Human Glutaminase Interacting Protein (GIP): a Potential Candidate for Anti-Cancer Drug Design

by

Monimoy Banerjee

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 6, 2011

Keywords: Human glutaminase interacting protein, NMR, Phage display, Immunocytochemistry

Approved by

Smita Mohanty, Chair, Associate Professor of Chemistry and Biochemistry
Tatiana I. Samoylova, Associate Research Professor of Scott-Ritchey Research Center and Pathobiology
Rik Blumenthal, Associate Professor of Chemistry and Biochemistry
Curtis Shannon, Professor of Chemistry and Biochemistry

Abstract

PDZ domains are one of the most ubiquitous protein-protein interaction modules found in living systems. The glutaminase interacting protein (GIP), also known as the Tax interacting protein 1 (TIP-1), is a PDZ domain containing protein, which plays pivotal roles in many aspects of cellular signaling, protein scaffolding, and modulation of tumor growth. Unlike other PDZ domain containing proteins, GIP contains only one PDZ domain. We detail here the over-expression, purification, and characterization of structure, function, dynamics and interactions of recombinant human GIP with different C-terminal target protein recognition sequence motifs by CD, fluorescence, and high-resolution solution NMR methods. The target protein C-terminal recognition motifs employed in our interaction studies are glutaminase L (LGA), β-catenin, and FAS. This is the first report of GIP recognition of the cell surface protein FAS, which belongs to the tumor necrosis factor (TNF) receptor family and mediates cell apoptosis. We have also shown interaction studies of GIP with the C-terminus of Kir 2.3 and HPV16 E6. Interestingly, these ligands interact with different residues of GIP and show similar binding affinities. Dissociation constant (K_D) values ranging from 1.97 μM to 3.66 μM indicate strong interactions between GIP and its binding partners. Similar to previously determined PDZ domain structures, GIP contains two α-helices and six β-strands. One of the most striking features of GIP is that, unlike other PDZ domain containing proteins, GIP shows apparent conformational change upon ligand binding. Interestingly, the

backbone dynamics data of both free GIP and GIP-glutaminase C-terminus complex show that binding of the glutaminase C-terminal recognition motif results in a decrease in the order parameter (S^2), indicating an increase in global flexibility of GIP in this complex.

Studies thus far show that GIP interacts with the C-terminus of its partner proteins. Using an f8-type phage displayed peptide library, we have discovered multiple peptides with novel internal sequence motifs that specifically bind to GIP. We identified S/T-X-V/L-D as the consensus GIP-binding motif and, using protein data base searches, found several cancer-related human proteins as potential GIP partners. NMR spectroscopy was used to investigate the mechanisms of novel internal motif recognition by GIP. Through mutagenesis, we found two amino acid residues within the peptides that are important for the interaction with GIP. Importantly, one of the peptides co-localizes with GIP within human glioma cells and inhibits their metabolism, indicating that GIP might be a potential target for anti-cancer therapeutic interventions.

Acknowledgements

First of all, I would like to express my highest appreciation to my research advisor Dr. Smita Mohanty for her guidance, support, encouragement, and help throughout my graduate study. Secondly, I would like to thank Dr. Tatiana Samoylova for allowing me to do phage display and biological work in her laboratory and her guidance, encouragement, support, and help. I would like to thank Dr. Curtis Shannon, Dr. Rik Blumenthal, and Dr. Narendra Singh for their help, support and constructive suggestion to my dissertation. I also want to thank my former and current lab mates: Dr. Chengdong Huang, Dr. Tianzhi Wang, Dr. David Zoetewey, Dr. Janarthan Krishnamurthy, Dr. Uma Katre, Dr. Shigeki Saito, Suman Mazumder, Mohiuddin Ovee, Priscilla Ward, Bhaven Sayania, Amit Kumar, Dr. Alexandre Samoylov, Nancy Morrison, Ann Cochran for their meaningful discussions and help. I would like to thank NMR Lab Managers: Dr Rajagopalon Bhaskaran, and Dr. Michael Meadows for their help. I would like to thank the chairman of the Department of Chemistry and Biochemistry; Prof. J. V. Ortiz for his support and help. I would like to thank all professors and the administration staff at the Department of Chemistry and Biochemistry, and all my friends at Auburn who directly or indirectly contribute to the success of my study. I would like to thank my parents, all my family members, and relatives for their support and help. Last but not least, I would like to thank Ishita for her continuous encouragement, love and moral support.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

Literature Review

1.1 PDZ domains

Cellular signaling systems are an important part of physiological function and are mediated largely by protein-protein interactions. One of the most well-known protein-protein interaction motifs, the PDZ (postsynaptic density protein, disc large, zona occludens) domain (*1*, *2*), is present in several hundred human proteins (*3*). These PDZ motifs are small interaction modules normally spanning 80-100 amino acid residues. PDZ domains mediate various signaling pathways (*4*), localize and cluster ion-channels and membrane receptors, maintain cell polarities (*5*), and are involved in scaffolding of multimeric complexes by recognizing the C-terminal amino acid sequence motif of the interacting proteins (*6*, *7*). These interactions localize membrane proteins to specific sub-cellular domains, enabling the assembly of supramolecular complexes. Many of these proteins possess several PDZ domains within the same protein or are associated in multimeric complexes which form a conglomeration of PDZ modules. PDZ-containing proteins interact with various proteins within cells and mediate various diseases. Therefore, investigation of the regulatory mechanisms of PDZ protein-protein interactions is important for understanding of biological functions of PDZ domains and for design of new therapeutics.

1.2 PDZ domains distributions

PDZ domains are primarily found in eukaryotic organisms (*8*). However, a related kind of PDZ domains has also been identified in yeast, bacteria and plants (9, *10*) while there is no evidence that this domain is involved in protein-protein interaction (*8*). One bacterial PDZ domain with a solved structure reveals that the domain has similarity as well as differences in structure or fold compared to canonical PDZ domains (*11*). Thus, many of these bacterial 'PDZ-like' domains are likely to be significantly different from canonical metazoan PDZ domains (*8*). PDZ domains can solely compose a single protein or along with some other domains. Also, more than one PDZ domain can be present in a single protein. Examples of PDZ domain(s) which are components of different proteins are shown in **Figure 1.1**.

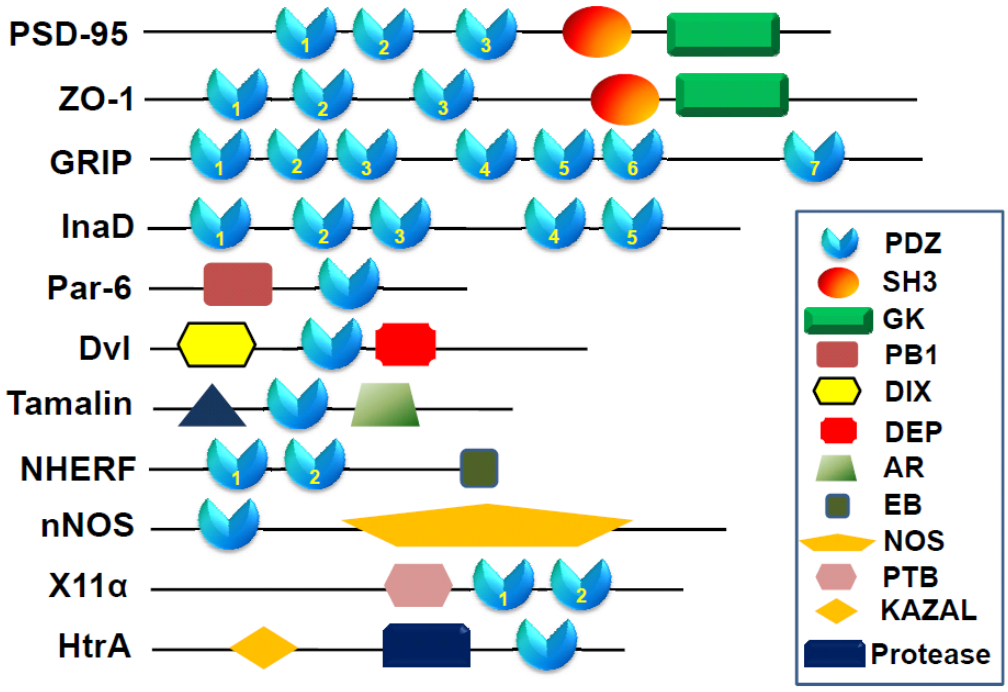

Figure 1.1 Examples of PDZ domain-containing proteins. Proteins are indicated by black lines scaled to the length of the primary sequence of the protein. Adapted from reference 8. The short form of the proteins are abbreviated here: Postsynaptic density protein-95 (PSD-95), Zonnula occludens 1 (ZO-1), Glutamate receptor interacting protein (GRIP), Activation no afterpotential D (InaD), Partitioning defective 6 (Par-6), Dishevilled (Dvl), Na^+/H^+ exchanger regulatory factor (NHERF), Neuronal nitric oxide synthase (nNOS), High temperature requirement A (HtrA).

1.3 Classification of PDZ domains

PDZ domains are very specific towards their binding proteins. They recognize the C-terminal sequence motifs of their target proteins. The length of these motifs is usually about five residues, but in rare cases specificity of interaction is defined by more than five residues (*12*). The residues within the PDZ-binding C-terminal sequence motif are classified as the followings: the C-terminal residue is the P_0 residue; subsequent residues towards the N-terminus are termed as P_{-1}, P_{-2}, P_{-3}, etc (*8*). It has been found that the P_0 and P_{-2} residues are the most important for PDZ domain recognition (*13*, *14*). These two residues are very selective in the binding mechanism. Based on how these PDZ domains recognize the C-terminal sequence motif of their target proteins, they are classified into three major classes: class I PDZ domains recognize the motif S/T-X-Φ-COOH (where Φ is a hydrophobic amino acid and X is any amino acid); class II PDZ domains recognize the motif Φ-X-Φ-COOH; and class III PDZ domains recognize the motif X-X-C-COOH. There are a few other PDZ domains that do not fall into any of these specific classes (**Table 1.1**) (*15*).

PDZ Domain	Concensus binding sequence	Ligand protein
Class I	P_{-3} -P_{-2} - P_{-1} -P_0 S/T-X-φ-COOH	
Syntrophin	E-S-T-V-COOH	Voltage-gated Na+ channel
PSD-95	E-T-D-V-COOH	Shaker-type K+ channel
GIP	E-S-M-V-COOH	Glutaminase-L
Class II	φ-X-φ-COOH	
hCASK	E-Y-Y-V-COOH	Neurexin
Erythrocyte p55	E-Y-F-I-COOH	Glycophorin C
Class III	X-X-C-COOH	
Mint-1	D-H-W-C-COOH	N-type Ca +2 Channel
SITAC	Y-X-C-COOH	L6 antigen
Other		
nNOS	G-D-X-V-COOH	MelR
MAGI PDZ2	S/T-W-V-COOH	Phage display
Engineered from SF6	K/R-Y-V-COOH	Synthesized peptide

Table 1.1 Examples of PDZ ligands. The short form of the proteins are abbreviated here: Postsynaptic density protein 95 (PSD-95), Glutaminase interacting protein (GIP), Human calcium/calmodulin-dependent serine kinase (hCASK), The munc-18 interacting-1(Mint-1), Neuronal nitric oxide synthase (nNOS), Membrane-associated guanylate kinase inverted (MAGI). Adapted from reference 14.

1.4 Structure and mechanism of interaction of PDZ domains

To date, more than **200** structures of PDZ domain, PDZ-binding partner complex structures have been solved by NMR or X-ray crystallography (*8, 16*). Structural characterization show that canonical PDZ domains consist of six β-strands (βA- βF) and two α-helices (αA and αB) (Figure 2A) (*17-19*). Between these two α helices (αA and αB), αa is shorter compare to αB. Secondary structures of the PDZ domains may vary in length, but the folds are conserved among all PDZ domains (*20-24*). Usually, most of the PDZ domains are monomer, but there are some reports of PDZ domain dimers (*25-30*). It is interesting that PDZ2 of zonula occludens-1 forms a dimer which creates a distinct interface. The charge–charge interactions created by residues in the PDZ dimer interface and residues of the Cx43 peptide give insight into the interaction specificity for the complex formation. This interaction regulates phosphorylation-mediated switch for the dynamics of the Cx43 gap junctions (*30*). Dimerization does not affect the target peptide binding *in vitro* (*8*). Structures of PDZ, PDZ-like, PDZ-PDZ dimer domains are shown in **Figure 1.2**.

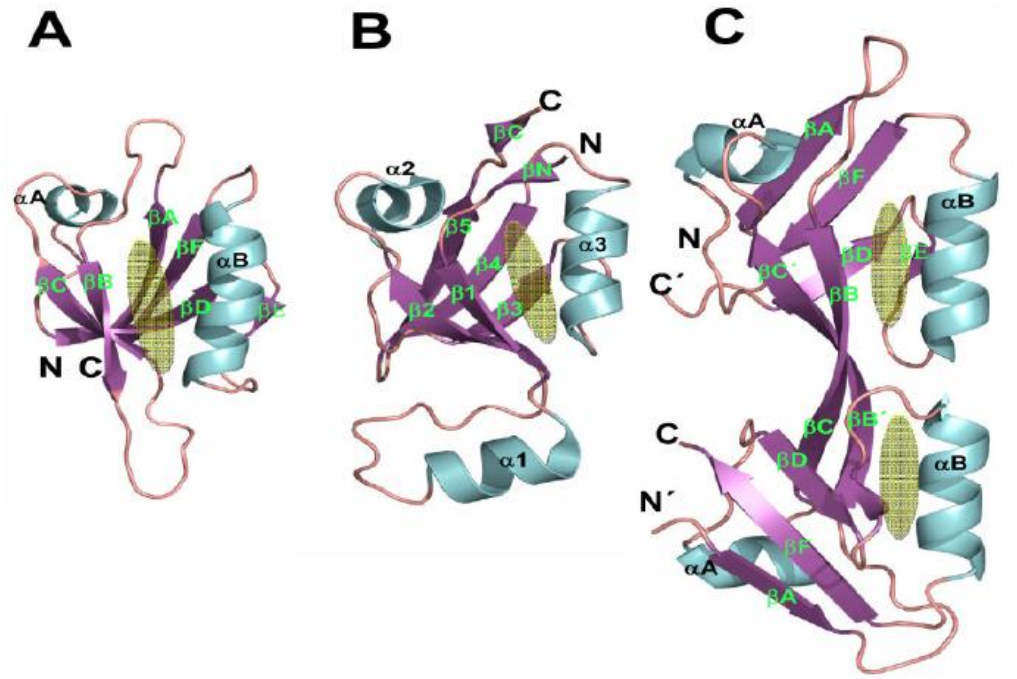

Figure 1.2 Structures of PDZ, PDZ-like, and PDZ-PDZ dimer domains. (A) Ribbon diagram of Dishevilled-1 (Dvl-1) PDZ (PDB code: 2KAW). (B) High temperature requirement A2 (HtrA2) PDZ (PDB code: 1LCY). (C) ZO-1 PDZ2 (PDB code:2RCZ). The binding site of each PDZ domain is shown by a yellow oval. Adapted from reference 8.

Recently, GIP and GIP-glutaminase L complex structures have been reported (*31*). It shows that GIP contains two extra small beta sheets other than canonical PDZ domains (**Figure 1.3**).

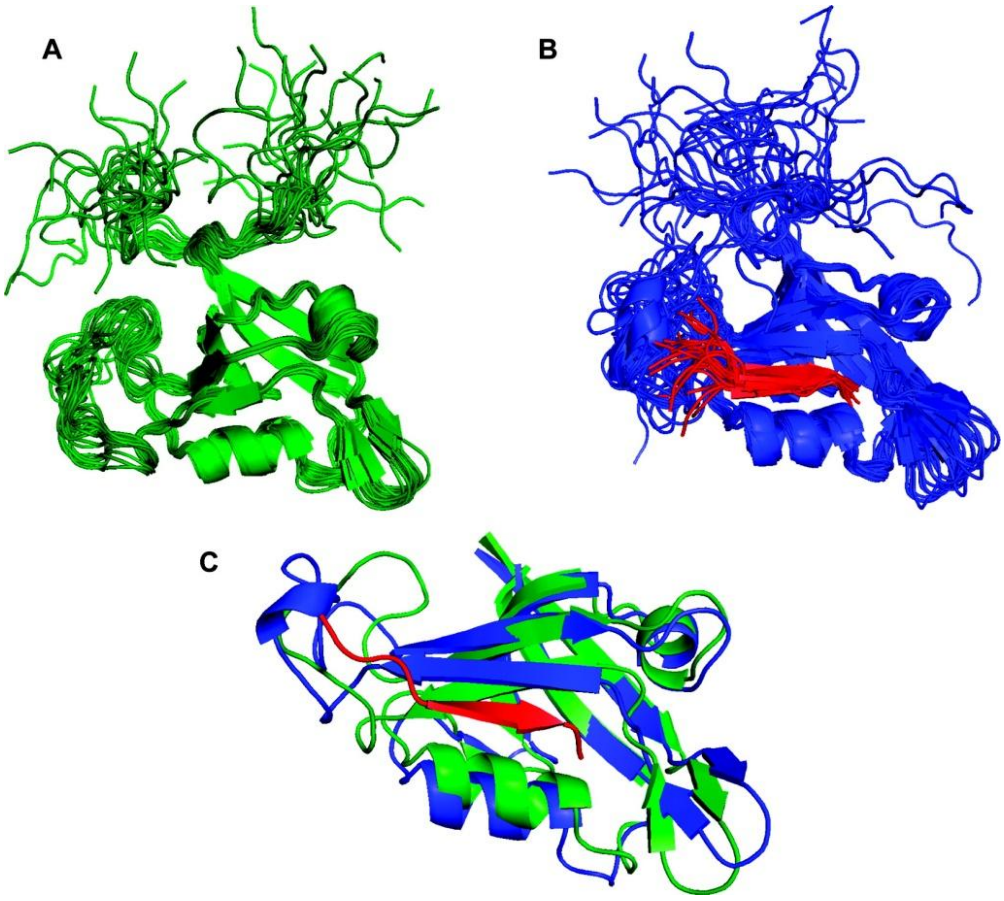

Figure 1.3 Ribbon diagrams of the ensemble of the 20 superimposed lowest-energy structures. (A) Free GIP is colored green. (B) Complexed GIP is colored blue, and glutaminase L peptide is red. (C) Overlays of free GIP and complexed GIP are colored green and blue, respectively, with the glutaminase L peptide colored red. Adapted from reference 31.

1.5 Mechanism of ligand recognition by PDZ domains

It has been reported that the PDZ domain binds the C-terminus of the interacting partner in an elongated groove as an antiparallel β-strand between the second α-helix and the second β-strand, termed the PDZ binding groove (*32, 33)*. The Gly-Leu-Gly-Phe (GLGF) motif is highly conserved among all PDZ domains and is also known as the carboxylate-binding loop. It is located within the βA- βB connecting loop and is important for hydrogen bond coordination of the C-terminal carboxylate group (COO-) *(3)*. The second Gly residue in this motif is conserved among all the canonical PDZ domain proteins while the first Gly residue can be replaced by other amino acid residues (*34*). The amide from the second glycine acts as a hydrogen-bond donor to the carboxylate group from the binding peptide. The second and the forth residues of the motif form a hydrophobic binding pocket of PDZ domain where P_0 of the peptide can sit (*35*). Coordination of the carboxylate group of the binding peptide and the strand to strand interactions between the peptide and the PDZ domain helps the peptide to sit into the PDZ binding groove (*15, 16*). In class I PDZ domains, the side chain of P_{-2} of the binding peptide forms a specific hydrogen bond with the N-3 nitrogen of the histidine residue which is present at the αB:1 position of the PDZ domain while the P_0 residue of the binding peptide positions itself into the hydrophobic binding pocket (*16*). Class II PDZ domains show a slightly different preference for their binding peptides at the P_{-2} position. It recognizes a hydrophobic residue due to the replacement of the histidine residue by a leucine or methionine residue in the αB:1 position of the PDZ domain (*36*).

PDZ-complex structures explain the importance of other than P_{-2} and P_0 residues. These interactions are unique for different PDZ domains. Thus, the ligand binding specificity of PDZ domains are primarily determined by the P_{-2} and P_0 residues of the binding peptide (*15, 16*). There are examples of other residues that are also involved in specific interactions with PDZ domains (*31, 37-39*). Na^+/H^+ exchanger regulatory factor (NHERF) PDZ1-Cystic fibrosis transmembrane conductance regulator (CFTR) interactions are good examples for such binding. Structural characterization of the NHERF PDZ1-CFTR interactions shows that Arg at P_{-1} plays important role in the interaction. It forms two salt bridges with $O^{\varepsilon 2}$ of Glu43 and two hydrogen bonds with the carbonyl oxygen of Asn22 (*39*). This kind of interaction is unusual for PDZ domain binding as generally P_{-1} residue is not important for PDZ domain interactions. **Figure 1.4** shows how a PDZ domain binds to its target proteins. It also shows the mode and mechanism of interaction at the molecular level.

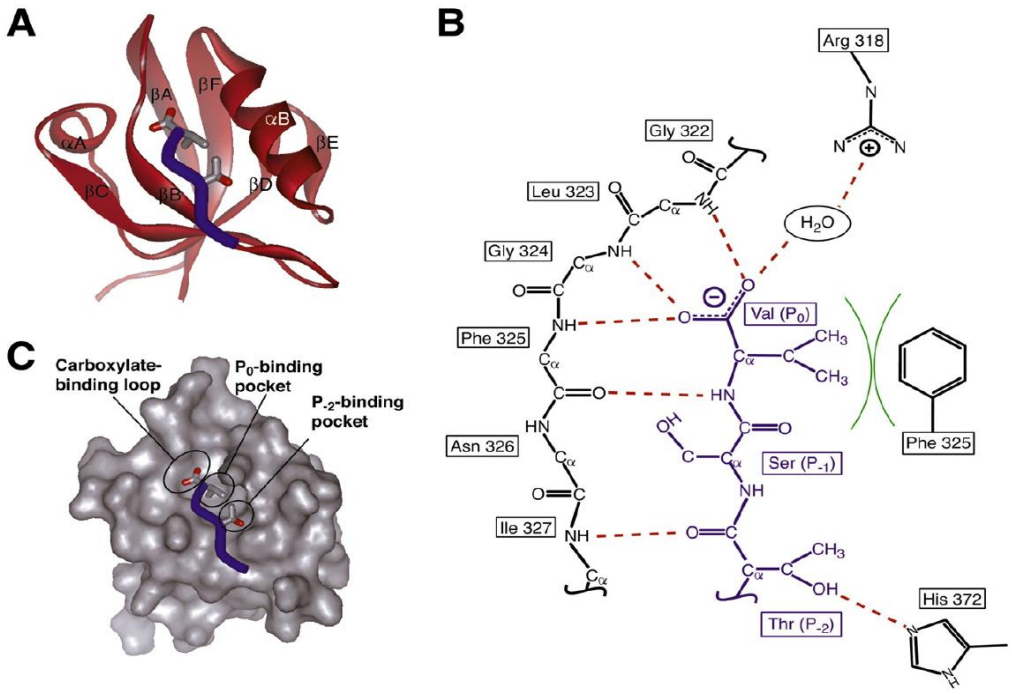

Figure 1.4 Structure of the PDZ domain and mechanism of peptide recognition. (A) Ribbon diagram of PSD-95 PDZ domain 3 (residues 306-394, shown in red) with a bound peptide (NH_2-KQTSV-COOH, shown in blue). Names of the β-strands and α-helices are indicated. The side chains of the peptide P_0 residue (val) and P_{-2} residue (Thr) are shown in stick form, as is the terminal carboxylate. (B) Diagram of the peptide-binding pocket. Residues in the PDZ-domain-binding pocket are shown in black; the peptide is shown in blue. Hydrogen bonds are drawn as red dotted lines, and hydrophobic packing is indicated by green arcs. (C) Solvent-accessible surface representation of the structure shown in (A) (probe radius=1.4 Å). The peptide is drawn as in (A), and key binding pockets are indicated by circles. Adapted from reference 15.

1.6 Protein-protein interaction detection methods

There are various methods available to study PDZ domain mediated interactions. NMR and X-ray crystallographic methods are used to find interactions at the molecular level. Several proteomic and protein arrays methods are available to determine new binding partners and interactions (*8*). These methods include yeast two-hybrid (Y2H), co-immuno precipitation (Co-IP), protein microarray, and screening peptide libraries. These methods have been widely used for studying PDZ domain interactions (*40-47*).

The Y2H method is used to find protein binding partners. To find PDZ domain binding partners, the C-terminal motif of the target protein is sub-cloned into a bait vector containing a DNA-binding domain and the PDZ domain is sub-cloned into the prey vector that contains the activation domain. Binding will cause expression of both fusion proteins in the same yeast cell that leads to transcriptional activation of a reporter gene (*44*). The major disadvantage of this method is that results are not always accurate as this method shows a high rate of false positives and false negatives (*42-43*). It has been reported that a newly made yeast two-hybrid array was used for peptide binding domains like PDZ domains and novel interactions were discovered rapidly. This method was used to find five PDZ domain interactor peptides (*40*). The Co-IP method is commonly used to confirm the interactions that were already found by Y2H , immunofluorescence co-localization (*8-50*), or phage display methods (*8, 41, 44, 46, 51-54*). Other biophysical techniques such as Circular Dichroism (CD), intrinsic fluorescence spectroscopy, Nuclear Magnetic Resonance Spectroscopy (NMR), fluorescence polarization can be used to confirm or provide additional data on potential interactions.

PDZ domain arrays are also used to determine interacting partners for a PDZ domain. In this case PDZ domains are expressed with different tags such as the His- or S-tag (*48-50*). It has been found that 20% of G protein-coupled receptors (GPCRs) proteins in the human genome have PDZ domain binding sites and PDZ domain arrays are being used to identify the GPCR protein partners (*8, 48-50*). There are two useful examples of different proteins identified by proteomic arrays with PDZ domains. An interaction between the purinergic P2Y1 receptors (P2Y1R) and Na^+/H^+ exchanger regulatory factor 2 (NHERF-2) proteins was obtained using this method (*48*). It can be an alternative to yeast two-hybrid library screening for protein-protein interaction approach. The PDZ protein NHERF-2 associates with a definite motif on the $P2Y_1R$ CT and regulates P2Y1R functional activity in a cell-specific manner. This is an example of GPCR regulation by a receptor-associated scaffold protein. These data can be used for development of antithrombotic drugs and another pharmaceuticals for the treatment of human diseases (*48*). The C termini of two different proteins (metabotropic glutamate receptor 1 (mGluR1) and metabotropic glutamate receptor 5 (mGluR5) from the glutamate receptor family were also screened against a PDZ domain proteomic array (*49*). These proteins specifically bind to the second PDZ domain of NHERF-2. Mutation to the mGluR5 C-terminus disrupts the interaction with the PDZ domain (*49*). Full-length mGluR5 interacts with full-length NHERF-2 in cells, as confirmed by co-immunoprecipitation and confocal microscopy experiments. Co-expression of both proteins causes GluR5-mediated cell death (*49*). Thus, the PDZ domain containing protein NHERF-2 regulates mGluR5 signaling (*49*). Another important approach for identification of interacting partners for a PDZ domain is phage display library screening.

1.7 Phage display technology

Phages are viruses which infect bacterial cells. Phage display is a commonly used method for displaying of combinatorial peptide libraries (*55*). Since PDZ domains recognize short motifs of their target proteins, combinatorial phage display libraries can be used for identification of the binding peptide sequences on PDZ target proteins. Phage display libraries are mixtures of genetically re-engineered phage particles. The viral capsid is composed of several coat proteins which can be genetically modified to display peptide sequences at various locations on the phage (*57-58*). It is possible to generate peptide libraries with up to 10^{11} random sequences on the viral particles by inserting randomized DNA sequences into the specific coat proteins (*56-58*). Filamentous phages such as M13 and fd are used frequently as vectors to create phage display libraries. Filamentous phages are about 1 μm long and 7 nm in diameter. The particles look like flexible rods and contain a helical 50 amino acid residue long major coat protein pVIII (*59, 60*). There are 2700 copies of this protein that encapsulate a single-stranded viral DNA (*58*). The particle has two tips that are composed of minor coat proteins. One side of the particle contains the minor coat proteins pIII and pVI and the other side of the particle has minor coat proteins pVII and pIX (*59*). Phage-display systems are classified based on how it displays coat proteins. Foreign peptides can be fused to any of coat proteins, pIII, pVIII or pVI. In the majority of phage display libraries foreign peptides are fused to pVIII or pIII coat proteins. Only short peptides (up to 9 amino acids) can be displayed on every copy of pVIII. To construct pVIII libraries with longer peptides, phage vectors with two genes (88 type vector) encoding for the protein are used. A different type of vector is also known which is called 8+8 system vector with two VIII

genes are on different genome. The wild-type gene in on the helper phage and the recombinant gene is present on a phagemid. Phagemid contains a filamentous phage replication origin. It only infects cells when helper phages are added (59). **Figure 1.5** shows filamentous phage structure.

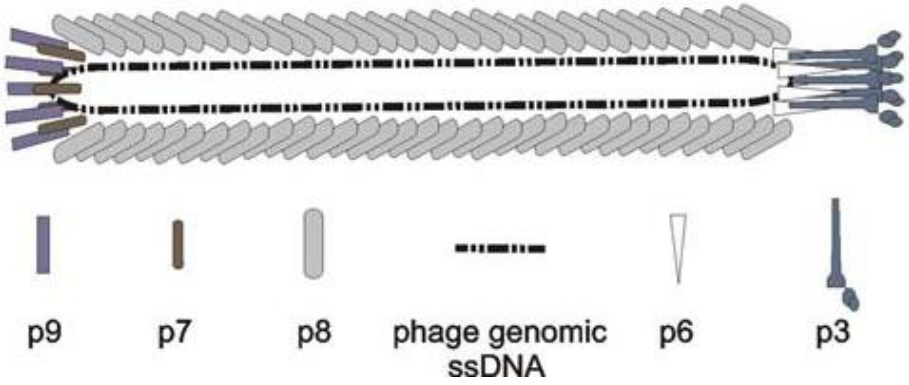

Figure 1.5 Filamentous phage structure. Filamentous phages are rod-like viruses with a circular ssDNA genome. The tube-like capsid is composed of several thousand copies of tightly packed major coat protein [gene VIII product (p8)], capped by five copies of p3 and p6 on one end and five copies of p7 and p9 on the opposite end of the phage particle. Adapted from reference 60.

Phage libraries of random peptides can be used to identify target peptides for a PDZ domain (*55*). The process involves several rounds of selection before the final selection of the binding peptide and sequencing its DNA (*59, 61, 62*). Briefly, the whole procedure is described as follows. First the phages are allowed to bind to the target in the phage selection process. Then, the unbound phages are washed away and bound phages are eluted. This process is repeated several times to enrich for fittest binding peptides to the target. After 3-4 rounds of selection, phage DNAs are sequenced and phage clones characterized individually. The amino acid sequences of the peptides that bind the target protein are determined by the corresponding coding sequence in the viral DNA. **Figure 1.6** shows steps in the phage display selection procedure.

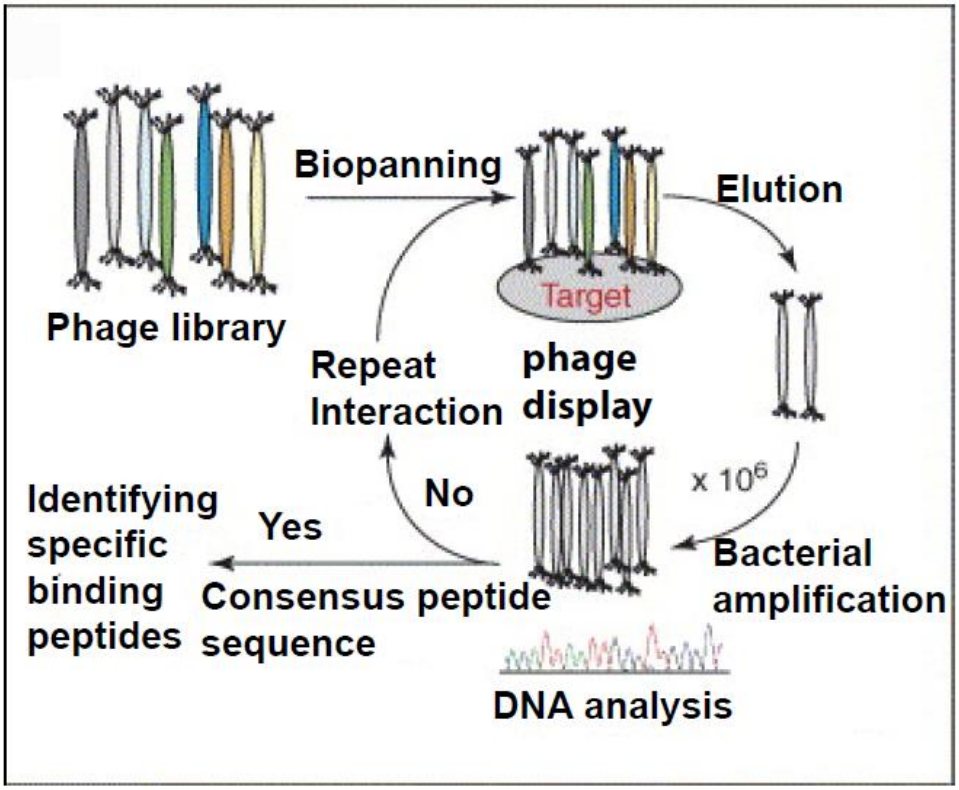

Figure 1.6 Phage display selection process. Sequential steps involved in selection of target-binding peptides are shown. Adapted from reference 60.

Because PDZ domains recognize the C-terminus of their target protein, several C-terminal phage displayed libraries have been used to find PDZ domain interacting partners (*55*, *63-67*). It has been found from the specificity profiling of 91 point mutants of a model PDZ domain that the binding site is highly specific and binding specificity can change under mutation of residues (*66*). The 3rd PDZ domain of the protein tyrosine phosphatase BAS-like (PTP-BL) binds to C-terminal phage displayed peptides as indicated by the phage binding assay. The ligand binding is governed by two cystein residues at the P_{-4} and P_{-1} positions and a valine residue at the P_0 position. Addition of reducing agent such as dithiothrietol interrupts binding indicating that the interaction is disulfide-bridge-dependent and a site-directed mutagenesis study shows that PTP-BL PDZ3 displays a preference for cyclic C-terminal ligands (*67*).

There are reports that show the PDZ domain of the Discs Large protein binds selected peptides with the consensus carboxyl terminal motif Glu-(Ser/Thr)-X-(Val/Ile) (where X represents any amino acid). On the contrary, PDZ domains of LIN-2, p55, and Tiam-1, select peptides with hydrophobic or aromatic side chains at the carboxyl terminal three residues. To explain this anomaly, the structure of the complex was determined and characterized (*63*). It was shown that the PDZ2 of tyrosine phosphatase PTP-BL protein that interact with class III-type ligands was modulated by the presence of PDZ1 while screening a random C-terminal peptide lambda phage display library. A long-range allosteric effect was seen, which explains the altered PDZ2 binding choice (*65*). There are also reports of internal motif recognition by PDZ domains (*4*, *24*, *68-70*). These internal motifs were found through phage display libraries or Y2H methods or by rational design. The use of phage display libraries to find internal motifs for a PDZ domain is still

uncommon and to date there are only three crystal structures available that show internal motif recognition by PDZ domains (*4*,*24*,*68*). There are no NMR structures available for internal motif recognition. The PDZ domain of the partitioning defective 6 (Par-6) protein interacts with an internal motif from the protein associated with Lin seven 1 (Pals1) protein (*68*). Structural investigation shows that the PDZ ligand-binding site is deformed to allow internal ligand binding to bypass the requirement for chain termination (68). However, in the case of nNOS-syntrophin internal mode of interaction shows that a sharp turn is seen which allows the ligand to bypass the steric requirement forced by the carboxylate group (*4*). **Figure 1.7** shows modes of PDZ domain C-terminal and internal motif recognition. Along with these, there are only a few other reports of PDZ domain internal motif recognition.

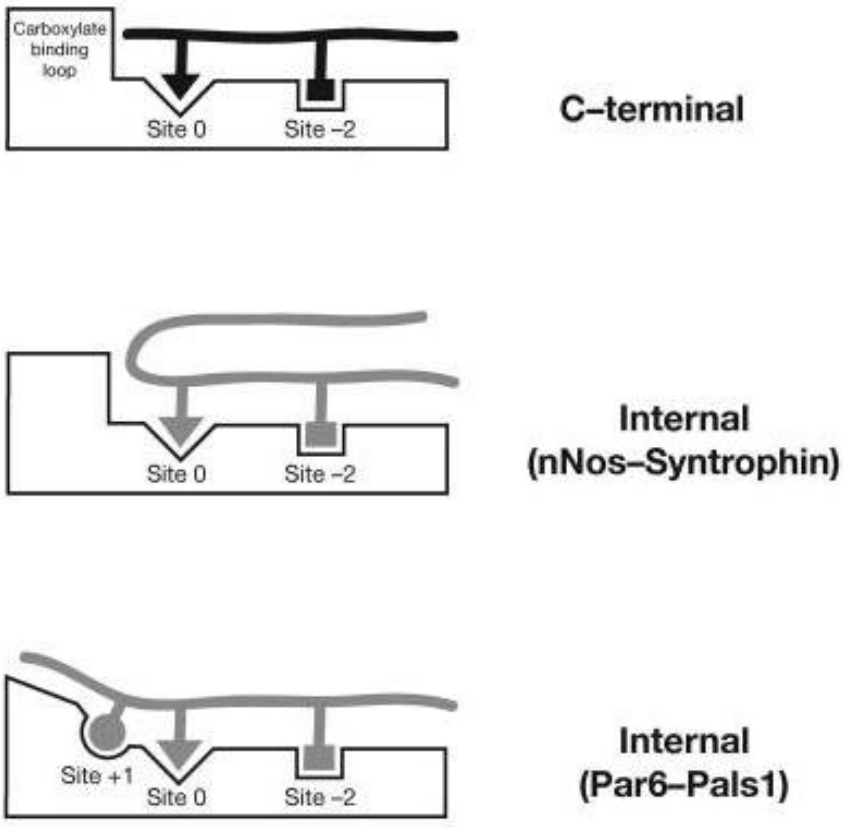

Figure 1.7 Modes of PDZ C-terminal and internal recognition. In PDZ C-terminal ligand recognition the carboxylate-binding loop enforces C-terminal binding by preventing extension past the P_0 residue. In the β-finger internal PDZ recognition mode of recognition, used by neuronal nitric oxide synthase (nNOS)-syntrophin and presumably disulfide containing ligands, a sharp turn in the ligand allows it to bypass the steric requirement imposed by the carboxylate-binding loop. The Pals1-Par-6 PDZ interaction represents a new type of internal interaction in which the carboxylate-binding loop is deformed to allow for residues past the P_0 residue. An interaction with the P_{+1} residue is critical for this mode of recognition. This figure is adapted from reference 68.

It has been found that Dishevelled (Dvl) PDZ domain binds to the internal motif of Inhibition of the Dvl and Axin complex (Idax) protein containing RKTGHQICKFRKC (KTXXXI) motif which is important for the interaction (*69*). This type of binding gives higher dissociation constant values compared to the canonical C-terminal motif recognition (*69*). The syntrophin PDZ domains bind to the C-terminal peptide sequences of its target proteins. Using a phage display library, it has been found that the syntrophin PDZ domain also binds cyclic peptide ligands which were displayed at the N-terminus of protein III of bacteriophage M13 (*70*). These peptides are very specific to the syntrophin PDZ domain and bind at the same C-terminal binding site. The cyclic peptides were formed through the cysteine residues. When the disulfide bond was disrupted, no interaction was seen (*70*), but the peptide conformation after binding was not determined.

1.8 PDZ domains are drug targets

It is well known that PDZ domains are involved in protein-protein interactions in various signaling pathways. Mutations in the PDZ domain can cause deregulation of these pathways and lead to various diseases. It has also been found that PDZ domains are involved in various types of cancers (*71-78*). PDZ domains recognize G protein-coupled receptors and ion channels, which are among the most important classes of drug targets for the pharmaceutical industry. Thus, developing small molecule inhibitors to compete with PDZ targets might be useful for the development of pharmaceutical agents (*71*). It has been found that the C-terminal last nine residues of the NR2B methyl-D-aspartate receptors (NMDAR) subunit affects the signaling pathways downstream of NMDAR activation, on *in vitro* excitotoxicity, and on *in vivo* ischemic brain damage. Interfering

with the interaction between NMDARs and postsynaptic density protein 95 (PSD-95) may interrupt signaling downstream from NMDARs leading to neuronal death. The study concludes that targeting NMDAR–PSD-95 interaction is a promising strategy for treating strokes in the future (*74*). Designing cancer therapeutics such as peptides, small molecules, and peptidomimetics based on the understanding of PDZ interactions might have potential therapeutic advantages (*8, 79-89*). There are several examples of PDZ domains as targets for various diseases. The Dishevelled (Dvl) PDZ domain is involved in the Wnt signaling pathway and abnormal activation of this process causes different types of cancers. Small molecules sulindac and sulindac sulfone bind to the Dvl PDZ domain and cause inhibition of abnormal Wnt signaling by obstructing the Dvl PDZ domain (*79*). It was also found that the NHERF1/EBP50 (Na^+/H^+ exchanger regulatory factor 1/ezrin-radixin-moesin-binding phosphoprotein 50) complex consists of two PDZ domain that regulate G protein-coupled receptors. NHERF1 was also found to interact with the phosphatase and tensing homolog (PTEN) tumor supressor. A drug has been developed that specifically affect the expression of NHERF1 in PTEN-positive breast cancer cells (*82*).

1.9 Basics of NMR

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful technique which can be used to determine structure for small molecules as well as macromolecules such as proteins and nucleic acids. In NMR spectroscopy, experiments are performed on the nuclei of atoms, not the electrons. Information is obtained about a particular nucleus from its chemical shift environment. Nuclei of certain isotopes contain spin. In several atoms

(such as ^{12}C) these spins are paired against each other. As a result, the nucleus of the atom has no overall spin. On the contrary, in some atoms (such as ^{1}H and ^{13}C) the nucleus does possess an overall spin. The net nuclear spin can be determined by number of neutrons and number of protons. If both numbers are even, then the nucleus has no spin. If the total number is odd, then the nucleus has a half-integer spin (1/2, 3/2 etc.). If both of them are odd, then the nucleus has an integer spin (1, 2 etc.). Thus, the overall spin (I) is important. A nucleus with spin 1/2 will have 2 possible orientations. In the absence of an external magnetic field, these orientations are of equal energy. If a magnetic field is applied, then the energy levels split **(Figure 1.8)**.

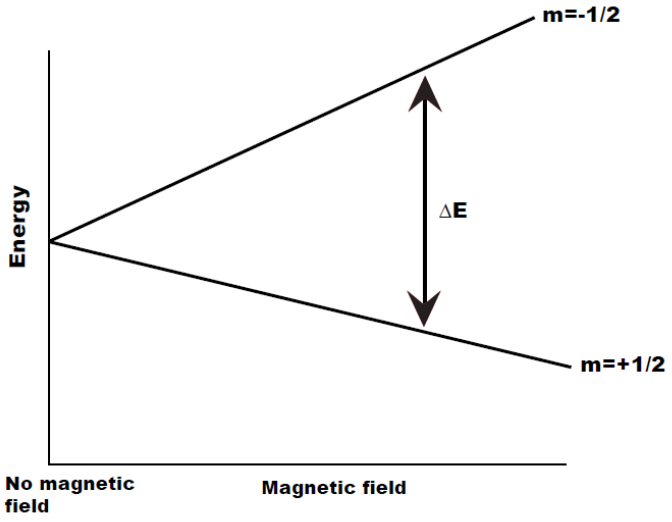

Figure 1.8 Energy splitting as a function of magnetic field strength.

When the nucleus is in a magnetic field, the initial populations of the energy levels are determined by the Boltzmann distribution. The lower energy level will contain slightly more nuclei than the higher energy level. It is possible to excite these nuclei into the higher level with electromagnetic radiation. The frequency of radiation required is determined by the difference in energy between the energy levels. Chemically different nuclei in a molecule do not experience the external magnetic field identically. Electrons shield the nucleus and in that way reduce the effective magnetic field. As a result, energy of a lower frequency is required to induce the transitions between spin states. On the contrary, when electrons are withdrawn from a nucleus, the nucleus is deshielded and feels a stronger magnetic field requiring more energy to achieve resonance. Normally, nuclei attached to the electron withdrawing groups tend to resonate at higher frequencies. The frequency of the resonance expressed with reference to a standard compound, which is defined to be at 0 ppm (for most experiments this is tetramethylsilane (TMS)). This is called chemical shift and is expressed in parts per million (ppm). It is independent on frequency of the spectrometer. The process of determining the frequencies of the nuclei in a particular molecule is called resonance assignment.

With the development of computational power, it is possible to obtain multidimensional spectra. In one dimensional NMR spectrum a single chemical shift coordinate along X axis is available. In two dimensional experiments, both the X and the Y axis have chemical shift scales. The 2D spectra looks like a map. By looking at the peaks and matching them to the X and Y axes, information of chemical shifts of the residues obtained. Several 2D experiments are available to obtain useful information of a protein structure. 2D NMR experiments can easily be expanded to multidimensional

spectroscopy by the appropriate combination of several 2D NMR experiments. In the 3D experiments, both X, Y, and Z axis have chemical shift scales. These experiments are widely used for protein structure determination.

1.9.1 Protein NMR

NMR spectroscopy is one of the two high resolution techniques to determine the three-dimensional structures of protein, DNA, and RNA. This technique also provides information regarding protein secondary structures, dynamics, structural and functional aspects of proteins-ligand binding. Biomolecular NMR spectroscopy is a key technique for studying the mechanisms of cellular function because it can determine both the structure and dynamics of protein or complexes in solution under near physiological conditions. NMR is a versatile technique that not only provides protein structure at atomic resolution, but also useful information regarding protein motions in solution. NMR also plays an important role in studying weak and transient interactions, which are associated with biological functions. Three-dimensional NMR is used for backbone resonance assignments of a protein, and there are several experiments available. Using these experiments, it is possible to identify and assign all the amino acid residues in a protein and complete backbone assignments. The first step in protein NMR structure determination is to have all the amino acid residues assigned in the heteronuclear single quantum coherence (HSQC) spectrum which is also called the finger print region of a protein. Once the HSQC is completely assigned, it is possible to visualize the change in chemical shift positions of each residue involved in the binding. This is an important tool to find out protein-ligand interactions. This process is useful for ligand binding studies, as

it will show which amino acid residues are involved in the binding or in conformational change if there is any. Additionally, backbone resonance assignments can provide information on secondary structure of a protein.

The backbone resonance assignments by solution NMR can be divided into several parts: (1) sample preparation, (2) recording and processing of NMR data, and (3) sequential resonance assignments.

For biomolecular NMR structural studies, sample preparation is very important. Milligram quantities of pure protein are required for all 3D NMR heteronulcear experiments data collection. For any multidimensional heteronuclear NMR experiments uniform labeling of the protein with ^{15}N, ^{13}C and/or ^{2}H is required. As ^{13}C, ^{15}N, and ^{2}H have very low natural abundance, isotope labeling is required during protein expression. To achieve this, $^{15}NH_4Cl$ and ^{13}C-glucose are used as the sole nitrogen and carbon sources respectively. Deuterium oxide (D_2O) is used in the minimal media instead of H_2O in the culture while expressing protein in *E. Coli* host. There are three pairs of 3D NMR experiments used for backbone assignments. These are: HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA, HNCO, and HN(CA)CO. All these experiments correlate backbone amide proton, and nitrogen atoms to Cα, Cβ, and CO atoms. For example, HNCACB correlates each amide proton and nitrogen with the Cα and Cβ chemical shift of its own residue and of the preceding residue in the sequence, while the CBCA(CO)NH correlates each amide proton and nitrogen only with the Cα and Cβ chemical shifts of the previous residue (**Figure 1.9**). The HNCO and HN(CA)CO work in a similar manner, but they give the chemical shifts of the carbonyls carbons rather than alpha carbons. The HNCA, HN(CO)CA contains the chemical shifts of the alpha carbon of its own and the previous

residue, respectively (**Figure 1.10**). All of these experiments should be obtained to find the complete backbone assignments of the protein.

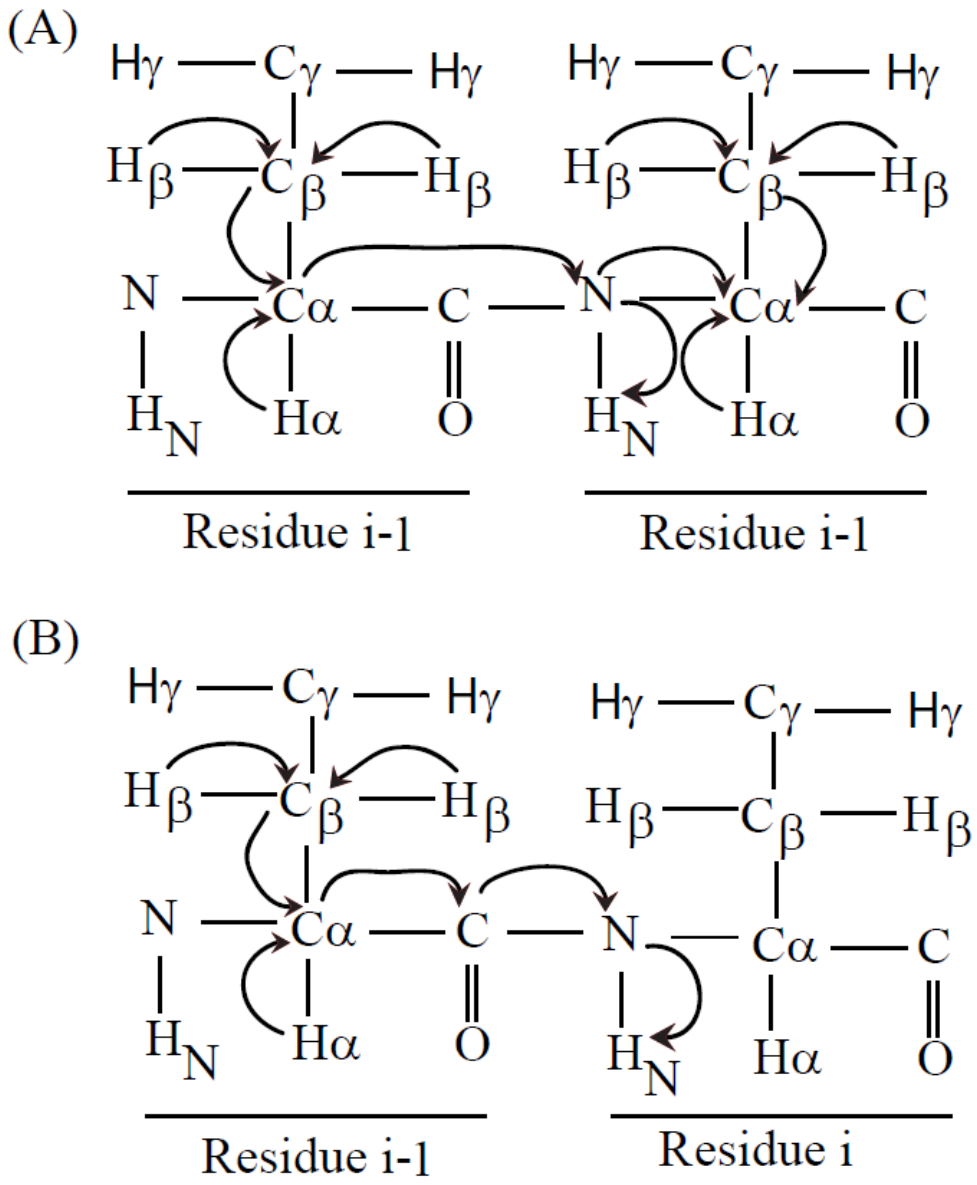

Figure 1.9 Magnetization coherence transfer schemes of 3D NMR experiment for protein backbone experiments. (A)HNCACB (B) CBCACONH.

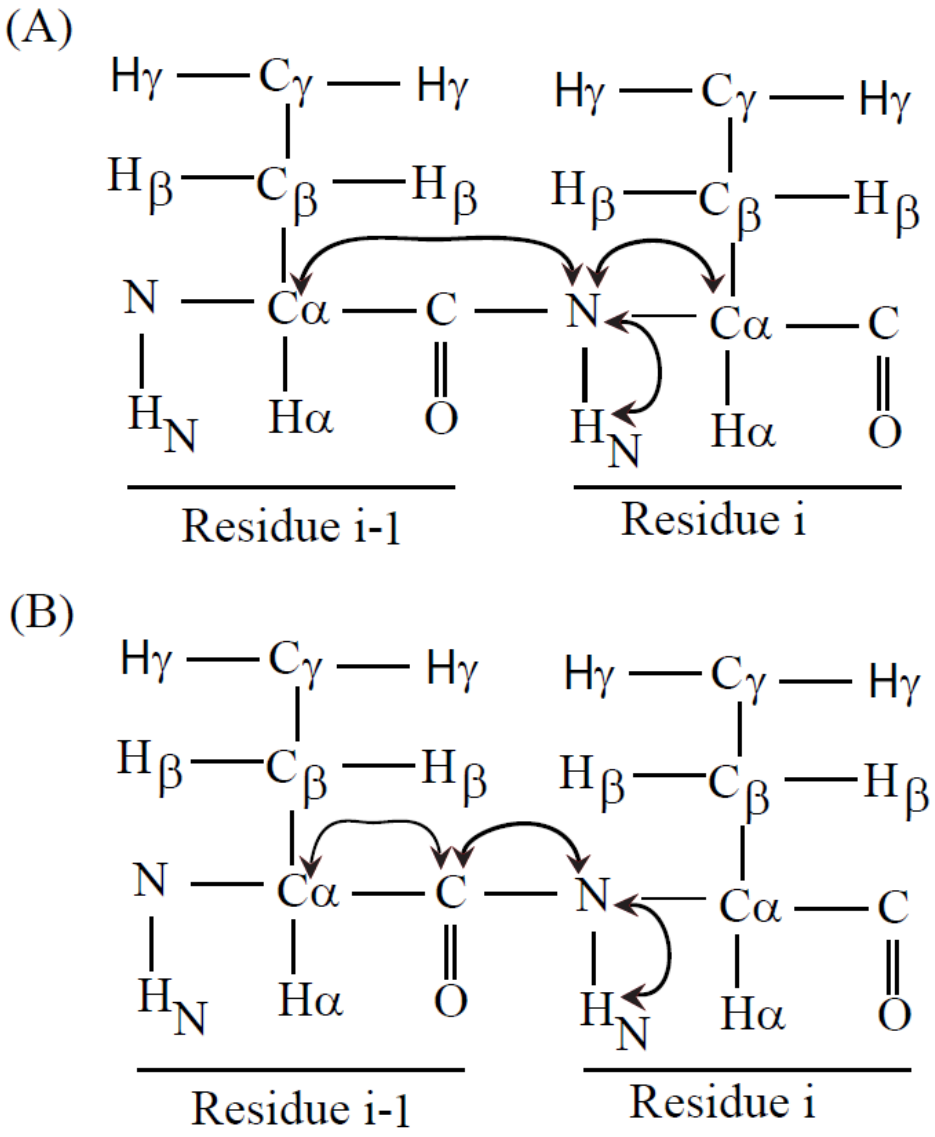

Figure 1.10 Magnetization coherence transfer schemes of 3D NMR experiment for protein backbone experiments. (A)HNCA (B) HN(CO)CA.

1.10 NMR as a tool for mapping protein-ligand binding interfaces

NMR is a powerful method and is regularly used for mapping the interaction of a protein with ligands, such as peptides, proteins, nucleic acids, small molecules or metal ions (90). The finger print region in 2D HSQC experiment is very sensitive to change in temperature, pH, salt, addition of ligand etc. Sequential titration of a ligand to the ^{15}N labeled protein can be monitored by 2D HSQC experiment. The chemical shift perturbation of each residue due to ligand binding can be monitored. Thus ligand binding site may be determined by mapping the chemical shift perturbation to the protein primary sequence. Based on the chemical shift change, a chemical shift perturbation plot can be generated which can provide information on the regions of protein involved in the binding. The disadvantage of this method is that amino acid residues not directly involved in ligand binding may have significant perturbation due to conformational change.

The ModelTitr program is used to obtain dissociation constant values for each of the residues in the HSQC spectra of a protein upon ligand binding according to the chemical shifts associated with ^{1}H or ^{15}N (*91, 92*). NMR is often used for the determination of dissociation constants of a protein to its substrate. In this method, equilibrium association or dissociation constant values are calculated based on nonlinear least-squares fitting of the chemical shifts data versus ligand concentration assuming protein to ligand ratio of 1:1 (*91*). Using this method either the amide proton or nitrogen or a weighted grouping of the two types of shifts can be analyzed.

1.11 Protein dynamics and PDZ domains

In NMR experiments, the signal corresponds to the sample coming back to the equilibrium state from the excited state. This process is called relaxation. Backbone dynamics measurements for a protein include three different processes: ^{15}N longitudinal relaxation rate (R_1) or ($1/T_1$), ^{15}N transverse relaxation rate (R_2) or ($1/T_2$), and the heteronuclear nuclear overhauser effect (^{1}H-^{15}N NOE).

1.11.1 ^{15}N longitudinal relaxation (T_1)

Longitudinal or spin lattice relaxation is the process by which the excited magnetization state returns to its ground state along the Z-axis, which is the same as the external magnetic field. This is called spin lattice relaxation because energy exchanges between spin and lattice. It is defined as the time required for the longitudinal magnetization to recover ~63% [1-(1/e)] of its initial value after being flipped into the magnetic transverse plane by a 90° radiofrequency pulse (**Figure 1.11**). The relaxation time, T_1 is dependent on the gyromagnetic ratio of the nucleus and the mobility of the lattice. The recovery of spin lattice can be expressed as an exponential equation: $M_t = M_0 [1-\exp(-t/T_1)]$.

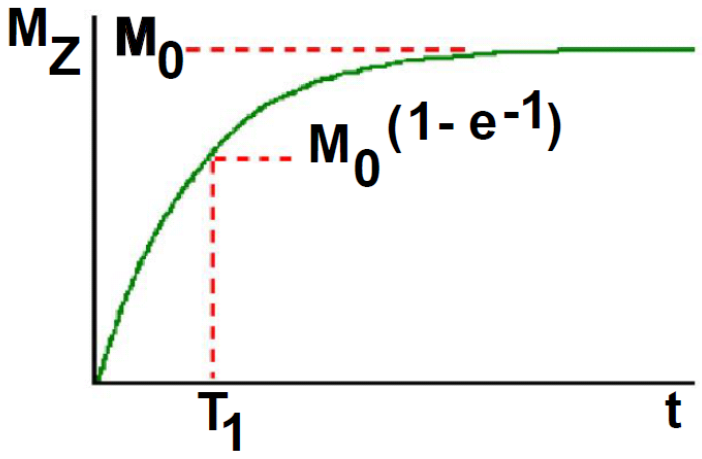

Figure 1.11 T_1 relaxation curve.

1.11.2 ^{15}N Transverse relaxation (T_2)

Transverse relaxation is the process by which the excited magnetization vector decays perpendicular to the direction of external magnetic field along X-Y plane. This process is also called spin-spin relaxation decay. It is defined as the time it takes for the magnetic resonance signal to reach 37% (1/e) of its initial value after its generation by tipping the longitudinal magnetization towards the magnetic transverse plane (**Figure 1.12**). *T_2* relaxation generally proceeds more rapidly than T_1 recovery. The magnitude can be expressed by the following exponential eq. $M_t = M_0 \exp(-t/T_2)$, where t is time.

Thermodynamically these processes are different. T_1 is an enthalpy-driven process while T_2 is an entropy-driven process.

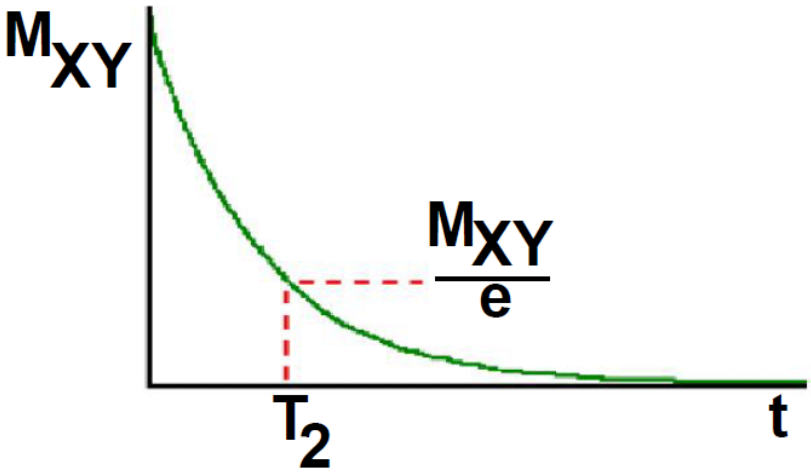

Figure 1.12 T_2 relaxation curve.

1.11.3 ^{1}H-^{15}N Heteronuclear NOE

Nuclear overhauser effect (NOE) is defined by the change in intensity of one resonance when the spin transition of another is perturbed from its equilibrium positions. This perturbation can be saturation or inversion. The two spin states interact through dipole-dipole interaction. For better understanding, two spin system I_1 and I_2 and two different states α and β as energy level diagram are shown in **Figure 1.13**.

According to the **Figure 1.13**, perturbation of one spin (i.g. I_1) causes change in populations of the α and β state, but relaxation will force I_1 back to the equilibrium state. In this mechanism spin I_1 will relax without affecting spin I_2. On the other hand, when the I_1 polarization going back to the equilibrium from saturation, the Wo system will force the neighboring spin to move away from the equilibrium position and will cause a decrease in the α, β population difference. As a result, there will be decrease in I_2 intensity, after a 90^{o} pulse. This is called a negative NOE effect. On the contrary, the W_2 mechanism will force the population difference of the unperturbed spin I_2 to increase. This result causes an increase in signal intensity which is called positive NOE effect. Generally, the experiment is repeated with and without saturation. The spectrum without saturation is subtracted from the irradiated spectrum. The small intensity changes observed are due to the NOE effect.

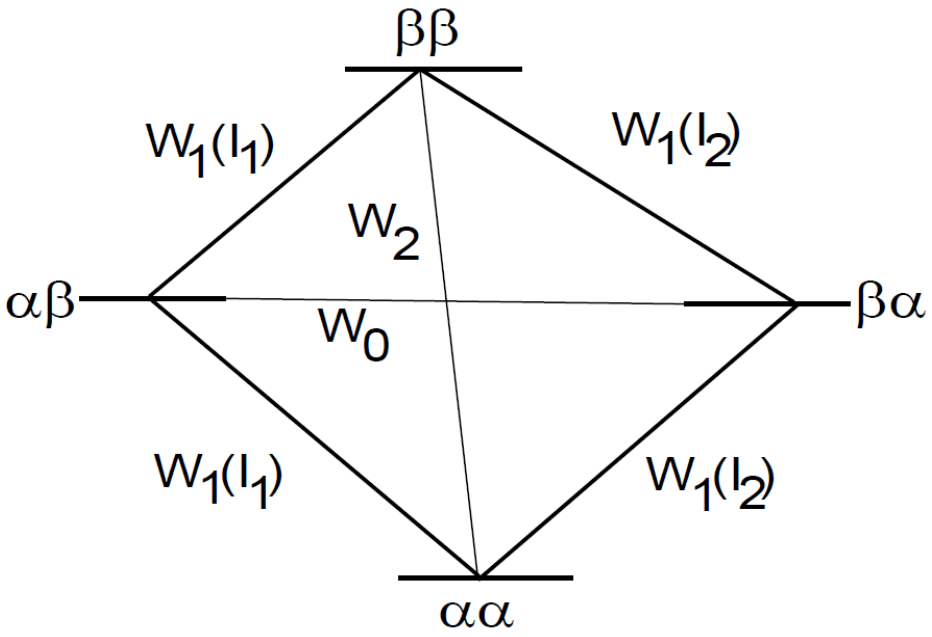

Figure 1.13 Energy diagram for a dipolar-coupled two-spin system. The four states are αα, αβ, βα, and ββ; the zero- single- and double-quantum transitions are represented by W_0, W_1 and W_2, respectively. Drawn according to reference 93.

1.11.4 Dynamics studies of PDZ domains

Binding or release of a ligand by proteins is accompanied by conformational changes, which is a dynamic process (*94, 95*). Thus, to understand protein function, a dynamics study for a particular protein is very critical. Protein flexibility can cause unexpected ligand binding (*95-96*). There are several techniques presently being used for protein dynamics studies among which NMR is a special use for this purpose. NMR covers full range of time scale that is necessary for dynamics studies of a protein (*94*). The spectral density function J(w) is related to three different relaxation rates. It also depends on the S^2 which is termed the order parameter (*97, 98*). The value of the order parameter varies from 1 to 0 which corresponds to completely restricted and completely unrestricted protein motion, respectively. Protein flexibility is important for understanding how a drug molecule works in biological systems, binding orientation, binding kinetics and metabolism. Protein flexibility is also critical for ligand induced binding in a protein for drug design (*96*). There is correlation between a protein's dynamics and its function as shown in **Figure 1.14**.

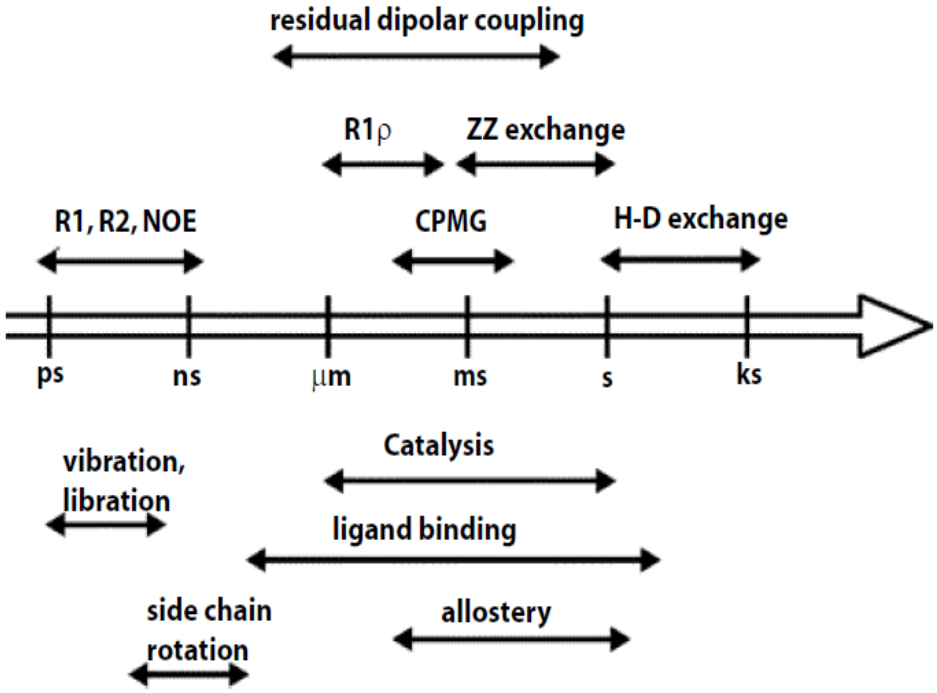

Figure 1.14 Solution NMR techniques cover the complete range of dynamic events in enzymes. Adapted from reference 94.

There are reports available for PDZ domain and PDZ-ligand dynamics (*99*). Generally, dynamic properties of PDZ domains show a regular pattern; secondary structures are more rigid and loop structures are more flexible (*99*). It has been found that dynamic fluctuations in proteins contribute to allosteric signal transduction (*100*). Dynamics behavior of the PDZ2 domain from tyrosine phosphatase PTP-BL upon target peptide binding show that Gly26, which is involved directly in the binding process by forming a hydrogen-bond with the C-terminal carboxyl group of the peptide shows a dramatic decrease in its S^2 value (*101*). The overall dynamics show rigidity upon ligand binding (*101*). All of the PDZ domains found so far show that upon ligand binding the protein becomes more rigid. It would be interesting to find a PDZ domain whose dynamics show that it become more flexible upon ligand binding.

1.12 Objective of this work

PDZ domains are one of the most important protein-protein interaction modules that are involved in localization, clustering of ion channels, and receptors in living entities. These domains are also involved in various signaling pathways which lead to diseases by interacting with its target proteins. PDZ domains interact with the C-termini of their target proteins. GIP is a multifunctional protein containing a class-I PDZ domain. GIP interacts with several target proteins, and the target proteins are involved in various signaling pathways, mostly various cancerous pathways. GIP interacts with them and regulates their pathways. Thus, this protein is involved in various cancerous pathways by interacting with its various target proteins. GIP is primarily composed of a single PDZ domain, and it is a small protein involved in various cancerous pathways. Thus, it is an

excellent target for cancer drug design. To design a successful cancer drug, it is important that we understand the complete network for GIP. To design a drug, it is necessary to characterize all the proteins which are involved in GIP binding to minimize side effects. My thesis is directed towards this goal. To achieve the goal, first characterization of structure, function, and dynamics studies are required. The first part of this thesis (chapter-2) covers over expression, purification, and biophysical characterization of the protein, followed by structural investigation. Then we have done interaction studies of GIP with the C-termini of its several target proteins (chapter-2). We have found new target proteins of GIP. Dynamics studies of GIP and GIP-ligand complexes were carried out to understand the protein motion or flexibility upon ligand binding. The objective of the second part of the thesis explains how GIP can be used as a target for the development of peptide based cancer therapeutics. For that purpose, it was necessary to find the complete network of GIP. So far, GIP has been found to interact with the C-termini of its target proteins. To advance our understanding of GIP-ligand interaction, analysis was extended to find out novel motif peptides by a phage display library screening. Interestingly, all the peptides contain internal binding motifs and are present in several human proteins which are involved in various cancerous pathways. The role of GIP in cancer is still unknown. To learn more about the role of GIP in human cancer, GIP was expressed in human glioma cells. Co-localization studies of GIP with the target peptide were carried out. Finally, one of the peptides was tested in a cell proliferation assay to evaluate its potential use as an anti-cancer therapeutic. The aim of this work is to elucidate a mechanism of interaction of GIP with its target peptides and their biological functions. The obtained data might be useful for understanding how other PDZ domains

function. This study might have broad application in biomedical research and in the design of drugs.

1.13 References

1. Ponting, C. P., and Phillips, C. (1995) DHR domains in syntrophins neuronal NO synthases and other intracellular proteins. *Trends Biochem. Sci. 20*, 102–103.
2. Kennedy, M. B. (1995) Origin of PDZ (DHR, GLGF) domains. *Trends Biochem. Sci. 20*, 350.
3. Sheng, M., and Sala, C. (2001) PDZ domains and the organization of supramolecular complexes. *Annu. Re*V. *Neurosci. 24*, 1–29.
4. Hillier, B. J., Christopherson, K. S., Prehoda, K. E., Bredt, D. S., and Lim, W. A. (1999) Unexpected Modes of PDZ Domain Scaffolding Revealed by Structure of nNOS-Syntrophin Complex. *Science 284*, 812–815.
5. Bhat, M. A., Lzaddost, S., Lu, Y., Cho, K-O., Choi, K-W., and Bellen, H. J. (1999) Discs Lost, a Novel Multi-PDZ Domain Protein, Establishes and Maintains Epithelial Polarity. *Cell 96*, 833–345.
6. Kornau, H. C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science 269*, 1737–1740.
7. Kim, E., Niethammer, M., Rothschild, A., Jan, Y. N., and Sheng, M. (1995) Clustering of Shaker-type K+ channels by interaction with a family of membrane-associated guanylate kinases. *Nature 378*, 85–88.
8. Lee, H-J., and Zheng., J. J. (2010) PDZ domains and their binding partners: structure,specificity, and modification. *Cell Comm. and Signal. 8, 1-18.*
9. Pallen, M. J., and Ponting, C. P. (1997) PDZ domains in bacterial proteins (letter). *Mol. Microbiol. 26*, 411-413.

10. Ponting, C. P. (1997) Evidence for PDZ domains in bacteria, yeast, and plants. *Protein Sci. 6*, 464-468.

11. Liao, D. I., Qian, J., Chisholm, D. A., Jordan, D. B. and Diner, B. A. (2000) Crystal structures of the photosystem II D1 C-terminal processing protease. *Nat. Struct. Biol. 7*, 749-753.

12. Niethammer, M., Valtschanoff, J. G., Kapoor, T. M., Allison, D. W., Weinberg, T. M., Craig, A. M., and Sheng, M. (1998) CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90. *Neuron 20*, 693-707.

13. 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) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science 275*, 73-77.

14. Schultz, J., Hoffmüller, U., Krause, G., Ashurst, J., Macias, M. J., Schmieder, P., Schneider-Mergener, J., and Oschkinat, H. (1998) Specific interactions between the syntrophin PDZ domain and voltage-gated sodium channels. *Nat. Struct. Biol. 5*, 19-24.

15. Harris, B. Z., and Lim, W.A. (2001) Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Sci. 114*, 3219-3231.

16. Roh, M. H., Margolis, B. (2003) Composition and function of PDZ protein complexes during cell polarization. *Am J Physiol Renal Physiol*., *285*, F377-F387.

17. Fanning, A. S., Anderson, J. M. (1996) Protein-protein interactions: PDZ domain networks. *Curr Biol*. *6*, 1385-1388.

18. Lee, H-J., Wang, N. X., Shao, Y., Zheng, J. J. (2009) Identification of tripeptides recognized by the PDZ domain of Dishevelled. *Bioorg. Med. Chem*. *17*, 1701-1708.

19. Lee, H. J., Wang, N. X., Shi, D. L,, and Zheng, J. J. (2009) Sulindac inhibits canonical Wnt signaling by blocking the PDZ domain of the protein Dishevelled. *Angew Chem Int. Ed. Engl. 48*, 6448-6452.

20. Morais Cabral, J. H., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., Liddington, R. C. (1996) Crystal structure of a PDZ domain. *Nature 382*, 649-652.

21. Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J., Lee, A. L. (2009) Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA*. *106*, 18249-18254.

22. Lee, H-J., Wang, N. X., Shi, D. L., Zheng, J. J. (2009) Sulindac inhibits canonical Wnt signaling by blocking the PDZ domain of the protein Dishevelled. *Angew Chem. Int. Ed. Engl. 48*, 6448-6452.

23. Wong, H. C., Bourdelas, A., Krauss, A., Lee, H-J., Shao, Y., Wu, D., Mlodzik, M., Shi, D. L., Zheng J. (2003) Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled. *Mol Cell 12*, 1251-1260.

24. Zhang, Y., Appleton, B. A., Wiesmann, C., Lau, T., Costa, M., Hannoush, R. N., Sidhu, S. S. (2009) Inhibition of Wnt signaling by Dishevelled PDZ peptides. *Nat Chem. Biol. 5*, 217-219.

25. Im, Y. J., Lee, J-H., Park, S. H., Park, S. J., Rho, S. H., Kang, G. B., Kim, E., Eom, S. H. (2003) Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. *J Biol Chem*. *278*, 48099-48104.

26. Im, Y. J., Lee, J-H., Park, S. H., Park, S. J., Rho, S. H., Kang, G. B., Kim, E., Eom, S. H. (2003) Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. *J Biol Chem*. *278*, 48099-48104.

27. Utepbergenov, D. I., Fanning, A. S., Anderson, J. M. (2006) Dimerization of the scaffolding protein ZO-1 through the second PDZ domain. *J Biol Chem 28*, 24671-24677.

28. Wu, J., Yang, Y., Zhang, J., Ji, P., Du, W., Jiang, P., Xie, D., Huang, H., Wu, M., Zhang, G., Wu, J., Shi, Y. (2007) Domain-swapped dimerization of the second PDZ domain of ZO2 may provide a structural basis for the polymerization of claudins. *J Biol Chem 282, 35988-35999.*

29. Fanning, A. S., Lye, M. F., Anderson, J. M., Arnon, L. (2007) Domain swapping within PDZ2 is responsible for dimerization of ZO proteins. *J. Biol. Chem. 282*, 37710-37716.

30. Chen, J., Pan, L., Wei, Z., Zhao, Y., Zhang, M. (2008) Domain-swapped dimerization of ZO-1 PDZ2 generates specific and regulatory connexin43-binding sites. *EMBO J. 27*, 2113-2123.

31. Zoetewey, D. L., Ovee, M., Banerjee, M., Bhaskaran, R., and Mohanty, S. (2011) Promiscuous binding at the crossroads of numerous cancer pathways: insight from

the binding of glutaminase interacting protein with glutaminase L, *Biochemistry 50,* 3528-3539.

32. Zhang, Q., Fan, J-S., and Zhang, M. (2001) Interdomain Chaperoning between PSD-95, Dlg, and Zo-1 (PDZ) Domains of Glutamate Receptor-interacting Proteins. *J. Biol. Chem. 276,* 43216–43220.

33. Banerjee, M., Huang, C., Marquez, J., and Mohanty, S. (2008) Probing the Structure and Function of Human Glutaminase-Interacting Protein: A Possible Target for Drug Design. *Biochemistry 47,* 9208–9219.

34. Roh, M. H., Margolis, B. (2003) Composition and function of PDZ protein complexes during cell polarization. *Am J Physiol Renal Physiol., 285,* 377-387.

35. Morais Cabral, J. H., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., Liddington, R. C. (1996) Crystal structure of a PDZ domain. *Nature, 382,* 649-652.

36. Daniels, D. L., Cohen, A. R., Anderson, J. M. and Brünger, A. T. (1998) Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition. *Nat. Struct. Biol. 5,* 317-325.

37. Doyle, D. A., Lee, A., Lewis, J., Kim, E., Sheng, M. and MacKinnon, R. (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell 85,* 1067-1076.

38. Tochio, H., Zhang, Q., Mandal, P., Li, M. and Zhang, M. (1999) Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. *Nat. Struct. Biol. 6,* 417-421.

39. Karthikeyan, S., Leung, T. and Ladias, J. (2001b) Structural basis of the Na+/H+ exchanger regulatory factor PDZ1 interaction with the carboxylterminal region of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem. 276*, 19683-19686.

40. Song, E., Gao, S., Tian, R., Ma, S., Huang, H., Guo, J., Li, Y., Zhang, L., Gao, Y. (2006) A high efficiency strategy for binding property characterization of peptide binding domains. *Mol Cell Proteomics 5*, 1368-1381.

41. Hering, H., Sheng, M. (2002) Direct interaction of Frizzled-1, -2, -4, and -7 with PDZ domains of PSD-95. *FEBS Lett*. *521*, 185-189.

42. Uetz, P. (2002) Two-hybrid arrays. *Curr Opin Chem Biol*. *6*, 57-62.

43. Soares, H. D., Williams, S. A., Snyder, P. J., Gao, F., Stiger, T., Rohlff, C., Herath, A., Sunderland, T., Putnam, K., White, W. F. (2004) Proteomic approaches in drug discovery and development. *Int Rev Neurobiol*. *61*, 97-126.

44. Lee, H-W., Ko, J., Kim, E. (2006) Analysis of PDZ domain interactions using yeast two-hybrid and coimmunoprecipitation assays. *Methods Mol Biol*. *332*, 233-244.

45. Gisler, S. M., Kittanakom, S., Fuster, D., Wong, V., Bertic, M., Radanovic, T., Hall, R. A., Murer, H., Biber, J., Markovich, D., Moe, O. W., Stagljar, I. (2008) Monitoring proteinprotein interactions between the mammalian integral membrane transporters and PDZ-interacting partners using a modified splitubiquitin membrane yeast two-hybrid system. *Mol Cell Proteomics*. *7*, 1362-1377.

46. Von, N. P., Ismail, M., Gardin, C., Suila, H., Zara, I., Belgrano, A., Valle, G., Carpen, O., Faulkner, G. (2009) A class III PDZ binding motif in the myotilin and FATZ families binds enigma family proteins: a common link for Z-disc myopathies. *Mol Cell Biol. 29*, 822-834.

47. Wolf-Yadlin, A., Sevecka, M., MacBeath, G. (2009) Dissecting protein function and signaling using protein microarrays. *Curr Opin Chem Biol. 13*, 398-405.

48. Fam, S. R., Paquet, M., Castleberry, A. M., Oller, H., Lee, C. J., Traynelis, S. F., Smith, Y., Yun, C. C., Hall, R. A. (2005) P2Y1 receptor signaling is controlled by interaction with the PDZ scaffold NHERF-2. *Proc Natl Acad Sci USA. 102*, 8042-8047.

49. Paquet, M., Asay, M. J., Fam, S. R., Inuzuka, H., Castleberry, A. M., Oller, H., Smith, Y., Yun, C. C., Traynelis, S.F., Hall, R.A. (2006) The PDZ scaffold NHERF-2 interacts with mGluR5 and regulates receptor activity. *J Biol Chem. 281*, 29949-29961.

50. He, J., Bellini, M., Inuzuka, H., Xu, J., Xiong, Y., Yang, X., Castleberry, A. M., Hall, R. A. (2006) Proteomic analysis of beta1-adrenergic receptor interactions with PDZ scaffold proteins. *J Biol Chem. 281*, 2820-2827.

51. Lee, H-W., Choi, J., Shin, H., Kim, K., Yang, J., Na, M., Choi, S. Y., Kang, G. B., Eom, S. H., Kim, H., Kim, E. Preso, (2008) a novel PSD-95-interacting FERM and PDZ domain protein that regulates dendritic spine morphogenesis. *J Neurosci. 28*, 14546-14556.

52. Zhang, J., Cheng, S., Xiong, Y., Ma, Y., Luo, D., Jeromin, A., Zhang, H., He, J. (2008) A novel association of mGluR1a with the PDZ scaffold protein CAL modulates receptor activity. *FEBS Lett*. *582*, 4117-4124.

53. Meerschaert, K., Remue, E., De, G. A., Staes, A., Boucherie, C., Gevaert, K., Vandekerckhove, J., Kleiman, L., Gettemans, J. (2008) The tandem PDZ protein Syntenin interacts with the aminoacyl tRNA synthetase complex in a lysyl-tRNA synthetase-dependent manner. *J Proteome Res*. *7*, 4962-4973.

54. Gee, H. Y., Kim, Y. W., Jo, M. J., Namkung, W., Kim, J. Y., Park, H. W., Kim, K. S., Kim, H., Baba, A., Yang, J., Kim, E., Kim, K. H., Lee, M. G. (2009) Synaptic scaffolding molecule binds to and regulates vasoactive intestinal polypeptide type-1 receptor in epithelial cells. *Gastroenterology*. *137*, 607- 617.

55. Fuh, G., Pisabarro, M. T., Li, Y., Quan, C., Lasky, L. A., and Sidhu, S. S. (2000) Analysis of PDZ Domain-Ligand Interactions Using Carboxyl-terminal Phage Display. *J Biol Chem*. *275*, 21486–21491.

56. Merzlyak, A., and Lee, S-W. (2006) Phage as templates for hybrid materials and mediators for nanomaterial synthesis. *Current Opinion in Chemical Biology*, *10*, 246–252.

57. Smith, G. P., Petrenko, V. A. (1997) Phage display. Chem Rev. *97*, 391-410.

58. Petrenko, V. A., Smith, G. P., Gong, X., Quinn, T. (1996) A library of organic landscapes on filamentous phage. *Protein Eng*. 9, 797-801.

59. George P. Smith and Valery A. Petrenko (1997) Phage Display. *Chem. Rev. 97*, 391−410.

60. Marvin, D. A.; Hale, R. D.; Nave, C.; Citterich, M. H. (1994) Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and Ike. *J. Mol. Biol. 235*, 260-286.

61. Konthur, Z., and Crameri, R. (2003) High-throughput applications of phage display in proteomic analyses. *Targets **2***, 261-270.

62. Sidhu, S. S., Bader, G. D., Boone, C. (2003) Functional genomics of intracellular peptide recognition domains with combinatorial biology methods. *Curr Opin Chem Biol*. *7*, 97-102.

63. Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M., Cantley, L. C. (1997). Recognition of unique carboxyl terminal motifs by distinct PDZ domains. *Science 275*, 73-77.

64. Kurakin, A., Swistowski, A., Wu, S. C., Bredesen, D. E. (2007) The PDZ domain as a complex adaptive system. *PLoS ONE 2*, e953.

65. Berk, L. C., van den, L. E., Walma, T., Vuister, G. W., Dente, L., Hendriks, W. J. (2007) An allosteric intramolecular PDZ-PDZ interaction modulates PTP-BL PDZ2 binding specificity. *Biochemistry 46*, 13629-13637.

66. Tonikian, R., Zhang, Y., Sazinsky, S. L., Currell, B., Yeh, J. H., Reva, B., Held, H. A., Appleton, B. A., Evangelista, M., Wu, Y., Xin, X., Chan, A. C., Seshagiri, S., Lasky, L. A., Sander, C., Boone, C., Bader, G. D., Sidhu, S. S. (2008) A specificity map for the PDZ domain family. *PLoS Biol*. *6*, e239.

67. Berk, L.C., van den, L. E., Harmsen, E., Dente, L., Hendriks, W. J. (2005) Redoxregulated affinity of the third PDZ domain in the phosphotyrosine phosphatase PTP-BL for cysteine-containing target peptides. *FEBS J. 272*, 3306.

68. Penkert, R. R., DiVittorio, H. M., and Prehoda, K. E. (2004) Internal recognition through PDZ domain plasticity in the Par-6-Pals1 complex. *Nat Struct Mol Biol 11*, 1122-1127.

69. London, T.B.C., Lee, H-J, Shao, Y and Zheng, J. (2004) Interaction between the internal motif KTXXXI of Idax and mDvl PDZ domain. *Biochem. Biophy. Res. Comm. 322*, 326-332.

70. Gee, S. H., Sekely, S.A., Lombardoi, C., Kurakin, A., Froehner, S.C., and Kay, B. K. (1998) Cyclic Peptides as Non-carboxyl-terminal Ligands of Syntrophin PDZ Domains J. Biol. Chem. *273*, 21980–21987.

71. Wang, N. X., Lee, H. J., Zheng, J. J. (2008) Therapeutic use of PDZ protein-protein interaction antagonism. *Drug News Perspect. 21*, 137-141.

72. Cushing, P. R., Fellows, A., Villone, D., Boisguerin, P., Madden, D. R. (2008) The relative binding affinities of PDZ partners for CFTR: a biochemical basis for efficient endocytic recycling. *Biochemistry 47*, 10084-10098.

73. Wong, H. C., Bourdelas, A., Krauss, A., Lee, H-J., Shao, Y., Wu, D., Mlodzik, M., Shi, D. L., Zheng, J. (2003) Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled. *Mol Cell 12*, 1251-1260.

74. Aarts, M., Liu, Y., Liu, L., Besshoh, S., Arundine, M., Gurd, J. W., Wang, Y. T., Salter, M. W., Tymianski, M. (2002) Treatment of ischemic brain damage by perturbing NMDA receptor-PSD-95 protein interactions. *Science 298*, 846-850.

75. Dev, K. K. (2004) Making protein interactions druggable: targeting PDZ domains. *Nat Rev Drug Discov. 3*, 1047-1056.

76. Wolde, M., Fellows, A., Cheng, J., Kivenson, A., Coutermarsh, B., Talebian, L., Karlson, K., Piserchio, A., Mierke, D. F., Stanton, B. A., Guggino, W. B., Madden, D. R. (2007) Targeting CAL as a negative regulator of DeltaF508-CFTR cell-surface expression: an RNA interference and structure-based mutagenetic approach. *J. Biol. Chem*. *282*, 8099-8109.

77. Georgescu, M. M., Morales, F. C., Molina, J. R., Hayashi, Y. (2008) Roles of NHERF1/EBP50 in cancer. *Curr. Mol. Med*. *8*, 459-468.

78. Georgescu, M. M. (2008). NHERF1: molecular brake on the PI3K pathway in breast cancer. *Breast Cancer Res*. *10*, 106.

79. Lee, H-J., Wang, N. X., Shi, D. L., Zheng, J. J. (2009) Sulindac inhibits canonical Wnt signaling by blocking the PDZ domain of the protein Dishevelled. *Angew Chem Int Ed Engl.*, *48*, 6448-6452.

80. Law, A. B., Fuentes, E. J., Lee, A. L. (2009) Conservation of side-chain dynamics within a protein family. *J. Am. Chem. Soc*., *131*, 6322-6323.

81. Wolde, M., Fellows, A., Cheng, J., Kivenson, A., Coutermarsh, B., Talebian, L., Karlson, K., Piserchio, A., Mierke, D. F., Stanton, B. A., Guggino, W. B., Madden, D.R. (2007) Targeting CAL as a negative regulator of DeltaF508-CFTR cell-surface expression: an RNA interference and structure-based mutagenetic approach. *J Biol Chem.*, *282*, 8099-8109.

82. Georgescu, M. M. (2008) NHERF1: molecular brake on the PI3K pathway in breast cancer. *Breast Cancer Res.*, *10*, 106.

83. Dev, K. K., Nakanishi, S., Henley, J. M. (2004) The PDZ domain of PICK1 differentially accepts protein kinase C-alpha and GluR2 as interacting ligands. *J. Biol. Chem., 279*,41393-41397.

84. Shan, J., Shi, D. L., Wang, J., Zheng, J. (2005) Identification of a specific inhibitor of the dishevelled PDZ domain. *Biochemistry, 44*, 15495-15503.

85. Hammond, M. C., Harris, B. Z., Lim, W. A., Bartlett, P. A. (2006) Beta strand peptidomimetics as potent PDZ domain ligands. *Chem Biol., 13*, 1247-1251.

86. Fujii, N., You, L., Xu, Z., Uematsu, K., Shan, J., He, B., Mikami, I., Edmondson, L. R., Neale, G., Zheng, J., Guy, R. K., Jablons, D. M. (2007) An antagonist of disheveled protein-protein interaction suppresses beta-catenin-dependent tumor cell growth. *Cancer Res. 67*, 573-579.

87. Shan, J., Zheng, J. J. (2009) Optimizing Dvl PDZ domain inhibitor by exploring chemical space. *J Comput Aided Mol Des., 23*, 37-47.

88. Udugamasooriya, G., Saro, D., Spaller, M. R. (2005) Bridged peptide macrocycles as ligands for PDZ domain proteins. *Org Lett. 7*, 1203-1206.

89. Udugamasooriya, D. G., Sharma, S. C., Spaller, M. R. (2008) A chemical library approach to organic-modified peptide ligands for PDZ domain proteins: a synthetic, thermodynamic and structural investigation. *Chembiochem 9*, 1587-1589.

90. Zuiderweg, E. R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry 41*, 1-7.

91. Johnson, P.E., Tomme, P., Joshi, M.D., and McIntosh, L.P., (1996) Structure of the N-terminal cellulose-binding domain of Cellulomonas fimi CenC determined by nuclear magnetic resonance spectroscopy. *Biochemistry 35*, 13895-13906.

92. Zhou M., Harlan J.E., Wade, W.S., Crosby, S., Ravichandran, K.S., Burakoof, S.J. and Fesik, S,W. (1996) Binding affinities of tyrosine phosphorylated peptides to the COOH terminal SH2 and NH2-terminal phosphotyrosine binding domains of Shc. *J. Biol. Chem., 271*, 31119-31123.

93. Malcolm. H. Levitt, Spin Dynamics, John Wiley & Sons Ltd., 2001.

94. Boehr, D. D., Dyson, H. J., and Peter E. (2006) An NMR Perspective on Enzyme Dynamics. *Chem. Rev., 106*, 3055-3079.

95. Lee, A. L., Urbauer, J. L., Wand, A. J. (1997) Improved labeling strategy for 13C relaxation measurements of methyl groups in proteins *J. Biomol. NMR, 9*, 437.

96. Teague, S. J. (2003) Implications of protein flexibility for *drug discovery*. *Nat. Re*V. *Drug Disco*V*ery*, *2*, 527.

97. Lipari, G., Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules I. Theory and range of validity. *J. Am. Chem. Soc., 104*, 4546.

98. Lipari, G.; Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results *J. Am. Chem. Soc., 104*, 4559.

99. Walma, T., Spronk, C. A. E.M., Tessari, M., Aelen, J., Schepens, J., Hendriks, W., and Vuister, G. W. (2002) Structure, Dynamics and Binding Characteristics of the Second PDZ Domain of PTP-BL. *J. Mol. Biol. 316*, 1101-1110.

100. Ernesto J. Fuentes1,4, Channing J. Der and Andrew L. Lee. (2004) Ligand-dependent Dynamics and Intramolecular Signaling in a PDZ Domain. *J. Mol. Biol. 335*, 1105–1115.

101. S. Gianni, T. Walma, A. Arcovito, N. Calosci, A. Bellelli, Ake Engstrom, C. Travaglini-Allocatelli, M. Brunori, P. Jemth, and G. W. Vuister. (2006) Demonstration of Long-Range Interactions in a PDZ Domain by NMR, Kinetics, and Protein Engineering. *Structure 14*, 1801–1809.

CHAPTER TWO

Characterization of Structure, Function, and Dynamics Studies of Human Glutaminase Interacting Protein (GIP)

2.1 Introduction

2.1.1 GIP

PDZ domain containing proteins are involved in a number of important biological pathways. One such example is the human glutaminase-interacting protein (GIP). Human GIP is a small soluble multifunctional protein containing a single PDZ domain. This protein was named glutaminase-interacting protein because it was originally identified in a yeast two-hybrid genetic selection system while looking for interactors of glutaminase in the human brain (*1*). GIP interacts with the C-terminus of various target partners, including Glutaminase L (LGA). Glutaminase L is responsible for synaptic transmission and regulation of cerebral concentrations of glutamine and the neurotransmitter glutamate (*1*). Glutaminase is also important for the normal activity of the central nervous system (*2*); this enzyme also plays an important role in tumor development (*2-5*). A partial GIP cDNA clone was previously identified from human lymphocytes through a yeast two-hybrid screening and named TIP-1, for Tax-interacting protein 1, as it is a binding partner for the viral oncoprotein Tax (*6*).

2.1.2 Interacting partners for GIP

The full-length human GIP contains 124 amino acid residues. GIP contains a single PDZ domain which is composed of residues 11-112 of the protein sequence (*1*). GIP contributes to the determination of the sub-cellular distribution and localization of LGA and/or regulation of its function (*7*, *8*). The mechanism by which GIP and LGA interact in the brain is not completely clear. It has been found that GIP and LGA co-localize in astrocytes and neurons (*7*). This study suggests that GIP plays an important role as a scaffolding PDZ protein in the mammalian brain (*7*). In addition to glutaminase L, a number of other target partners have been reported, implicating GIP in key biological processes. Like other PDZ domain containing proteins, GIP also interacts with the C-terminus of the binding proteins. These binding proteins are involved in various signaling pathways. The other reported GIP target partners are the viral oncoproteins human T-cell leukemia virus type I (HTLV-I) Tax (*6*), human paillomavirus 16 (HPV16) E6 (*9*), and the Rho-activator rhotekin (*10*), which involves GIP in the Rho signaling pathway, and the potassium channel Kir 2.3, where GIP regulates channel expression in the plasma membrane of the renal epithelia (*11*). GIP has also been shown to participate in the regulation of transcription. GIP also binds to the C-terminus of β-catenin (*12-14*). β-catenin mediates Wnt signaling involved in the regulation of cell-cell adhesion during embryogenesis; however, deregulation of this signal transduction pathway frequently leads to the formation of various cancerous tumors (*12-14*). Overexpression of GIP reduces the proliferation and anchorage-independent growth of colorectal cancer cells (*12*). GIP is not only involved in the CNS (Central Nervous System) and various human cancers; we discovered that it also interacts with the C-terminal region of a cell-surface

protein FAS, which belongs to the TNF (Tumor Necrosis Factor) receptor family (*15*) and mediates cell apoptosis (*16, 17*).

2.1.3 PDZ domain of GIP

Most of the PDZ domains fit into three major classes (Class I to Class III). GIP belongs to class I PDZ domain which recognizes the C-terminal sequence S/T-X-Φ-COOH of the target proteins (X denoting any amino acid and Φ representing a hydrophobic residue). The amino acid residues at positions 0 and -2 of the peptide (position 0 referring to the C-terminal residue) play a key role in the specificity and affinity of the interaction (*12*). A free carboxylate group at the extreme C-terminus of the target protein sequence motif is important for the interaction process. Generally, class I PDZ domains contain 6β strands (β1-β6) and 2α helices (α1-α2). It has been reported that the PDZ domain binds the C-terminus of the interacting partner in an elongated groove as an antiparallel β-strand between the second α-helix and the second β-strand, which is called the PDZ binding groove (*18*). The Gly-Leu-Gly-Phe (GLGF) motif is conserved among all the PDZ domain containing proteins. Very often, it is located within the β1-β2 connecting loop. This loop is important for hydrogen bond coordination of the C-terminal carboxylate group (COO-) (*19*). The GLGF motif is also called carboxylate binding loop. In case of GIP, GLGF motif is replaced by ILGF motif.

2.1.4 GIP and drug design

PDZ domain containing proteins are potential targets for drug discovery, as they have a well-defined binding site. GIP interacts with a number of target proteins which are

involved in various cancerous pathways. GIP is one of the smallest members of the PDZ domain containing family. GIP contains only one PDZ domain which represents its full primary structure except the N and C termini. Therefore, GIP could be an excellent target for structure-based drug design. For successful drug design, it is important to characterize the protein completely. Structure, function, dynamics, and interaction studies of GIP with the C-terminal recognition motif of different binding partners will provide us insight into the mechanisms of action of this multifunctional protein, which is definitely a necessary prelude for successful drug design. We report here the production of pure recombinant GIP and the characterization of structure, function, and dynamics by CD, fluorescence and high-resolution solution NMR.

2.2 Materials and methods

2.2.1 Cloning and over-expression of GIP

GIP which encodes a 1.3 kb cDNA was originally cloned into the pJG4-5 vector (*20*). Cloning and optimization of over-expression of GIP was done previously in our lab (*21*). The procedures are described here. The 13.7 kDa GIP protein was expressed without any additional residues. For this purpose, the open reading frame (ORF) sequence was amplified and cloned into the NdeI/BamHI (Invitrogen) sites of the pET-3c vector (Novagen), using the primers mentioned here: forward, 5'-AGC-AGG-GTC-CAT-ATG-TCC-TAC-ATC-CCG-3' and reverse, 5'-CGG-CAG-GCA-GGA-TCC-GCA-GAT-GGT-GG-3'. The sequence of the GIP cDNA in the expression vector was confirmed by sequencing. Recombinant GIP protein was then expressed in *E. coli* BL21 DE3pLys cells (Novagen). Saturated LB-ampicillin culture was diluted in LB media and grown at 37°C

to an A_{600} nm of 0.5. Expression was induced with 1 mM IPTG (Gold Biotechnology, Inc.) and cells were harvested by centrifugation after incubation at 30°C for 4 hours. The bacterial cells were suspended in phosphate buffer at pH 8 which contains 200 mM NaCl, 4 mM EDTA, 4% glycerol, 1mM PMSF and lysed using sonication. After centrifugation at 14,000 rpm for 30 minutes both supernatant and pellet-containing GIP inclusion bodies were collected. For production of ^{15}N, ^{13}C-labeled proteins, bacterial cells were grown in minimal media containing ^{15}N-labeled ammonium chloride and either regular or ^{13}C-labeled glucose (Isotec).

2.2.2 Purification of GIP

Soluble GIP was purified in a single step by size exclusion chromatography with a Sephacryl S-100 column (GE Healthcare) fitted to an FPLC system using 20mM phosphate buffer containing 150 mM NaCl, 1 mM EDTA and 0.1 % NaN_3 as the mobile phase.

2.2.3 Fluorescence

All fluorescence spectra were recorded on Perkin Elmer LS 55 Luminescence spectrofluorometer. Tryptophan fluorescence was excited at 280 nm for ligand titration. Emission spectra were recorded over the range 300 nm – 500 nm with 1 nm steps. All fluorescence spectra were recorded at room temperature at pH 6.5. The buffer used in the titration and quenching experiments was 20 mM phosphate buffer (pH 6.5), 150 mM NaCl, 0.1 mM EDTA, and 0.01% NaN_3. In each titration experiment the concentration of the protein was corrected for volume dilution, and observed emission was corrected for

absorbance of the quencher. Emission from the control was considered by recording subtraction spectra between sample and control probes.

2.2.4 Circular dichroism (CD)

All circular dichroism (CD) experiments were performed on a Jasco J-810 automatic recording spectropolarimeter. Both Far-UV and near-UV CD spectra were measured in a 0.05 cm quartz cell at room temperature. The buffer used was 20 mM phosphate buffer (pH 6.5). The protein concentration was 20 μM. The ligand concentrations were ranged from 2-30 μM. Data was averaged over 100 scans for each protein sample and over 50 scans for each control sample. Response time was 1s and scan speed was 100 nm min^{-1}. To characterize GIP interaction with target partner protein, titration studies were carried out. The concentration of GIP was calculated by measuring the absorbance at 280 nm. For tertiary structure determination 800 μM GIP was used.

2.2.5 Nuclear magnetic resonance (NMR)

All NMR data were collected at 298K on a Bruker Avance 600 MHz spectrometer equipped with triple resonance H/C/N TCI cryoprobe in the Department of Chemistry & Biochemistry, Auburn University and processed using NMRPipe (*22*) and analyzed using NMRView (*23*). The sample contained ~800 μM uniformly $^{15}N/^{13}C$ labeled GIP, 50 mM phosphate buffer (pH 6.5) containing 5 % D_2O, 1 mM EDTA, and 0.01% (w/v) NaN_3.. The following spectra were used for the backbone assignment: 2D {$^{1}H,^{15}N$}-HSQC, 3D HNCACB, 3D CC(CO)NH, 3D CBCACONH, 3D HCCONH, 3D HNHA, 3D HNCO and 3D HNCACO, 3D ^{15}N edited HSQC-TOCSY, 3D ^{15}N edited HSQC-NOESY. 2D

{^{1}H, ^{15}N,}-HSQC were obtained using the suitable Bruker pulse programs. $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{13}CO chemical shifts were used to determine the secondary structure using the programs PSSI (*24*) and PsiCSI (*25, 26*). Ligands titration experiments were monitored by a series of two-dimensional ^{15}N-edited HSQC experiments.

Interaction studies were carried out by titration of 100 µM GIP with three different C-terminal target sequences: the C-terminus of glutaminase (KENLESMV-COOH), β-catenin (FDTDL-COOH), (FDTDL-$CONH_2$), FAS (NEIQSLV-COOH), Kir 2.3 and HPV16 E6. The amide chemical shift perturbations (Δδ) were calculated as $\Delta\delta = |\Delta\delta^{15}N|/f + |\Delta\delta^{1}H|$ (*27*, *28*). The introduction of the f factor and its value were justified by the difference in the spectral widths of the backbone ^{15}N resonances and the ^{1}H signals (^{15}N range, 131.8-100.4 ppm = 31.4 ppm; ^{1}H range, 10.1-6.5 ppm = 3.6; correction factor f=31.4/3.6 = 8.7). Thus, the correction factor f = 8.7 was used in order to give roughly equal weighting for each of the ^{1}H and ^{15}N chemical shift changes. ModelTitr program (*29*) was used to calculate dissociation constant values of various residues of GIP. The target peptides were obtained with > 95 % purity from Genemed Synthesis (CA), Synbiosci (CA) or Chi Scientific (MA). The 5-20 mM stock solutions of the above peptides were prepared in 10 mM phosphate buffer at pH 6.5 for NMR, CD and fluorescence titration experiments.

2.2.6 NMR relaxation measurements

All NMR relaxation data were collected at 298K on a Bruker Avance 600 MHz spectrometer equipped with triple resonance H/C/N TCI cryoprobe in the Department of Chemistry & Biochemistry, Auburn. Samples for relaxation measurements contained 550

μM protein in 50 mM phosphate buffer (pH 6.5) (containing 1 mM EDTA and 0.01% (w/v) NaN_3) prepared using 10% D_2O and 90% H_2O (at 25 °C). The relaxation studies were carried out on a 1:2 GIP-glutaminase C-terminal peptide complex. Relaxation rates and heteronuclear ^1H-^{15}N NOE enhancements were calculated from peak heights of the resonances in a series of ^1H-^{15}N HSQC spectra. NMRPipe (*22*) and NMRView (*23*) were used for peak picking and peak intensity measurements. The MODEL FREE program was employed for data analysis. Errors in ^1H-^{15}N NOE were expressed as the standard deviation of six pairs of repeated ^1H-^{15}N NOE experiments. The longitudinal and transverse relaxation rate constant, R_1 and R_2, were obtained by a two-parameter nonlinear optimization of a single-exponential function using the Monte Carlo nonlinear least-square curve-fitting algorithm. Errors in relaxation rates were obtained by Monte Carlo analysis. For the Bruker Avance 600 instrument, standard Bruker pulse programs were used to measure T_1 and T_2. T_1 data were acquired using relaxation delays of 10, 50, 100, 200, 300, 400, 500, 600 and backbone amide ^{15}N T_2 values were obtained similarly: 17, 34, 51, 68, 85, 102, 119, 136, 152, 153 ms. The same saturation time without pulse was used for control experiments. T_1 and T_2 values were converted to R_1 and R_2. R_1 and R_2 are rate constants for T_1 and T_2 respectively. The length of the proton saturation for NOE measurements was 5s. The NOE control experiment was collected with a 5s recycle delay.

2.2.7 ModelFree analysis

Relaxation parameters were used in programs developed by Palmer et al. (*30, 31*); to analyze ^{15}N-backbone dynamics. Five different models were used to determine the

order parameter (S^2) which gives the degree of spatial restriction of the ^{1}H-^{15}N bond vector, with values ranging from zero for isotropic internal motions to unity for completely restricted motion. This represents dynamics on the pico to nanosecond time scale. The effects of millisecond-microsecond time scale motions (*Rex*) are included in the third and fourth models. τ_c (total correlation time) were determined by R_2/R_1.

2.3 Results

2.3.1 Protein expression

Unlabeled GIP was expressed as described above in bacteria using standard LB media. ^{15}N-labeled and ^{15}N, ^{13}C-labeled GIP were expressed in *E. coli* cells grown in minimal media containing ^{15}N-labeled ammonium chloride and either regular or ^{13}C-labeled glucose.

2.3.2 Protein purification

GIP was purified to homogeneity in a single step by size-exclusion chromatography on a Sephacryl S-100 column (GE Healthcare) (**Figure 2.1**). Recombinant GIP was successfully isotope-enriched (both ^{15}N and ^{15}N, ^{13}C) and purified to homogeneity with the above protocol. About 56 mg of pure uniformly labeled GIP (^{15}N or ^{15}N, ^{13}C) were obtained per liter of bacterial culture.

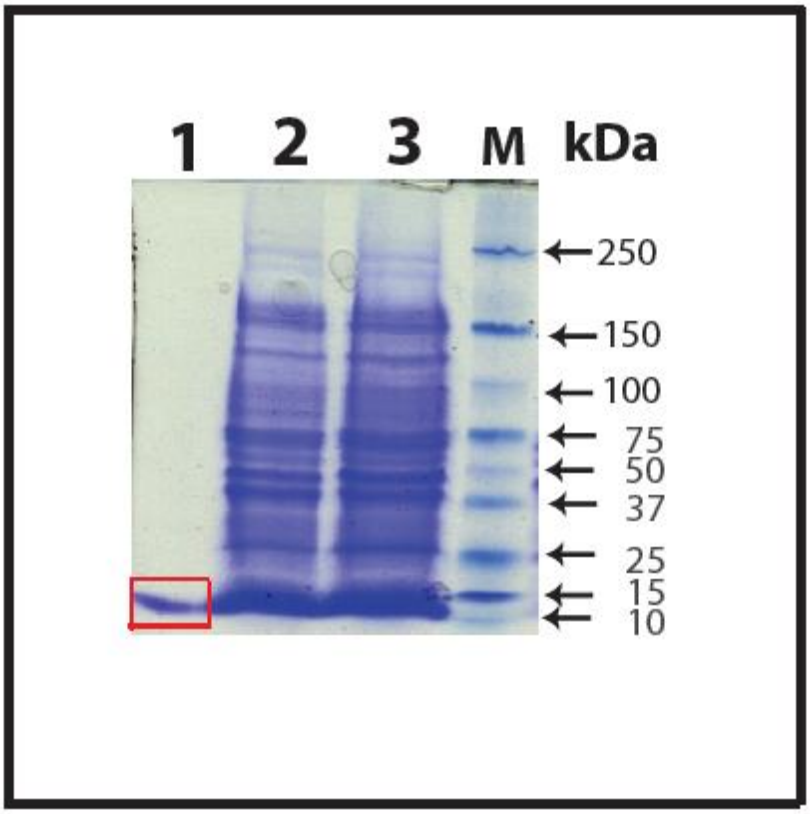

Figure 2.1 SDS-PAGE analysis of GIP purified by size exclusion chromatography (Lane 1). Lane 2 and 3 represent bacterial cell lysate before purification of GIP. Lane M represents a protein marker (M).

2.3.3 Characterization of GIP

2.3.3.1 Mass determination by ESI-MS mass spectroscopy

The molecular weight of GIP was verified by mass spectroscopy. A sharp peak at 13.7 kDa was observed that indicates pure ^{15}N labeled GIP (**Figure 2.2**).

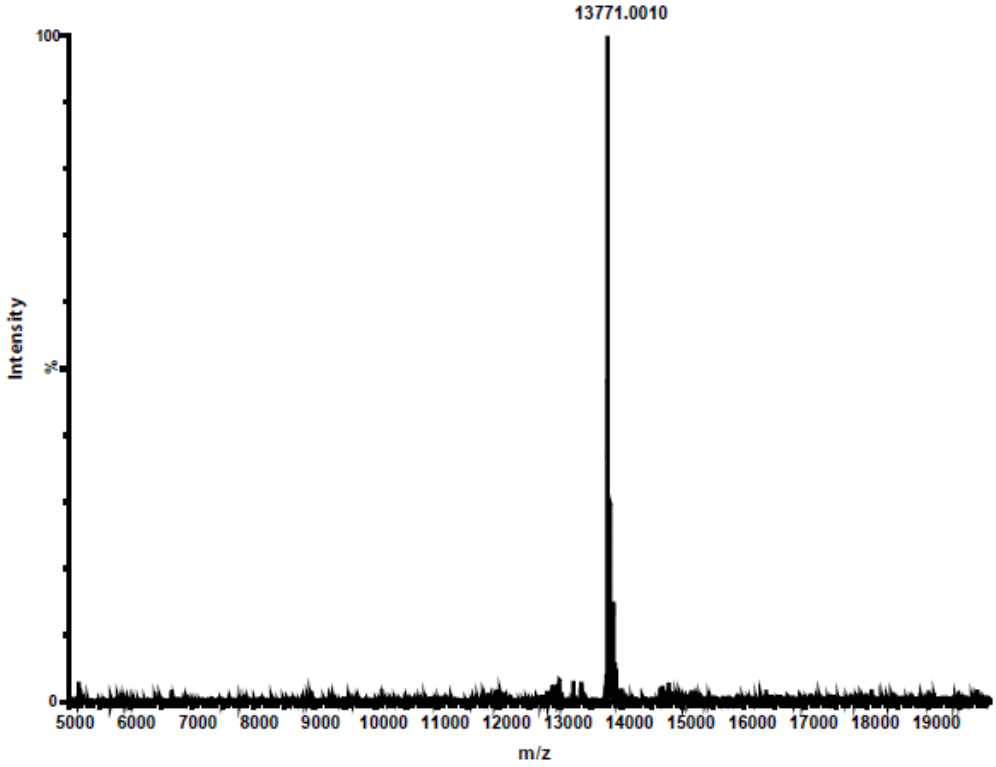

Figure 2.2 ESI-MS mass spectrum of GIP. A sharp peak at 13.7 kDa was observed that indicates pure GIP.

2.3.3.2 Characterization of GIP by far-UV and near-UV CD Spectroscopy

Circular dichroism (CD) spectroscopy measures the difference in absorption of left- and right- circularly polarized light when it passes through an optically active or chiral sample. Spectra in the far UV wavelength ranges from ~190 to 250 nm provide information on the polypeptide backbone conformations of proteins. Thus far UV CD provides important information about the secondary structures of an unknown protein. Far UV CD is a powerful technique that provides useful information in the change of secondary structural elements of a protein upon ligand binding. Deconvolution of a CD spectrum gives qualitative contents of α helix, β sheet and random coil (32). CD Spectroscopy has immense importance in the pharmaceutical industry for drug-binding studies providing information on significant structural rearrangements in the protein or greatly enhance or diminish stability of a protein (33). Overall, CD spectroscopy gives information on secondary structural analysis, protein folding, and stability. The secondary structure of GIP was characterized by far-UV CD spectroscopy in phosphate buffer at pH 6.5 (**Figure 2.3**). Deconvolution of CD data shows that GIP contains 8.5% α-helix and 35.5% β-strand in phosphate buffer at pH 6.5.

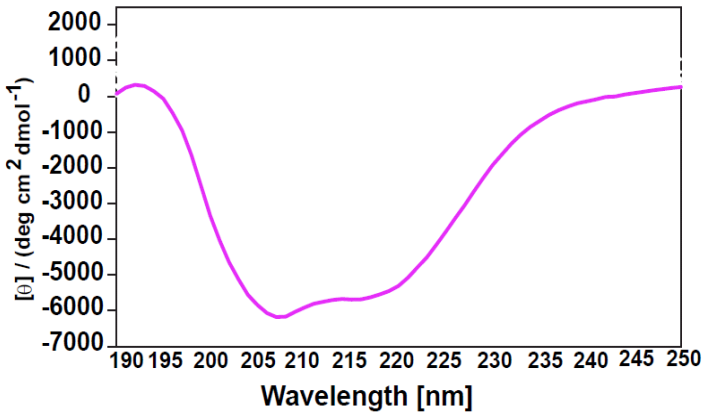

Figure 2.3 Far UV-CD spectrum of GIP in 20 mM phosphate buffer (pH 6.5) at 25°C. It shows that GIP has secondary structures.

CD spectroscopy can be used to find the tertiary structure of a protein. In this case near UV CD is being used. The near UV CD regions are from 250 nm-350 nm. The choromophores used for this region are aromatic amino acids and disulfide bonds. The regions from 250-270 nm contain the signals for phenylalanine residues, signals from 270-290 nm correspond to tyrosine residues, and signals in the region from 280-300 nm are for tryptophan residues. Disulfide bonds give broad weak signals over the entire near-UV spectrum. Sometimes it is possible that the protein has a well defined secondary structure but no tertiary structure.

Presence of Near-UV signals is an indication of a well folded protein structure. Thus this region is very sensitive to change in environments of a protein. The main difficulty associated with near-UV CD is that the process gives weak signals. Thus, large quantities of protein are needed for this type of experiment. Near-UV CD data indicated that recombinant GIP has tertiary structure, an indication of properly folded protein (**Figure 2.4**).

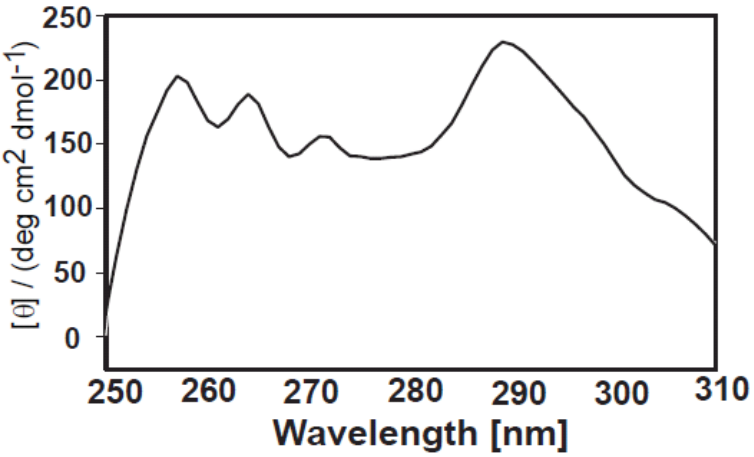

Figure 2.4 Near UV CD spectrum of GIP in 20 mM phosphate buffer (pH 6.5) at 25^{o}C. This spectrum shows that GIP has tertiary structure.

2.3.3.3 Intrinsic tryptophan fluorescence

Fluorescence is the emission from molecules that have been excited to higher energy states by the absorption of electromagnetic radiation. The amino acids tryptophan and tyrosine are known to demonstrate a strong fluorescence around 300-350 nm depending in the polarity of the local environment upon excitation around 280 nm. Protein fluorescence may be used as a diagnostic of the conformational state of a protein. GIP has a tryptophan residue. The fluorescence emission spectrum of GIP (containing a tryptophan residue) was recorded upon excitation at 280 nm. The emission maximum was observed at a wavelength of 353 nm (characteristic of solvent-exposed tryptophan) (**Figure 2.5**).

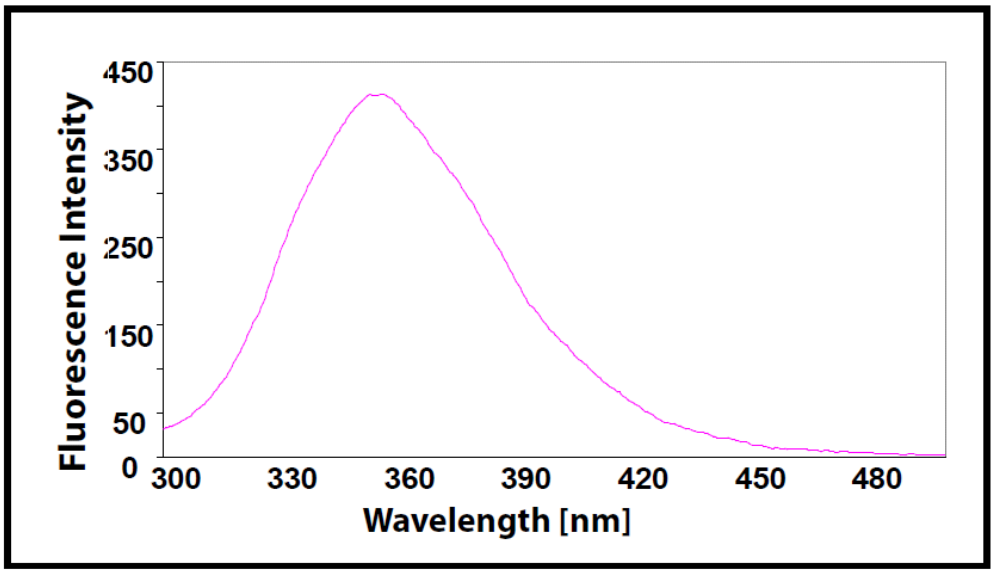

Figure 2.5 Intrinsic tryptophan fluorescence of GIP in 20 mM phosphate buffer (pH 6.5) at 25^{o}C.

2.3.3.4 Characterization of GIP by NMR Spectroscopy

NMR is an extremely robust technique for monitoring the conformational changes in the protein sample caused by pH, temperature, salt or ligand. 2D {^{1}H, ^{15}N}-HSQC experiments were performed in phosphate buffer at pH 6.5. It revealed that the backbone amide resonances from GIP are nicely dispersed. (**Figure 2.6**). The nice dispersion of resonances in the spectrum confirmed that GIP is a well-folded protein.

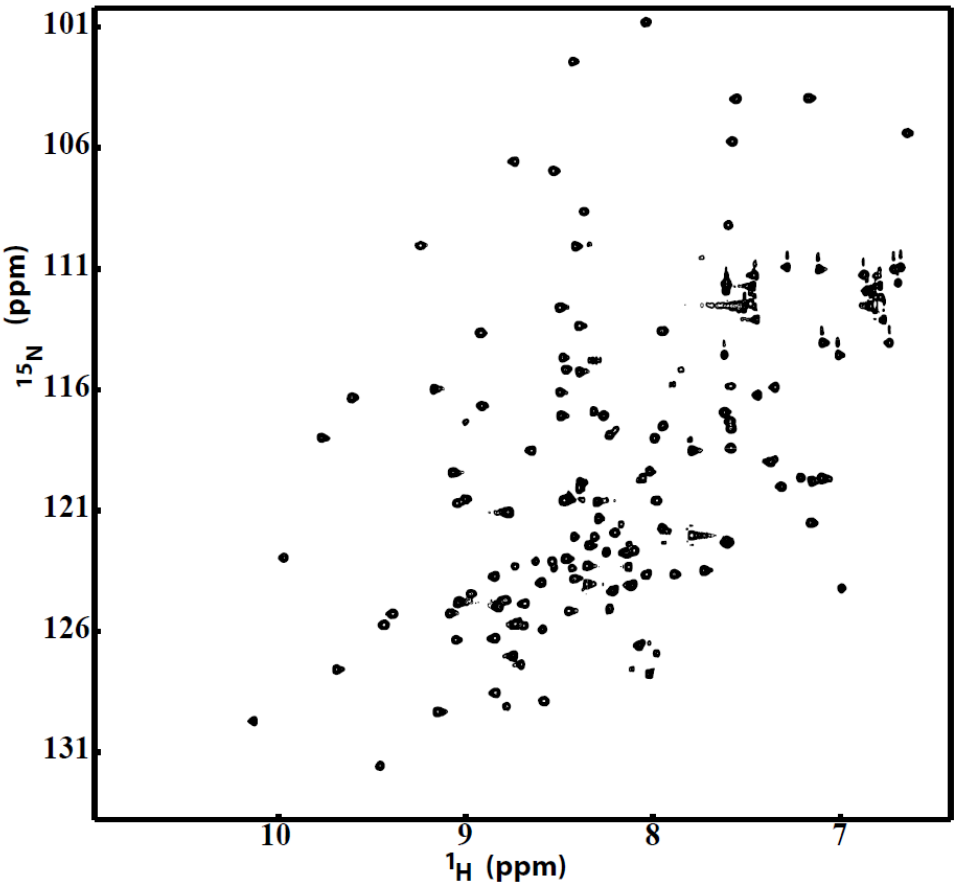

Figure 2.6 HSQC spectrum of GIP in 50 mM phosphate buffer (pH 6.5) at 25^{o}C.. It shows that GIP is properly folded.

2.3.4 Interaction studies of GIP with C-terminus ligands by CD and fluorescence spectroscopy

2.3.4.1 Effect of ligand binding on GIP by CD

The effect of ligand binding on GIP secondary structure was also investigated by CD spectroscopy. As most of the PDZ domain-containing proteins require five to eight C-terminal residues for ligand binding and specificity, GIP was titrated with the peptides comprised of the last 8 C-terminus residues of the ligand glutaminase (KENLESMV-COOH), the last 5 residues of the ligand β-catenin (FDTDL-COOH), and the last 7 residues of the ligand FAS (NEIQSLV-COOH).

All these peptides have very negligible contribution to the CD spectra, which is almost equivalent to the phosphate buffer. The titration of GIP with all the three peptides caused significant secondary structure changes in the CD spectra of GIP (**Figures 2.7-2.9**). However, unlike the glutaminase C-terminal peptide, a small amount of white precipitate formed during the titration of β-catenin and FAS C-terminal peptides with GIP. Deconvolution of the CD data of GIP-ligand complexes was performed and the secondary structure content was calculated using the program CDPro (*34*). The deconvolution results indicate that for all the three ligands, the percentage of helix and random coil structure content decreases with increasing concentration of ligand while the percentage of β-sheet content increases. Overall for the three complexes, the helix content is reduced by 30-60 %, random coil content is reduced by 5-10 % and the β-sheet structure content is increased by 18-27 %.

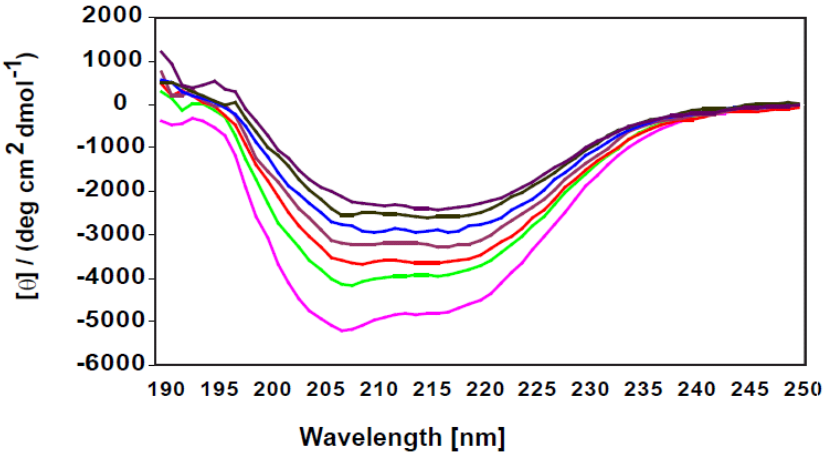

Figure 2.7 Changes in the far-UV CD spectra of 20 µM GIP (in 20 mM phosphate buffer atpH 6.5 at 25°C) upon addition of glutaminase (KENLESMV-COOH) measured from 1:0.00 to 1:1.09 molar ratio. Ratios of GIP to glutaminase are 1:0 (*magenta*), 1:0.18 (*green*), 1:0.36 (*red*), 1:0.55 (*brown*), 1:0.73 (*blue*), 1:0.91 (*black*), 1:1.09 (*brown*).

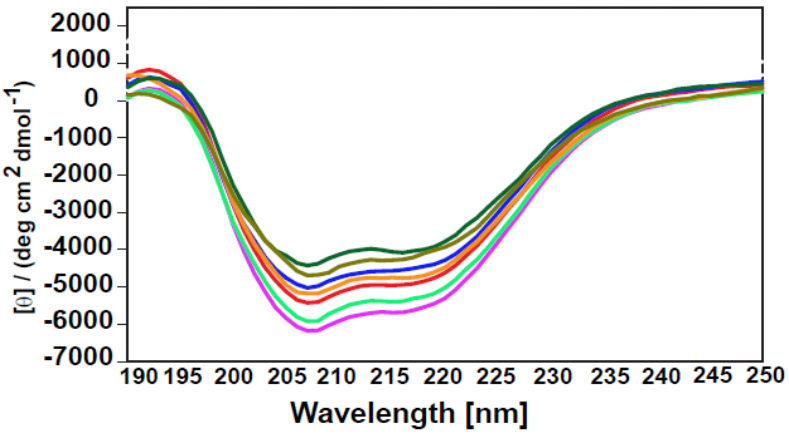

Figure 2.8 Changes in the far-UV CD spectra of 20 µM GIP (in 20 mM phosphate buffer at pH 6.5 at 25°C) upon addition of β-catenin (FDTDL-COOH) measured from 1:0.00 to 1:1.5 molar ratio. Ratios of GIP to β-catenin are 1:0 (*magenta*), 1:0.2 (*green*), 1:0.4 (*red*), 1:0.6 (*orange*), 1:0.8 (*blue*), 1:1 (*Gold*), 1:1.5 (*green*).

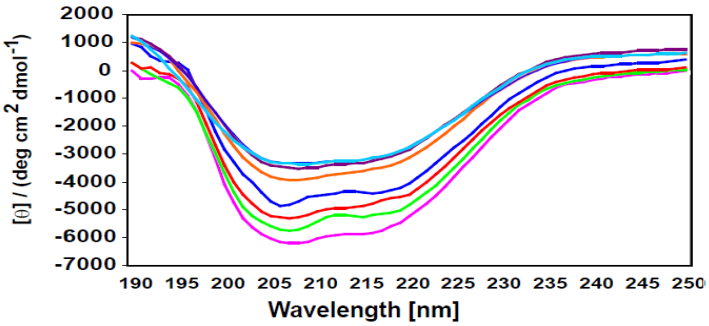

Figure 2.9 Changes in the far-UV CD spectra of 20 µM GIP (in 20 mM phosphate buffer atpH 6.5 at 25°C) upon addition of FAS (NEIQSLV-COOH) measured from 1:0.00 to 1:1.5 molar ratio. Ratios of GIP to FAS (bottom to top) are 1:0 (*magenta*), 1:0.2 (*green*), 1:0.4 (*red*), 1:0.6 (*blue*), 1:0.8 (*orange*), 1:1 (*brown*), 1:1.5 (*cyan*).

2.3.4.2 Fluorescence and homology based model structure of GIP

Tryptophan fluorescence of GIP was quenched effectively by C-terminus peptide ligand. For each sample, when the sample was exited at 280 nm, the emission maximum was observed at a wavelength of 353 nm (characteristic of solvent-exposed tryptophan) and did not shift during titration.

This result is consistent with the homology based model structure of GIP, in which the side chain of Trp83 is fully exposed to the solvent (**Figure 2.10**).

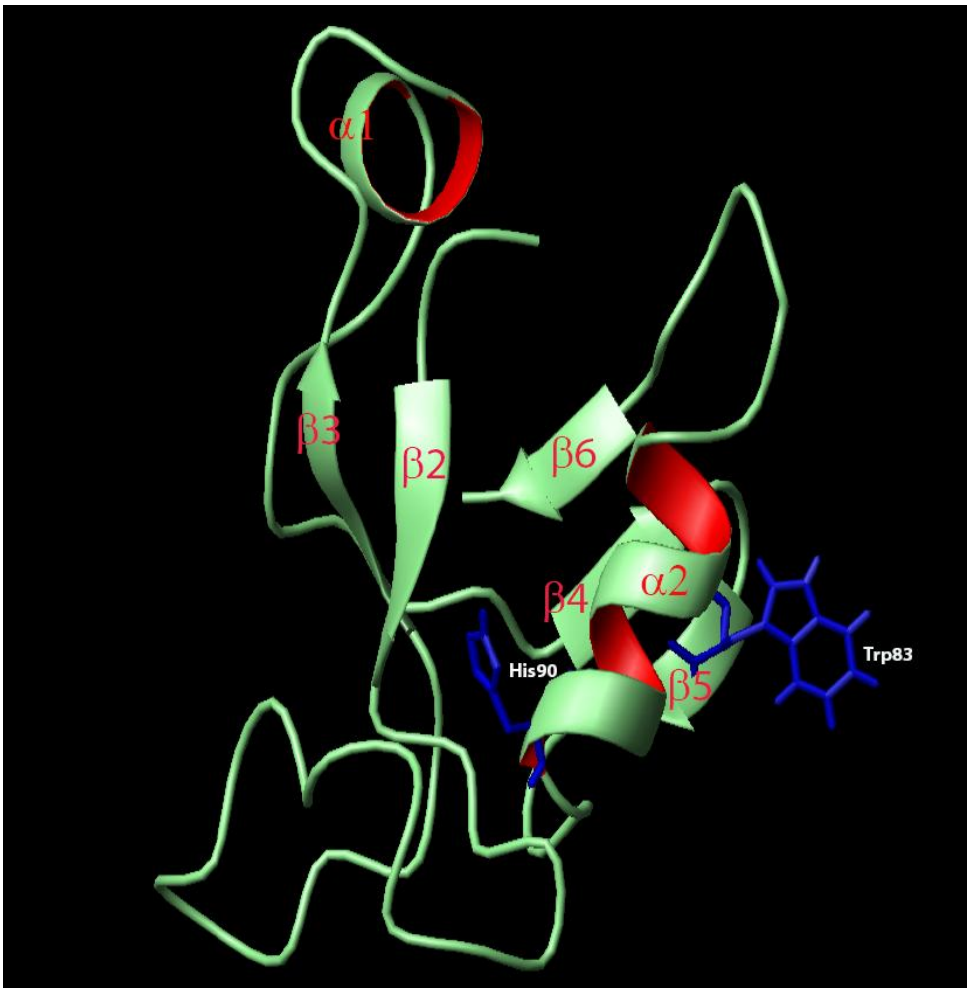

Figure 2.10 Homology based 3D Model structure of GIP. Secondary structure elements are indicated. The side chain of the tryptophane residue, Trp83, and Histidine residue, His90 are shown in blue and labeled. The model was prepared with the program MOLMOL.

2.3.4.3 Interactions of target protein recognition peptides and GIP by fluorescence spectroscopy

GIP contains only one tryptophan residue. However this tryptophan can serve as an intrinsic fluorescence probe providing important insight about possible conformation change upon ligand binding. The possible change in the native environment of tryptophan upon ligand binding can manifest itself in the quenching of the fluorescence of the tryptophan residue. Upon addition of ligands, the microenvironment of the tryptophan residue may change which will cause a change in the fluorescence spectrum. This is a useful technique to study protein-protein interactions. Upon protein-ligand interactions, fluorescence may be altered due to energy-transfer or conformational change of a protein. Dissociation constant values can easily be obtained in a straightforward manner for the interaction studies. This is a common technique to find out dissociation constant values for protein-ligand interactions.

Addition of target peptides to GIP resulted in a small but consistent decrease in fluorescence intensity. The dissociation constant K_D ($K_D = 1/K_a$) was determined using the Eq. 1. (*35*):

$$\log[C]_f = -\log[K_a] + \log[(F_0-F_c)/(F_c-F_i)] \quad \text{(Eq. 1)}$$

From the ordinate intercept of the double reciprocal plot of $F_0/(F_0–F_c)$ versus 1/[C], where F_0 and F_c are the fluorescence intensities of the free protein and that of the protein at a peptide concentration [C], F_i, the fluorescence intensity upon saturation of all the ligand binding sites was obtained. Assuming 1:1 ligand binding to protein, double logarithmic plots were drawn. In the plot of $\log[(F_0-F_c)/(F_c-F_i)]$ versus log[C], the

abscissa intercept yielded the K_D value (the dissociation constant) for the protein-ligand interactions, which is the reciprocal of K_a (the association constant).

The titration of C-terminal peptide sequence motifs of glutaminase (KENLESMV-COOH), β-catenin (FDTDL-COOH) and FAS (NEIQSLV-COOH) with GIP resulted in a reduction in fluorescence intensity and yielded dissociation constants of 1.66 μM, 1.94 μM and 2.64 μM respectively (**Figures 2.11-2.13**).

The free energy changes of the association were calculated by using the Eq. 2

$$\Delta G = -RT \ln K_a \qquad \text{(Eq. 2)}$$

where K_a is the association constant, T is temperature and R is universal gas constant.

The corresponding ΔG values for binding of glutaminase, β-catenin and FAS C -terminus to GIP are -32.97 kJ mol^{-1}, -32.58 kJ mol^{-1}, and -31.82 kJ mol^{-1} respectively, which indicate the spontaneous nature of these bindings.

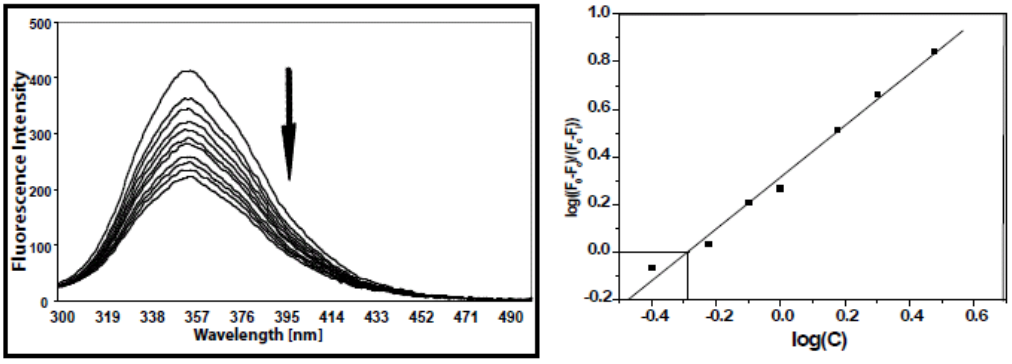

Figure 2.11 Double reciprocal plot of titration of 1 µM GIP (in 20 mM phosphate buffer (pH 6.5) at 25°C) with Glutaminase C-terminus. Inset represents the fluorescence emission plots corresponding to (top to bottom) 0 to 2.5 µM concentration of KENLESMV in 1 µM protein sample. (B) Double-log plot of glutaminase quenching of GIP fluorescence at 298 K.

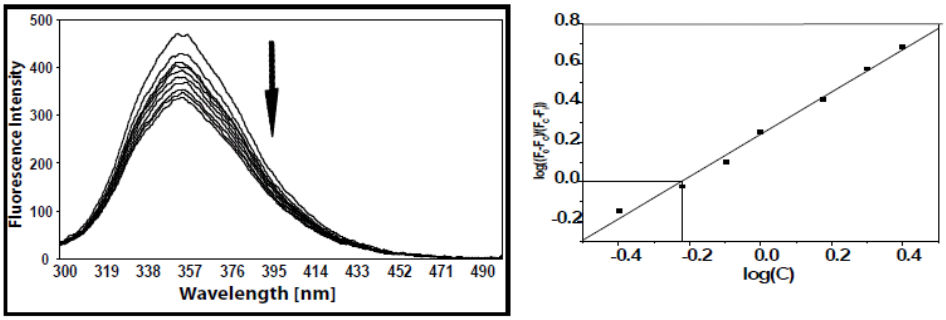

Figure 2.12 Fluorescence emission spectra of 1 µM GIP (in 20 mM phosphate buffer (pH 6.5) at 25°C) with C-terminal peptides of β-catenin (FDTDL). Double reciprocal plot of titration of GIP with β-Catenin C-terminus. Inset represents the fluorescence emission plots corresponding to (top to bottom) 0 to 5 µM concentration of FDTDL in 1 µM protein sample. (D) Double-log plot of β-catenin quenching of GIP fluorescence at 298 K.

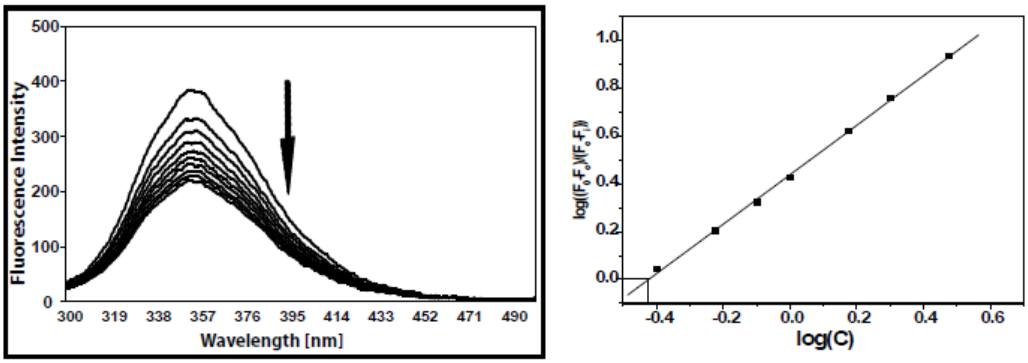

Figure 2.13 Fluorescence emission spectra of 1 µM GIP (in 20 mM phosphate buffer (pH 6.5) at 25^{o}C) with C-terminal peptides of FAS (NEIQSLV). Double reciprocal plot of titration of GIP with FAS C-terminus. Inset represents the fluorescence emission plots corresponding to (top to bottom) 0 to 3 µM concentration of NEIQSLV in 1 µM protein sample. (F) Double-log plot of FAS quenching of GIP fluorescence at 298 K.

2.3.5 GIP backbone assignment

Assignment of $^{1}H_{N}$, $^{1}H_{\alpha}$, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{13}CO backbone resonances was carried out based upon three-dimensional heteronuclear triple resonance experiments that used one- and two-bond scalar couplings to connect the atoms. Regions of 3D HNCACB spectrum is showing sequential connectivity for residues between V109 and K116 (**Figure 2.14**).

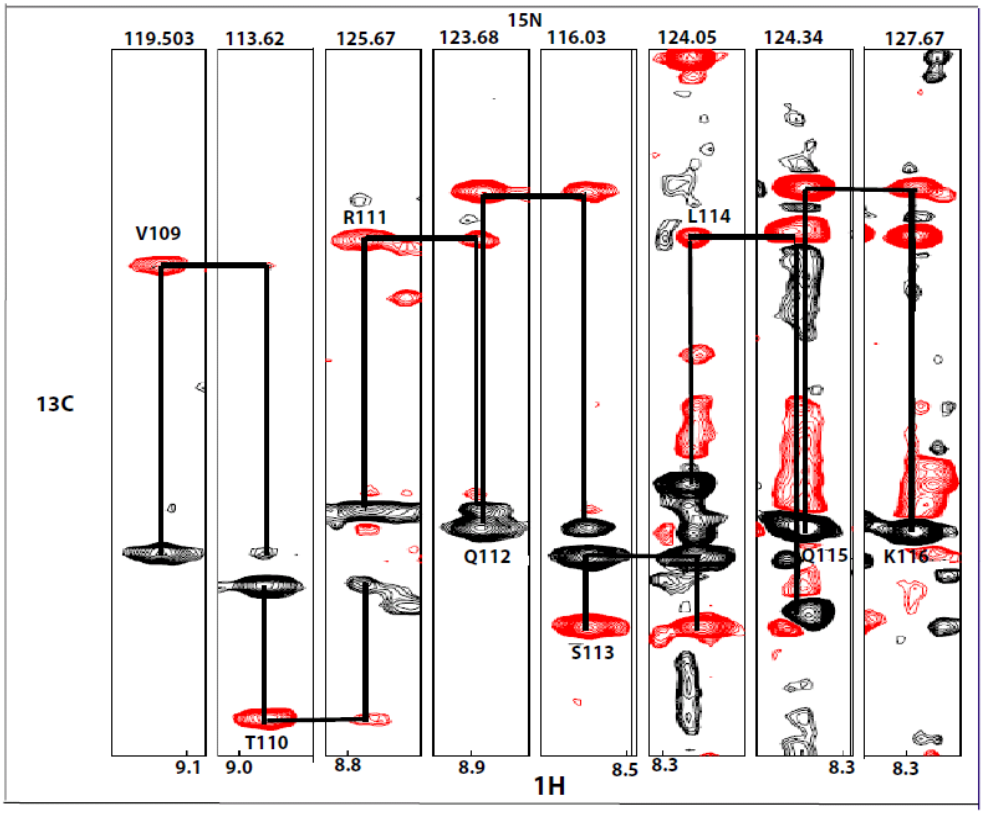

Figure 2.14 Regions of 3D HNCACB spectrum showing sequential connectivity for residues between V109 and K116.

The 2D {^{1}H, ^{15}N}-HSQC spectrum with single-letter amino- acid code with residue number is shown (**Figure 2.15***)*.

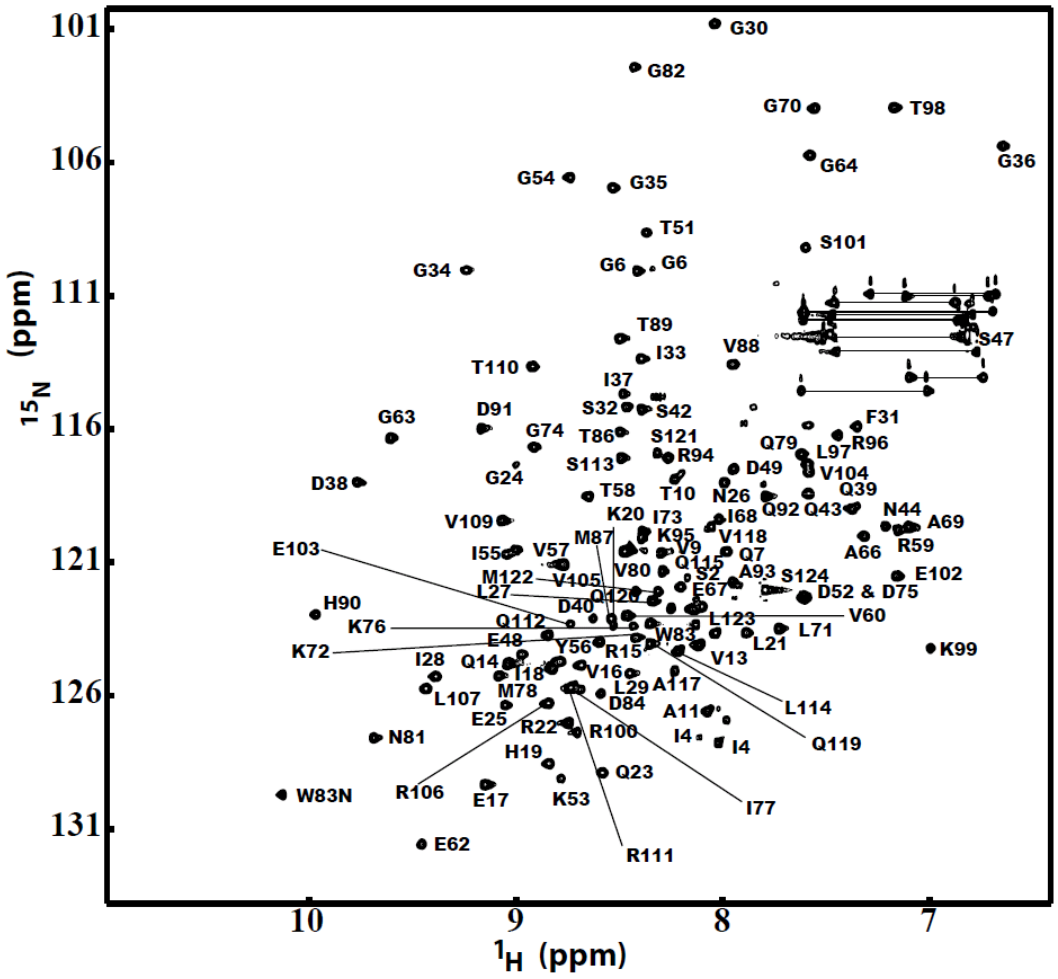

Figure 2.15 2D {^{1}H, ^{15}N}-HSQC spectrum of GIP in 50 mM phosphate buffer at pH 6.5 and 25°C. The assignment of the GIP is labeled in the HSQC spectrum.

2.3.6 Secondary Structure

The secondary structure of GIP was calculated in our lab based on the assigned chemical shifts using the program PSSI, which as well corrected the inconsistencies in ^{13}C and ^{15}N chemical shift references. The assigned chemical shifts were referenced according to the PSSI results and exported to the PsiCSI program for the secondary structure calculation. These results indicate that GIP consists of six β-strands and two α-helices with residues: 11-23 (β1), 27-35 (β2), 54-60 (β3), 64-70 (α1), 76-81 (β4), 83-88 (β5), 90-100 (α2) and 103-112 (β6) at pH 6.5 and 298 K (**Figure 2.16**).

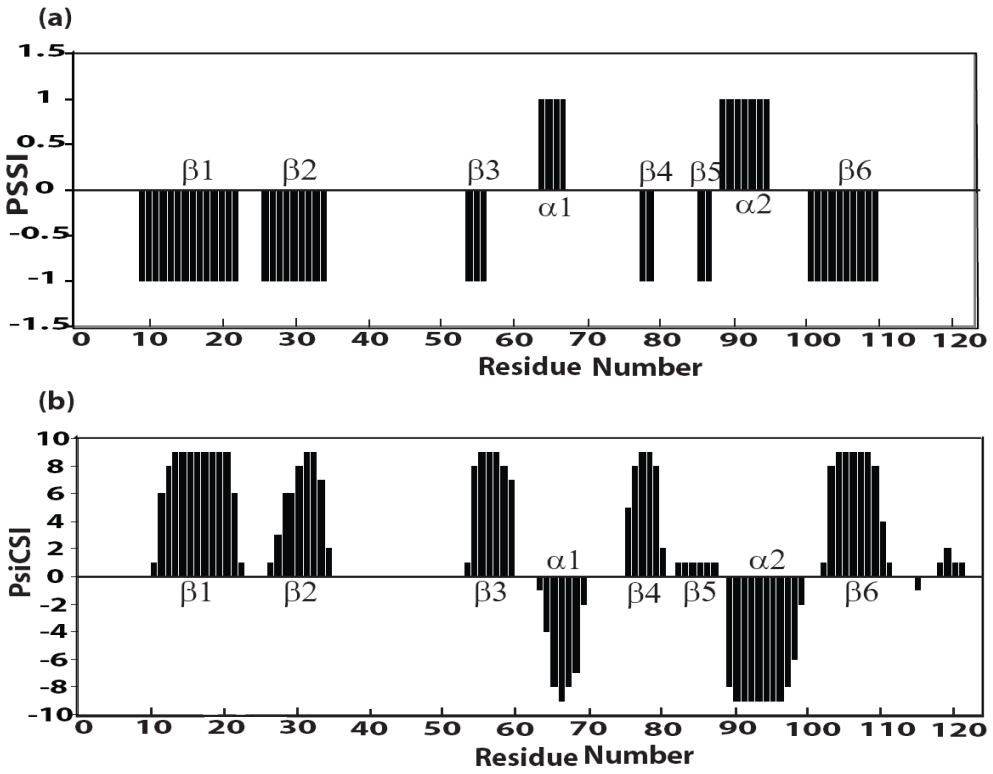

Figure 2.16 Consensus (using H_α, C_α and CO chemical shifts) chemical shift index(CSI) showing secondary structural elements of GIP using (a) PSSI and (b) PSICSI programs in our lab. Adapted from reference 21.

2.3.7 Interaction of GIP with C-terminal target protein sequence motifs by {^{1}H, ^{15}N}-HSQC NMR

NMR is a versatile technique for monitoring structure-activity relationships (SAR) for protein-protein or protein-ligand interactions (*36*). In the 2D {^{1}H, ^{15}N}-HSQC spectra, the chemical shifts of residue/s change upon ligand binding. This region is called the finger print region of a protein and is very sensitive. It provides information on the protein conformation. Perturbation in the chemical shifts from the original positions upon ligand binding or change in temperature or pH indicates conformational change of a protein. To investigate the binding and possible conformational change in GIP, titration studies were carried out with the three C-terminal peptides that act as surrogate to the target protein while monitoring the fingerprint region of the protein in the 2D {^{1}H, ^{15}N}-HSQC spectra. NMR titration studies were performed for GIP with excess ($\geq$ 70 times) ligand. Analysis of the HSQC spectra of the titration studies indicate that significant changes in the chemical shift positions occur for the complex/es. An overlay of the HSQC spectra collected in the absence and presence of different concentrations of target peptides are shown in **Figures 2.17 – 2.21**. The chemical shift positions of most residues in the free GIP were perturbed in the HSQC spectra upon binding to C-terminal target protein recognition sequence motifs. Dissociation constant (K_D) values for various residues of GIP were calculated assuming 1:1 binding stoichiometry. Estimates for dissociation constants (K_D), measured using fluorescence and NMR techniques, have ranged from low micromolar to mid micromolar indicating moderate affinities. It should be noted here that depending upon techniques used, and the initial protein concentration required for a particular technique, dissociation constant varies (*37-39*).

We observed that the free FAS peptide precipitated out from the buffer slowly with time at all concentrations. Some precipitation was observed during the CD and HSQC titration experiments of both β-catenin and FAS peptides with GIP, although the signal to noise ratio (S/N) of the protein was not significantly affected. However, we did not find any precipitation (by visual inspection) during titration studies carried out by fluorescence measurements which may be due to the less protein used in fluorescence experiments. Unlike the FAS C-terminal peptide, β-catenin C-terminal peptide was stable in aqueous buffer solution, although freshly prepared ligand was always used for all titration studies. It is important to mention here that C-terminal amidated β-catenin peptide (FDTDL-$CONH_2$) does not interact with GIP (data not shown) while the unprotected C-terminal peptide (FDTDL-COOH) does bind to GIP.

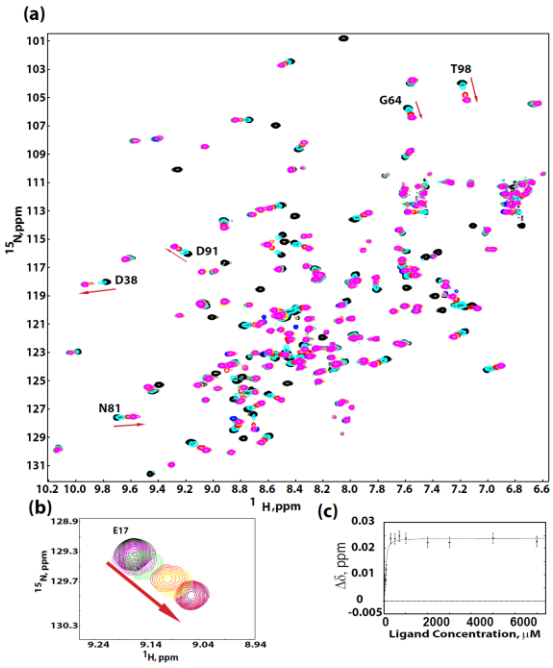

Figure 2.17 Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of glutaminase (KENLESMV) to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25^{o}C. (A) The 2D {^{1}H, ^{15}N}-HSQC spectrum demonstrating chemical shift perturbations of residues upon titration of glutaminase (KENLESMV) to GIP. Ratios of GIP to glutaminase ranges from 1:0 to 1:70. (B) Expanded regions of the spectrum demonstrating chemical shift perturbations of residue E17 upon titration of GIP with glutaminase (KENLESMV). Ratios of GIP to glutaminase are 1:0 (*black*), 1:0.2 (*magenta*), 1:0.4 (*green*), 1:0.8 (*red*), 1:1 (*yellow*), 1:50 (*blue*), 1:70 (*red*). (C) NMR titration curve for the titration of GIP with glutaminase (KENLESMV). The plot shows the changes in the chemical shift of E17 induced by the addition of glutaminase (KENLESMV), versus the peptide concentration. Dashed line is the titration curve as fitted by the program modelTitr from NMRPipe. The apparent dissociation constants K_D corresponding to residue E17, was determined by fitting the chemical shift change of the residue. The determined K_D value is 20.7 ± 5.8 µM.

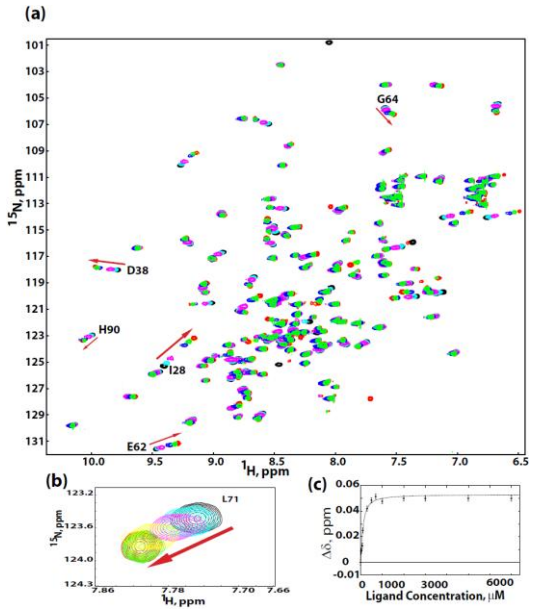

Figure 2.18 Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of β-catenin (FDTDL-COOH) to 100 μM GIP in 50 mM phosphate buffer (pH 6.5) at 25^{o}C. (A) The 2D {^{1}H, ^{15}N}-HSQC spectrum demonstrating chemical shift perturbations of residues upon titration of β-catenin (FDTDLCOOH) to 100 μM GIP. Ratios of GIP to β-catenin (FDTDL) ranges from 1:0 to 1:90. (B) Expanded regions of the spectrum demonstrating chemical shift perturbations of residue L71 upon titration of GIP with β-catenin (FDTDL). Ratios of GIP to β-catenin (FDTDL) are 1:0 (black), 1:0.4 (*cyan*), 1:1 (*magenta*), 1:20 (*yellow*), 1:70 (*red*), 1:90 (*green*). (C) NMR titration curve for the titration of GIP with β-catenin. The plot shows the changes in the chemical shift of L71 induced by the addition of β-catenin, versus the peptide concentration. Dashed line is the titration curve as fitted by the program modelTitr from NMRPipe. The apparent dissociation constants K_D, corresponding to residue L71, was determined by fitting the chemical shift change of the residue. The determined K_D value is 21.6 ± 3.5 μM.

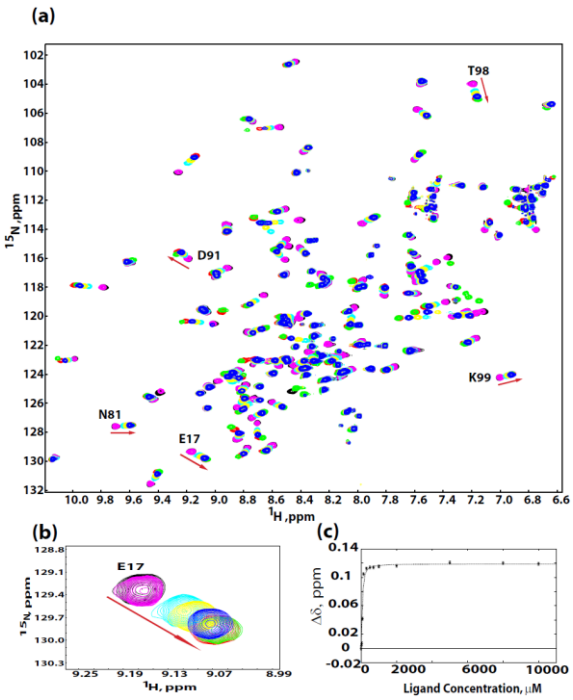

Figure 2.19 Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of FAS (NEIQSLV) to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25°C . (A) The 2D {^{1}H, ^{15}N}-HSQC spectrum demonstrating chemical shift perturbations of residues upon titration of FAS (NEIQSLV) to GIP. Ratios of GIP to FAS (NEIQSLV) range from 1:0 to 1:100. (B) Expanded regions of the spectrum demonstrating chemical shift perturbations of residue E17 upon titration of GIP with FAS (NEIQSLV). Ratios of GIP to FAS (NEIQSLV) are 1:0 (*black*), 1:0.4 (*magenta*), 1:3 (*cyan*), 1:5 (*yellow*), 1:10 (*blue*), 1:50 (*red*), 1:100 (*green*). (C) NMR titration curve for the titration of GIP with FAS (NEIQSLV). The plot shows the changes in the chemical shift of E67 induced by the addition of FAS (NEIQSLV), versus the peptide concentration. Dashed line is the titration curve as fitted by the program modelTitr from NMRPipe. The apparent dissociation constants K_D corresponding to residue E67, was determined by fitting the chemical shift change of the residue. The determined K_D value is 19.7 ± 2.1 µM.

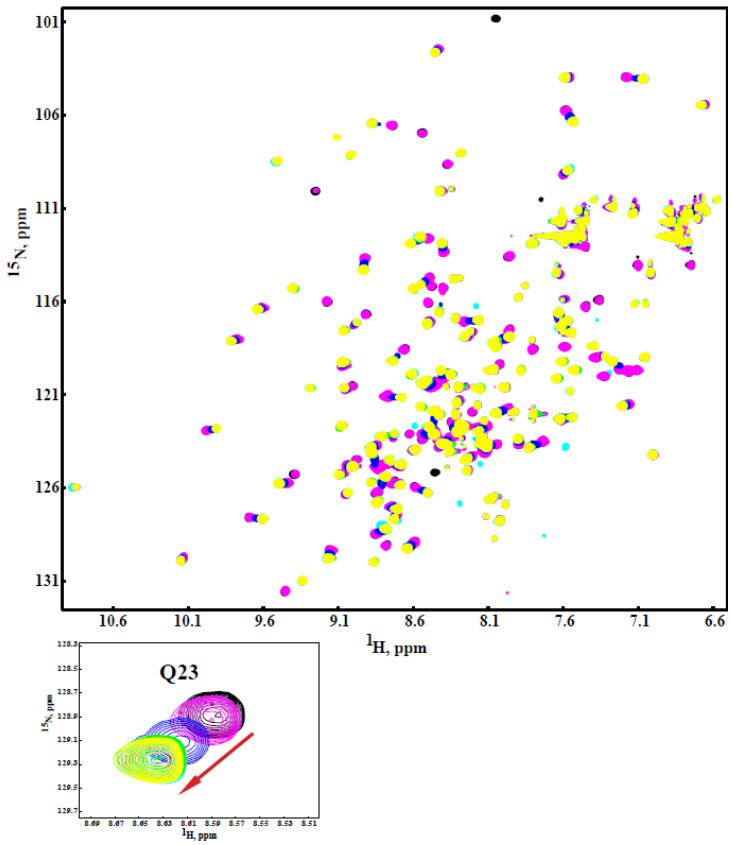

Figure 2.20 Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of Kir 2.3 to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25°C. (A) The 2D {^{1}H, ^{15}N}-HSQC spectrum demonstrating chemical shift perturbations of residues upon titration of Kir 2.3 to GIP. Ratios of GIP to Kir 2.3 range from 1:0 to 1:100. (B) Expanded regions of the spectrum demonstrating chemical shift perturbations of residue Q23 upon titration of GIP with Kir 2.3. Ratios of GIP to Kir 2.3 are 1:0 (*black*), 1:0.4 (*magenta*), 1:3 (*cyan*), 1:5 (*Blue*), 1:10 (*yellow*), 1:50 (*green*).

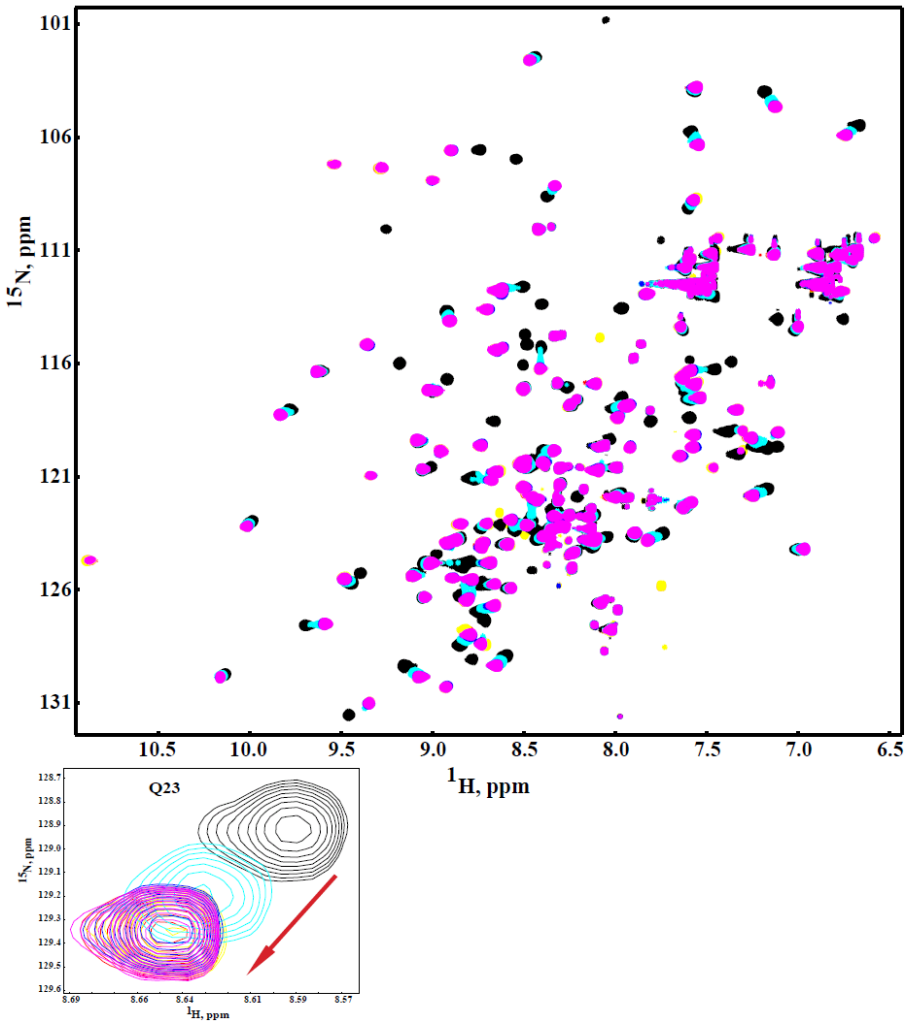

Figure 2.21 Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of HPV 16 E6 to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25°C. (A) The 2D {^{1}H, ^{15}N}-HSQC spectrum demonstrating chemical shift perturbations of residues upon titration of HPV 16 E6 to GIP. Ratios of GIP to HPV 16 E6 range from 1:0 to 1:100. (B) Expanded regions of the spectrum demonstrating chemical shift perturbations of residue Q23 upon titration of GIP with HPV 16 E6. Ratios of GIP to HPV 16 E6 are 1:0 (*black*), 1:0.4 (*cyan*), 1:3 (*yellow*), 1:10 (*blue*), 1:50 (*magenta*).

2.3.8 Chemical shift perturbation of GIP upon binding to glutaminase, β-catenin and FAS C-termini

Chemical shift mapping is a powerful method to investigate possible protein-ligand interactions by NMR. To investigate GIP-ligand binding, we studied the interaction of GIP with C-terminal sequence motifs of target proteins by NMR. In the {^{1}H, ^{15}N}-HSQC spectra, the amide proton and nitrogen resonances of most residues shifted gradually with increasing ligand concentration, indicating that the complexes were in fast exchange on the NMR timescale except for few residues which were found to be in intermediate exchange. Glutaminase C-terminal motif caused pronounced chemical shift changes in the {^{1}H, ^{15}N}-HSQC spectrum of GIP. Residues with large chemical shift changes (greater than 0.1 ppm) of their backbone amide group were found to cluster in several regions of the GIP sequence while interacting with glutaminase L, in particular, Ile18-Arg22, Gly34-Ile37, Ser42-Gln43, Tyr56-Glu62, Leu71, Thr86, Arg94-Arg101, and Phe46 (**Figure 2.22A**). However, residues Gly34, Gln39, Ile55, Tyr56, Glu62, Met78 and Met87 showed chemical shift changes greater than 0.2 ppm (highly shifted amide groups). The resonances of several residues particularly Ile28, Leu29, Gly30, Ile33 and Gln39 located in the β2 strand showed severe broadening, indicating the presence of conformational exchange phenomena. The largest chemical shift changes also occurred in the long stretch Ile28-Gln39, which corresponds to the β2 strand. Analysis of the chemical shift differences between the GIP and the GIP-β-catenin complex revealed that the largest changes in the amide resonance frequencies occur for Arg106 and the residues in the polypeptide segments Ile28-Gln39, Ile55-Gly76 and Thr86-Arg96 which correspond to β2, β3, α1, β5, α2 (**Figure 2.22B**).

Residues Leu29 and Gly30 also show conformational exchange. Chemical shift perturbation map of GIP-FAS complex indicated that many residues are affected from different regions of the protein. The chemical shift difference between the GIP and GIP-FAS complex revealed that the largest changes in the amide resonance frequencies occur for Glu17-Gln23, Phe31-Gln39, Glu48-Thr58, Ser61-Glu62, Glu67, Ala69, Arg94-Glu102, Val105-Arg106, Arg111, and Lys116 (**Figure 2. 22C**). Residues Gly30, Ile28, and Leu29 show conformational exchange. These results indicate that the protein undergoes a conformational change upon binding to ligands.

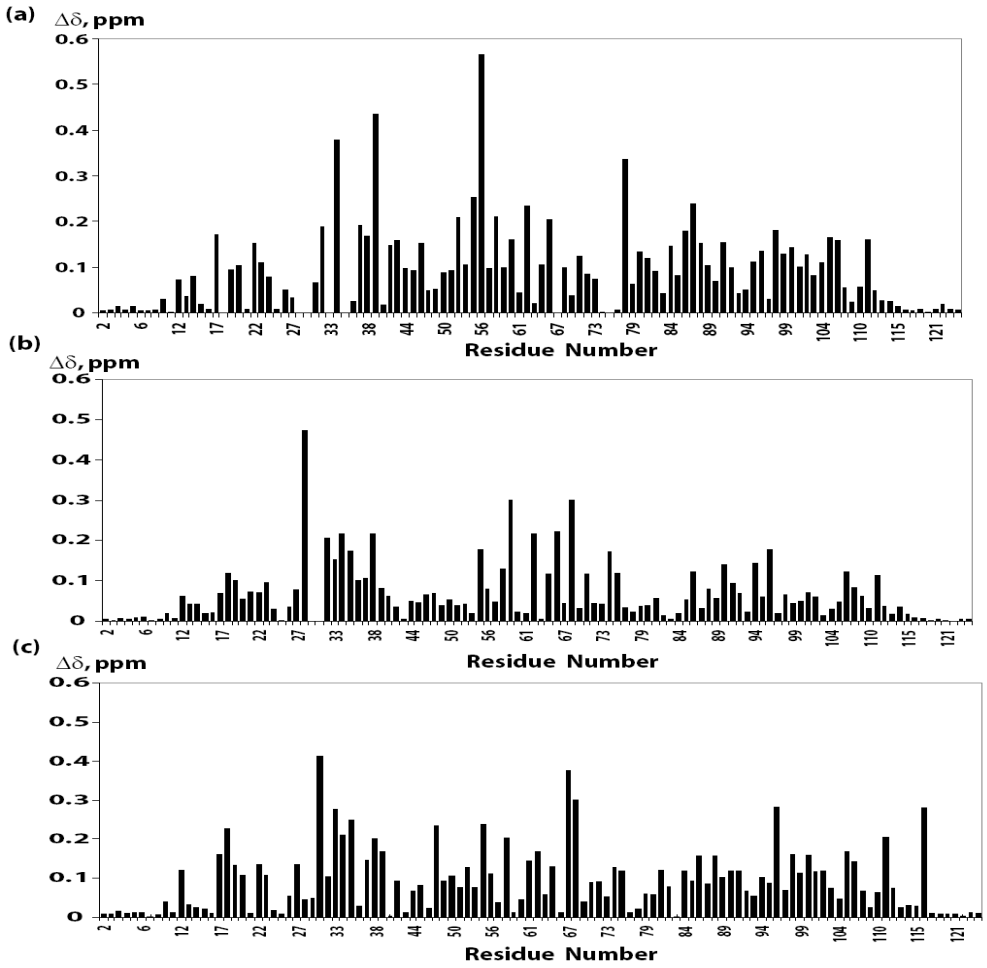

Figure 2.22 Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with glutaminase, β-catenin and FAS C-terminus. The combined ^{1}H and ^{15}N chemical shift differences, calculated using the equation $\Delta\delta = |\Delta\delta^{15}N|/f + |\Delta\delta^{1}H|$, where f is the scaling factor (f=8.7) used to normalize the ^{1}H and ^{15}N chemical shifts, were plotted against residue number. (A) Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding the glutaminase C-terminus. (B) Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding the β-catenin C-terminus. (C) Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding the FAS C-terminus.

Kir 2.3 caused pronounced chemical shift changes in the {^{1}H, ^{15}N}-HSQC spectrum of GIP. Residues with large chemical shift changes (greater than 0.2 ppm) of their backbone amide group were found to cluster in several regions of the GIP sequence while interacting with kir 2.3 peptide, in particular, Ile28- Gly35, Thr86- Gln92 (**Figure 2.23A**). However, residues Leu29, Gly30, Gly34, Glu67, Gly74, val88, Asp91 showed chemical shift changes greater than 0.2 ppm (highly shifted amide groups). The resonances of several residues particularly Ile28, Leu29, Gly30, Ile33 located in the β2 strand showed severe broadening, indicating the presence of conformational exchange phenomena. The largest chemical shift changes also occurred in the β2 strand and the α2 helix.

When HPV 16 E6 was added, systematic changes in the spectra occurred and the beginning and end points of the HPV 16 E6 titration, where the chemical shift perturbation is most easily identified. The signals fall into three categories. First, many signals show substantial changes in chemical shift (*i.e.* the peak positions change by significantly more than the total peak line width). Residues with large chemical shift changes (greater than 0.2 ppm) of their backbone amide group were found to cluster in several regions of the GIP sequence while interacting with HPV 16 E6 peptide, in particular, Ile28-Ile37, Thr86-Arg96 (**Figure 2.23B**). However, residues Ile28, Leu29, Gly30, Ile 33, Gly34, Gly35, Asp40, Lys53, Thr86, val88, Asp91, Gln92 and Arg96 showed chemical shift changes greater than 0.2 ppm (highly shifted amide groups). The resonances of several residues particularly Leu29, Gly30, Ile33 located in the β2 strand showed severe broadening, indicating the presence of conformational exchange phenomena.

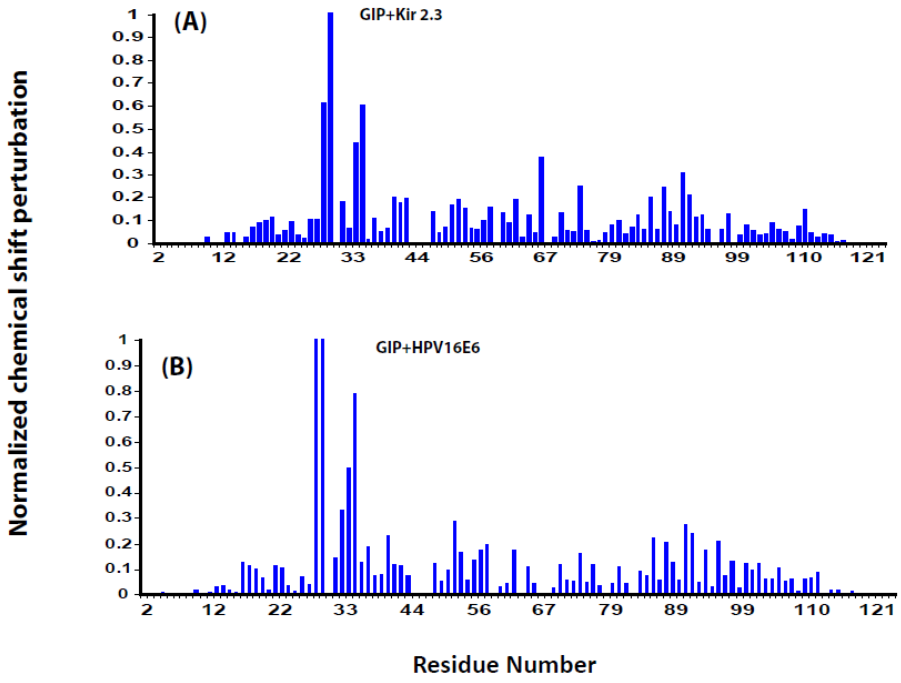

Figure 2.23 Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the kir 2.3 and HPV 16 E6 C-terminus. The combined ^{1}H and ^{15}N chemical shift differences, calculated using the equation $\Delta\delta = |\Delta\delta^{15}N|/f + |\Delta\delta^{1}H|$, where f is the scaling factor (f=8.7) used to normalize the ^{1}H and ^{15}N chemical shifts, were plotted against residue number. (A) Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding the kir 2.3 C-terminus. (B) Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding the HPV 16 E6 C-terminus.

2.3.9 Molecular model and refinement

A molecular model was constructed for the residues ranging from Leu29 to Leu108 using known coordinates of most similar non-identical protein second PDZ domain of human scribble protein of known structure (PDB code 1whaA). Human scribble protein shares 43 % sequence identity with GIP. The rough model was constructed with SWISS-MODEL (*38, 39*). The three-dimensional visualization was performed on a UNIX workstation using the MOLMOL program (*40*).

2.3.10 Dynamics of GIP and GIP-glutaminase peptide complex from ^{15}N relaxation measurements

^{15}N relaxation data of the backbone amide nitrogens in both free and complexed forms, including steady-state {^{1}H-^{15}N} NOE intensities and R_1 and R_2 relaxation rates as a function of GIP sequence were calculated. Overlapped peaks and uncharacterized residues are excluded from the data. The backbone ^{15}N relaxation parameters are broadly similar. The residues in the defined secondary structure exhibit highly restricted mobility, while residues at both termini of the protein are flexible, which is consistent with our secondary structure model. There is a little change in the ^{1}H-^{15}N NOE value upon glutaminase peptide binding.

To further elucidate the binding mechanism of the Glutaminase L peptide to GIP, we carried out backbone dynamics studies. The order parameters {S^2} from both free and complex forms of GIP were calculated using: steady-state ^{1}H-^{15}N NOE intensities, R_1 and R_2 relaxation rates by Modelfree analysis. R_1 (**Figure 2.24A**) and R_2 (**Figure 2.24B**) relaxation rates and steady state ^{1}H-^{15}N NOE (**Figure 2.24C**) intensities for GIP are

determined. The order parameters $\{S^2\}$ of GIP were calculated using: steady-state ^{1}H-^{15}N NOE intensities, rate constant for T_1 (R_1) and rate constant for T_2 (R_2)relaxation rates by Modelfree analysis (**Figure 2.25**).

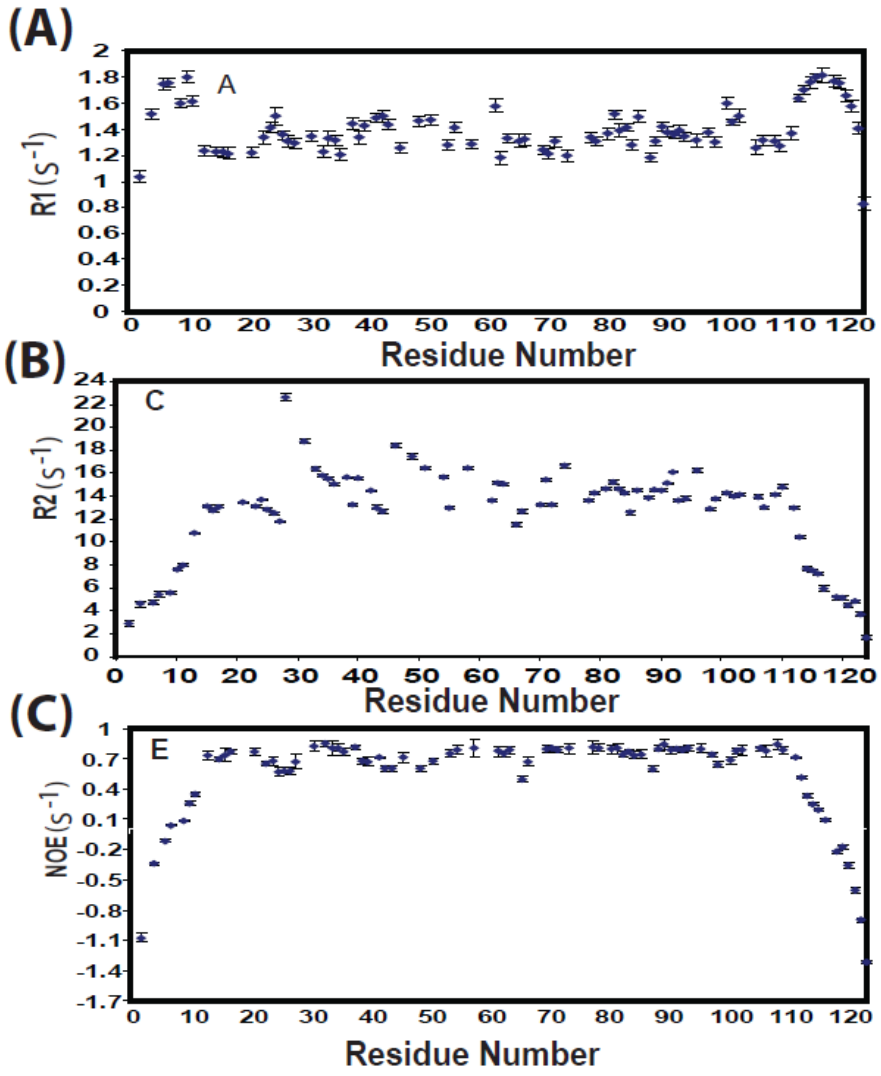

Figure 2.24 Plots of R_1, R_2, NOE as a function of residue number. Measured R_1 values are shown in panels A R_2 values are shown in panels B, and ^{1}H-^{15}N NOE data are shown in panels C.

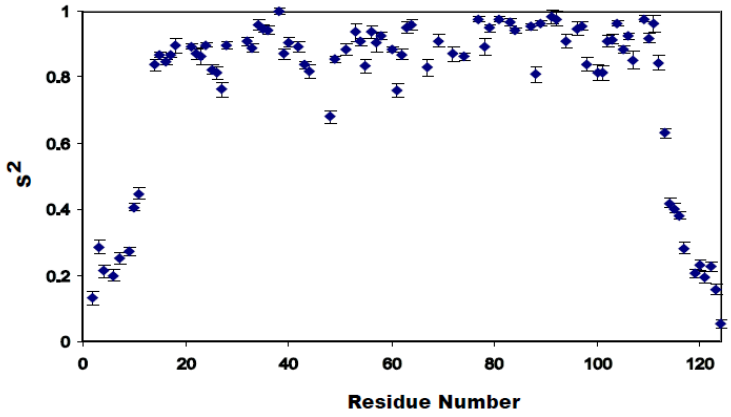

Figure 2.25 Plot S^2 as a function of residue number for GIP.

Similarly, R_1 (=1/T_1) (**Figure 2.26A**) and R_2 (=1/T_2) (**Figure 2.26B**) relaxation rates and steady state ^{1}H-^{15}N NOE (**Figure 2.26C**) intensities for GIP-glutaminase L complexes are determined. The order parameters {S^2} of GIP-glutaminase L complexes were calculated using: steady-state ^{1}H-^{15}N NOE intensities, R_1 and R_2 relaxation rates by Modelfree analysis (**Figure 2.27**). Overlapped peaks and residues that could not be characterized because of their low intensity or absence in the HSQC spectra were excluded from our data analysis. Residues M1, P5, P8, V12, L21, P41, P45, K50, D52, P65, D75, and V80 were excluded from both free and bound states. Residues K20, L29, G30, V57, R59, I68, A69, I73, M87, K95, V105, and V118 were excluded from only the free state. Residues V13, N26, F31, G35, I37, D40, Q43, E48, D49, Y56, S61, Q72, W83, M85, T86, and A93 were excluded from only the bound state. Of these, L29, G30, F31, G35, D40, Q43, E48, and D49 are residues that form part of the binding pocket, including the ILGF motif (canonical GLGF) and the β2-β3 loop could not be measured. This was because they were too close to the intermediate-exchange regime to provide sufficient intensity required for observation in the NMR dynamics data. The backbone ^{15}N relaxation parameters are broadly similar for both the free and complex states. The residues in the defined secondary structure exhibit a relatively restricted mobility, while residues at both termini of the protein are flexible which is consistent with both our secondary structure model and general NOE pattern. The 1:2 ratio was chosen for studying the dynamic interaction between GIP and Glutaminase L peptide. It is interesting to note that the residues that encompass the binding loop ILGF (consensus GLGF) and most of β2 disappear in the titration process.

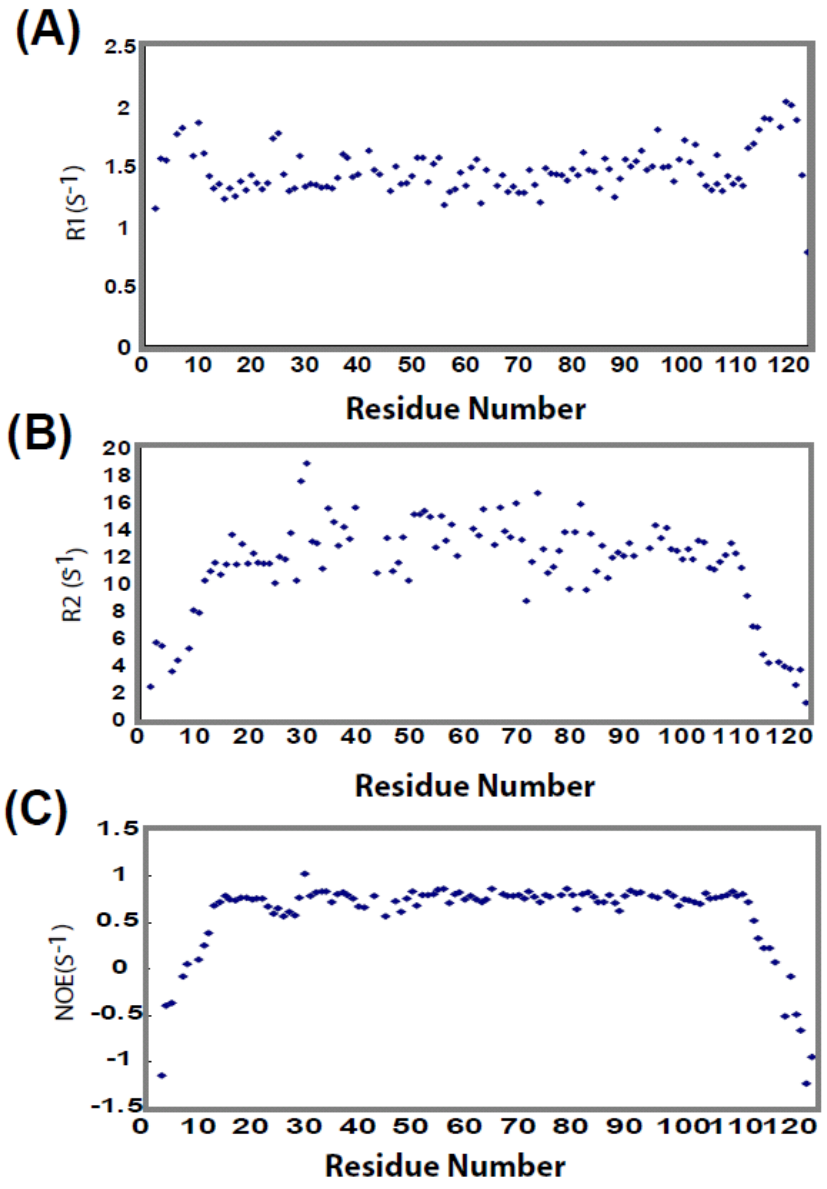

Figure 2.26 Plots of R_1, R_2, NOE of GIP-glutaminase complex as a function of residue number. Measured R_1 values are shown in panel A, R_2 values are shown in panelB, and ^{l}H-^{15}N NOE data are shown in panel C.

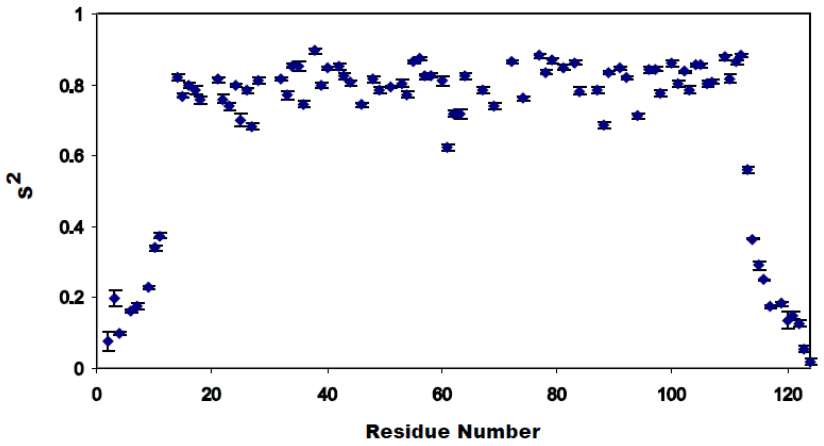

Figure 2.27 Plot of S^2 of GIP-Glutaminase L complex as a function of residue number.

Fascinatingly, the overall correlation times of the free and complex forms of the GIP were found to be 10.1 and 9.3 nsec, respectively. Dynamics of GIP and GIP-Glutaminase complex were carried out to understand the function of GIP. The dynamic properties of GIP show a regular pattern; secondary structures are more rigid and loop structures are more flexible. Interestingly, residues Gly34, Asp38, Gly64, and Gln91 exhibit S^2 values lower than 1 in the complex form of the protein. Hence, it appears that GIP induces a conformational change whereby residues that have lower flexibility in the free state are released from their restricted environment in the complex state.

In the free GIP, residues Ile28, Phe31, Ile33, Gly35, Gly36, Asp40, Phe46, Asp49, Thr51, Gly54, Thr58, Gly63, Leu71, Gly74, Gly82, Val88, Gln92, Arg96, Ser101 of GIP showed conformational exchanges ($R_{ex}>1$) on the micro- to millisecond timescale. Upon binding to the C-terminal glutaminase peptide, residues Gly24, Gly34, Gly36, Gln39, Ser42, Phe46, Thr51, Gly54, Val57, Thr58, Gly63, Ala66, Leu71, Met85, Asp91, Leu97, Lys99, and Thr110 showed conformational exchanges (*Rex>1*) on the micro- to millisecond time scale.

2.4 Discussion

Because GIP plays a pivotal role in many cellular signaling pathways, it is considered an excellent drug target (*41*). For successful drug design, understanding the mechanisms of protein-protein interactions, binding, recognition and specificity are very critical.

Recombinant GIP has been produced in minimal media with a yield of 56 mg/L for structural, functional, and protein-protein interaction studies. The single tryptophan residue present in the GIP protein is a convenient feature for fluorescence studies. Tryptophan acts as an ideal intrinsic fluorescence probe, as it has a large molar absorption coefficient, and its fluorescence intensity (I_F) and wavelength of maximum intensity (λ_{max}) are sensitive to the microenvironment of the indole group (*42*). The λ_{max} of the tryptophan residue of GIP was red-shifted to 353 nm, characteristic of a solvent-exposed tryptophan side chain. Fluorescence studies for all the three ligands show consistent quenching of intrinsic tryptophan fluorescence intensity upon binding, which is indicative of a change in the environment of the fluorophore. The binding of the ligand to the protein may directly affect the fluorescence of a tryptophan residue with the ligand acting as a quencher, or by physically interacting with the fluorophore and thereby changing the polarity of its environment and/or its accessibility to solvent. Alternatively, the ligand may bind at a site on protein remote from the tryptophan residue but can induce a protein conformational change, altering the microenvironment of the tryptophan (*43*). From our data, it could be concluded that the solvent exposed tryptophan of GIP was not directly involved in ligand binding; rather the ligands induced a conformational change of the GIP and thus altered the microenvironment of the tryptophan. The dissociation constants determined from the fluorescence intensity measurements revealed that the three ligands bind to GIP with similar affinities in the range of 1-10 µM. These moderate affinities are probably suitable for regulatory functions (*44*). Our results show that GIP has several target partners, and we hypothesize that it might have more target partners (**Figure 2.28**).

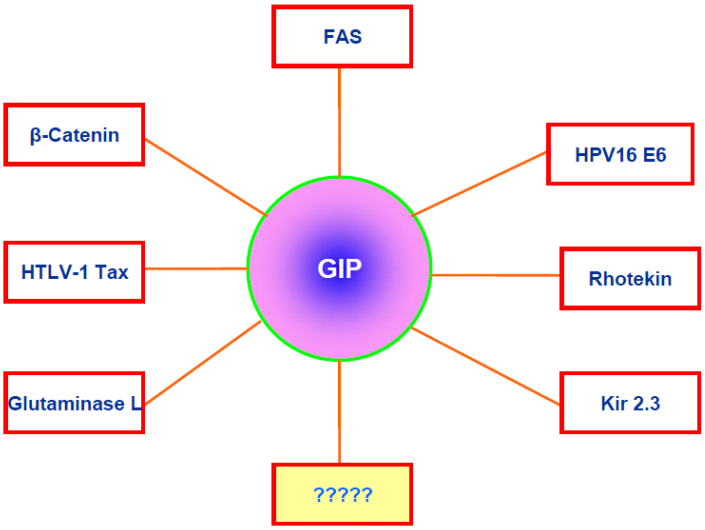

Figure 2.28 GIP tree: various interacting partners of GIP.

Deconvolution of CD data show that the amount of helix and random coil content in GIP decreases with increasing concentration of ligand while the amount of β-sheet increases. Thus the qualitative CD structural analysis is in excellent agreement with 2D HSQC titration data, which show chemical shift perturbation of GIP upon ligand binding.

The secondary structure calculations in our lab from PSSI and PSICSI show that the structure of GIP comprises of 6 β-strands (β1- β6) and 2 α helices (α1 and α2). The three-dimensional model of GIP shows that the major part of the protein consists of a five-strand antiparallel β-sandwich formed by a three-strand β-sheet (β4, β5 and β6) and a two-strand β-sheet (β2 and β3). In the three-dimensional model, α2 orients parallel to β2, thus forming a groove, where the ligands for PDZ domain bind. The mapping of ligand-induced chemical shift perturbations clearly indicates that the ligands bind to the canonical binding pocket, the β2 and α2 groove of GIP.

One important difference between the PDZ domain of GIP and other known PDZ domains is the substitution of the first glycine in the carboxylate-binding motif GLGF by Ile. The GLGF motif is highly conserved among PDZ domain. The second glycine in the GLGF motif is absolutely conserved, but a serine, threonine, or proline replaces the first glycine in a minority of PDZ domains (*19*). The exception to GLGF motif is the PDZ domain of PDZ2 from Human Phosphatase hPTP1E with SLGF motif (*45*). The strict requirement of a C-terminal free carboxyl group of the target protein for binding to PDZ domain is confirmed by the observation that C-terminal protected β-catenin (FDTDL-$CONH_2$) does not bind to GIP, since there is no change in the 2D HSQC titration data (data not shown). According to the peptide binding mechanism of class I PDZ domain (*46*), His90 in the α2 helix should contribute significantly to the binding of ligands C-

termini i.e. the N-3 nitrogen of His residue might be forming a specific hydrogen bond with the hydroxylated side chains of either a serine or threonine residue, thus probably giving rise to a preference for a particular residue at P_{-2} position of class I PDZ domains. In fact, our HSQC titration results confirm that the chemical shift of His90 is affected greatly by the binding to C-termini recognition motifs of the target proteins.

It is clear from our HSQC titration data that chemical shifts for many of the amide signals of GIP change upon complex formation, which suggests that a significant conformational change in GIP might be occuring upon binding of ligands. How does the PDZ domain of GIP recognize different target sequences? What is the mode of interaction and recognition of these targets by the PDZ domain of GIP? A detailed structure-function investigation is needed to answer these questions.

It is known that proteins do exhibit conformational flexibility while performing their biological function. Knowledge of conformational flexibility on various time scales is often essential to understand the function and mode of interaction of a protein. Conformational fluctuation in the millisecond time scale often plays an important role in protein-protein and protein-ligand interactions. Dynamics of the GIP and the GIP-C-terminal glutaminase peptide complex were carried out to understand the interactions between GIP and C-terminal glutaminase peptide. ^{15}N NMR relaxation data have been used to characterize the backbone dynamics of the human glutaminase interacting protein in its free and glutaminase C-terminal peptide bound state. Fast motions in the nano- to picosecond time scale are biologically very important because of their strong effect on the conformational entropy of the system.

Where S^2 could be measured in both free GIP and the GIP-glutaminase L complex, some residues show either substantial increases or decreases in backbone flexibility. In general, residues at the binding site tend to become more ordered, while residues peripheral to the binding site in GIP become more disordered, with a few exceptions. The dynamic properties of GIP show a regular pattern; secondary structures are more rigid and loop structures are more flexible. It also shows that, overall there is an increase of GIP flexibility upon glutaminase C-terminus binding. This, again from another aspect, corroborates our conclusion that GIP is induced to a conformational change after binding to glutaminase recognition sequence motif. Interestingly, this result is contrary to a previous study which proposed a global compaction of PDZ domain after the encounter of the target ligand (*47*). In contrast to the overall increase in flexibility observed in the binding of the Glutaminase L peptide to GIP, previous dynamic measurements of ligand-binding on other PDZ domains have shown more local effects (*48*). It showed that upon ligand binding, while there was a small increase in S^2 values throughout the protein, the differences were most pronounced in the β2-β3 loop and the α2 helix. Thus, the general trend is that the overall complex structure becomes more flexible which is distributed throughout the protein. Furthermore, the increase in flexibility throughout GIP also includes the flexible regions of the protein such as the βa-βb hairpin and β2-β3 loops as well as both termini.

2.5 Conclusions

The studies performed include steps towards understanding the structure, function, dynamics, and mode of interactions of a PDZ domain-containing

multifunctional human protein. GIP was overexpressed and purified to homogeneity by size exclusion chromatography. Our results show that GIP interacts with C-terminus of its several target proteins. GIP undergoes a conformational change upon ligand binding. Structure, function, dynamics, and interaction studies of GIP with different binding partners are critical to provide the insight into the mechanisms and role of this PDZ domain containing human protein. Our results show that GIP is involved in various signaling pathways, by recognition of C-terminus of its different binding proteins. GIP is a small protein and primarily composed of a single PDZ domain. GIP is also involved in the cancerous pathways by interacting with its binding proteins. Thus GIP can be used as a potential target for structure based anti-cancer drug design. Indeed, GIP regulates many of the signaling processes through its PDZ domain. It has been found that dynamics of GIP do not show any local change, but rather it shows that overall the protein becomes more flexible upon C-terminus of glutaminase L binding. To get an insight into cellular signaling, it is important that we understand protein-protein interactions, and how these interactions regulate the well-being of a cell.

2.6 References

1. Olalla, L., Aledo, J.C., Bannenberg, G., and Marquez, J. (2001) The C-terminus of human glutaminase L mediates association with PDZ domain-containing proteins. *FEBS Lett. 488*, 116-122.
2. Ruiz-Bellido, M.A., Aledo, J.C., Marquez, J., and Nunez De Castro, I. (2000) Inhibition of glutaminase expression by antisense mRNA decreases growth and tumourigenicity of tumour cells. *Biochem. J. 348*, 257-261.
3. Medina, M.A. (2001) Glutamine and cancer. *J. Nutr. 131*, 2539S-2542S.
4. Maity, P., Chakraborty, S., and Bhattacharya, P. (2000) Neovascularisation offers a new perspective to glutamine related therapy. *Indian J. Exp. Biol. 38*, 88-90.
5. Lora, J., Alonso, F. J., Segura, J. A., Lobo, C., Márquez, J., and Matés, J. M. (2004) Antisense glutaminase inhibition decreases glutathione antioxidant capacity and increases apoptosis in Ehrlich ascitic tumour cells. *Eur. J. Biochem. 271*, 4298-4306.
6. Rousset, R., Fabre, S., Desbois, C., Bantignies, F., and Jalinot, P. (1998) The C-terminus of the HTLV-1 Tax oncoprotein mediates interaction with the PDZ domain of cellular proteins. *Oncogene 16*, 643-654.
7. Olalla, L., Gutierrez, A., Jiménez, A. J., López-Téllez, J. F., Khan, Z. U., Pérez, J., Alonso, F. J., de la Rosa V, Campos-Sandoval, J. A., Segura, J. A., Aledo, J. C., and Márquez, J. (2008) Expression of the scaffolding PDZ protein glutaminase-interacting protein in mammalian brain. *J. Neurosci Res. 86*, 281-92.

8. Márquez, J., López, de la Oliva, A. R., Matés, J. M., Segura, J. A., and Alonso, F. J (2006) Glutaminase: A multifaceted protein not only involved in generating glutamate. *Neurochem Int. 48*, 465-471.

9. Hampson, L., Li, C., Oliver, A.W., Kitchener, H.C., and Hampson, I.N. (2004) The PDZ protein Tip-1 is a gain of function target of the HPV16 E6 oncoprotein. *Int. J. Oncol. 25*, 1249-1256.

10. Reynaud, C., Fabre, S., and Jalinot, P. (2000) The PDZ Protein TIP-1 Interacts with the Rho Effector Rhotekin and Is Involved in Rho Signaling to the Serum Response Element. *J. Biol. Chem. 275*, 33962-33968.

11. Alewine, C., Olsen, O., Wade, J. B., and Welling, P. A. (2006) TIP-1 Has PDZ Scaffold Antagonist Activity. *Mol. Biol. Cell. 17*, 4200- 4211.

12. Kanamori, M., Sandy, P., Marzinotto, S., Benetti, R., Kai, C., Hayashizaki, Y., Schneider, C., and Suzuki, H. (2003) The PDZ Protein Tax-interacting Protein-1 Inhibits β-Catenin Transcriptional Activity and Growth of Colorectal Cancer Cells. *J. Biol. Chem. 278*, 38758-3876.

13. Huber, A. H., Nelson, W. J., and Weis, W. I. (1997) Three-Dimensional Structure of the Armadillo Repeat Region of β-Catenin. *Cell 90*, 871-882.

14. Shibata, T., Chuma, M., Kokubu, A., Sakamoto, M., and Hirohashi, S. (2003) EBP50, a β-Catenin-Associating Protein, Enhances Wnt Signaling and Is Over-expressed in Hepatocellular Carcinoma. *Hepatology 38*, 178-186.

15. Itoh, N., and Nagata, S. (1993) A Novel Protein Domain Required for Apoptosis. *J. Biol. Chem. 268*, 10932-10937.

16. Saras, J., Engstrom, U., Gonez, L. J., and Heldin, C- H. (1997) Characterization of the Interactions between PDZ Domains of the Protein-tyrosine Phosphatase PTPL1 and the Carboxyl-terminal Tail of Fas. *J. Biol. Chem. 272*, 20979–20981.

17. Nagata, S. (1997) Apoptosis by Death Factor. *Cell 88*, 355-365.

18. Zhang, Q., Fan, J-S., and Zhang, M. (2001) Interdomain Chaperoning between PSD-95, Dlg, and Zo-1 (PDZ) Domains of Glutamate Receptor-interacting Proteins. *J. Biol. Chem. 276*, 43216-43220.

19. Sheng, M., and Sala, C. (2001) PDZ domains and the organization of supramolecular complexes. *Annu. Re*v. *Neurosci. 24*, 1–29.

20. Aledo, J. C., Rosado, A., Olalla, L., Campos, J. A., and Marquez, J. (2001) Overexpression, Purification, and Characterization of Glutaminase-Interacting Protein, a PDZ-Domain Protein from Human Brain. *Protein. Expr. Purif. 23*, 411–418.

21. Banerjee, M., Huang, C., Marquez, J., and Mohanty, S. (2008) Probing the structure and function of human glutaminase-interacting protein: a possible target for drug design, *Biochemistry 47*, 9208-9219.

22. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multifunctional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.

23. Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR 4*, 603–614.

24. Wang, Y., and Jardetzky, O. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci. 11*, 852-861.

25. Hung, L. H., and Samudrala, R. (2003) PROTINFO: secondary and tertiary protein structure prediction. *Nucleic Acids Res. 31*, 3296-3299.

26. Hung, L. H., and Samudrala, R. (2003) Accurate and automated classification of protein secondary structure with PsiCSI. *Protein Sci. 12*, 288-295.

27. Coudevylle, N., Antoine, M., Bouguet-Bonnet, S., Mutzenhardt, P., Boschi-Muller, S., Branlant, G., and Cung, M-T. (2007) Solution Structure and Backbone Dynamics of the Reduced Form and an Oxidized Form of E. coli Methionine Sulfoxide Reductase A (MsrA): Structural Insight of the MsrA Catalytic Cycle. *J. Mol. Biol. 366*, 193-206.

28. Williamson, R. A., Carr, M. D., Frenkiel, T. A., Feeney, J., and freedman, R. B. (1997) Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinase-2 by NMR chemical shift perturbation. *Biochemistry 36*, 13882-13889.

29. Johnson, P.E., Tomme, P., Joshi, M.D., and McIntosh, L.P. (1996) Interaction of soluble cellooligosaccharides with the N-Terminal cellulose-Binding Domain of *Cellulomonas fimi* CenC. 2. NMR and Ultraviolet Absorption Spectroscopy. *Biochemistry 35*, 13895-13906.

30. Palmer, 3rd, A. G., Rance, M. and Wright, P. E. (1991) Intramolecular motions of a zinc finger DNA-binding domain from xfin characterized by proton-detected natural abundance 13C heteronuclear NMR spectroscopy. *J. Am. Chem. Soc. 113*, 4371–4380.

31. Mandel, A. M., Akke, M. and Palmer, 3rd, A. G. (1996) Dynamics of ribonuclease H: Temperature dependence of motions on multiple time scales *Biochemistry 35*, 16009–16023.

32. R.W. Woody, Theory of circular dichroism of proteins. In: G.D. Fasman, Editor, *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York (1996), pp. 25–67.

33. D.J. Rodi, R.W. Janes, H.J. Sangasnee, A. Soares, R.A. Holton, B.A. Wallace and L. Makowski , Screening of a library of phage-displayed peptides identifies human Bcl-2 as a Taxol-binding protein. *J Mol Biol* 285 (1999), pp. 197–203.

34. Sreerama, N., and Woody, R. W. (2000) Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochemistry 39*, 252–260.

35. Chipman, D. M., Grisaro, V., and Sharon, N. (1967) The binding of oligosaccharides containing *N*-acetylglucosamine and *N*-acetylmuramic acid to lysozyme. *J. Biol. Chem. 242*, 4388-4394.

36. Mao, H., Hajduk, P. J., Craig, R., Bell, R., Borre, T., and Fesik, S. W. (2001) Rational Design of Diflunisal Analogues with Reduced Affinity for Human Serum Albumin. *J. Am. Chem. Soc. 123*, 10429–10435.

37. Cavanagh, J. F., Wayne, J., Palmer, A. G., III. and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and practice*, Academic Press, Inc., San Diego, CA.

38. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res. 31*, 3381-3385.

39. Guex, N., and Peitsch, M. C. (1997) SWISS_MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis 18*, 2714-2723.

40. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph. 14*, 51–55.

41. Gunther, J., Bergner, A., Hendlich, M., and Klebe, G. (2003) Utilising Structural Knowledge in Drug Design Strategies: Applications using Relibase. *J. Mol. Biol. 326*, 621-636.

42. Alston, R. W., Urbanikova, L., Sevcik, J., Lasagna, M., Reinhart, G. D., Scholtz, J. M., and Pace, C. N. (2004) Contribution of Single Tryptophan Residues to the Fluorescence and Stability of Ribonuclease Sa. *Biophys. J. 87*, 4036-4047.

43. Ladokhin, A. S. (2000) *in Encyclopedia of Analytical Chemistry: Fluorescence Spectroscopy in Peptide and Protein Analysis* (Meyers, R.A., Ed), pp. 5762–5779, John Wiley & Sons, Chichester.

44. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckerman, R. N., and Lim, W. A. (1998) Exploring the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science 282*, 2088-2092.

45. Kozlov, G., and Gehring K. (2000) Solution Structure of the PDZ2 Domain from Human Phosphatase hPTP1E and Its Interaction with C-Terminal Peptides from the Fas Receptor. *Biochemistry 39*, 2572-2580.

46. Tochio, H., Zhang, Q., Mandal, P., Li, M., and Zhang, M. (1999) Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. *Nat. Struct. Biol. 6*, 417-421.

47. Jemth, P. & Gianni, S. (2007) PDZ Domains: Folding and Binding. *Biochemistry 46*, 8701-8707.

48. Gianni, S., Walma, T., Arcovito, A., Calosci, N., Bellelli, A., Engstrom, A.,Travaglini-Allocatelli, C., Brunori, M., Jemth, P. & Vuister, G. W. (2006) Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure 14*, 1801-9.

CHAPTER THREE

Structural and Functional Characterization of GIP Interaction with Novel Internal Motif Peptides: Potential Candidates for Development of Anti-Cancer Therapeutics

3.1 Introduction

3.1.1 GIP and anti-cancer drug design

GIP is involved in various cancerous pathways. Thus, it can be used as a target for anti-cancer therapeutics. GIP is a small protein and primarily composed of a single PDZ domain. GIP has a well-defined binding site. Thus, it is an excellent candidate for anti-cancer drug design. Development of drugs targeting a specific protein is achievable when its network is fully characterized to minimize side-effects. To extend the GIP interaction network, we used an f8-type phage displayed peptide library to screen for new GIP-binding peptides that may lead to identification of new partner proteins. Such peptides may serve as leads for development of novel anti-cancer therapeutics that specifically target GIP.

3.1.2 PDZ domain and phage display technology

Phage display was used for identification of C-terminal interacting motifs for several PDZ domains (*1*,*2*, *3*). For example, a phage display library with peptides fused to the carboxyl terminus of the M13 major coat protein was used for two different PDZ domains from a membrane-associated guanylate kinase to find consensus sequences (*1*).

Bacteriophage T7 was used to identify binding sequences for the 10th PDZ domain (PDZ10) of the multi-PDZ domain protein 1 (MUPP1) and the third PDZ domain (PDZ3) of postsynaptic density-95 (PSD-95) protein (*4*).

The use of phage display libraries for finding internal motifs for a PDZ domain is uncommon. To date there are a few crystal structures available that show internal motif recognition by PDZ domains (*5*, *6*, *7*). We show that GIP selectively binds internal motif consensus sequences from a highly diverse peptide library fused to the α-helical position of P8 (*8*). We describe the analysis of PDZ domain binding specificities that utilizes this display system. Together with structural analysis, we establish the internal motif phage display as a powerful new method for mapping PDZ domain ligand specificity.

3.1.3 Methods for characterization of PDZ-containing proteins at cellular level: immunocytochemistry, intracellular colocalization, and MTT assay

Immunocytochemistry is a commonly used technique for detection of proteins in cells by the use of protein-specific antibody. The presence of a particular protein as well as the expression level of the protein can be demonstrated by this method.

Colocalization is a common term in cell biology. It is defined as the overlap between different fluorescent labels with different wavelengths to see if the different targets are located in the same region of the cell. Using fluorescence microscopy, different targets containing fluorescent tags can be located within cells, and their colocalization can be established. This technique was used for different PDZ domains for localization of the protein inside various cells. It has been reported that in some cases PDZ domain-containing proteins form complexes to perform their function. Once they

form the complex, signal transduction pathways can be regulated. Mutation in the PDZ domain may cause no formation of the complex, and consequently, its function will be also disrupted. There are reports indicating that many other proteins help to form a complex and mutations in them that cause disruptions in complex formation (*9, 10*).

Tetrazolium salts are commonly used in cell biology to assay cell metabolism (*11*). MTT is a colorimetric assay for measuring the activity of mitochondrial enzymes present in metabolically active cells. The method is based on the metabolic activity of viable cells and measures cellular oxidative metabolism. The dye is cleaved to a colored product by the activity of NAD(P)H-dependent dehydrogenase, and this indicates the level of energy metabolism in cells. The color development (yellow to blue) is proportional to the number of metabolically active cells. This method is used to determine cellular toxicity of a tested compound. Cytotoxicity, cell viability, and proliferation of living cells have been studied using the MTT method (*12, 13, 14*). The use of this method helps design drugs and will be useful for disruption of signaling pathways where GIP is involved.

3.2 Materials and Methods

3.2.1 Protein expression and purification

The procedure for expression and purification of GIP was carried out as described previously (*15*). Briefly, the recombinant pET-3c/GIP plasmid was transformed into *Escherichia coli* (*E. coli*) BL21 DE3pLys cells, and expression was performed in LB media for unlabeled and M9 minimal media containing ^{13}C-labeled glucose and/or^{15}N-labeled ammonium chloride for structural studies. Overnight bacterial culture was

diluted 1:25 with fresh minimal media and was grown at 37 °C until OD_{600} = 0.4-0.5. Expression was induced with 1 mM IPTG at 30° C. After 15 h of continuous incubation, the bacterial cells were harvested by centrifugation. The harvested cells were lysed by sonication in lysis buffer (50 mM sodium phosphate, 200 mM NaCl, 4 mM EDTA, 4% glycerol, 1 mM PMSF, pH 8.0). After centrifugation, the supernatant containing GIP was collected and retained for further purification. GIP was purified by a single-step size exclusion chromatography on a Sephacryl S-100 column (GE Healthcare). Pooled fractions of the pure protein were concentrated using millipore concentrator (with a 3000 MWCO) and the buffer was exchanged using the same concentrator with a buffer containing 50 mM sodium phosphate at pH 6.5, 1 mM EDTA and 0.01% (w/v) NaN_3.

3.2.2 Screening of the phage displayed peptide library

For identification of GIP-binding peptides, a pVIII 9-mer phage display library was used (*16*). The library contains $2x10^9$ different phage clones with multivalently displayed foreign peptides, providing incredible diversity for finding target proteins in non-stringent conditions (*17*). Prior to the library selection, GIP was purified as described above and dialyzed against 0.1M phosphate buffer at pH 8.0. Two wells of a Medisorp (Nunc, thermo scientific, catalog # 467320) 96-well plate were coated with the purified protein at a 100 µg/ml concentration overnight at 4°C. The protein-coated well was blocked with 1% Bovine Serum Albumin (BSA) in Tris buffered saline (TBS) for 1 h at room temperature. To select for the target-binding phage, an aliquot of 10^9 colony forming units (cfu) of the library was added to the well for additional 1h incubation at room temperature. After incubation, unbound phages were discarded and the wells were

washed 10 times with TBS containing 0.1% Tween 20 (TBST). The bound phages were eluted with 0.2 M glycine, pH 2.2 for 10 minutes. The eluate was immediately neutralized using 1M Tris-HCl, pH 9.1. The eluted phages were amplified in *E. coli* K91BluKan bacteria, purified, and titered for the next round of selection. In rounds two and three, 10^{10} cfu aliquots were used in the selection procedures. After the third round, phage DNA in the area of the *gpVIII* gene was PCR amplified from 33 random phage-infected bacterial colonies, purified and sequenced. Sequences of GIP-binding peptides were deduced from phage DNA sequences using Chromas software.

3.2.3 Phage binding assay

Medisorp 96-well plates were coated with GIP at 70 µg/ml, at 4^0C overnight. The wells were blocked with 1% BSA in TBS for 1h at room temperature. An additional set of uncoated wells was also blocked for the negative control. The wells were washed with washing buffer (0.1% Tween 20 in TBS buffer, pH 7.0) two times. Each selected phage clone was amplified individually and added at $5x10^6$ cfu/well to the GIP-coated wells for 1 hr incubation at room temperature. After incubation, the wells were washed 10 times with TBST washing buffer. Bound phage were recovered by adding 25 µl of lysis buffer (2.5% CHAPS, 0.1% BSA in TBS buffer, pH 7.0) to the wells for 10 minutes at room temperature. After that, freshly prepared *E. coli* starved cells (125 µl/well) were added to the wells for 15 minutes to allow phage infection. Next, 180 µl of NZY broth (pH 7.5) containing 0.4 µg/ml tetracycline was added to each well and the plates were placed in a 37^o C incubator for 45 minutes. Finally, the content of each well was plated on NZY

plates containing 20 µg/ml of tetracycline for overnight incubation at 37^{o} C. To quantify the phage, bacterial colonies were counted by a colony counter next morning.

3.2.4 GIP-peptide titration by NMR

Interaction studies were carried out by titration of 100 µM GIP with peptides containing several different internal sequences: ESSVDLLDG, ASSSVDDMA, GTNLDGLDG, GSSLDVTDN, GSGTDLDAS, and GSSAAVTDN. The target peptides were obtained with > 95 % purity from Chi Scientific (MA). The 10 mM stock solutions of the above peptides were prepared in 10 mM phosphate buffer at pH 6.5. The amide chemical shift perturbations (Δδ) were calculated as $\Delta\delta = |\Delta\delta^{15}N|/f + |\Delta\delta^{1}H|$ (*15*). The introduction of the f factor and its value were justified by the difference in the spectral widths of the backbone ^{15}N resonances and the ^{1}H signals (^{15}N range, 131.5-100.8 ppm = 30.7 ppm; ^{1}H range, 10.1-6.6 ppm = 3.5; correction factor f=30.7/3.5 = 8.7). Thus, the correction factor f = 8.7 was used in order to give roughly equal weighting for each of the ^{1}H and ^{15}N chemical shift changes. For each titration, peptide from stock solution and after a ratio of 10:1, solid peptide was added in increasing amounts to an excess that approached saturation with peptide: protein ratios between 40:1 and 140:1 for the above peptides.

3.2.5 GIP-peptide models by solution NMR

To get the model structure of GIP in complex with ESSVDLLDG and GSGTDLDAS peptides, I performed the following experiments: 2D TOCSY (*18*) and 2D ROESY (*19*) on each peptide, 2D ^{15}N/^{13}C F1, selectively filtered NOES (*20*), 3D ^{13}C-

edited/filtered HSQC-NOESY and 3D ^{15}N-edited/filtered HSQC-NOESY (*21*) on each peptide/protein complex. The sample contained ~400 μM uniformly ^{15}N/^{13}C labeled GIP, 8 mM ESSVDLLDG or 16 mM GSGTDLDAS, 50 mM phosphate buffer containing 5 % D_2O (pH 6.5), 1 mM EDTA and 0.01% (w/v) NaN_3. Dr. David Zoetewey analyzed the NOESY data and carried out structure calculations including ARIA refinement. He found peptide-peptide and peptide-protein NOEs that were added to the set of previously determined protein NOEs from free GIP for structure calculation using ARIA (*22*). On an iterative basis, the structures were evaluated and refinements made to the ARIA inputs using VMD (*23*) to visualize the structures. For the final ensemble of structures, out of total of 200 starting structures, the 25 with lowest energy were chosen for water refinement. Of those, the 20 with the lowest energy were selected for analysis (*24*).

3.2.6 Immunocytochemical localization of GIP in cancer cells

Human glioma cells D54 MG, U87 MG, and GBM8 were plated onto 4-chamber slides (Nunc, Naperville, IL) at the density of $3x10^4$ cells/chamber in Dulbecco's modified Eagle's medium (DMEM F-12) supplemented with 10% fetal bovine serum (FBS) and grown under 5% CO_2 at 37^{o}C for 24 hrs. For immunocytochemical localization of GIP, the cells were fixed by 1% paraformaldehyde in PBS for 30 min, and then permeabilized with 0.5% Triton X-100 in PBS for 25 min at room temperature. The cover slips were blocked with MACS buffer (0.5% BSA, 2mM EDTA in PBS, pH 7.2) for 1h. The cells were incubated with primary anti-GIP mouse monoclonal IgG (Novus Biologicals, Littleton, CO) diluted 1:15 in PBS with 1% BSA, overnight at 4^{o}C. The primary antibody was removed and the slides were rinsed with PBS. Secondary goat anti-

mouse Alexa 488-conjugated antibody (Novus Biologicals, Littleton, CO) diluted 1:40 in PBS/BSA was added, and incubation was carried out for 1 h at room temperature. Unbound secondary antibody was removed by washing with PBS. Slides were mounted with cover slips using Vectashield DAPI mounting medium (Vector Laboratories, Inc., Burlingame, CA). Fluorescence images were acquired with an Olympus BH-2 fluorescence microscope equipped with Nikon Digital Sight DS-L1 camera.

3.2.7 Peptide internalization and co-localization with GIP in D54 MG cancer cells

To demonstrate the ability of the peptide to be internalized by human glioma D54 MG cells, human glioma cells D54 MG cells were plated onto 4-chamber slides (Nunc, Naperville, IL) at the density of $3x10^4$ cells/chamber in Dulbecco's modified Eagle's medium (DMEM F-12) supplemented with 10% fetal bovine serum (FBS) and grown under 5% CO_2 at 37^oC for 24 hrs. The cells were treated with TAMRA-labeled ESSVDLLDGGG$(R)_7$ peptide at 1 µM for 25 min. After incubation, the cells were washed three times with PBS and fixed with 1% paraformaldehyde for 15 minutes. Fixed cells were mounted with cover slips using Vectashield DAPI mounting medium. The slides were evaluated by fluorescence microscopy.

For GIP-peptide co-localization studies, cells plated on chamber slides as above were treated with TAMRA-labeled ESSVDLLDGGG$(R)_7$ peptide at 1 µM for 25 min and fixed with 1% paraformaldehyde for 15 min, followed by three PBS rinses. Fixed cells were permeabilized with 0.5% Triton X-100 in PBS for 25 min, rinsed three times with PBS, and blocked with MACS buffer for 1h at room temperature. The cells were then incubated with primary anti-GIP mouse monoclonal IgG antibody overnight at 4 °C

and washed three times with PBS as in the previous section. Fluorescein Alexa 488 anti-mouse secondary antibody was then added and incubated for 1h at room temperature. After washing the cells three times with PBS, the cells were mounted and evaluated by fluorescence microscopy.

3.2.8 MTT assay

The effect of a synthetic peptide (ESSVDLLDG conjugated to R_7 (RRRRRRR) to penetrate the cell membrane) on D54 MG cells were examined by an assay that utilizes MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (Sigma-Aldrich, St. Louis, MO) salt. This assay measures cellular oxidative metabolism. The dye is cleaved to a colored product by the activity of NAD(P)H-dependent dehydrogenase, and this indicates the level of energy metabolism in cells. The color development (yellow to blue) is proportional to the number of metabolically active cells. For these experiments, D54 MG cells were plated on 96-well culture plates at a density of 3 x 10^3cells/well and cultured overnight in DMEM F-12 medium (Mediatech Inc, Manassas, VA) containing 10% fetal bovine serum at 37 °C. Next day, the peptide was added to the cells at 10, 20, 40, 50, 75, 100, 200 µM concentrations. The cells were incubated at 37 °C until the total treatment time reached 16 h. After that, 10% volume of MTT stock solution (5 mg/ml) was added to the cell cultures for four hours for color development. The converted dye was then solubilized, and the absorbance was measured at 550 nm. Each data point was normalized against the control cells.

3.3 Results

3.3.1 Identification of GIP-binding peptides by phage display

GIP-binding peptides were selected from a f8-type 9-mer phage displayed peptide library (*25*) that displays 4000 copies of the foreign nonamers in the N-terminal part of the major coat protein pVIII of fd phage. The library was constructed in the laboratory of Dr. Petrenko (CVM, AU) by replacement of amino acids 2–5 of pVIII with random nonamers. The library allows selection of highly homologous families of peptides in non-stringent conditions due to its multivalency and avidity effect (*17*) with easily recognizable binding motifs (*26*). The selection process was consisting of several steps. First, the phages are allowed to bind to GIP. Then, the unbound phages are washed away and bound ones were eluted and bacterial amplified. The selection process was repeated three times. To reveal GIP-binding motifs, the gene *gpVIII* DNAs were amplified by PCR (using the following primers: f8s-20: 5´-CAAAGCCTCCGTAGCCGTTG-3´, f8as-20: 5´-CATTCCACAGACAGCCCTCA-3´) from 33 phage clones, sequenced, and translated into 18 unique peptide sequences. Based on sequence alignment, they were placed into two groups (**Table 3.1**). Group 1 contained peptides with S/T-S-V/L-D as a common motif. Interestingly, this motif was identified in different positions within the nine amino acid peptide sequences, including 2-5, 3-6, and 4-7. Group 2 contained a three residue N-L-D motif, which occupied positions from 2-4 and 3-5 within the peptides. An additional sequence, GSGTDLDAS, was also identified. Comparative analysis of all sequences revealed S/T-X-V/L-D to be the consensus motif.

Group 1	Group 2	Other
GGSSVDSEG	DSNLDVSVE	GSGTDLDAS
ESSVDLLDG	VSNLDTTND	
GSSSVDVDG	GTNLDGNGD	
AISSVDSMG	GSMNLDVQS	
ESSVDMIGD	GGNLDVNVG	
GSSVDLVGD	DGNLDSYDN	
AYESSVDDN		
ASSSVDDMA		
GSSLDVTSE		
GSSLDVTDN		
GYETSLDSN		

Table 3.1 Peptide sequences identified for GIP binding from a phage display library placed into two main groups. Group 1 contained peptides with S/T-S-V/L-D as a common motif which occupied position from 2-5, 3-6, and 4-7. Group 2 contained a three residue N-L-D motif, which occupied positions from 2-4 and 3-5 within the peptides. An additional sequence, GSGTDLDAS, was also identified.

The specificity of the selected phage clones to GIP was confirmed through a phage-binding assay by comparing relative binding of individual phage clones to the target protein in comparison with the controls, BSA or empty wells of the plastic plates used for phage selection. As an additional control for binding specificity, the above assay was repeated with phage f8-5, the vector that does not display any fusion foreign peptides (*16*). Equal numbers of individual phage clones were added to the wells containing either GIP or the above controls followed by incubation and quantification of the bound phage by titering in the host *E.coli* cells K91BK. It was observed that GIP-selected clones bind to GIP at numbers ranging from 5 to 278 (cfu), but do not bind either to BSA or to the plastic. The vector phage alone did not bind to GIP.

3.3.2 GIP binding to internal motif peptides monitored by NMR spectroscopy

Several GIP-specific peptides revealed in the selection experiments were synthesized to assess their interactions with GIP using NMR spectroscopy. Five peptides representing motifs with either S/T or V/L amino acids in positions P_{-2} or P_0 according to standard PDZ nomenclature (*27*), were selected for the NMR studies.

Chemical shift mapping is a powerful method to investigate possible protein ligand interactions by NMR. The 2D ^{1}H-^{15}N HSQC spectrum provides the fingerprint region of a protein. This NMR experiment is a sensitive technique to study protein-ligand interactions in solution (*27, 28, 29*). Any perturbation in the chemical shift resonances from their original positions in this region indicates a change in the local environment of the affected residues within a protein (*15*). To elucidate a molecular mechanism of GIP-ligand binding, we studied the interaction of GIP with selected peptides by 2D HSQC titration experiments. The amide proton and nitrogen resonances in the HSQC spectra were followed for each titration point. Resonances from most of the residues of GIP followed fast exchange kinetics on the NMR time scale as observed by gradual and systematic changes in their chemical shift positions (**Figure 3.1**). A few specific residues such as Leu29 and Gly30 followed intermediate exchange kinetics as seen by the disappearance of these peaks (**Figure 3.1**). The decrease in peak intensity of these residues is due to exchange between amide resonances of free and bound GIP. Residues Leu29 and Gly30 are part of the ILGF binding loop that makes specific hydrogen bonds to the negatively charged terminal carboxylate group of the partner protein with a C-terminal recognition motif during binding (*27*). This causes large chemical shift

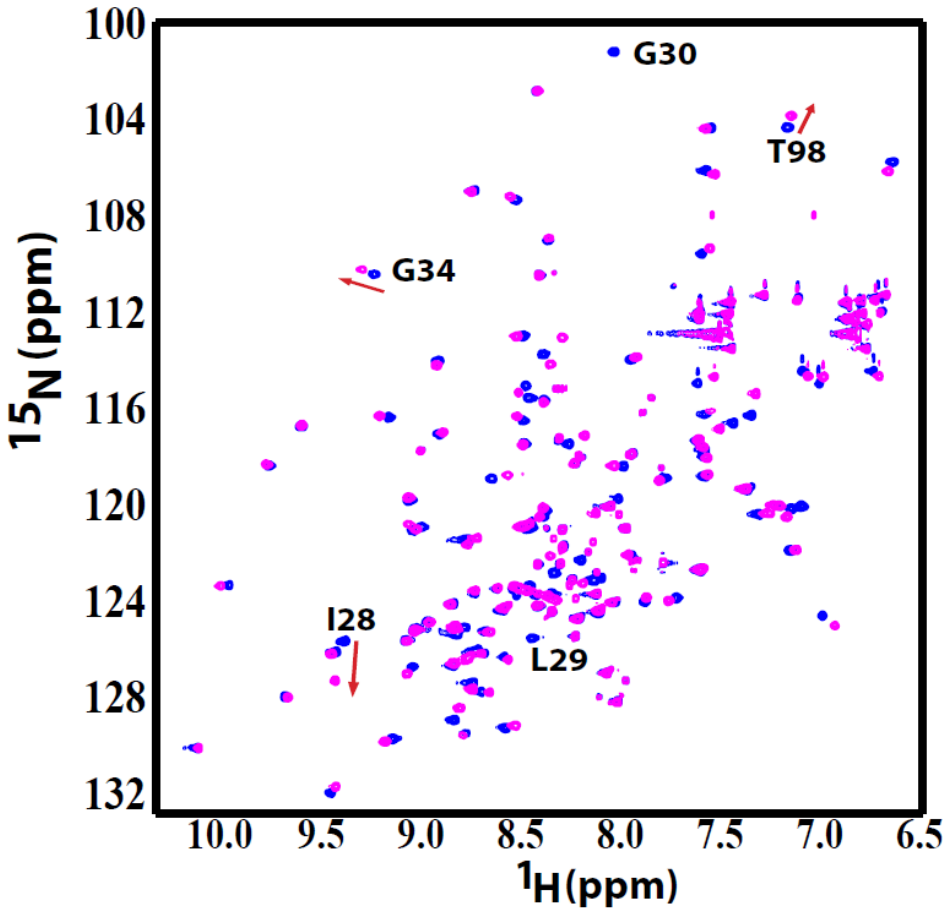

Figure 3.1 GIP PDZ domain shows direct interaction with the GSSLDVTDN internal motif peptide. Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of GSSLDVTDN peptide to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25^{o}C. ^{15}N-HSQC spectra of ^{15}N-labeled GIP PDZ domain alone (blue) and with GSSLDVTDN peptide (magenta).

perturbations in these residues (*30*). For our titration experiments, the magnitude of changes in the chemical shifts of residues in GIP can be correlated to the relative proximity to the peptide in the complex.

3.3.3 Chemical shift perturbation of GIP upon binding to internal motif peptide ligands

The chemical shift perturbation for each residue was calculated from the chemical shift changes of both ^{1}H and ^{15}N nuclei. When internal motif peptides were added, systematic changes of the amide resonances occurred in the titration spectra (**Figure 3.2**) and the significant chemical shift perturbations were grouped into three categories: medium shifted (>0.1 ppm), large shifted (>0.2 ppm), and intermediate exchange (**Table 3.2**).

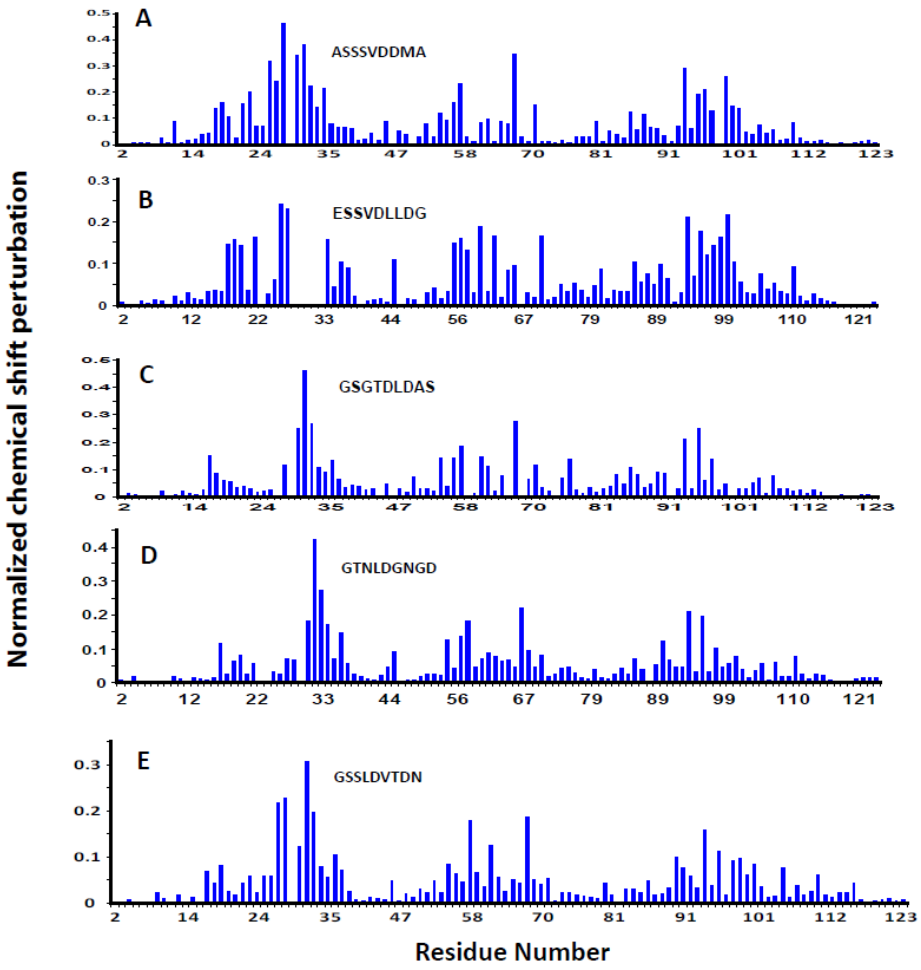

Figure 3.2 Chemical shift perturbation profiles of GIP upon binding to various internal motif peptides. Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding each of the following internal motif peptides: (A) ASSSVDDMA, (B) ESSVDLLDG, (C) GSGTDLDAS, (D) GTNLDGNGD, (E) GSSLDVTDN. Δδ is calculated as described in the **Methods section**.

Peptides	Medium shifted (>0.1 ppm) residues/regions	Large shifted (>0.2 ppm) residues/regions	Residues in Intermediate exchange
ASSSVDDMA	Ile18, His19, Lys20, Arg22, Gln23, Gly34, Ile55, Val57, Leu71, Thr86, Val88, Arg96, Thr98, Ser101, Glu102	Asn26, Leu27, Ile28, Phe31, Ser32, Ile33, Gly35, Thr58, Glu67, Arg94, Leu97, Arg100	Leu29, Gly30
ESSVDLLDG	Ile18, His19, Lys20, Arg22, Gly34, Gly36, Phe46, Tyr56, Val57, Thr58, Val60, Glu62, Glu67, Leu71, Thr86, Arg96, Leu97, Thr98, Lys99, Ser101	Asn26, Leu27, Arg94, Arg100	Ile28, Leu29, Gly30, Phe31, Ser32, Ile33
GSGTDLDAS	Glu17, Ile28, Gly34, Gly36, Ile55, Val57, Thr58, Ser61, Glu62, Leu71, Ile77, Thr86, Thr98	Phe31, Ser32, Ile33, Glu67, Arg94, Arg96	Leu29, Gly30
GTNLDGNGD	Glu17, Phe31, Gly34, Gly36, Ile55, Val57, Thr58, His90, Arg96, Thr98	Ser32, Ile33, Glu67, Arg94	Leu29, Gly30
GSSLDVTDN	Phe31, Ile33, Gly36, Thr58, Ser61, Glu67, His90, Arg94, Arg96, Thr98, Lys99	Leu27, Ile28, Ser32	Leu29, Gly30

Table 3.2 Chemical shift perturbations of GIP residues upon binding to internal motif peptides.

The magnitudes of the amide chemical shift changes upon binding to different internal motif peptides are mapped on to the ribbon diagram of GIP as indicated by different colors (**Figure 3.3**). The three dimensional structure of GIP was developed in our lab previously (*27*).

Chemical shift perturbation analysis shows that the ILGF loop, β2 strand, and α2 helix are the regions of GIP that are most affected. It also shows that residues in the region Ile18-Gln23, Ile55-Glu62, and Glu67 which belong to the β1, βa, and β3 strands as well as the α1 helix are also affected, but to a lesser extent. This observation suggests that the peptides with internal binding motifs bind to the same binding site nestled between the β2 strand and α2 helix of GIP as the canonical C-terminal motif. This binding is allostrically driven, reminiscent of the way GIP binds to C-terminal motifs (*27*).

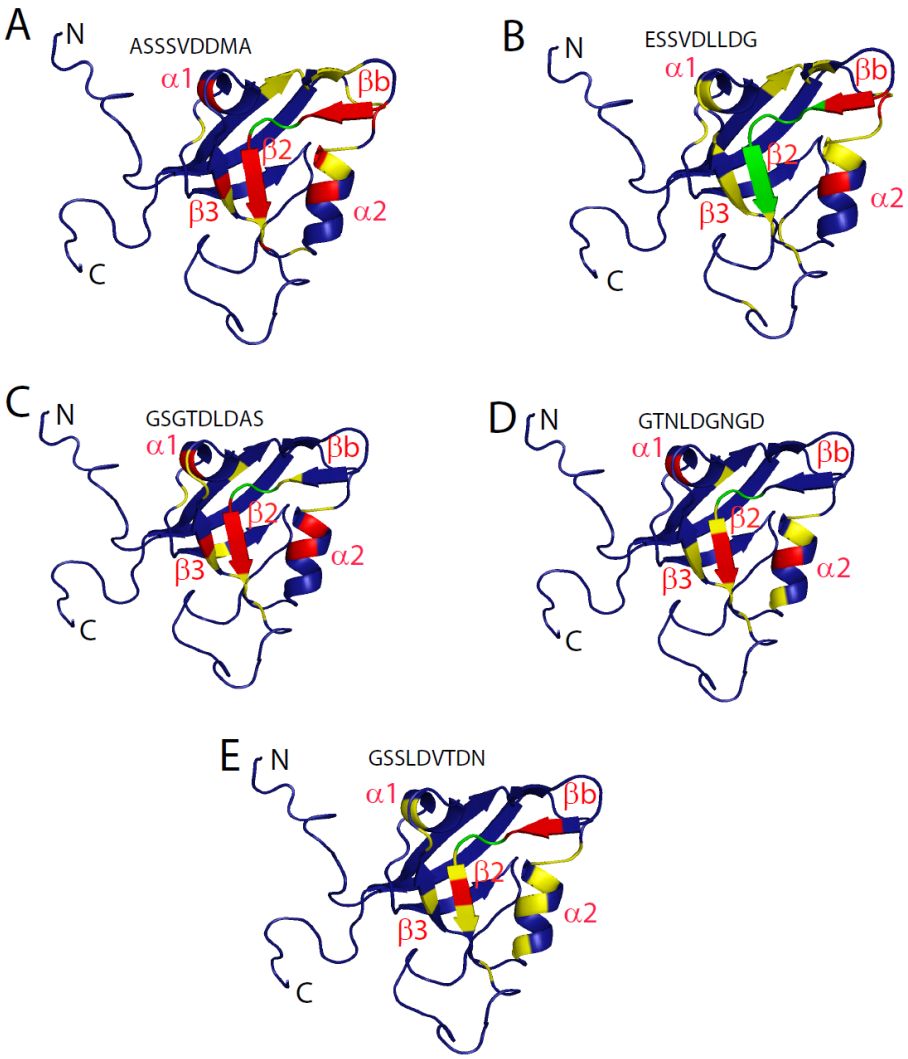

Figure 3.3 The magnitude of the amide chemical shift changes is represented in different colors on a ribbon diagram of GIP bound to the various internal motif peptides, (**A**) ASSSVDDMA, (**B**) ESSVDLLDG, (**C**) GSGTDLDAS, (**D**) GTNLDGNGD, (**E**) GSSLDVTDN. Red indicates $\Delta\delta > 0.2$ ppm, yellow indicates $0.2\text{ppm} > \Delta\delta > 0.1$ ppm, blue indicates $0.1\text{ ppm} > \Delta\delta$, and green indicates intermediate exchange. Selected secondary structural elements are labeled in red. The free GIP structure was developed in our lab previously. Free GIP structure was adapted from reference 27.

3.3.4 Role of the residues at P_0 and P_{+1} of the peptide in GIP-peptide binding

To analyze the role of specific residues in the internal motif recognition by GIP, we created a double alanine substitution for LD in the GSS**LD**VTDN peptide. NMR titrations were performed to determine how these substitutions affect the binding to GIP. GIP was titrated with various concentrations of the alanine substituted peptide GSS**AA**VTDN. Compared with the wild-type GSSLDVTDN peptide (**Figure 3.1**), the chemical shift perturbation is negligible for the AA substitution (**Figure 3.4**). This indicates that any interaction between the peptide and GIP was completely eliminated. Interestingly, the observation that GSS**AA**VTDN peptide has virtually no binding to GIP suggests that both L and D are important for optimal interactions. Titrations with each of the identified peptides show that Leu29 and Gly30 are always in intermediate exchange for both residues (**Table 3.2**). Since LD or VD is present in each peptide and Leu29 and Gly30 are in intermediate exchange for the titration of each peptide, this supports our hypothesis that LD interacts directly with Leu29 and Gly30 of the ILGF carboxylate binding loop as a mimic of a hydrophobic C-terminal residue from a canonical C-terminal motif.

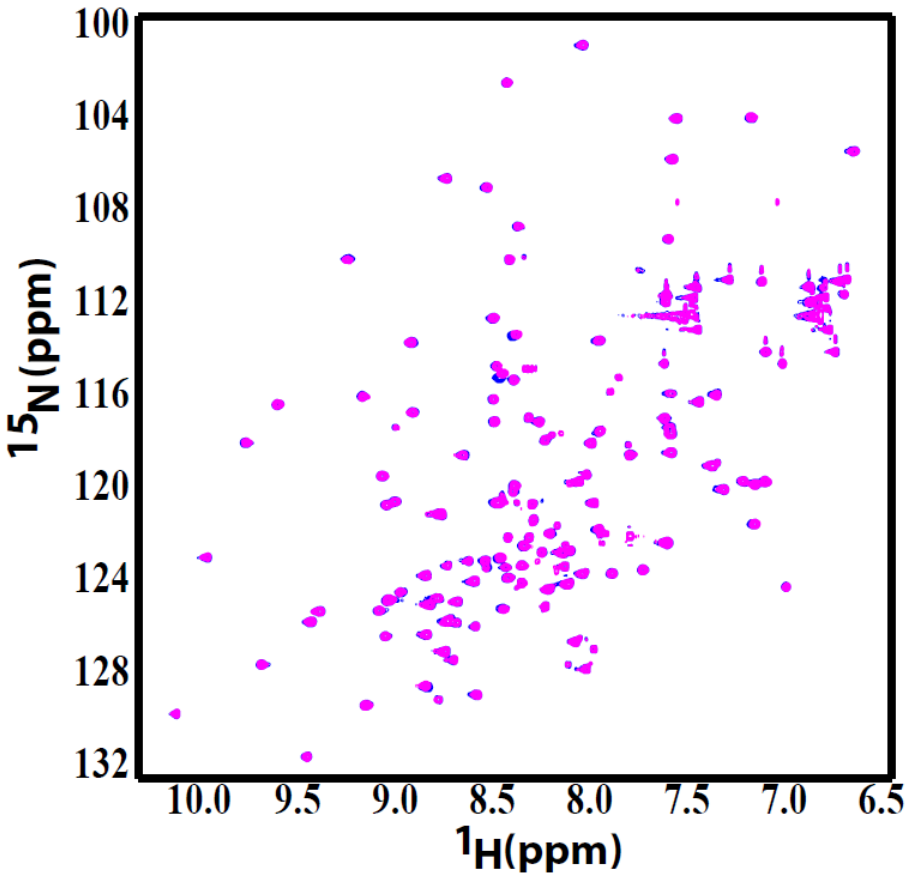

Figure 3.4 GIP PDZ domain does not bind with the double-substituted GSSAAVTDN peptide. Changes of 2D {^{1}H, ^{15}N}-HSQC spectrum upon addition of GSSAAVTDN peptide to 100 µM GIP in 50 mM phosphate buffer (pH 6.5) at 25^{o}C . ^{15}N-HSQC spectra of ^{15}N-labeled GIP PDZ domain alone (blue) and with GSSAAVTDN peptide (magenta).

3.3.5 Mechanism of internal motif recognition by GIP from model structure

The structure of GIP bound to internal motif peptides were determined by David Zoetewey in our lab. Internal motifs bound by β-strand addition to form an antiparallel β-sheet with the β2 strand from GIP like C-terminus peptide binding. Internal peptides bind to GIP as an anti-parallel β-stand through this process, and the hydrophobic residue at P_0 buries into the hydrophobic pocket created by Leu29, Phe31, Leu97 and Ile33 residues from GIP. For ESSVDLLDG and GSGTDLDAS the P_0 is taken by V4 and L6 respectively, which are hydrophobic residues. Both side chains fit into the hydrophobic pocket created by Leu29, Phe31, Leu97 and Ile33 residues from GIP. It was found that S or T at the P_{-2} position forms a hydrogen bond with His90 at the α2:1 position in GIP. S2 from ESSVDLLDG (**Figure 3.5A,B**) and T4 from GSGTDLDAS (**Figure 3.5C,D**) form hydrogen bonds with His90 in GIP. S2 and T4 which are at the P_{-2} position from V4 and L6 at the P_0 position in the ESSVDLLDG and GSGTDLDAS peptides respectively. Interestingly, other S or T is not involved in the binding process. Both Leu29 and Gly30 amides within the ILGF loop of GIP make hydrogen bonds to a carboxylate group from the side-chain of D at P_{+1} for an internal motif. For canonical C-terminal peptide binding this takes place with the C-terminal carboxylate of the interacting partner (*27*). For the internal recognition motifs, the side chain of aspartate at the P_{+1} position acts as a substitute for the C-terminus. Sidechain carboxylate group of D5 and D7 in ESSVDLLDG and GSGTDLDAS play this role.

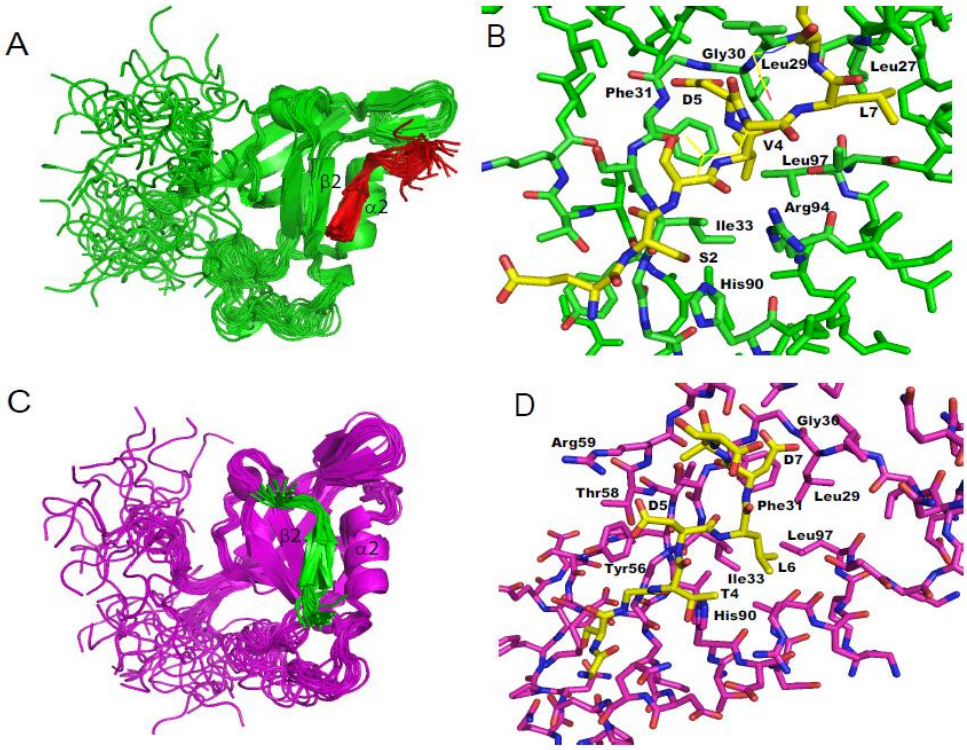

Figure 3.5 Model structures of GIP-ESSVDLLDG and GIP-GSGTDLDAS complexes were determined in collaboration with Dr. David Zoetewey in Mohanty's lab. (**A**) Ensemble of 20 lowest energy structures of the GIP-ESSVDLLDG complex. (**B**) D at P_{+1} forms hydrogen bonds with Leu29 & Gly30 HN and S at P_{+2} with His90. V at P_0 buries itself into a hydrophobic pocket created by Leu29, Phe31, Ile33 and Leu97. (**C**) Ensemble of 20 lowest energy structures of the GIP-GSGTDLDAS complex. (**D**) D at P_{+1} forms hydrogen bonds with Leu29 & Gly30 HN and T at P_{+2} with His90. L at P_0 buries itself into a hydrophobic pocket created by Leu29, Phe31, Ile33 and Leu97. Adapted from reference 24.

3.3.6 GIP- binding internal motif peptides as potential anti-cancer agents

While the exact role of GIP in cancer is unknown, the protein was detected in several types of cancer cells and represents a promising target for anti-cancer therapeutics (*31, 32, 33*). We have screened several human glioma cell lines to express GIP. We found that GIP expresses in D54 MG (**Figure 3.6A**) and U87 MG (**Figure 3.6B**) glioma cells. We did not find expression of GIP in GBM8 human cancer stem cell (**Figure 3.6C**). Based on studies by others (personal communication with Dr. Zhaozhong Han, Vanderbuilt University), GIP expression level could be increased by irradiation. To explore this possibility, cells were irradiated by X-rays at different doses (3 Gy, 4 Gy, and 5 Gy), but we did not see substantial increase in GIP expression.

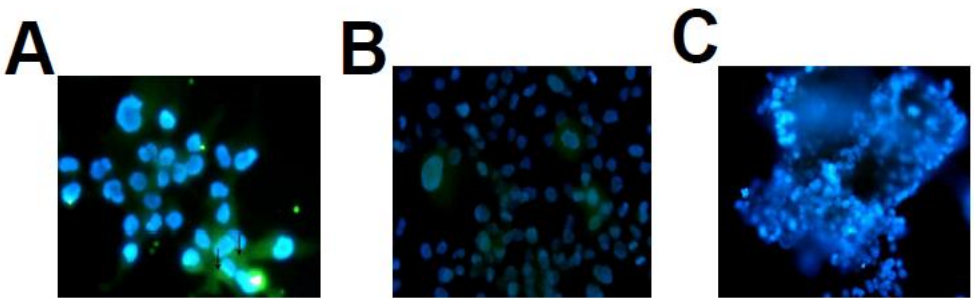

Figure 3.6 GIP detection in glioma cells. (A) GIP was detected in cytoplasm of the D54 MG cells using primary anti-GIP antibody followed by secondary antibody conjugated to Alexa 488 (green). (B) GIP was detected in cytoplasm of the U87 MG cells using primary anti-GIP antibody followed by secondary antibody conjugated to Alexa 488 (green). (C) No expression of GIP in GBM8 human glioma stem cells was detected. Cell nuclei were stained with DAPI (blue).

Our searches of protein databases (UniProt) indicated multiple cancer-related proteins containing the internal motif identified in this study through phage library selections (**Table 3.3**). To investigate the potentials of the phage-selected GIP-binding peptides for development of anti-cancer therapeutics, we studied intracellular distribution of the peptides in human D54 MG glioma cells. The cells were treated with a synthetic ESSVDLLDG peptide fused to a cell-penetrating peptide G_2R_7 (GGRRRRRRR) labeled with TAMRA. By fluorescence microscopy, the labeled peptide was shown to be uniformly distributed in the cytoplasm of the glioma cells (**Figure 3.7A**). Next, cultures of D54 MG cells were treated with the TAMARA-labeled peptide followed by GIP staining with an anti-GIP antibody detected with a secondary antibody conjugated to Alexa 488 (**Figure 3.7**). Both, the peptide (red) as well as GIP (green), were detected in the cytoplasm of the cells supporting the possibility that an interaction between the two may take place.

Protein (synonyms)	**Uniprot accession number**	**Function (s)**	**Peptide motif**	**Subcellular localization**
Ankyrin-B	**Q01484**	Attaches integral membrane proteins to cytoskeletal elements (*34*).	GSSVDS	Cytoplasm, Cell membrane
mediator complex subunit 1(MED1)	**Q15648**	Component of the mediator complex, a coactivator involved in the regulated transcription of nearly all RNA polymerase II-dependent gene (*35*). A decrease of RB18A/MED1 expression in human melanoma cells increases their tumorigenic phenotype (*36*).	GSSVDS	Nucleus
CYLD cylindromatosis (turban tumor syndrome)/ **Probable ubiquitin carboxyl-terminal hydrolase CYLD**	**Q9NQC7**	Negatively regulates activation of the transcription factor NF-kappaB by specific tumour-necrosis factor receptors (TNFRs). Loss of the deubiquitinating activity of CYLD correlates with tumorigenesis (*37*).	GSSVDS	Cytoplasm, perinuclear region, cell membrane
MYO18B myosin XVIIIB	**Q8IUG5**	May play an important role as a tumor suppressor in the development of lung cancer (*38*).	ESSVDD	Cytoplasm, nucleus
autophagy related protein 7 (ATG7)	**O95352**	ATG7 functions at the apex of a signalling system that is required for autophagy, the	ESSVDL	Cytoplasm

		process whereby the cell directs the degradation of cytoplasmic components in response to reduced nutrients in the environment (*39*).		
Rho GTPase -activating protein 9 (ARHGAP9)	**Q9BRR9**	ARHGAP9 is involved in regulating adhesion of hematopoietic cells to the extracellular matrix (*40*).	ESSVDL	Intracellular, which includes the nuleus and cytoplasm.
X-ray repair complementing defective repair in CHO cells 4/X-ray repair cross complementing protein 4 (XRCC4)	**Q13426**	Involved in DNA non-homologous end joining (NHEJ) required for double-strand break repair and V(D)J recombination (*41*).	ISSLDV	Nucleus
RSF1 remodeling and spacing factor 1	**Q96T23**	Required for assembly of regular nucleosome arrays by the RSF chromatin-remodeling complex (*42*).	ETSLDS	Nucleus
HIV Tat specific factor 1	**O43719**	Functions as a general transcription factor playing a role in the process of transcriptional elongation (*43*).	GTNLDGN	Nucleus

Table 3.3 Human proteins that have homologies with GIP-binding peptides identified with phage display.

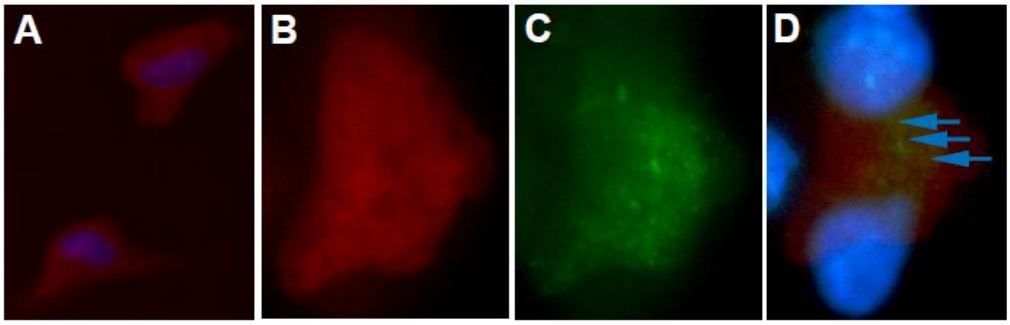

Figure 3.7 Localization of GIP and a GIP-binding peptide ESSVDLLDG in D54 MG human glioma cells (**A**) Internalization of the peptide by glioma cells. The peptide was labeled with TAMRA and is shown inside the cells as red fluorescence. (**B**) ESSVDLLDG peptide was localized in the cytoplasm of the glioma cell (red). (**C**) GIP was localized in the cytoplasm of the same cell using primary anti-GIP antibody followed by secondary antibody conjugated to Alexa 488 (green). (**D**) Merged image of (**B**) and (**C**) showing colocalization (indicated with arrows) of GIP and ESSVDLLDG peptide within the cell. Cell nuclei were stained with DAPI (blue).

To probe the potential anti-cancer activity of the ESSVDLLDG peptide, the glioma cells were treated with the peptide at concentrations ranging from 10 to 200 µM for 16 h and their metabolism was measured by the MTT assay. The cell metabolism was suppressed in a dose dependent manner with increasing peptide concentrations (**Figure 3.8**). The peptide concentration required to suppress 50% of the cell metabolism (IC_{50}) was found to be equal to 69±10 µM.

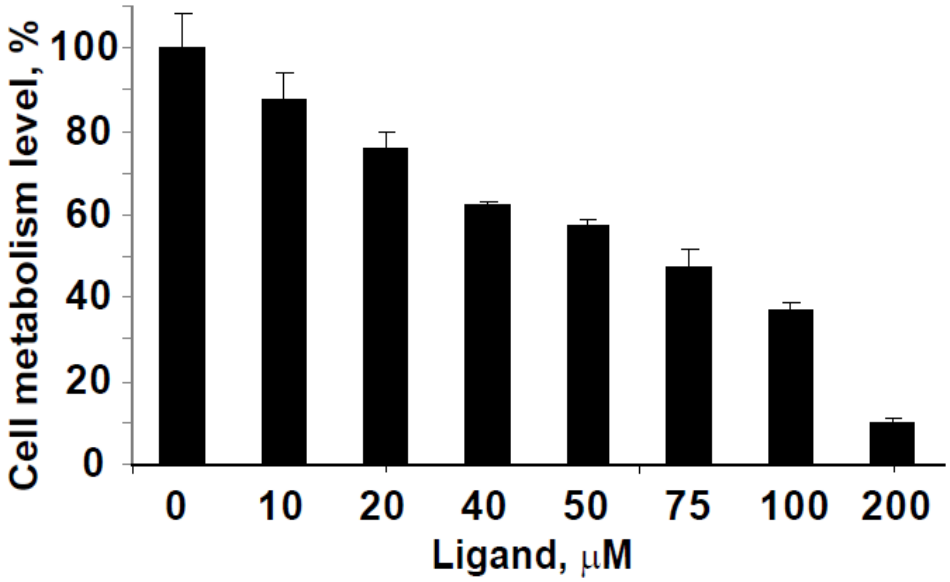

Figure 3.8 Effect of the ESSVDLLD peptide on the cellular metabolism of D54 MG human glioma cells. The peptide was added to the cells at different concentrations ranging from 10-200 μM. Cell metabolism was determined using the MTT assay and expressed as a percentage of the mean absorbance measured in untreated control cell cultures.

3.4 Discussion

In this study, a phage landscape library f8/9 with multivalently displayed foreign nonamers was used for identification of binding motifs for GIP, a single PDZ domain containing protein. The library used here was diverse, composed of two billion different phage clones. A randomized DNA segment was inserted into the N-terminus of the gene *gpVIII* that encodes the major phage coat protein (*25*). PDZ-binding phage clones were isolated from the library in three successive rounds of biopanning. In GIP-phage binding assay, all of the selected phage clones were confirmed to be specific to GIP. Analysis of the peptide sequences led to identification of a consensus internal-binding motif S/T-X-L/V-D. In the majority of previously reported phage display studies on PDZ-binding motifs, the identified peptide ligands were C-terminal recognition motifs (*1, 44*). To our knowledge, this is the first report of GIP recognition of internal binding motifs. In the selected sequences, the P_{-2} position was always occupied by S or T, which were followed by variable amino acids in position $P_{-1.}$ Position P_0 was always occupied by V or L, but not by I. This might indicate that steric factors are involved in the binding, thus, only the symmetric V or L side chains fit into the hydrophobic cavity, but not the asymmetric I side chain. Aspartate in the P_{+1} position was absolutely required.

The complex model structures of GIP and internal motif peptides (ESSVDLLDG and GSGTDLDAS) developed by Dr. David Zoetewey in our lab show that the ligands bind in the β-strand addition mechanism. It also shows that S or T at P_{-2} form a hydrogen bond with His90, and V or L at P_0 binds within the hydrophobic pocket created by Leu29, Phe31, Ile33 and Leu97. These mechanisms are similar with C-terminal recognition by GIP. Chemical shift perturbation data show that these residues show large chemical shift

change. Chemical shift perturbations data of GIP and internal motif peptides show similar patterns of binding within GIP. In case of internal motif, P_0 is not the C-terminus with a free carboxylate group. Both Leu29 and Gly30 amides within the ILGF loop of GIP make hydrogen bonds to either a canonical C-terminal carboxylate group or carboxylate group from the side-chain of D at P_{+1} for an internal motif. Interaction patterns show that the ILGF loop appears to be somewhat flexible. It moves in to bind to a canonical C-terminal carboxylate group or moves out to bind to a carboxylate group from D at P_{+1} of an internal motif. In comparison to ESSVDLLDG, GSGTDLDAS forms a slightly more extended antiparallel β-sheet with β2 of GIP. ESSVDLLDG has both VD and LD pairs in its sequence, but only VD binds to the ILGF loop because it contains S at the relative position P_{-2}. The His90 at α2:1 is responsible for the formation of hydrogen bond with side chain hydroxyl group of S/T at P_{-2} of the interacting partner. A hydrophobic residue at P_0 and D at P_{+1} is present in the internal motif. Our analysis of the identified phage sequences shows that D is absolutely conserved among all the internal binding motifs. Each internal binding synthetic peptide did bind to GIP as monitored by NMR titrations. Thus, while E is also negatively charged, it appears that its side-chain is too large to accommodate the binding pocket in an energetically favorable mode.

Interestingly, endonuclein, a cell cycle regulated WD-repeat protein, was recently reported to interact with GIP (*45*). Endonuclein does not contain a canonical C-terminal PDZ binding motif, but contains the sequence EISGLDL (387-393) within its five WD repeats. WD repeats are β-sheet domains that contain multiple β-hairpin turns. It is possible that endonuclein interacts with GIP through this region that serves as an internal

motif. If confirmed, this would be the first independent example of an interaction of GIP with a non-canonical internal motif.

To test the anti-cancer therapeutic potentials of the identified peptides, several biological assays were carried out with cancer cells. GIP was shown to have the same subcellular localization (**Figure 3.8**) as the synthetic peptide, ESSVDLLDG. Most importantly, the peptide was able to inhibit metabolism of the glioma cells. Additional experiments will be needed to validate that the negative effect of the peptide on glioma cells is the result of the peptide-GIP interaction. If confirmed, the peptides might be used in design of novel anti-cancer therapeutics targeting GIP.

Using protein database searches, we have identified several proteins with the S/T-X-V/L-D internal motif that were previously shown to be involved in various cancer pathways and tumorigenesis (**Table 3.3**). For example, reduced expression of the mediator complex subunit 1 (MED 1) protein containing the above motif was associated with a more pronounced tumorigenic phenotype in human melanoma cells (*36*). The CYLD gene that encodes the cylindromatosis 1 protein also has this motif and was found to be down-regulated in human hepatocellular carcinoma cells and involved in their apoptotic resistance and increasing evidence indicates that CYLD deficiency may promote the development of various cancers (*46*). Another S/T-X-V/L-D internal motif-containing protein, MYO18B, was reported to act as a tumor suppressor in the development of lung cancer (*47*). The MYO18B protein was also shown to play an essential role in ovarian cancer (*48*).

3.5 Conclusions

In conclusion, we have found that GIP interacts with new internal motifs and the potential target proteins that contain these motifs. GIP is most likely involved in many important cancer pathways through interactions with some or all of these potential target proteins. It is possible that GIP might be involved in such cancer pathways through interactions with the proteins containing motifs identified in this study. Our results from the binding of the internal motif peptides to GIP and their negative effect on the metabolism of human glioma cells demonstrate potentials for the development of anti-cancer strategies targeting GIP.

3.6 References:

1. Fuh, G., Pisabarro, M. T., Li, Y., Quan, C., Lasky, L. A., and Sidhu, S. S. (2000) Analysis of PDZ Domain-Ligand Interactions Using Carboxyl-terminal Phage Display. *J Biol Chem. 275*, 21486–21491.
2. Tonikian, R., Zhang, Y., Sazinsky, S. L., Currell, B., Yeh, J. H., Reva, B., Held, H. A., Appleton, B. A., Evangelista, M., Wu, Y., Xin, X., Chan, A. C., Seshagiri, S., Lasky, L. A., Sander, C., Boone, C., Bader, G. D., and Sidhu, S. S. (2008) A specificity map for the PDZ domain family. *PLoS Biol 6*, e239.
3. Lee, H. J., and Zheng, J. J. (2010) PDZ domains and their binding partners: Structure, specificity, and modifications. *Cell Commun Signal 8*, 8.
4. Sharma, S. C., Memic,A., Rupasinghe, C. N., Duc, A-C. E., and Spaller, M.R. (2009) T7 phage display as a method of peptide ligand discovery for PDZ domain proteins. *Peptide Sci. 92*, 183-193.
5. Zhang, Y., Appleton, B. A., Wiesmann, C., Lau, T., Costa, M., Hannoush, R. N., and Sidhu, S. S. (2009) Inhibition of Wnt/β-catenin signaling by peptides that target dishevelled PDZ domains. *Nat Chem Biol.* 5, 217-219.
6. Hillier, B. J., Christopherson, K. S., Prehoda, K. E., Bredt, D. S., and Lim, W. A. (1999) Unexpected Modes of PDZ Domain Scaffolding Revealed by Structure of nNOS-Syntrophin Complex. *Science 284*, 812–815.
7. Penkert, R. R., DiVittorio, H. M., and Prehoda, K. E. (2004) Internal recognition through PDZ domain plasticity in the Par-6-Pals1 complex. *Nat Struct Mol Biol 11*, 1122-1127.

8. Petrenko, V. A., Smith, G. P., Mazooji, M. M., and Quinn, T. (2002) Alfa-helically constrained phage display library. *Protein Eng 15*, 943-950.

9. Hung, T-J., and Kemphues, K. J. (1999) PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in Caenorhabditis elegans embryos. *Development 126*, 127-135.

10. Torres, R., Firestein, B. L., Dong, H., Staudinger, J., Olson, E. N., Huganir, R. L., Bredt, D. S., Gale, N. W., and Yancopoulos, G. D. (1998) PDZ Proteins Bind, Cluster, and Synaptically Colocalize with Eph Receptors and Their Ephrin Ligands. *Neuron 21*, 1453-1463.

11. Altman, F. P. (1976) Tetrazohum salts and formazans. Prog. Histochem. *Cytochem. 9*, 1–56.

12. Vellonen, K. S., Honkakoski, P., and Urtti, A. (2004) Substrates and inhibitors of efflux proteins interfere with the MTT assay in cells and may lead to underestimation of drug toxicity. *Eur. J. Pharm. Sci. 23*,181–188.

13. Laaksonen, T., Santos, H., Vihola, H., Salonen, J., Riikonen, J., Heikkilä, T., Peltonen, L., Kumar, N., Murzin, D. Y., Lehto, V-P., and Hirvonen, J. (2007) Failure of MTT as a Toxicity Testing Agent for Mesoporous Silicon Microparticles. *Chem. Res. Toxicol. 20*, 1913–1918.

14. Zhang, Y., Appleton, B. A., Wiesmann, C., Lau, T., Costa, M., Hannoush, R. N., and Sidhu, S. S. (2009) Inhibition of Wnt signaling by Dishevelled PDZ peptides *Nature Chemical Biology 5*, 217-219.

15. Banerjee, M., Huang, C., Marquez, J., and Mohanty, S. (2008) Probing the structure and function of human glutaminase-interacting protein: a possible target for drug design. *Biochemistry 47*, 9208-9219.

16. Petrenko, V. A., Smith, G. P., Mazooji, M. M., and Quinn, T. (2002) Alpha-helically constrained phage display library. *Protein Eng 15*, 943-950.

17. Petrenko, V. A., Smith, G. P., Gong, X., and Quinn, T. (1996) A library of organic landscapes on filamentous phage. *Protein Eng 9*, 797-801.

18. Cavanagh, J., and Rance, M. (1990) Sensitivity improvement in isotropic mixing (TOCSY) experiments. *Journal of Magnetic Resonance (1969) 88*, 72-85.

19. Bax, A., and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance (1969) 65*, 355-360.

20. Otting, G., and Wüthrich, K. (1989) Extended heteronuclear editing of 2D 1H NMR spectra of isotope-labeled proteins, using the X([omega]1, [omega]2) double half filter. *Journal of Magnetic Resonance (1969) 85*, 586-594.

21. Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxB RNA Complex. *Journal of the American Chemical Society 119*, 6711-6721.

22. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics 19*, 315-316.

23. Bishop, T. C. (2009) VDNA: the virtual DNA plug-in for VMD. *Bioinformatics 25*, 3187-3188.

24. Banerjee, M., Zoetewy, D., Petranko, V., Samoylova, T., and Mohanty, S. (2011) Structural and functional insights into the interaction of GIP with novel internal motif peptides: potential anti-cancer therapeutics. Submitted.

25. Kuzmicheva, G. A., Jayanna, P. K., Sorokulova, I. B., and Petrenko, V. A. (2009) Diversity and censoring of landscape phage libraries. *Protein Eng Des Sel 22*, 9-18.

26. Sidhu, S. S. Phage Display in Biotechnology and Drug Discovery. Drug Discovery Series, in: A. Carmen (Ed.), vol. 3. 2005, CRC Press, Taylor & Francis Group: Boca Raton, London, New York, Singapore, 748pp.

27. Zoetewey, D. L., Ovee, M., Banerjee, M., Bhaskaran, R., and Mohanty, S. (2011) Promiscuous binding at the crossroads of numerous cancer pathways: insight from the binding of glutaminase interacting protein with glutaminase L. *Biochemistry 50*, 3528-3539.

28. Kornau, H. C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science 269*, 1737-17402.

29. Katre, U. V., Mazumder, S., Prusti, R. K., and Mohanty, S. (2009) Ligand binding turns moth pheromone-binding protein into a pH sensor: effect on the Antheraea polyphemus PBP1 conformation. *J Biol Chem 284*, 32167-32177.

30. Schultz, J., Hoffmuller, U., Krause, G., Ashurst, J., Macias, M. J., Schmieder, P., Schneider-Mergener, J., and Oschkinat, H. (1998) Specific interactions between the syntrophin PDZ domain and voltage-gated sodium channels. *Nat Struct Biol 5*, 19-24.

31. Kanamori, M., Sandy, P., Marzinotto, S., Benetti, R., Kai, C., Hayashizaki, Y., Schneider, C., and Suzuki, H. (2003) The PDZ protein tax-interacting protein-1 inhibits beta-catenin transcriptional activity and growth of colorectal cancer cells. *J Biol Chem 278*, 38758-38764.

32. Wang, H., Yan, H., Fu, A., Han, M., Hallahan, D., and Han, Z. (2010) TIP-1 translocation onto the cell plasma membrane is a molecular biomarker of tumor response to ionizing radiation. *PLoS One 5*, e12051.

33. Hariri, G., Yan, H., Wang, H., Han, Z., and Hallahan, D. E. (2010) Radiation-guided drug delivery to mouse models of lung cancer. *Clin Cancer Res 16*, 4968-4977.

34. Mohler, P. J., Schott, J.-J., Gramolini, A. O., Dilly, K. W., Guatimosim, S., duBell, W. H., Song, L.-S., Haurogne, K., Kyndt, F., Ali, M. E., Rogers, T. B., Lederer, W. J., Escande, D., Marec, H. L., and Bennett, V. (2003) Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. *Nature 421*, 634-639.

35. Ndong, J. d. L. C., Jean, D., Rousselet, N., and Frade, R. (2009) Down-regulation of the expression of RB18A/MED1, a cofactor of transcription, triggers strong tumorigenic phenotype of human melanoma cells. *International Journal of Cancer 124*, 2597-2606.

36. Yuan, C.-X., Ito, M., Fondell, J. D., Fu, Z.-Y., and Roeder, R. G. (1998) The TRAP220 component of a thyroid hormone receptor- associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-

dependent fashion. *Proceedings of the National Academy of Sciences of the United States of America 95*, 7939-7944.

37. Trompouki, E., Hatzivassiliou, E., Tsichritzis, T., Farmer, H., Ashworth, A., and Mosialos, G. (2003) CYLD is a deubiquitinating enzyme that negatively regulates NF-[kappa]B activation by TNFR family members. *Nature 424*, 793-796.

38. Nishioka, M., Kohno, T., Tani, M., Yanaihara, N., Tomizawa, Y., Otsuka, A., Sasaki, S., Kobayashi, K., Niki, T., Maeshima, A., Sekido, Y., Minna, J. D., Sone, S., and Yokota, J. (2002) MYO18B, a candidate tumor suppressor gene at chromosome 22q12.1, deleted, mutated, and methylated in human lung cancer. *Proceedings of the National Academy of Sciences of the United States of America 99*, 12269-12274.

39. Schulman, B. A., and Harper, J. W. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol 10*, 319-331.

40. Furukawa, Y., Kawasoe, T., Daigo, Y., Nishiwaki, T., Ishiguro, H., Takahashi, M., Kitayama, J., and Nakamura, Y. (2001) Isolation of a Novel Human Gene, ARHGAP9, Encoding a Rho-GTPase Activating Protein. *Biochemical and Biophysical Research Communications 284*, 643-649.

41. Li, Z., Otevrel, T., Gao, Y., Cheng, H.-L., Seed, B., Stamato, T. D., Taccioli, G. E., and Alt, F. W. (1995) The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell 83*, 1079-1089.

42. Shamay, M., Barak, O., Doitsh, G., Ben-Dor, I., and Shaul, Y. (2002) Hepatitis B virus pX interacts with HBXAP, a PHD finger protein to coactivate transcription.

J Biol Chem **277**, 9982-9988

43. Zhou, Q., and Sharp, P. A. (1996) Tat-SF1: cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science 274*, 605-610.

44. Sharma, S. C., Memic, A., Rupasinghe, C. N., Duc, A. C., and Spaller, M. R. (2009) T7 phage display as a method of peptide ligand discovery for PDZ domain proteins. *Biopolymers 92*, 183-193.

45. Ludvigsen, M., Ostergaard, M., Vorum, H., Jacobsen, C., and Honore, B. (2009) Identification and characterization of endonuclein binding proteins: evidence of modulatory effects on signal transduction and chaperone activity. *BMC Biochem 10*, 34.

46. Sun, S. C. (2010) CYLD: a tumor suppressor deubiquitinase regulating NF-kappaB activation and diverse biological processes. *Cell Death Differ 17*, 25-34.

47. Nishioka, M., Kohno, T., Tani, M., Yanaihara, N., Tomizawa, Y., Otsuka, A., Sasaki, S., Kobayashi, K., Niki, T., Maeshima, A., Sekido, Y., Minna, J. D., Sone, S., and Yokota, J. (2002) MYO18B, a candidate tumor suppressor gene at chromosome 22q12.1, deleted, mutated, and methylated in human lung cancer. *Proceedings of the National Academy of Sciences of the United States of America 99*, 12269-12274.

48. Yanaihara, N., Nishioka, M., Kohno, T., Otsuka, A., Okamoto, A., Ochiai, K., Tanaka, T., and Yokota, J. (2004) Reduced expression of MYO18B, a candidate tumor-suppressor gene on chromosome arm 22q, in ovarian cancer. *Int J Cancer 112*, 150-154.

CHAPTER FOUR

Conclusions and Future Directions

The studies performed include steps towards understanding the structure, function, dynamics, and mode of interactions of a PDZ domain-containing multifunctional human protein. GIP PDZ domain interacts with the C-termini of their target proteins. GIP interacts with several target proteins, and the target proteins are involved in various signaling pathways, mostly various cancerous pathways. GIP interacts with them and regulates their pathways. To design a drug, it is necessary to characterize all the proteins which are involved in GIP binding to minimize side effects. For that purpose, it was necessary to find the complete network of GIP. So far, GIP has been found to interact with the C-termini of its target proteins. To advance our understanding of GIP-ligand interaction, analysis was extended to identify novel motif peptides by selection from a phage display library. Interestingly, all the peptides contain internal binding motifs and present in several human proteins which are involved in various cancerous pathways. It is possible that GIP might be involved in such cancer pathways through interactions with the proteins containing motifs identified in this study. We have found that GIP interacts with both C-terminal and internal motif target proteins. Our results from the binding of the internal motif peptides to GIP and their negative effect on the metabolism of human glioma cells demonstrate potentials for the development of anti-cancer strategies targeting GIP.

This conclusion from my dissertation shows future directions on studies of GIP. First of all, we have found internal motifs for GIP and also the potential target proteins

for GIP. We have to confirm the interactions between GIP and the target proteins. We also need to check the signaling pathways how GIP regulates this interactions. When these pathways are well understood, it will be easier to determine the role of GIP in cancer. Secondly, we have found that the internal motifs are capable of reducing the metabolism of human glioma cells. Thus, the internal motifs should be tested on cancer cells whether these peptides can be used as anti-cancer drugs or not. We have also found how GIP interacts with the C-terminal or the internal motif of its target proteins. Therefore, it is possible to identify small drug-like molecules for GIP and check their roles as anti-cancer therapeutics as a target of GIP.