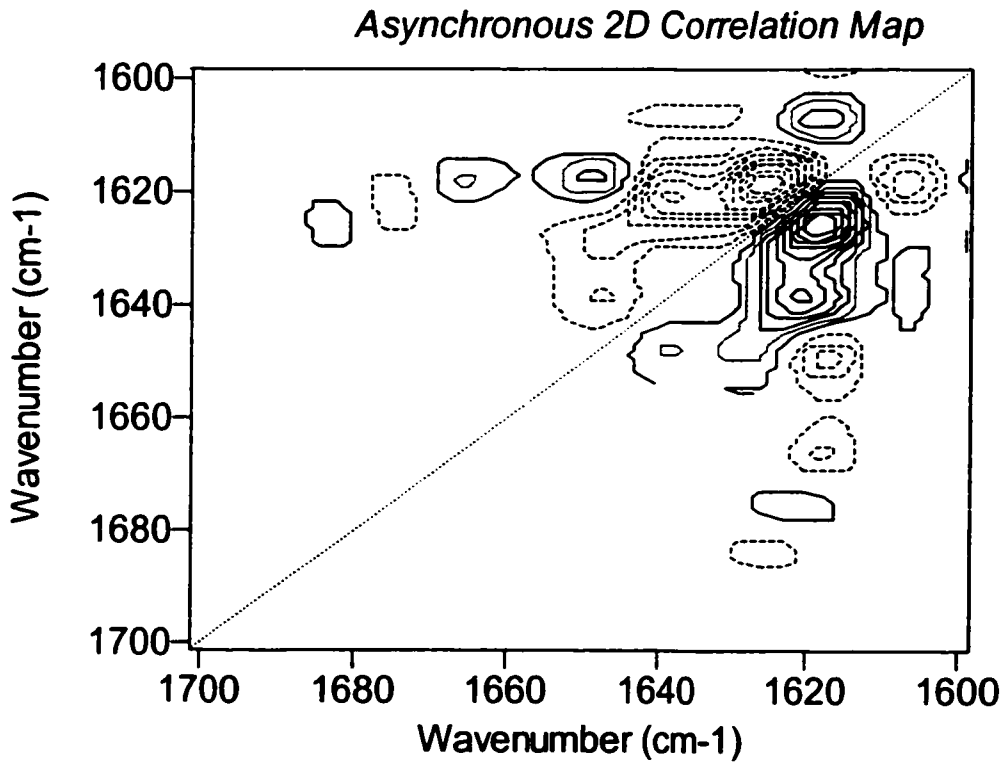

Figure 2.19 Asynchronous map of the denaturation of free PDZ over the temperature range of 27-59°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

Table 2.4 Correlation table summarizing the 2D observations for the PDZ domain at 27-59°C

	ν_1				
ν_2		1684cm^{-1}	1675cm^{-1}	1639cm^{-1}	1628cm^{-1}
	1618 cm^{-1}	+ / ~	− / − ←	− / − ←	− / − ←
	1628 cm^{-1}	+ / − →	+ / ~	+ / ~	
	1639 cm^{-1}	− / ~	~		
	1675 cm^{-1}	~ / ~			
	Sequence of events: β-sheets and β-turns are disrupted prior to aggregation				

Observations in the summary table are assigned as follows:

synchronous/asynchronous
+ positive correlation peak
- negative correlation peak
~ no peak observed

← ν_1 changes before ν_2
→ ν_1 changes after ν_2

at 65 to 89°C. Autopeaks appear at 1618, 1653 and 1684cm^{-1} indicating that the intensity of these peaks is affected as the temperature increases. 1618 and 1684cm^{-1} represent aggregation while 1653cm^{-1} corresponds to random, unordered structures. A broad negative cross-peak appears at 1618 and 1653cm^{-1} indicating that these are inversely related. Subsequently, a negative cross-peak is also found at 1653 and 1684cm^{-1}. A positive cross-peak appears at 1618 and 1684cm^{-1} which is expected since any changes in either aggregation band should be synchronized.

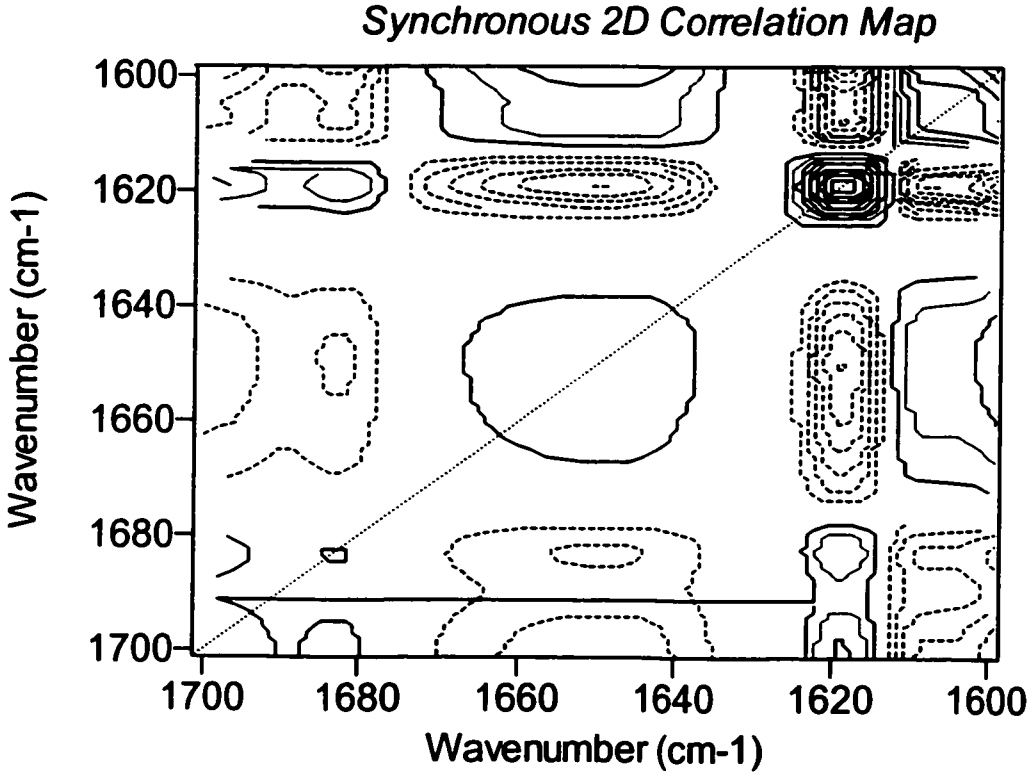

Figure 2.20 Synchronous map of the denaturation of free PDZ over the temperature range of 65-89°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

The corresponding asynchronous map in Figure 2.21 shows a negative cross-peak between 1618 and 1653cm^{-1}. Since this cross-peak was also negative in the synchronous map, 2D correlation rules indicate that the peak at 1618cm^{-1} changes before 1653cm^{-1}. Therefore, at high temperatures, intermolecular aggregation is disrupted and these strands form random, unordered structures.

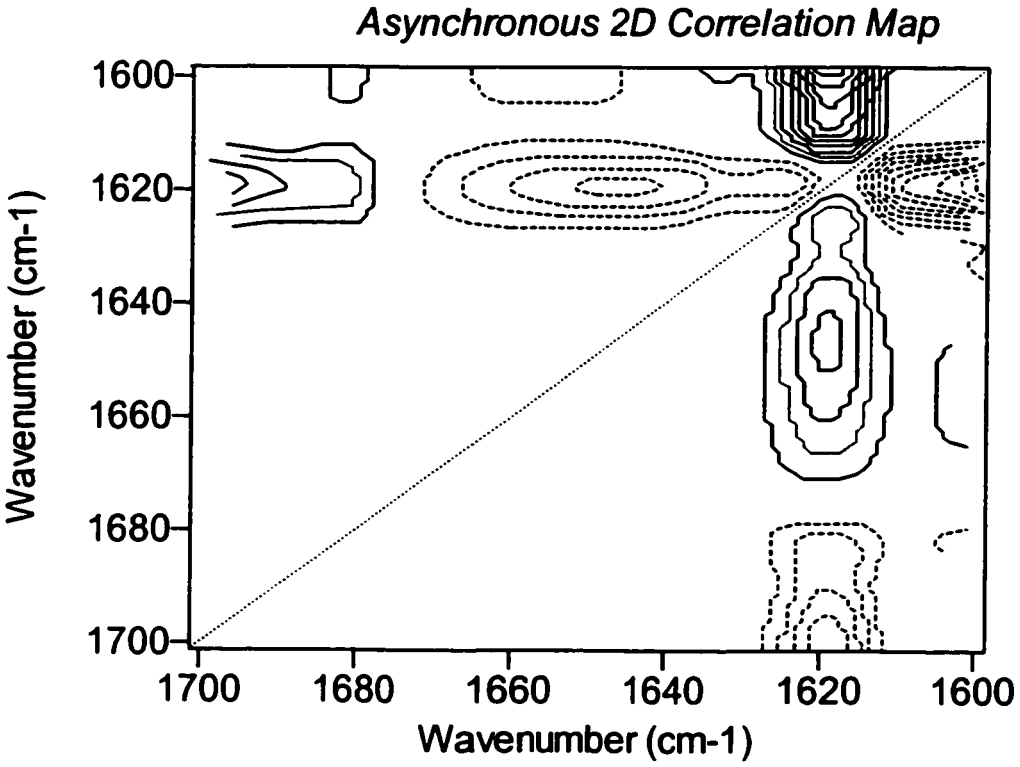

Figure 2.21 Asynchronous map of the denaturation of free PDZ over the temperature range of 65-89°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

The synchronous 2D contour map for the PDZ-ENEQVSAV complex at temperatures of 37-69°C is shown in Figure 2.22. The temperature range examined was shifted by +10°C to account for the increased T_m. The synchronous map is essentially identical to that of the free PDZ domain, and shows an auto-peak along the diagonal at

1618cm^{-1} and a broad auto-peak over the β-sheet range, 1624-1640cm^{-1}. Therefore, from observations in the 1D spectra (Figure 2.7), aggregation (1618cm^{-1}) increases and β-sheets (1624-1640cm^{-1}) decrease with increasing temperatures. A positive cross-peak appears at 1618/1684 cm^{-1} which implies that both aggregation bands increase in intensity with increasing temperatures. It is noteworthy, that the 1684cm^{-1} band was barely visible in the 1D stacked spectra (see Figure 2.7); therefore, the resolution of over-lapping bands using 2D analysis is evident. A negative cross-peak at 1618/1675cm^{-1} indicates that as the aggregation band increases the turn-like structures decrease. A broad negative cross-peak is observed between the 1618cm^{-1} and the β-sheet range (1624-1640cm^{-1}) indicating that as aggregation increases, the β-sheets are decreasing in intensity.

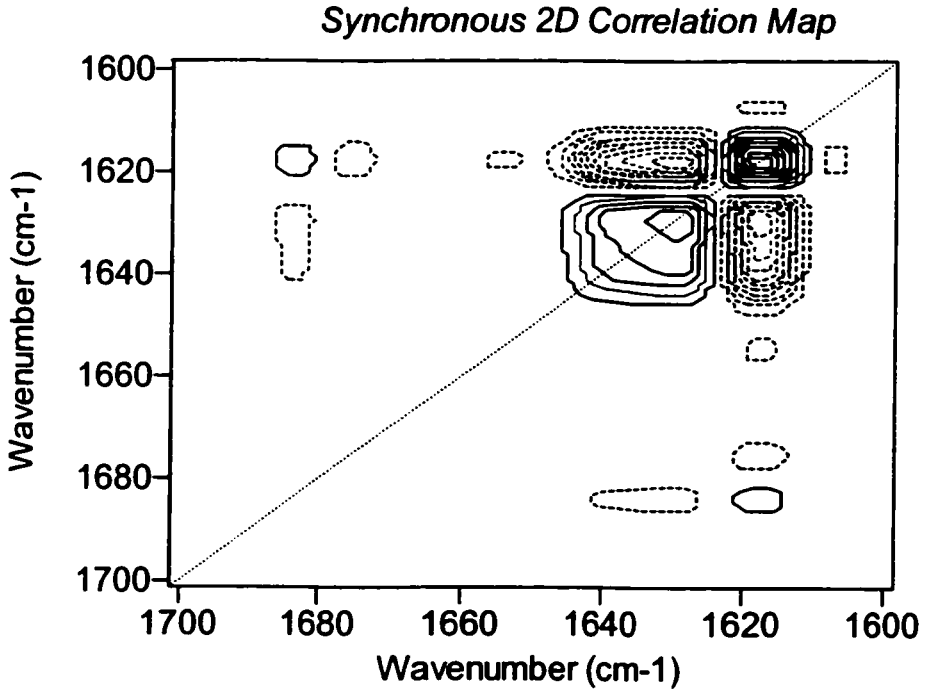

Figure 2.22 Synchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 37-69°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

A negative cross-peak is also observed between both β-sheet bands (1628, 1639cm^{-1}) and 1684cm^{-1} also indicating that as aggregation increases, β-sheets are decreasing.

The asynchronous 2D contour map for the PDZ-ENEQVSAV complex at temperatures of 37-69°C is shown in Figure 2.23. Cross-peaks in the asynchronous map appear for multiple combinations of peaks. A broad, negative cross-peak occurs between 1618cm^{-1} and the β-sheet range 1624-1640cm^{-1}. The observation of a negative cross-peak in the same region in the synchronous plot, indicates that the β-sheets (1624-1640cm^{-1}) decrease before aggregation appears (1618cm^{-1}). A positive cross-peak occurs

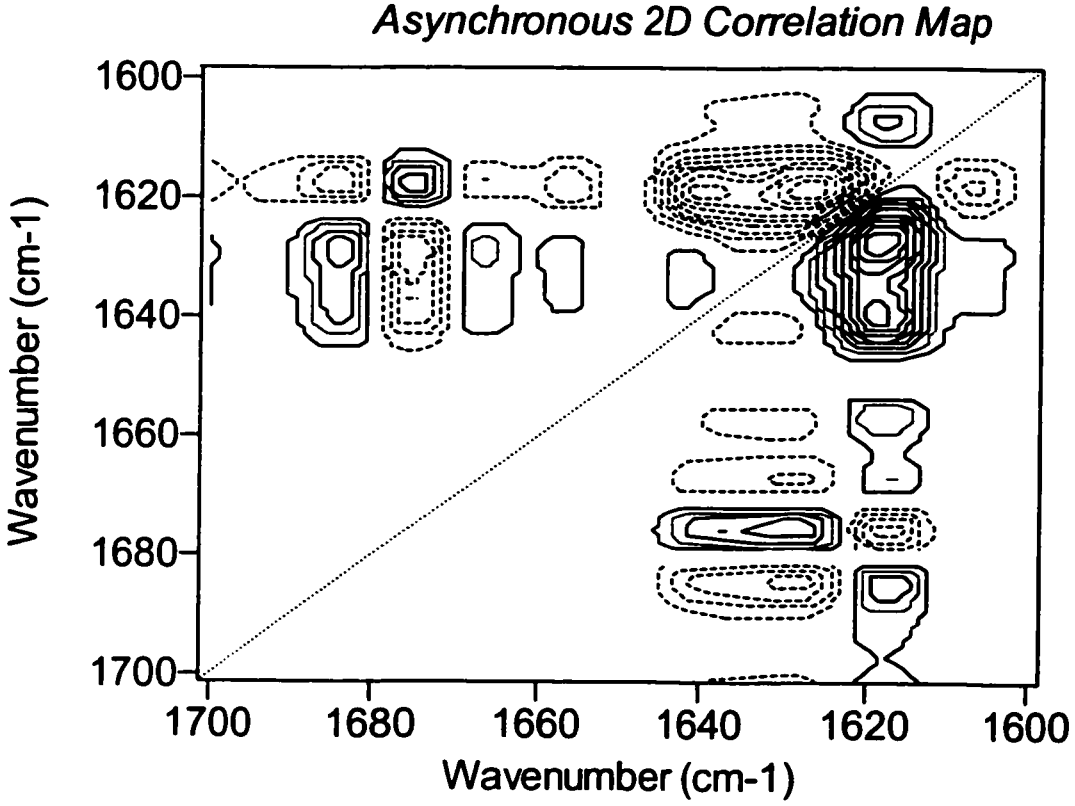

Figure 2.23 Asynchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 37-69°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

at 1618cm^{-1} and 1675cm^{-1}. The observation of a negative cross-peak in the same region in the synchronous plot, indicates that the β-turns (1675cm^{-1}) decrease after aggregation appears (1618cm^{-1}). Interestingly, there is a negative band between 1618 and 1684cm^{-1} indicating an unexplained, out-of-phase relation between both aggregation bands. Other peaks are observed in the asynchronous contour plot, however since they are not observed in the synchronous plot, it is not possible to assign any temporal relationships (Noda, 1990). Table 2.5 summarizes the sequence of events that occurs during the unfolding of the PDZ-ENEQVSAV complex.

Table 2.5 Correlation table summarizing the 2D observations for the PDZ-ENEQVSAV complex at 37-69°C

		ν_1			
		1684cm^{-1}	1675cm^{-1}	1639cm^{-1}	1628cm^{-1}
ν_2	1618 cm^{-1}	+/− →	−/+ →	−/− ←	−/− ←
	1628 cm^{-1}	−/+ →	~/−	+/~	
	1639 cm^{-1}	−/+ →	~/−		
	1675 cm^{-1}	~/~			
	Sequence of events: β-sheets unfold prior to aggregation. Turns are disrupted last.				

Observations in the summary table are assigned as follows:

synchronous/asynchronous
+ positive correlation peak
- negative correlation peak
~ no peak observed

← ν_1 changes before ν_2
→ ν_1 changes after ν_2

FT-IR spectra of the PDZ-ENEQVSAV complex at temperatures of 65 to 89°C were also analyzed by 2D in order to monitor secondary structure changes that occur after

the protein unfolded. The synchronous map in Figure 2.24 shows autopeaks at 1653 and 1618cm^{-1}. Negative cross-peaks appear between 1618 and 1653cm^{-1} and between 1653 and 1684cm^{-1}, indicating an inverse relationship between either aggregation band (1618cm^{-1}, 1684cm^{-1}) and unordered, random structures (1653cm^{-1}). The asynchronous map in Figure 2.25 shows a negative cross peak between 1618 and 1653cm^{-1}. Since this cross-peak is also negative in the synchronous map, 2D correlation rules dictate that aggregation (1618cm^{-1}) is disrupted before increases in unordered structures (1653cm^{-1}) occur.

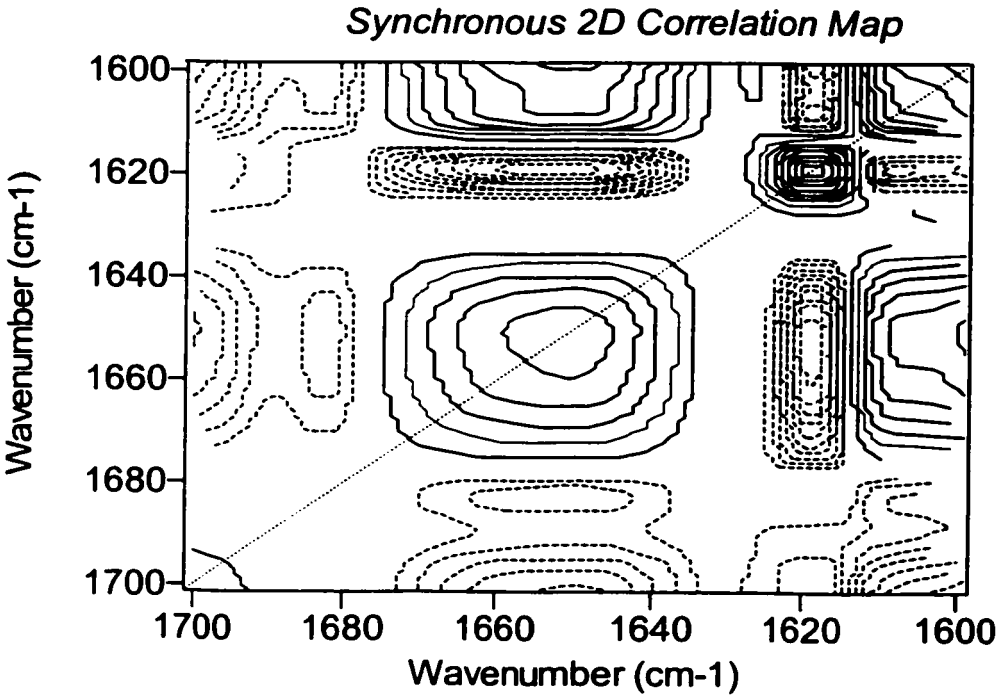

Figure 2.24 Synchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 65-89°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

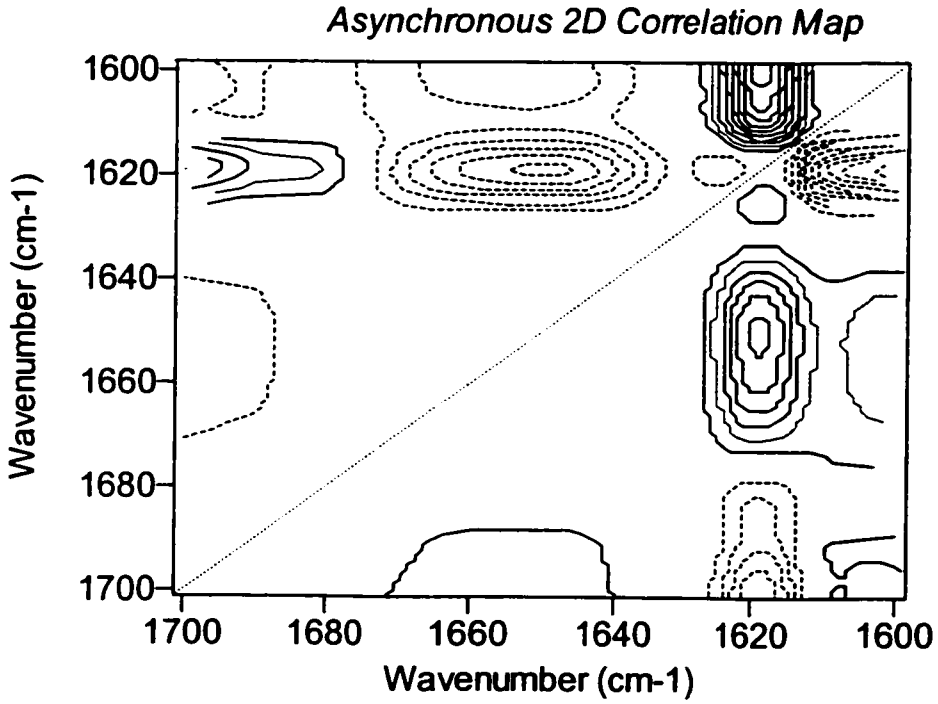

Figure 2.25 Asynchronous map of the denaturation of the PDZ-ENEQVSAV over the temperature range of 65-89°C (FSD bandwidth set to 15 with an enhancement factor of 1.5) Solid lines indicate positive peaks, dashed lines indicate negative peaks

2.5 Discussion

Curve-fitting and thermal denaturation

The second PDZ of hPTP1E recognizes peptides containing the C-terminal motif X-Ser/Thr-X-Val (Kozlov *et al.*, 2000). Three peptides each containing eight residues were selected in order to study the effect of substituting the -2 position from Ser to Cys or Tyr. The peptides sequences were ENEQVSAV, ENEQVCAV and KDDEVYYV. Although the remaining residues in the Tyr-containing peptide were not identical, its use was justified based on previous literature reports. The side chain of residues in the -1 position point out into solution, therefore substitutions at this position show very little effect on binding affinity (Fuh *et al.*, 2000). All peptides contained a C-terminal Val and a Val in the -3 position. The side chains of residues at -4 and beyond do not have

significant influence on specificity (Sheng and Sala, 2001) although reports suggest that interactions may occur with the amide nitrogens up to position –6 (Kozlov *et al.*, 2000).

Curve fitting of the FT-IR spectra detected an interaction within the β-sheet region of the protein, which is the expected binding target for the incoming peptide. The relative increase in the 1628cm^{-1} band can be attributed to the formation of an additional anti-parallel β-strand when the peptide is bound. However, the spectral resolution of the curve-fitted spectra was not sufficient to detect individual bonds of interest such as the hydrogen bond between Ser and His-71. Therefore Tyr and Cys residues were substituted in the –2 position since, their side-chains generate well-resolved FT-IR signals. His-71 was not observed in the FT-IR spectra since it produces only a very weak band at 1596cm^{-1} (Venyaminov and Kalnin, 1990) Careful subtraction of the spectra that would be required, was not possible using the current experimental setup.

Peptides with Tyr in the -2 position have been reported to bind to Class II PDZ domains based on hydrophobic interactions (Sheng and Sala, 2001). It was evaluated here for possible hydrogen bonding to His-71 since it has a lower pK_a than Ser. Also, Tyr is sensitive to its micro-environment and has been used to monitor changes in both secondary and tertiary structure of proteins (Fabian *et al.*, 1994) (Reinstadler *et al.*, 1996). The characteristic 1515cm^{-1} band of Tyr, due to C-C stretching in the phenol ring, can shift from 1514.5 to 1515.4 cm^{-1} as it becomes exposed to the surrounding solvent (Reinstadler *et al.*, 1996). It has been previously shown to shift as far as 1512cm^{-1} in aggregated concanavalin A due to interactions with nearby charged groups (Arrondo *et al.*, 1988). No FT-IR evidence for binding of KDDEVYYV with PDZ was obtained, in

part due to the interference from Tyr-36 and the inability of the spectrometer to resolve each of the Tyr signals.

Cys was selected in the -2 position of the peptide as a possible probe for hydrogen bonding to His-71. Hydrogen bonding of thiols to nucleophilic bases decreases the SH stretching frequency and the integrated area of the SH band increases due to a greater change in the dipole moment of the S-H bond (Moh *et al.*, 1987). Studying the SH stretching region (2500-2600cm^{-1}) did not show any peaks from the PDZ-ENEQVCAV complex. This does not necessarily rule out binding between Cys and His-71, since random interactions can cause heterogenous broadening of the SH peak (Bare *et al.*, 1975). The SH stretching region of human met-hemoglobin (Figure 2.13) demonstrates the ability to detect subtle changes in the local environment of cysteine residues. The intense ν(SH) band of human hemoglobin at 2555cm^{-1} can be attribute to the hydrogen bonding of Cys α-104 with the backbone carbonyl of Leu-100, while no hydrogen bonding partner exists for Cys β-93 (Moh *et al.* 1987).

The most striking difference between the stacked thermal spectra of free PDZ and those of the ENEQSVAV-PDZ complex is the resistance to thermal denaturation with the ENEQVSAV-PDZ complex. The increased resistance of the peptide-protein complex to thermal denaturation can be attributed to a binding interaction (Pace and McGrath, 1980), (Surewicz *et al.*, 1990). In fact, indirect methods that detect protein-ligand binding by measuring a change in conformational stability of the protein have been reported as early as 1890 by O'Sullivan and Thompson (Pace and McGrath, 1980). This approach has been recently automated and applied to screen large numbers of proteins by a high-throughput method called SUPREX (stability of unpurified proteins

from rates of H/D exchange) (Ghaemmaghami *et al.*, 2000). From a thermodynamic perspective, ligand binding favors the native state of the folded protein and therefore shifts the equilibrium of unfolding towards the folded state (Pace and McGrath, 1980). The observed T_m for the combinations of the PDZ domain with alternate peptides ENEQVCAV and KDDEVYYV showed only increases within the experimental error (1 to 2°C) above that of the free protein. For proteins that undergo reversible unfolding with temperature, the effect of ligand binding on the temperature at which unfolding occurs is related by equation 1 (Pace and McGrath, 1980):

$$(1) \qquad \Delta T_m = \frac{T_m T_o R}{\Delta H_D} \ln(1+K_B[S])$$

Where T_m is for the protein-ligand complex, $T_0=T_m$ of free protein, ΔH_D=enthalpy of denaturation of the protein-ligand complex, and R=Rydberg gas constant. The observed ΔT will increase as the concentration of free ligand [S] is increased or as the binding constant (K_B) becomes stronger. The negligible stabilizing effect of the peptides ENEQVCAV and KDDEVYYV at 1.2mM can therefore be attributed to weak binding constants. Analysis of the T_m values is shown in Appendix B, however, the K_D's reported should only be used as estimates since the unfolding was irreversible. Neither the free PDZ nor any PDZ-peptide combinations showed reversible thermal unfolding. Upon cooling back to 25°C the amide I′ band did not return to its original shape and evidence of aggregation remained.

In addition to comparing the T_m for free PDZ and its peptide complexes, the thermal denaturation can be used to help make structural assignments to the overlapping peaks in the FT-IR spectrum. α–Helices and unordered structures, whose peak maxima

are often too close to assign with certainty (Surewicz *et al.*, 1993), (Dong *et al.*, 1990), may be differentiated in the denatured spectra since the α-helix peak should begin to disappear as the protein unfolds, while the unordered bands would likely increase in intensity. While the stacked spectra do not clearly show the loss of helical features, the denatured spectra are dominated by a broad peak centered at 1645-1650cm^{-1} which is indicative of unordered structures. PDZ-ENEQVSAV shows some residual helical absorption at 1661cm^{-1} however this was not resolved by 2D analysis. Peptide binding involves α-helix #2, through His-71, and may prevent this helix from unfolding at high temperatures. However, examination of the CD spectra (Figures 2.16 and 2.17) reveals that helical absorption persisted for the free and peptide bound PDZ up to 75°C. This may be due to partial unfolding followed by aggregation which could trap helical structures before they unfolded. Therefore, the thermally denatured spectra may not truly represent the unfolded protein since the concentration used (1mM) may have promoted protein aggregation.

2D Analysis

The 2D correlation analysis provides a more powerful approach for FT-IR data interpretation. The ability to resolve overlapping peaks not observed in the 1D spectrum was demonstrated by the appearance of a cross-peak at 1684cm^{-1} (Figures 2.18, 2.22 and 2.23). This represents the high frequency component of the aggregation bands. In addition, the strong correlation between 1684 and 1618cm^{-1} further reinforces the assignment of both of these bands to intermolecular aggregation.

The further benefit of 2D analysis was evident through the assignment of unfolding events (as shown in Tables 2.4 and 2.5). β-Sheets unfolded and β-turns were

disrupted prior to aggregation of free PDZ. Therefore, the denatured state of free PDZ must consist of mainly unordered structures, intermolecular aggregated strands and perhaps some residual helical structures. This was observed in the 1D temperature plot (Figure 2.3) which shows a featureless Amide 'I region at temperatures beyond the T_m. Interestingly, the PDZ-ENEQVSAV complex showed a different unfolding sequence: β-sheets unfold prior to aggregation, while turns were disrupted last. Since helical bands were not resolved by the 2D analysis in either free PDZ or the PDZ-ENEQVSAV complex , it suggests that these structures did not unfold over the temperature range of 25 - 89°C, perhaps due to aggregation.

2.6 Conclusions

The combination of spectral data obtained by FT-IR and CD confirms a previously observed (Ekiel *et al.*, 1998) interaction between the target peptide and the PDZ domain. X-ray experiments report only a small change in the PDZ conformation upon binding of a target peptide (Doyle *et al.*, 1996), making detection by lower resolution spectroscopic methods more difficult. However, the combination of 2D correlation analysis with temperature perturbation as well as curve-fitting, enabled the observation of a peptide binding interaction. Peptide binding is characterized by an increase in the characteristic frequency of a short β-structure (Fabian *et al.*, 1993), while 2D analysis demonstrated the ability to differentiate the thermal unfolding events between free PDZ and a stable PDZ-peptide complex. Measurement of T_m was most diagnostic of peptide binding.

CHAPTER 3 - PROBING PEPTIDE BINDING TO THE SECOND PDZ DOMAIN OF hPTP1E USING ELECTROSPRAY MASS SPECTROMETRY

Abstract

The binding of three different peptides to the isolated recombinant second PDZ domain of hPTP1E was monitored using electrospray mass spectrometry. Peaks arising from the formation of a 1:1 complex were observed with the combination of the peptide ENEQVSAV and the PDZ domain. Combinations of PDZ with the peptides ENEQVCAV and KDDEVYYV did not show any substantial evidence of complex formation. The detection of the PDZ-ENEQVSAV complex by the mass analyzer was found to be dependent on the needle voltage and heated capillary temperature. Evidence for the formation of a stable complex between ENEQVSAV and the PDZ domain was further probed using H/D exchange. The rate of exchange was slower in the ENEQVSAV-PDZ complex as compared to free PDZ. Also, the total number of exchanged hydrogens in the PDZ-peptide complex was less than for the free PDZ domain. The results are consistent with the formation of a stable complex that shows resistance to solvent penetration and subsequent H/D exchange. This study demonstrates the usefulness of electrospray mass spectrometry for studying non-covalent interactions of proteins.

3.1 Introduction

Mass spectrometry has become a highly utilized instrumental technique for studying proteins. It is routinely used in proteomics, to fulfill the need for rapid identification of large numbers of proteins and for detecting post-translational modifications, such as phosphorylation (Moseley, 1990). This rapid expansion into protein science is due to the development of new methods of ionization, namely, matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). Both of these techniques are capable of forming intact molecular ions of biomolecules such as DNA, carbohydrates and proteins.

The advent of electrospray mass spectrometry also permits the study of proteins and their non-covalent complexes. The 'gentle' mechanism of electrospray ionization is well suited to proteins and non-covalent complexes whose three-dimensional structure is of limited stability. Low temperature and low ionization energy have been shown to preserve the intact structure of non-covalent complexes for detection by the mass analyzer (Ganguly *et al.*, 1992); (Baca and Kent, 1992). Hydrogen bonds and the hydrophobic and ionic interactions that hold weak protein complexes together are maintained as the protein is transferred to the gas-phase (Loo, 1995). Electrospray has been successfully applied to the of study protein binding interactions, such as binding of model peptides to the SH2 domain (Anderegg and Wagner, 1995) or binding of the *ras* protein to GDP or GTP (Huang *et al.*, 1993). Electrospray mass spectrometry is an invaluable tool for studying proteins since it can detect picogram quantities, it can identify large biomolecules through accurate molecular weight determination, and it provides information about the protein conformation from the observed charge state

distribution. Most proteins contain numerous basic amino acid residues and as a result form multiply charged molecules whose mass spectra contain multiple peaks at specific mass-to-charge ratios. The distribution of these peaks is dependent on the conformation of the protein, since unfolding may expose additional basic amino acids, which results in the observation of additional charge-states (Katta and Chait, 1991); (Konermann and Douglas, 1997).

Mass spectrometry is used here to probe the formation of a PDZ-peptide complex. The second PDZ domain of hPTP1E is known to bind to the C-terminus of the fas receptor or short peptides (8-15 residues) containing X-Ser/Thr-X-Val in the C-terminal sequences (Kozlov *et al.*, 2000), (Ekiel *et al.*, 1998). The last four or five amide-linkages of the incoming peptide form hydrogen bonds to a β-sheet in an anti-parallel orientation (Kozlov *et al.*, 2000). An additional hydrogen bond forms between the hydroxyl group of Ser at the –2 position in the peptide and the imidazole nitrogen of His-71 of the PDZ domain (Ekiel *et al.*, 1998). The binding interaction between the second PDZ of hPTP1E and the peptide ENEQVSAV was evaluated by direct observation of complex in the mass spectrum. The non-covalent complex that formed as a result of peptide binding remained intact during the ionization process and was detected by the mass analyzer. Combinations with alternate peptides ENEQVCAV and KDDEVYYV did not show any clear evidence of complex formation.

Additional work using H/D exchange was used to evaluate the stability of the PDZ-ENEQVSAV complex. Deuterium is useful as an isotopic label since incorporation of deuterons into the amide backbone can be measured by mass spectrometry. The rate of H/D exchange in proteins is sensitive to changes that may affect solvent exposure or

conformational stability of the protein (Wang and Tang, 1996). A reduction in H/D exchange rate has been previously shown to detect binding between a SH2 domain and model peptides (Anderegg and Wagner, 1995).

3.2 Experimental Procedures

3.2.1 Materials

The recombinant isolated second PDZ domain of hPTP1E as well as the peptides ENEQVSAV, ENEQVCAV and KDDEVYYV were obtained from Dr. Irena Ekiel (Biotechnology Research Insitute, Montreal, Canada). The peptide KDDEVYYV was obtained in crude form and required purification by HPLC on an HP1100 (Hewlett Packard) with a reversed-phase C18 column 150x4.6mm (Waters). The PDZ domain and the two other peptides had been purified as described previously (Ekiel *et al.*, 1998). HPLC grade acetonitrile was obtained from EM Merck, 99.9% ^2H_2O was obtained from Aldrich, 99% TFA was obtained from Lancaster and Nanopure 18.2 MΩ water was obtained from a Millipore system. DTT and DTPA were purchased from Sigma and ICN, respectively. Ammonium acetate and sodium phosphate were obtained from Aldrich and Fluka, respectively.

3.2.2 Methods

All mass spectra were obtained using a Thermo-Finnigan LcQ Deca Ion Trap equipped with an electrospray ionization source. Mass spectrometry was used to confirm the identity of the three peptides used for binding studies. In order to facilitate the

formation of positive ions in the sample solutions, peptide samples were dissolved in 40% ACN containing 0.1% TFA. Equimolar DTT and DTPA were added to the ENEQVCAV peptide to ensure a reduced cysteine sulfhydryl. Peptide solutions were infused using a Hamilton syringe directly into the ESI interface at 10μL/min (1mM). For each peptide, a tune file was generated to optimize the mass spectrometer parameters, and the peptide mass spectra were recorded and averaged over 60s in positive centroid mode to obtain the $[M+H]^+$ peak.

Electrospray ionization was used in positive mode to generate a mass spectrum for PDZ and related complexes. Both the normal-mass range (50-2000m/z) and the high-mass range (100-4000m/z) of the instrument were used. The PDZ domain was exchanged by ultrafiltration with 50mM ammonium acetate at pH 6.9 using Centricon-3 tubes (Millipore). The free PDZ domain was diluted to 55μM with the appropriate buffer and infused directly into the sample inlet at 10μL/min using a 100μL Hamilton syringe. The PDZ and peptide complexes were prepared by adding equal volumes of 0.2mM peptide and 0.1 mM protein to give final concentrations of 100μM peptide and 50μM PDZ. These samples were also infused directly into the sample inlet at 10μL/min. The PDZ domain and the PDZ-peptide samples were not combined with any organic solvents since this would likely alter binding affinities. M_r calculations were preformed using Xcalibur Bioworks software (ThermoFinnigan).

The H/D exchange was performed by concentrating a 200μL aliquot of 0.05mM PDZ domain in H_2O to 5μL using a rotary Speedvac (Savant Instruments) and then diluting the same to 200μL with D_2O at 4°C. Aliquots (15μL) were removed at 1, 2, 5, 7.5, 10, 30, 60, 1440 and 2880 min and their mass was recorded. A total of 10 scans

were averaged for each sampling time-point. The observed mass ($M_{r\ obs}$) in H_2O was used for time zero. The protein samples were maintained at 4°C throughout the experiment. The experiments were repeated under the same conditions for the PDZ-ENEQVSAV complex in order to compare the rate of H/D exchange of free and peptide-bound PDZ. 100μL of 0.1mM PDZ was combined with a 100μL of 0.2mM peptide and then evaporated to 5μL. This ensured that the complex was formed in solution prior to initiating the exchange process. The sample was diluted to 200μL with D_2O to initiate exchange.

3.3 Results

Peptide Mass Spectra

The mass spectra of the peptides, ENEQVSAV, ENEQVCAV and KDDEVYYV, are shown in Figures 3.1 to 3.3, respectively. The main peak in Figure 3.1 at $[M+H]^+$=875.4 confirms the identity of the singly charged peptide ENEQVSAV (exact M_r = 874.4). Figure 3.2 shows a main peak at $[M+H]^+$=891.4, which confirms the presence of the reduced form of the singly charged peptide ENEQVCAV (exact M_r = 890.4); a second peak appears at + 22 m/z (m/z=913.5), which corresponds to a sodium adduct. It was important to confirm that the cysteine sulfhydryl was reduced since it was proposed that hydrogen bonding could occur between the thiol of cysteine and the imidazole side chain of His-71 in the PDZ binding groove. Figure 3.3 shows the mass spectrum of purified KDDEVYYV peptide at $[M+H]^+$=1030.5 (exact M_r = 1029.5).

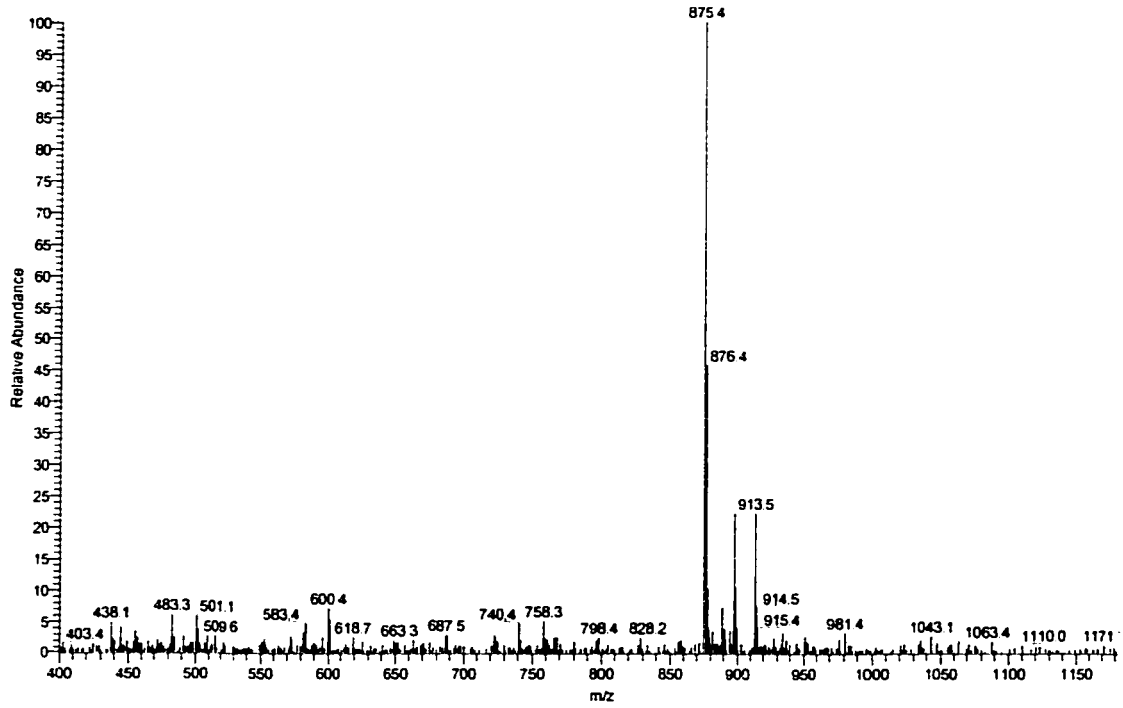

Figure 3.1 ESI mass spectrum of 1mM ENEQVSAV peptide in 40:60 ACN/0.1% TFA. The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 350°C.

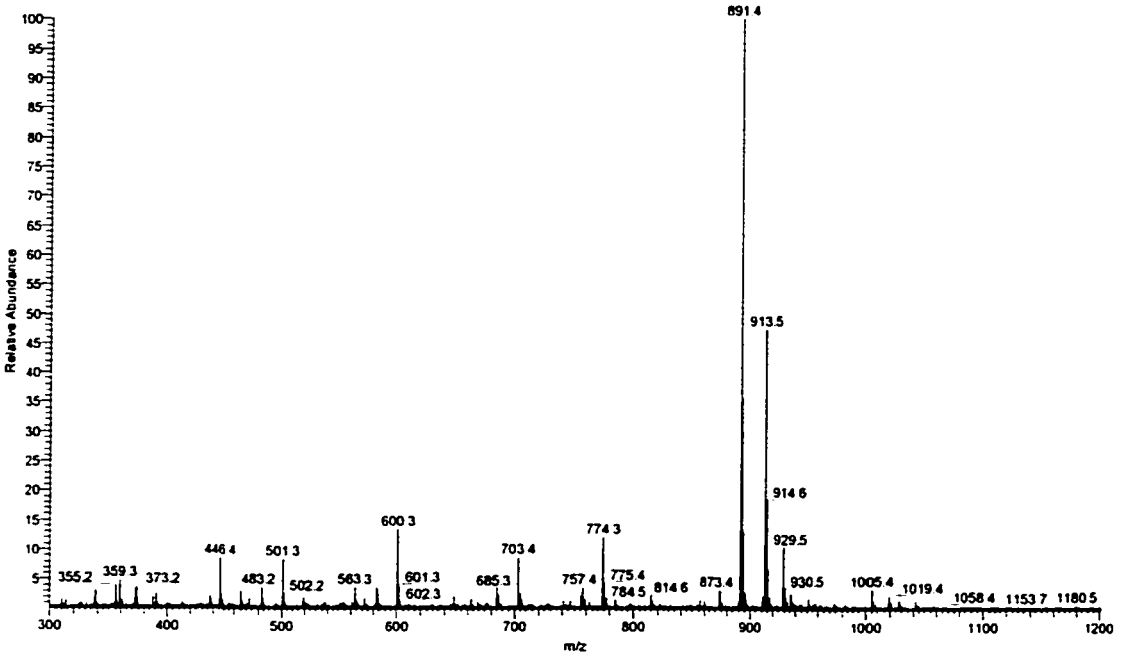

Figure 3.2 ESI mass spectrum of 1mM ENEQVCAV peptide in 40:60 ACN/0.1% TFA. The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 350°C.

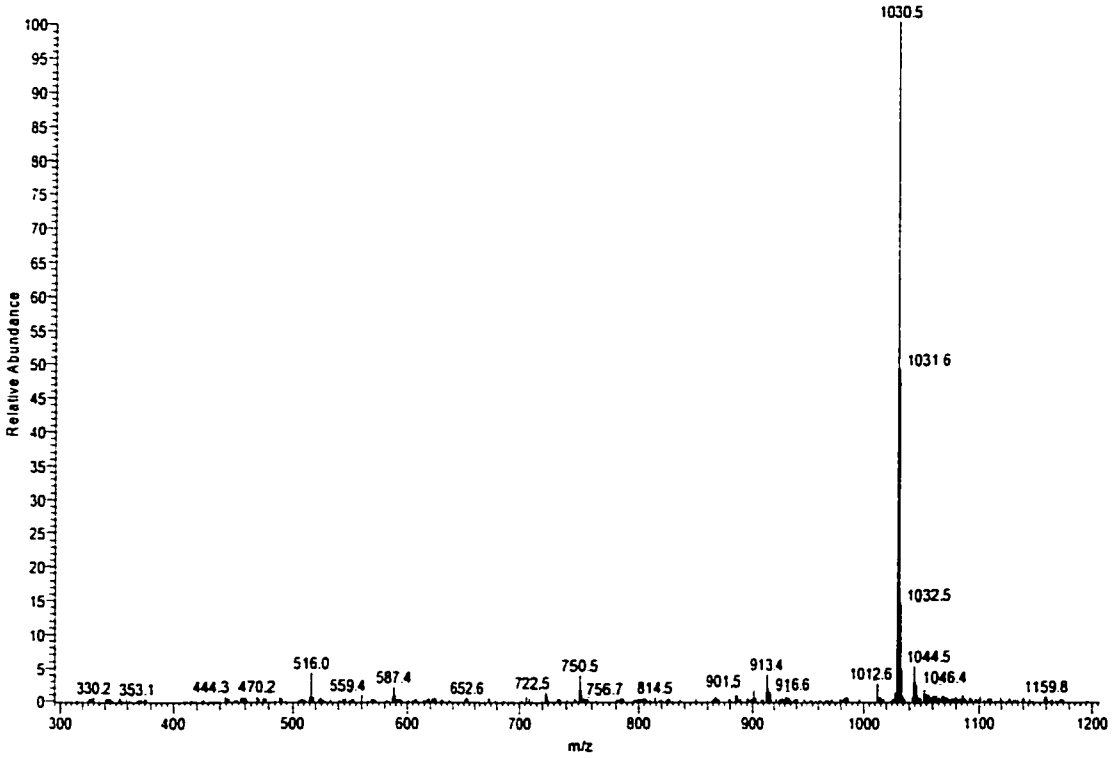

Figure 3.3 ESI mass spectrum of 1mM KDDEVYYV peptide in 40:60 ACN/0.1% TFA. The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 350°C.

PDZ Mass Spectra

The second PDZ domain from hPTP1E is a small protein of 96 amino acids. Based on the amino acid sequence, the theoretical mass of 10,006.2 daltons exceeds the scan range of the mass spectrometer. However, since the protein is multiply charged in the electrospray source of the mass spectrometer, it was possible to detect well resolved charge states of +3 to +9. In order to observe the maximum number of charge states, a high-mass scan range was used to observe the free PDZ domain and PDZ-peptide complexes. Ammonium acetate (pH 6.9) was used as a buffer since sodium phosphate buffer is not compatible with the mass spectrometer. Phosphate salts are non-volatile and cause clogging of the small internal diameter of the heated capillary as the solvent stream is evaporated. In addition, potassium or sodium ions can form adducts, known as

catenization, with various ionized groups present on the protein. This complicates mass determinations, especially with multiply charged proteins ions. Figure 3.4 shows the +6 and +7 charge states in the ESI mass spectrum of PDZ in sodium phosphate buffer (pH 6.9). The +6 charge-state is composed of multiple peaks which corresponds to a theoretical distribution of seven possible combinations of Na^+ and H^+ cations: $(M+6H)^{6+}$, $(M+Na+5H)^{6+}$, $(M+2Na+4H)^{6+}$, $(M+3Na+3H)^{6+}$, $(M+4Na+2H)^{6+}$, $(M+5Na+1H)^{6+}$, $(M+6Na)^{6+}$. Similarly, the +7 charge-state has eight theoretical combinations $(M+7H)^{7+}$, $(M+Na+6H)^{7+}$, $(M+2Na+5H)^{7+}$, $(M+3Na+4H)^{7+}$, $(M+4Na+3H)^{7+}$, $(M+5Na+2H)^{7+}$, $(M+6Na+H)^{7+}$ and $(M+7Na)^{7+}$.

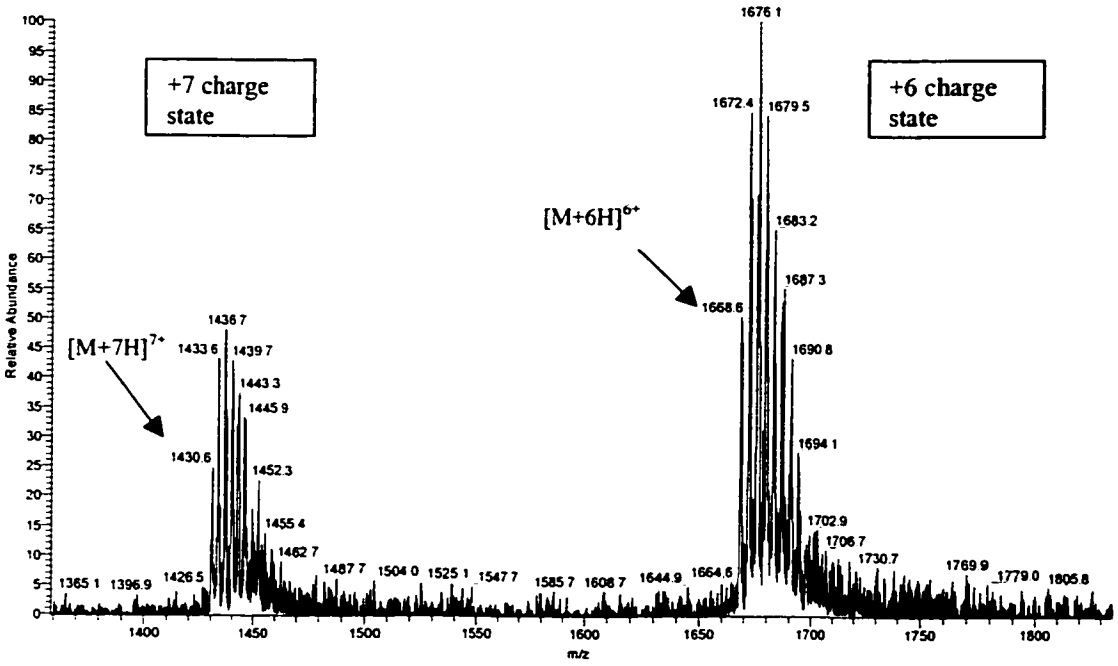

Figure 3.4 ESI mass spectrum of 55μM PDZ in 50mM phosphate buffer (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 4.0 kV and heated capillary set to 210°C.

The mass spectrum of PDZ in 50mM ammonium acetate buffer, pH 6.9, in Figure 3.5 (high-mass range), shows a charge state distribution centered at +6. Additional intense peaks are visible corresponding to charge states +7, +5, +4 and +3. The corresponding mass spectrum of PDZ with the peptide ENEQVSAV is shown in Figure 3.6, and reveals a charge-state distribution centered at +5.

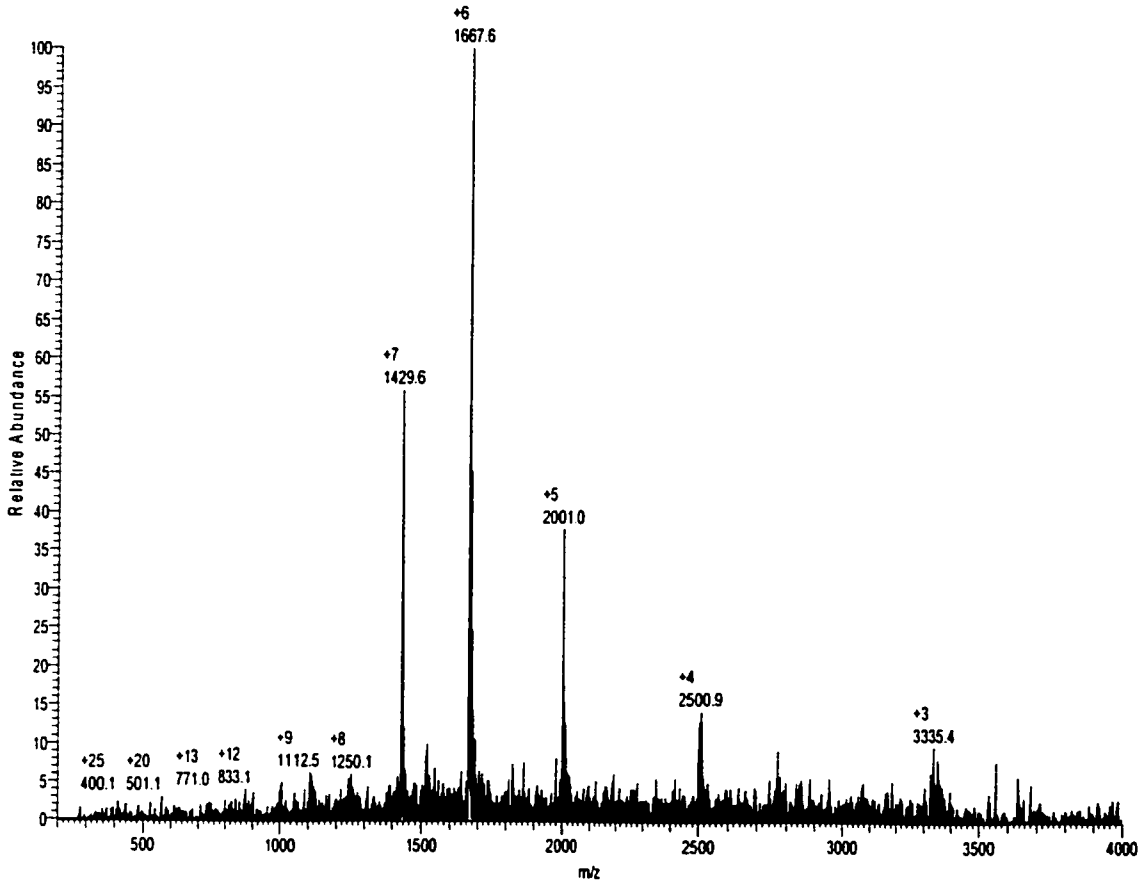

Figure 3.5 High- range ESI mass spectrum of 55μM PDZ domain in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 210°C.

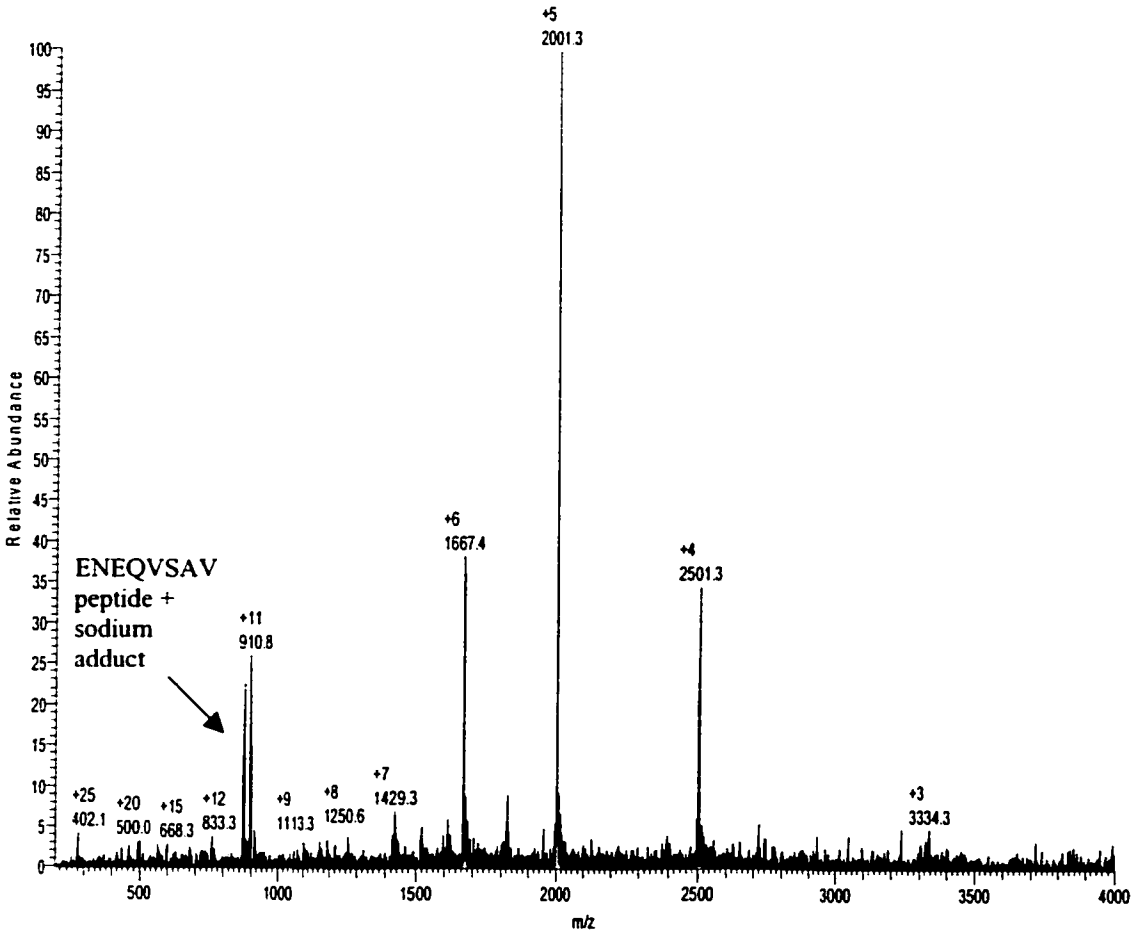

Figure 3.6 High-range ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 210°C.

For improved accuracy, the molecular weights were calculated from spectra collected using the normal-mass range (75-2000m/z) since the resolution is two-fold higher than in the high-mass range. Figure 3.7 shows the normal-range mass spectrum for the PDZ domain using parameters that were optimized for the highest sensitivity. The molecular weight of the protein was calculated using the peaks corresponding to charge states +6 and +7, and the observed mass of 10,006.3 +/- 0.1 Da for PDZ is in excellent agreement with the theoretical mass of 10,006.2 Da.

To observe peaks that arise from the PDZ-peptide complex, the electrospray ionization conditions had to be modified. The ionization conditions that provided the strongest signal for free PDZ (Figure 3.7), did not give rise to a detectable signal for the PDZ-peptide complex, as shown in Figure 3.8. Previous work by Anderegg and Wagner (1995) demonstrated that non-covalent complexes are disrupted by elevated needle voltages and orifice potentials. Also, Hopfgartner *et al.* (1999), using supramolecular organometallic complexes, showed that decreasing the temperature of the heated capillary was necessary to maintain non-covalent complexes intact during the ionization process.

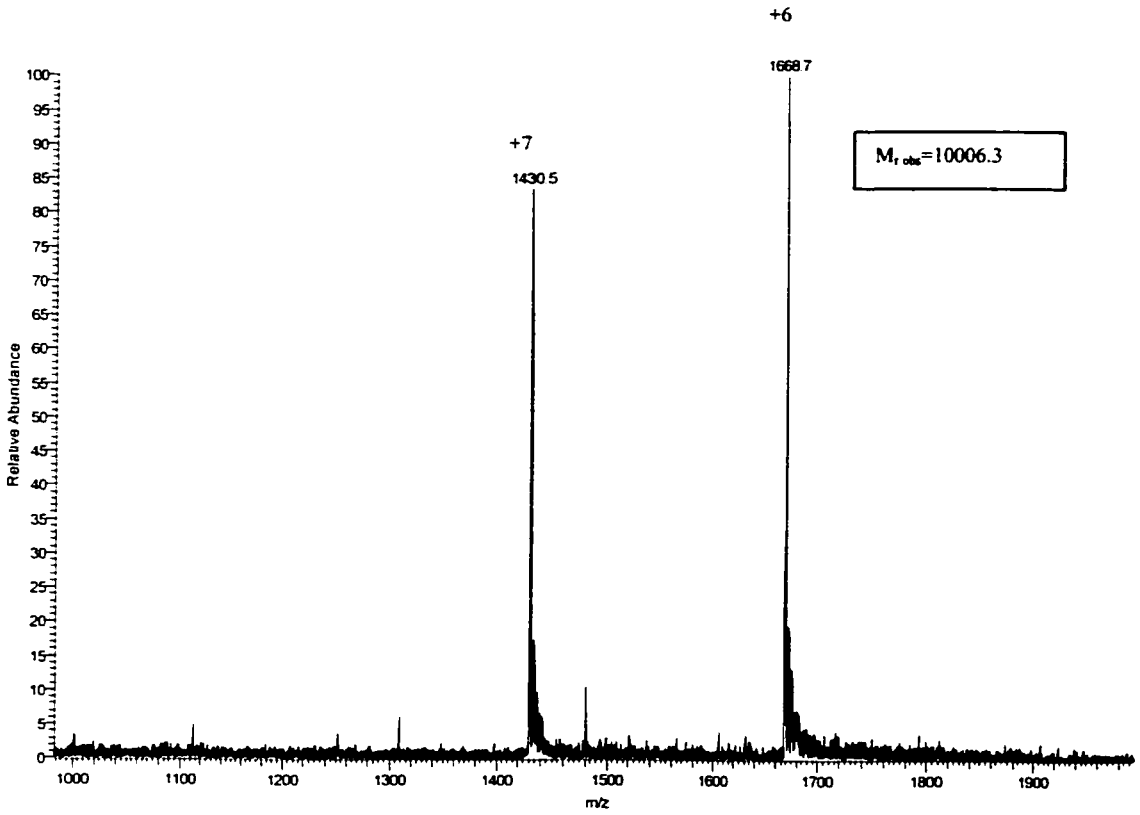

Figure 3.7 ESI mass spectrum of 55μM PDZ in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 210°C.

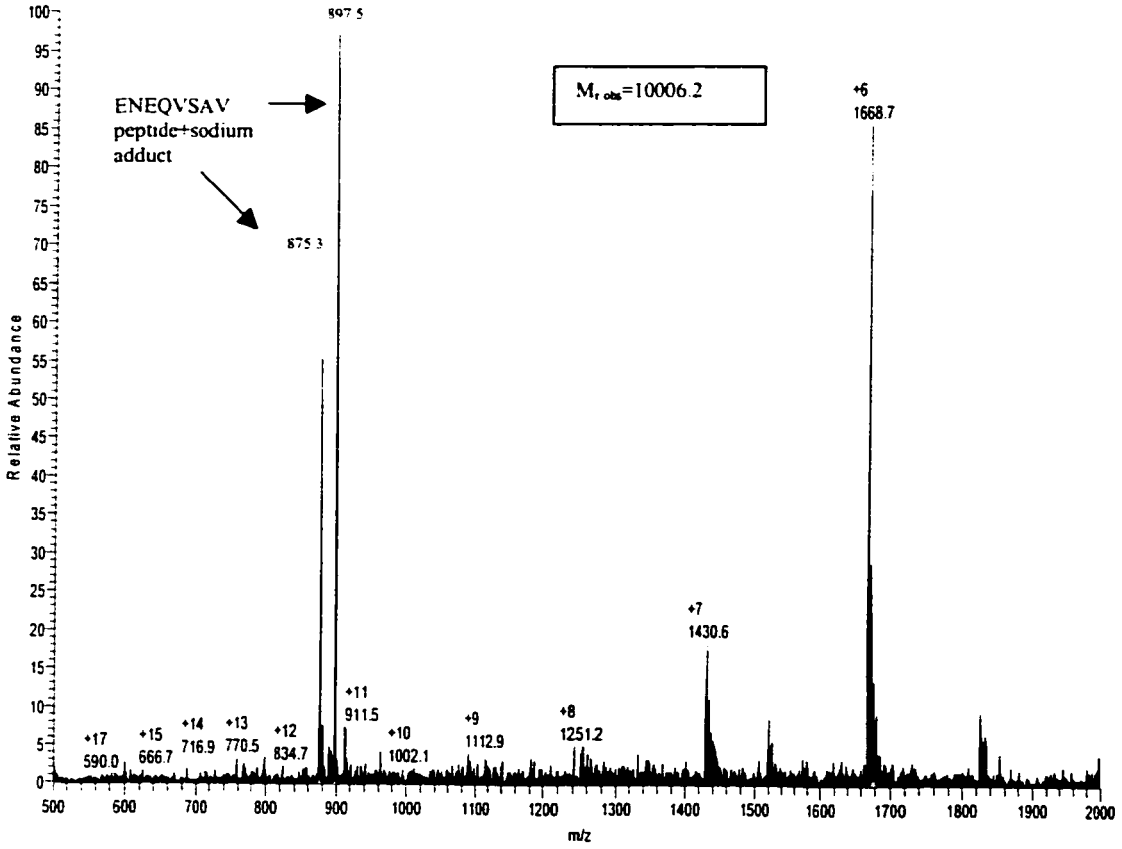

Figure 3.8 ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 4.0kV and heated capillary set to 210°C.

It is therefore important that the energy applied to ionize and desolvate the sample be minimized to prevent non-covalent complexes from breaking apart due to molecular collisions in the source (Anderegg and Wagner, 1995). The needle voltage and capillary temperature were found to have the most effect on the PDZ-peptide signal. These two parameters were optimized in order to obtain a signal that could be attributed to a PDZ-ENEQVSAV complex (Figure 3.9). The normal range mass spectra showed a decrease in the intensity of the +7 charge-state when the ENEQVSAV peptide is added.

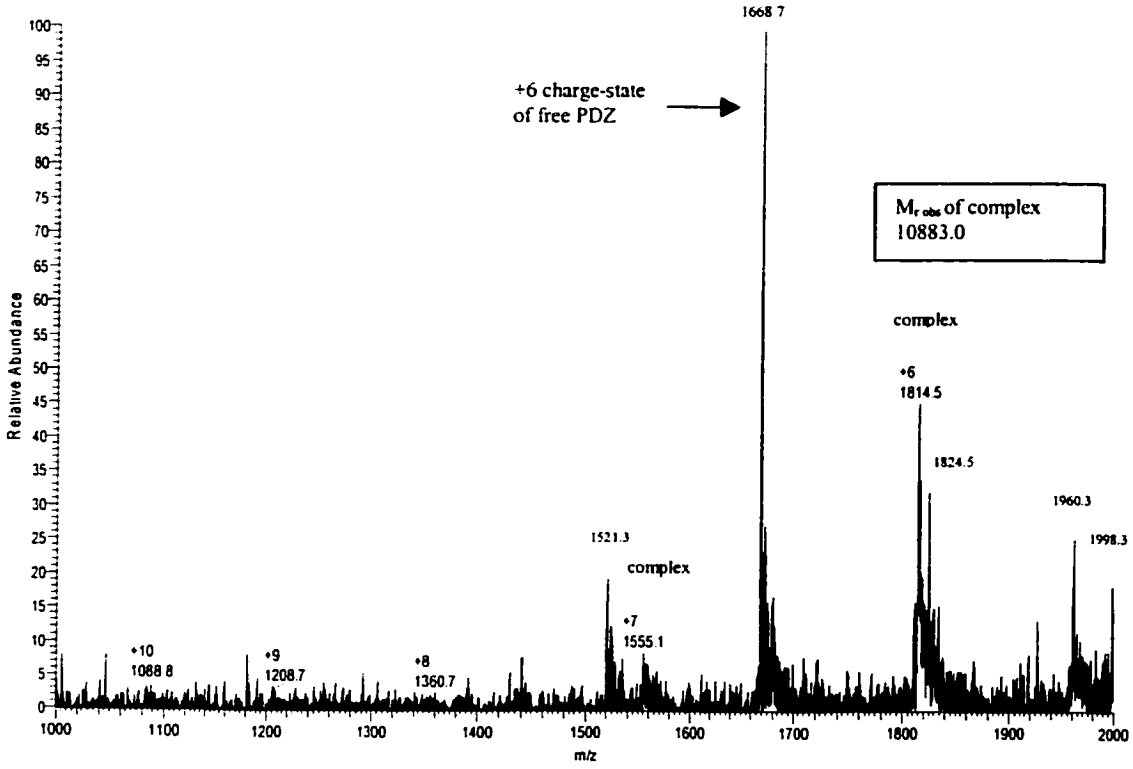

Figure 3.9 ESI mass spectrum of 50μM PDZ +100μM ENEQVSAV peptide in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 3.0kV and heated capillary set to 110°C.

Two alternate peptides, ENEQVCAV and KDDEVYYV, were evaluated for their binding to PDZ. A mixture of the KDDEVYYV peptide and PDZ did not show any peaks in its mass spectrum under the same conditions used in Figure 3.9, that could be attributed to the formation of a complex (Figure 3.10). The mass spectrum of a mixture of the ENEQVCAV peptide and PDZ showed very small peaks (only slightly greater than background noise) that corresponded to the +6 and +7 charge states of the ENEQVCAV-PDZ complex (Figure 3.11). Although reproducible, these were very weak, suggesting that the ENEQVCAV peptide has lower binding affinity for PDZ compared to ENEQVSAV.

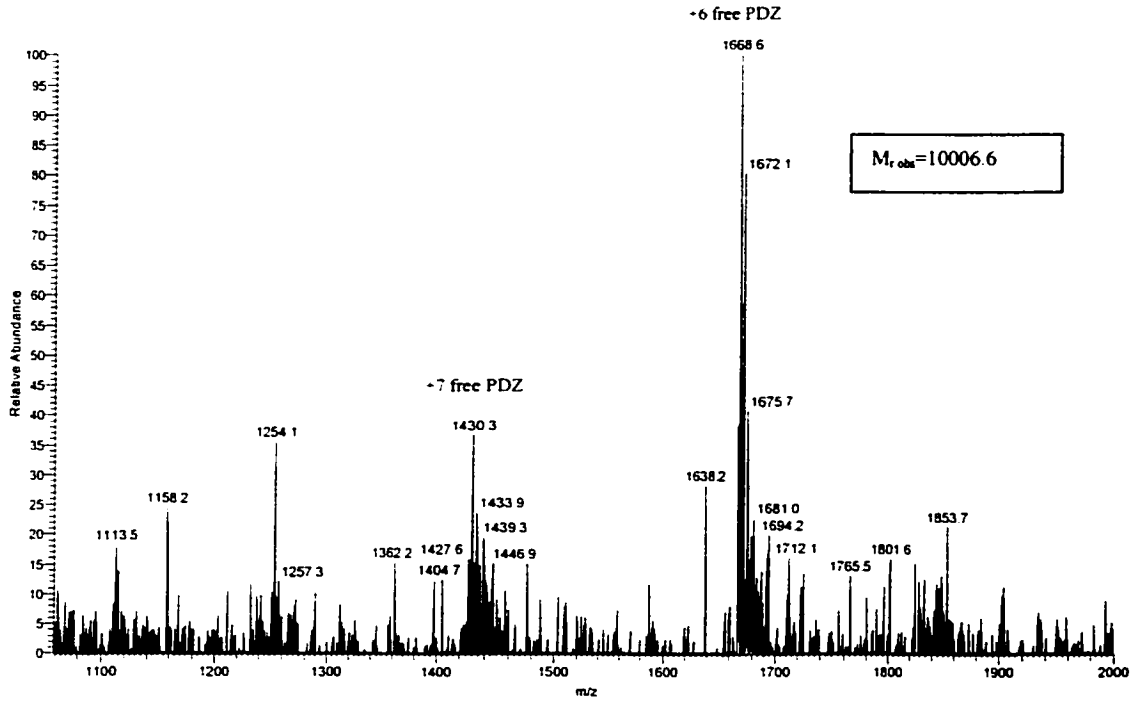

Figure 3.10 ESI mass spectrum of 50μM PDZ +100μM KDDEVYYV peptide in 50mM ammonium acetate (pH 6.9). The sample was infused at 10uL/min with a needle voltage of 3.0kV and heated capillary set to 110°C.

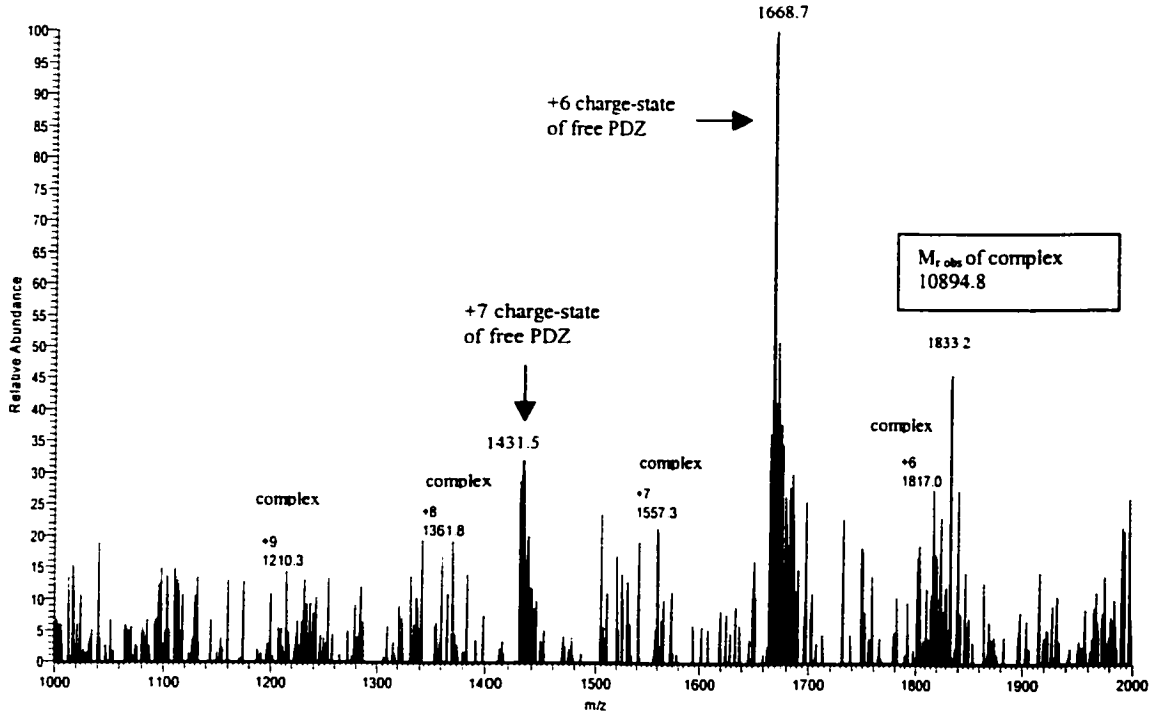

Figure 3.11 ESI mass spectrum of 50μM PDZ +100μM ENEQVCAV peptide in 50mM ammonium acetate (pH 6.9). The sample was infused at 10μL/min with a needle voltage of 3.0kV and heated capillary set to 110°C.

The calculated masses for the PDZ protein and related complexes with the peptides ENEQVSAV, ENEQVCAV and KDDEVYYV are shown in Table 3.1 The ionization conditions are also summarized.

Table 3.1 Effect of the electrospray ionization parameters on the species observed in the mass spectra of samples containing PDZ and peptides[a]

Sample	Needle voltage (kV)	Capillary Temp. (°C)	Theoretical M_r	Observed M_r
PDZ	4.0	210	10006.2	10006.3 ± 0.1
PDZ+ENEQVSAV	4.0	210	10881.1	10006.2 ± 1.4
PDZ+ENEQVSAV	3.0	110	10881.1	10883.0 ± 3.2
PDZ+ENEQVCAV	4.0	210	10897.1	10005.2 ± 1.2
PDZ+ENEQVCAV	3.0	110	10897.1	10894.8 ± 5.1
PDZ+KDDEVYYV	4.0	210	11036.3	10006.4 ± 2.1
PDZ+KDDEVYYV	3.0	110	11036.3	10006.6 ± 1.2

[a] Samples contained 50μM PDZ and 100μM peptide in 50mM ammonium acetate, pH 6.9. Concentration of free PDZ sample was 55μM in 50mM ammonium acetate, pH 6.9.

Based on the results summarized in Table 3.1, it is evident that electrospray conditions are critical when trying to detect the formation of non-covalent PDZ complexes. The reduction of the applied voltage and capillary temperature causes the appearance of peaks that correspond to the formation of the PDZ-ENEQVSAV complex, which has the expected stoichiometry of 1:1 (Kozlov *et al.*, 2000). The reduction of the heated capillary temperature, which is used to assist the evaporation process, reduces the kinetic energy of the PDZ-ENEQVSAV complex and disruptive molecular collisions.

Higher temperatures favor entropy, which in turn favors the break up of a protein complex. It is not, however, believed that the higher temperatures cause a partial denaturation, which would preclude binding, since the charge state distribution of free PDZ remained unchanged over of the capillary temperature range of 110-210°C (data not shown). Very low capillary temperatures (below 100°C) caused the background noise to increase dramatically. Since the protein samples were dissolved in a 100% aqueous medium, temperatures greater than 100°C were necessary for desolvation. The use of organic/aqueous mixtures such as methanol:water or acetonitrile:water could potentially allow for lower capillary temperatures. However, caution must be observed since the binding affinity may be altered in non-aqueous media.

H/D Exchange

The H/D exchange is shown in Figure 3.12 as a plot of mass vs D_2O exposure time. The plot contains H/D exchange kinetics for both free PDZ and the PDZ-ENEQVSAV complex at 4°C. The results show a rapid increase in mass, which begins to plateau at approximately 10min, with very little additional increase occurring even after an incubation period of 48h, suggesting that the exchange is complete. The mass increase of PDZ in the PDZ-ENEQVSAV complex is slower than that of free PDZ indicating protection against exchange. Both mass curves increase rapidly in the first minute. The rapid exchange likely represents non-hydrogen bonded, solvent exposed amides which are unaffected by peptide binding. Within the first 60min of D_2O exposure, the mass of free PDZ is higher by 35 (±9) amu than that of the complex. The exchange reaches a plateau and after 48h, the difference decreases to 18 (±5) amu. Back exchange of free

PDZ is observed between 1h and 48h, suggesting that some degree of equilibrium is reached or atmospheric moisture is absorbed by the sample during incubation. The exchange profile for both curves are similar suggesting perhaps that only a localized region of PDZ is affected by peptide binding.

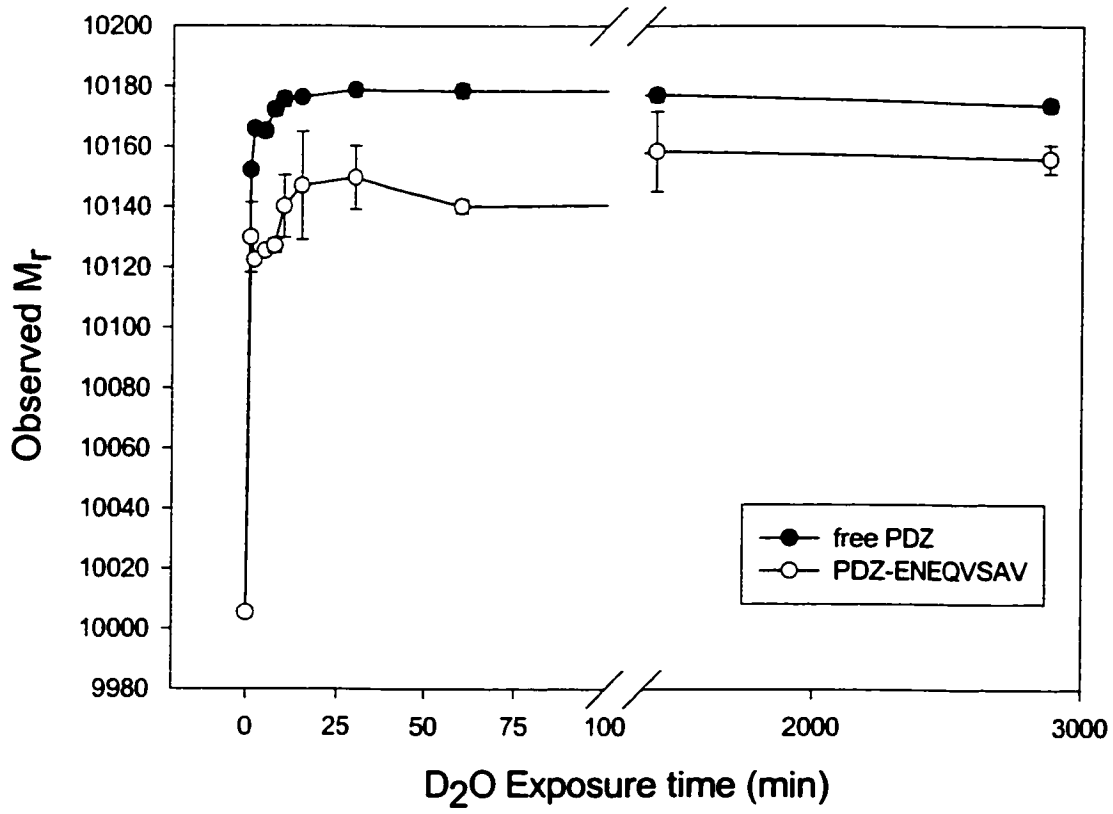

Figure 3.12 Plot of observed mass vs D_2O exposure time for PDZ (50μM) and PDZ-ENEQVSAV (50μM + 100μM) at 4°C. Samples were infused at 10μL/min with a needle voltage of 4.0 kV and heated capillary set to 210°C.

It is important to note that H/D exchange was monitored by following the increased mass of +6 and +7 charge states of free PDZ for both samples. It is not necessary to observe peaks for the PDZ-peptide complex in the mass spectrum since the signals from the free protein alone can be used to monitor H/D exchange (Anderegg and

Wagner, 1995). At each time-point, an aliquot (15μL) was withdrawn from the deuterated peptide-PDZ mixture and immediately infused into the mass spectrometer. For optimal sensitivity and mass accuracy, the needle voltage was set to 4kV and the heated capillary was set to 210°C. Even if the complex dissociates during the ionization stage, the solvent has already been removed from the sample preventing further H/D exchange from occurring.

3.4 Discussion

Peptide Binding

Understanding the binding specificity of PDZ interactions is important for identifying potential biological targets. Mass spectrometry provided rapid analysis of three different PDZ-peptide mixtures and could be used to screen large numbers of peptides. The positive identification of a complex between PDZ and ENEQVSAV confirmed NMR results that indicated that the second PDZ of hPTP1E binds peptides containing the X-Ser-X-Val motif (Ekiel *et al.*, 1998). Substituting Ser to Cys or Tyr in the –2 position of the peptide lowered the binding affinity to PDZ. This is consistent with the FT-IR thermal denaturation data (Table 2.2) that showed much lower thermal stabilization of complexes using the peptides ENEQVCAV and KDDEVYYV. The lack of complexation with alternate peptides demonstrates the specificity of the PDZ-peptide interaction. Therefore, it is assumed that the substituted peptides containing Tyr or Cys in the –2 position are unable to form a stable hydrogen bond with the imidazole side chain of His-71. Tyr and Cys are bulkier than Ser, therefore steric effects could also account for the weaker binding.

Ser/Cys substitutions are common mutations used in protein chemistry. The active site of PTP1 contains Cys-215 and it has been demonstrated that the mutation of this single residue to Ser-215 (which differs only by one atom) causes a complete lack of enzymatic activity even though the phosphatase is still capable of binding its substrate (Guan and Dixon, 1991). While this demonstrates how a single atom replacement can affect catalytic function of an enzyme, PTP1 (C215S) did not appear to exhibit weaker substrate binding unlike that observed here.

Based on the mass spectrometry data, the binding affinity of the peptides can be arranged as ENEQVSAV > ENEQVCAV > KDDEVYYV. The MS signal obtained for a non-covalent complex is dependent on the strength of the complex but its relative intensity may be greatly underestimated due to disruptive forces during ionization. Since the degree of contribution from these forces cannot be determined, MS does not yield accurate estimates of K_D. Nonetheless, the MS data suggest that the K_D of PDZ for ENEQVSAV is 10^{-5} M or less, if we infer from the mass spectrum (Figure 3.9) that about half of the PDZ is bound at the concentrations used. The reported K_D by NMR is 30μM (Ekiel *et al.*, 1998). Under the same MS conditions, the K_D of ENEQVCAV falls between 10^{-4} and 10^{-3} M, if we assume less than 10% of PDZ is bound. Since complexation is undetected with the KDDEVYYV peptide, the K_D is assumed to be greater than 10^{-3} M, since complexation was not observed at the concentrations used. These estimates agree with the analysis thermal denaturation results in Appendix B.

The charge-state maximum of PDZ combined with ENEQVSAV was decreased from +6 to +5 (Figures 3.5 and 3.6). A hydrogen bond formed between His-71 and Ser may be responsible for the observed shift. In order for the imidazole group to act as an

H-acceptor it must remain unprotonated. In the absence of the peptide, His-71 may be protonated thereby increasing the charge state of the protein by +1. The reported pK_a of free His is 6.0 (Leningher, 1993); therefore, under the experimental conditions (pH 6.9) the protonation state of His-71 may vary in the presence and absence of the ligand. The normal-range mass spectrum of PDZ-ENEQVSAV (Figure 3.8) also shows a reduction in the intensity of the +7 charge state as compared to the free PDZ mass spectrum (Figure 3.7). This effect was even more pronounced in Figure 3.9, which contains the mass spectrum of the PDZ-ENEQVSAV complex obtained using optimized spray-voltage and heated capillary temperature for detection of non-covalent interactions. Loss of the +7 charge state intensity was not as evident in combinations with either the KDDEVYYV or ENEQVCAV peptide (Figures 3.10 and 3.11). Alternatively, the binding interaction may cause a shift in the charge-state distribution either through a conformational change in PDZ that buries a basic residue or the formation of a bond between a basic residue other than His-71 and the incoming peptide. Within the PDZ binding groove, Lys13 and Ser17 are believed to interact with the C-terminal carboxylate group of the incoming peptide (Kozlov *et al.*, 2000). Therefore, salt-bridge formation between Lys13 and the peptide C-terminus (Figure 1.5) may cause loss of a positive charge on Lys13. Arg-79, located on α–helix#2, is also reported to interact with the peptide backbone (Kozlov *et al.*, 2000) and therefore may change protonation state during peptide binding.

The observation of a non-covalent complex using the LcQ Deca (Thermo-Finnigan) was found to be strongly dependent on the needle voltage and the temperature of the heated capillary. These observations are likely to be instrument specific based on the design of the ionization source. The manufacturer recommends operating the heated

capillary, which is used to evaporate the spray droplets, at 150-350°C. While the source design is efficient for desolvation at high HPLC flow rates, it appears to cause dissociation of non-covalent complexes unless it is operated well below the manufacturer's guidelines. Dissociation of weak complexes during ionization must be considered since even under the optimized conditions for detection of the PDZ-ENEQVSAV complex, the signal due to the free PDZ was more intense (Figure 3.9). Therefore, it is necessary to use an alternate method such as H/D exchange or thermal denaturation to detect weak interactions. H/D exchange using the ENEQVCAV and KDDEVYYV peptides was not performed since sufficient PDZ was not available within the timeframe of these experiments.

H/D Exchange

The rate of H/D exchange in D_2O depends on numerous factors including protein structure, degree of hydrogen bonding, solvent accessibility, pH and temperature (Wang and Tang, 1996). Proteins show reduced H/D exchange rates compared to short peptides due to limited solvent penetration to the internal core. The first protons to exchange are the surface exposed amides and side chains, followed by solvent inaccessible protons (Nabet and Pezolet, 1997). The folded protein is a dynamic macro-molecule that is subject to transient opening and closing motions termed local and/or global unfolding (Loh *et al.*, 1993). Local unfolding occurs in discrete structural units while global unfolding involves the entire protein. The rate of exchange of buried hydrogen atoms depends mostly on the kinetics of unfolding since upon opening of the protein, the buried amides should exchange at similar rates (Englander *et al.*, 1992). It is also important to

consider intra-molecular hydrogen bonding within secondary and tertiary structures (Williams *et al.*, 1996) since these bonds must be broken prior to exchange of protons with the solvent.

Upon ligand binding the possible mechanisms of protection from exchange can include conformational changes, hydrogen bonding of otherwise freely exchanging protons or shielding of previously exposed protein surfaces. Prior structural studies of the third PDZ of PSD-95 using X-ray crystallography (Doyle *et al.*, 1996) showed only a minor conformational change in the orientation of a single His side chain during the binding process. Williams *et al.* (1996) demonstrated that antigen-antibody complexes show reduced H/D exchange rates even though the crystal structure is unaltered after binding. The H/D exchange results indicated that 35 (±5) hydrogen atoms are protected from exchange in the PDZ-ENEQVSAV complex after 60min of D_2O exposure at 4°C. After 48h, this amount decreased to 18 (±5). However, The number of hydrogen bonds expected to participate in peptide binding (Figure 1.5) is about 10.

H/D exchange of free PDZ and peptide bound PDZ were previously reported by Ekiel *et al.*, (1998) using NMR. The findings demonstrated that H/D exchange kinetics of 26 out of the 44 slowly-exchanging residues ($k_{exch} < 10^{-2}$ min^{-1}) showed a 10-fold or greater decrease in exchange rate when a model peptide (RNEIQSLV) was bound to PDZ (Ekiel *et al.*, 1998). These 26 residues are localized all secondary structural elements except α–helix#1 (residues 45-49). The most protected residue was Ser-21 found in β-strand#2 which binds to the peptide backbone (See Figure 1.6) (Ekiel *et al.*, 1998).

The MS data suggest that stabilization against H/D exchange occurred through mechanisms other than hydrogen bonding. Increased conformational stability slows the

transient global and or local unfolding which is necessary for solvent penetration and exchange. This effect has been reported to reduce the H/D exchange rate of SH2-peptide complexes (Anderegg and Wagner, 1995). It is likely that at least some of the conformations of the PDZ-peptide complex are more stable than the free PDZ domain towards transient folding/unfolding. Increased conformational stability is also supported by the FT-IR thermal denaturation data (Table 2.1), which shows an increased T_m for the PDZ-ENEQVSAV complex.

3.5 Conclusions

The mass spectrometry data confirmed the formation of a 1:1 complex between the second PDZ of hPTP1E and the peptide ENEQVSAV, which is consistent with the motif recognized by Class I PDZ domains (Sheng and Sala, 2001). The complex that formed was detectable through direct observation in the mass spectra. In addition, the stability of the complex that formed was dependent on the electrospray ionization parameters, consistent with non-covalent complexation (Ganem *et al.*, 1993). A very weak interaction was detected with the ENEQVCAV peptide. The KDDEVYYV peptide did not form a complex that was detectable by mass spectrometry. However, this does not necessarily rule out the formation of non-covalent complexation since it is possible that very weak complexes are not preserved during ionization. The formation of a complex between the second PDZ of hPTP1E and the ENEQVSAV peptide was also observed by a decrease in H/D exchange. This may be attributed to an increased conformational stability of the non-covalent complex that resists H/D exchange with the solvent. Both MS methods showed usefulness in studying protein binding.

CHAPTER 4 – GENERAL CONCLUSIONS AND FUTURE WORK

4.1 General Conclusions

The ability to probe binding between a protein and its target peptide was demonstrated using FT-IR spectroscopy and mass spectrometry. FT-IR and MS have been previously applied to study protein-ligand binding; however, there are currently no literature reports that demonstrate evidence of peptide binding to PDZ domains using either technique. The data confirms an interaction previously observed by NMR (Ekiel *et al.*, 1998) between the second PDZ of hPTP1E and a target peptide containing the C-terminal motif X-Ser/Thr-X-Val. This C-terminal motif is recognized by Class I PDZ domains, which are found in hPTP1E, PSD-95 and Syntrophin (Sheng and Cala, 2001). Peptides substituted in the –2 position with either Tyr or Cys, did not show binding under similar conditions. These results confirm previously reported findings by Fuh *et al.* (2000) that substitutions at the –2 position of the peptide influences the specificity of binding.

The PDZ-ENEQVSAV complex was directly observed by mass spectrometry using PDZ concentrations of ~ 50μM. The observed mass of the complex confirmed the expected stochiometry of 1:1 (Kozlov 2000). The stability of the complex was sensitive to the ionization conditions, which is often a characteristic of non-covalent complexes (Ganem *et al.* 1993). Reduction of the spray voltage and heated capillary temperature was found to have the greatest influence on the MS signal arising from the PDZ-ENEQVSAV complex. These ionization parameters control the extent of destructive molecular collisions in the ionization source that can break apart non-covalent complexes

(Anderegg and Wagner 1995). Two alternate peptides, KDDEVYYV and ENEQVSAV, did not show substantial evidence of complexation with PDZ regardless of the ionization conditions. Further investigations using MS to monitor H/D exchange of PDZ showed stabilization of the protein structure during peptide binding. In the presence of the ENEQVSAV peptide, the H/D exchange rate of PDZ was reduced and the total number of exchanged amides was lower even after an incubation period of 48h.

Thermal denaturation and H/D exchange were selected to monitor changes in the stability of the PDZ domain upon binding a target peptide. FT-IR monitored, thermally induced denaturation of the PDZ-peptide complex showed a 10°C increase in the T_m (57 ± 1°C) as compared of free PDZ (47± 1°C) at concentrations of about 1mM. Combinations of PDZ with alternate peptides KDDEVYYV and ENEQVCAV showed minor increases in T_m (48 and 49°C, respectively) indicating little or no interaction at the concentrations used. This supports the MS data, which revealed that a 1:1 complex formed between PDZ and ENEQVSAV. 2D analysis of the thermal denaturation spectra revealed an altered unfolding pattern of the PDZ-peptide complex. The PDZ-ENEQVSAV complex that formed was more stable and contained residual turn-like structures in the denatured state.

Curve fitted FT-IR spectra showed increased intensity in the low frequency β-sheet band ($1628cm^{-1}$) at the expense of the high frequency β-sheet component ($1639cm^{-1}$) in the PDZ-ENEQVSAV complex. The low frequency β-sheet band is characteristic of short β-type structures found in the exposed region of a β-sheet (Fabian *et al.*, 1993). Therefore, the observed changes in the FT-IR spectrum, suggest that binding of the ENEQVSAV peptide can be localized to the β-sheet component of PDZ

(Figure 2.5). Previous X-ray and NMR work that showed peptides bind to β-strand #2, forming an additional anti-parallel β-strand. Resolution to the individual amino residue level using FT-IR is only possible with small number of residues that provide characteristic IR absorption bands such as Tyr (Arrondo *et al.* 1988) and Cys (Moh *et al.* 1987). Peptides containing substitutions of these residues in the –2 position did not show any evidence of binding which is consistent with the MS data and thermal denaturation studies.

4.2 Future Work

All three techniques demonstrated the ability to detect an interaction, either directly or indirectly, between a peptide and the PDZ domain. The use of mass spectrometry provided the most rapid detection and confirmed the formation of a PDZ-peptide complex. Mass spectrometry should be further explored with the addition of a proteolytic cleavage step after H/D exchange. This approach would potentially localize the region of the PDZ domain that showed resistance to H/D exchange upon binding the ENEQVSAV peptide in a similar manner to findings reported by Ekiel *et al.*, (1998) using NMR. Localization of the protected amide bonds could then be used to locate the bonding pocket and the residues that are involved in, or at least affected by, peptide binding. This type of work was pioneered by Zhang and Smith (1993) and demonstrated MS as a high-resolution technique capable of selectively probing regions of ~ 2 -10 residues in proteins.

In order to gain further insight using FT-IR, binding to PDZ should be investigated using peptides that are labeled with ^{13}C. This would allow for the

simultaneous monitoring of the peptide and protein spectra, with a significant reduction in spectral overlap. This type of experiment should also be investigated using PDZ-PDZ dimers, which are known to occur between nNOS and PSD 95 (Hillier *et al.*, 1999).

The denaturation of PDZ and its peptide complex should be further investigated by CD spectroscopy at lower concentrations in order to distinguish unfolding and aggregation. This would be useful in determining if the residual helical structures observed in the denatured spectra had been prevented from unfolding due to protein aggregation at FT-IR concentrations (~1mM).

4.3 References

Anderegg R.J., and Wagner D.S. (1995) *Journal of the American Chemical Society 117*, 1374-1377.

Arrondo, J.L.R., Young, M.N., and Mantsch H.H. (1988) *Biochimica et Biophysica Acta 952*, 261-268.

Baca, M., and Kent, B.H. (1992) *Journal of the American Chemical Society 114*, 3992-3993.

Balbach, J., Forge, V., van Nuland, N.A., Winder, S.L., Hore, P.J., and Dobson, C.M. (1995) *Nature Structural Biology 2*, 865-871.

Bandekar J. (1992) *Biochimica et Biophysica Acta 1120*, 123-143.

Bare, G.H., Alben, J.O., and Bromberg, P.A. (1975) *Biochemistry 14*, 1578-1583.

Braiman, M.S., and Rothschild K.J. (1988) *Annual Reviews in Biophysics and Biophysical Chemistry 17*, 541-570.

Cao, T.T., Deacon, H.W., Reczek, D., Bretscher, A., and von Zastrow, W. (1999) *Nature 401*, 286-289.

Christopherson, K.S., Hillier, B.J., Lim, W.A., and Bredt, D.S. (1999) *Journal of Biological Chemistry 274*, 27467-27473.

Casal H.L., Kohler, U., and Mantsch, H.H. (1988) *Biochimica et Biophysica Acta 957*, 11-20.

Craven, S.E., and Bredt, D.S. (1998) *Cell 93*, 495-498.

Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties*. 2nd edition. WH Freeman and Company, New York.

Dill, K.A. (1990) *Biochemistry 29*, 7133-7155.

Dodge R.W., and Scheraga, H.A. (1996) *Biochemistry 35*,1548-1559.

Dong, A., Huang, P., and Caughey, W.S. (1990) *Biochemistry 29*, 3303-3308.

Dong A., Huang, P., and Caughey, W.S. (1992) *Biochemistry 31*, 182-189.

Doyle, D.A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon R. (1996) *Cell 85*, 1067-1076.

Ekiel, I., Banville, D., Shen, S.H., and Gehring, K. (1998) *Biochemistry Cell Biology 76*, 334-340.

Englander, S.W., Englander, J.J., McKinnie, R.E., Ackers, G.K., Turner, G.J., Westrick, J.A., and Gill, S.J. (1992) *Science 256*, 1684-1687.

Esposito, C. (2001) MSc Thesis, Concordia University, Montreal, Canada.

Fabian, H., Schultz, C., Naumann, D., Landt, O., Hahn, U., and Saenger W. (1993) *Journal of Molecular Biology 232*, 967-981.

Fabian, H., Schultz, C., Backmann, J., Hahn, U., Saenger, W., Mantsch, H.H., and Naumann, D. (1994) *Biochemistry 33*, 10725-10730.

Filosa, A., Wang, Y., Ismail, A.A., and English, A.M. (2001). *Biochemistry 40*, 8256-8263.

Fuh, G., Pisabarro, M.T., Li, Y., Quan, C., Lasky, L.A., and Sidhu, S.S. (2000) *Journal of Biological Chemistry 275*, 21486-21491.

Ganem, B., Li, Y.T., and Henion, J.D. (1993) *Tetrahedron Letters 34*, 1445-1448.

Ganguly, A.K., Pramanik, B.N., Tsarbopoulos, A., Covey, T.R., Huang, E., and Fuhrman, S.A. (1992) *Journal of the American Chemical Society 114*, 6560-6562.

Ghaemmaghami S., Fitzgerald M.C., and Oas, T.G. (2000) *Proceedings of the National Academy of Science of the USA 97*, 8296-8301.

Gill, S.C., and Von Hippel, P.H. (1989) *Analytical Biochemistry 182*, 319-326.

Goormaghtigh E., Cabiaux V., and Ruysschaert J.M. (1990) *European Journal of Biochemistry 193*, 409-420.

Gratzer, W.B., and Cowburn, D.A. (1969) *Nature 222*,426-431.

Greenfield, N., and Fasman, G.D. (1969) *Biochemistry 8*, 4108-4116.

Greenfield, N., Davidson, B., and Fasman, G.D. (1967) *Biochemistry 6*, 1630-1637

Guan, K.L., and Dixon, J.E. (1991) *Journal of Biological Chemistry 266*, 17026-17030

Heyn, M.P., and Weischet, W.O. (1975) *Biochemistry 14*, 2962-2968.

Hillier, B.J., Christopherson, K.S., Prehoda, K.E., Bredt, D.S., and Lim,W.A. (1999) *Science 284*, 812-815.

Hirao, K., Hata, Y., Ide, N., Takeuchi, M., Irie, M., Yao, I., Deguchi, M., Toyoda, A., Sudhof, T.C., and Takai, Y. (1998) *Journal of Biological Chemistry 273*, 21105-21110.

Holzbaur, I.E., English, A.M., and Ismail, A.A. (1990). *Biochemistry 35*, 5488-5494.

Hopfgartner, G., Vilbois, F., and Piguet, C. (1999) *Rapid Communications in Mass Spectrometry 13*, 302-306.

Huang, E.C., Pramanik, B.N., Tsarbopoulos, A., Reichert, P., Ganguly, A.K., Trotta, P.P., and Nagabhushan, T.L. (1993) *Journal of the American Society for Mass Spectrometry. 4*, 624-630.

Jackson, M., and Mantsch, H.H. (1992) *Biochimica et Biophysica Acta 1118*, 139-143.

Jackson, M., and Mantsch, H.H. (1995) *Critical Reviews in Biochemistry and Molecular Biology 30*, 95-120.

Katta, V., and Chait, B.T. (1991) *Journal of the American Chemical Society 113*, 8534-8535.

Kebarle, P. (2000) *Journal of Mass Spectrometry 35*, 804-817.

Konerman, L., and Douglas, D.J. (1997) *Biochemistry 36*, 12296-12302.

Kozlov, G., Gehring, K., and Ekiel, I. (2000) *Biochemistry 39*, 2572-2580.

Lehninger, A.L., Nelson, D.L., and Cox, M.M. (1993) *Principles of Biochemistry*. 2nd edition. Worth Publishers, New York.

Loo, J.A. (1995) *Journal of Mass Spectrometry 30*, 180-183.

Loh, S., Prehoda, K.E., Wang, J., and Markley, J.L. (1993) *Biochemistry 32*, 11022-11028.

Mantsch., H.H., Perczel, A., Hollosi, M., and Fasman, G.D. (1993) *Biopolymers 33*, 201-207.

Matsuyama, K., Sen, A.C., and Perrin, J.H. (1987) *The Journal of Pharmacy and Pharmacology 39*, 190-195.

Mizra, U.A., Cohen, S.L., and Chait, B.T. (1993) *Analytical Chemistry 65*, 1-6.

Moh, P.P., Fiamingo, F.G., and Alben, J.O. (1987) *Biochemistry 26*, 6243-6249.

Moseley, M.A. (1990) PhD Thesis. University of North Carolina at Chapel Hill, North Carolina, USA.

Nabet, A., and Pezolet, M. (1997) *Applied Spectroscopy 51*, 466-469.

Noda, I. (1990) *Applied Spectroscopy 37*, 433-438

Noda, I. (1993) *Applied Spectroscopy 47*, 1329-1336.

Oschkinat, H. (1999) *Nature Structural Biology 6*, 408-410.

Ozaki, Y., and Noda, I. (1996) *Journal of Near Infrared Spectroscopy 4*, 85-99.

Ozaki, Y., Liu, Y. and Noda, I. (1996) *Applied Spectroscopy 51*, 526-535.

Pace, N.C., and McGrath, T. (1980) *Journal of Biological Chemistry 255*, 3862-3865.

Ponting, C.P, Phillips. C., Davies, K.E., and Blake, D.J. (1997) *Bioessays 19*, 469-479.

Ramelow, K., Hubner, W., and Ackermann, T.H. (1998) *Analytical Biochemistry 257*, 1-11

Reinstadler, D., Fabian, H., Backmann, J., and Naumann, D. (1996) *Biochemistry 35*, 15822-15830.

Saras, J., Engstrom, U., Gonez, L.J., and Heldin, C.H. (1997) *Journal of Biological Chemistry 272*, 20979-20981.

Saras, J., and Helidin, C.H. (1996) *Trends in Biochemical Sciences 21*, 455-458.

Sarver, R.W., and Krueger, W.C. (1991) *Biochemistry 199*, 61-67.

Sato, T., Irie, S., Kitada, S., and Reed J.C. (1995) *Science 268*, 411-415.

Schellman, J.A. (1975) *Biopolymers 14*, 999-1018.

Sheng, M., and Sala, C. (2001) *Annual Reviews in Neuroscience 24*, 1-29.

Shortle D. (1996) *The FASEB Journal 10*, 27-34

Smith, D.L., Deng, Y., and Zhang, Z. (1997) *Journal of Mass Spectrometry 32*, 135-146.

Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M., and Cantley, L.C. (1997) *Science 275*, 73-77.

Surewicz, W.K., Leddy, J.J., and Mantsch, H.H. (1990) *Biochemistry 29*, 8106-8111.

Surewicz, W.K., Mantsch, H.H., and Chapman, D. (1993) *Biochemistry 32*, 389-394.

Tadesse, L., Nazarbaghi, R., and Walters, L. (1991) *Journal of the American Chemical Society 113*, 7037-7039.

Tochio, H., Zhang, Q., Mandal, P., Li, M., and Zhang, M. (1999) *Nature Structural Biology 6*, 417-421.

Tochio, H., Mok, Y.K., Zhang, Q., Kan, H.M., Bredt, D.S., and Zhang, M. (2000) *Journal of Molecular Biology 303*, 359-370.

Trewhella J., Liddle, W.K., Heidorn, D.B., and Strynadka, N. (1989) *Biochemistry 28*, 1294-1301.

Tsunoda, S., Sierralta, J., Sun, Y., Bodner, R., Suzuki, E., Becker, A., Socolich, M., and Zuker, C.S. (1997) *Nature 388*, 243-249.

van Stokkum, I.H.M., Lindsell, H., Hadden, J.M., Harris, P.I., Chapman, D., and Bloemendal, M. (1995) *Biochemistry 34*, 10508-10518.

van Stokkum, I.H., Spoelder, H.J., Bloemendal, M., van Grondelle, R., and Groen, F.C. (1990) *Analytical Biochemistry 191*, 110-118.

Veenstra, T.D., Tomlinson, A.J., Benson, L., Kumar, R., and Naylor, S. (1998) *Journal of the American Society of Mass Spectrometry 9*, 580-584.

Venyaminov, S.Y., and Kalnin, N.N. (1990) *Biopolymers 30*, 1243-1257.

Wang, F., Scapin, G., Blanchard, J.S., and Angeletti, R.H. (1998) *Protein Science 7*, 293-299.

Wang, F. and Tang, X. (1996) *Biochemistry 35*, 4069-4078.

Wang, P., Zhang, Q., Tochio, H., Fan, J.S., and Zhang, M. (2000) *European Journal of Biochemistry 267*, 3116-3112.

Wang, Y., Murayama, K., Myojo, Y., Tsenkova, R., Hayashi, N., and Ozaki, Y. (1998) *Journal of Physical Chemistry B 102*, 6655-6662.

Wiseman, T., Williston, S., Brandts, J.F., and Lin L.N. (1989) *Analytical Biochemistry 179*, 131-137.

Williams, D.C., Benjamin, D.C., Poljak, R.J., and Rule, G.S. (1996) *Journal of Molecular Biology 257*, 866-876.

Woody, R. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules.* (Fasman, G.D. ed) Plenum Press, New York.

Zhang, Z., and Smith, D.L. (1993) *Protein Science 2*, 522-531.

Zhang, Z., and Smith, D.L. (1996) *Protein Science 5*, 1282-1289.

Zhang, M., Fabian, H., Mantsch, H.H., and Vogel, H.J. (1994) *Biochemistry 33*, 10883-10888.

APPENDIX A

Primary amino acid sequence of the second PDZ domain of hPTP1E

1 B----------#1----------B B-
Pro-Lys-Pro-Gly-Asp-**Ile-Phe-Glu-Val-Glu** **Leu**-Ala-Lys-Asn-Asp-Asn-Ser-Leu-Gly-**Ile**

-----#2------B 31 B----------#3----------
Ser-Val-The-Gly-Gly-Val-Asn-Thr-Ser-Val Arg-His-Gly-Gly-**Ile-Tyr-Val-Lys-Ala-Val**

-B A----------#1---------A 51 B-----#4----------
Ile-Pro-Gln-Gly-**Ala-Ala-Glu-Ser-Asp**-Gly Arg-Ile-His-Lys-Gly-Asp-**Arg-Val-Leu-Ala**

---B B--#5-B A----------#2----------A
Val-Asn-**Gly-Val**-Ser-Leu-Glu-Gly-Ala-Thr **His-Lys-Gln-Ala-Val-Glu-Thr-Leu-Arg**-Asn

81 B----------#6----------B 91
Thr-Gly-Gln-Val-**Val-His-Leu-Leu-Leu-Glu** Lys-Gly-Gln-Ser-Pro-Thr

B = β-sheet
A = α-Helix

(Kozlov *et al.*, 2000)

APPENDIX B

Estimation of binding constants from melting temperatures

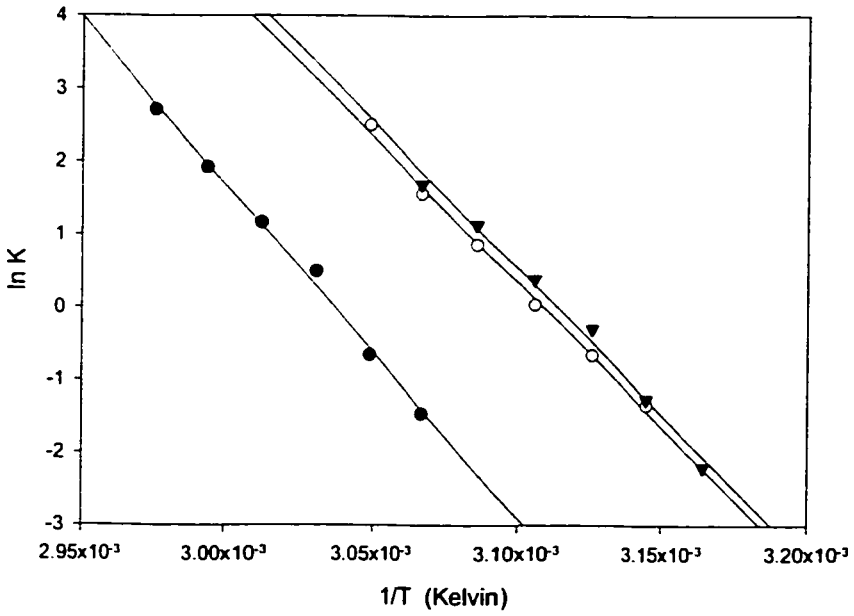

Figure B1. van't Hoff plot of β-sheet unfolding for combinations of PDZ with peptides ENEQVSAV (●), ENEQVCAV (○) and KDDEVYYV(▼)

Table B1. Estimates of binding affinity towards PDZ of three peptides based on T_m measurements and van't Hoff plots

Parameter	ENEQVSAV	ENEQVCAV	KDDEVYYV
ΔH_D (kcal/mol) [a]	91	79	80
ΔT_m	10	2	1
T_o (K)	320	320	320
T (K)	330	322	321
S (mol/L)	2×10^{-4} [b]	$7 - 12 \times 10^{-4}$ [c]	$7 - 12 \times 10^{-4}$ [c]
K_D [d]	3×10^{-6}	$10^{-4} - 10^{-3}$	10^{-3}

[a] Obtained from van't Hoff plot (slope= $-\Delta H/R$)

[b] Based on previous estimates of K_D

[c] Assumed less than half of protein is bound

[d] $1/K_D = K_B = \dfrac{e^{[(\Delta T_m \Delta H_D)/(T_o T R)]} - 1}{[S]}$