Structure, dynamics and interaction study of Glutaminase Interacting Protein (GIP) and its complex with Glutaminase L and Brain-specific Angiogenesis Inhibitor 2 (BAI2) peptide and characterization of subunit A of Heterodisulfide Reductase (HdrA) from *Methanothermobacter marburgensis*

by

Mohiuddin Ovee

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 3, 2013

Keywords: Glutaminase L, GIP, NMR, Hdr, FPLC

Approved by

Smita Mohanty, Major Professor & Co-Chair of dissertation committee, Associate Professor of Chemistry and Biochemistry
Eduardus Duin, Co-Chair, Associate Professor of Chemistry and Biochemistry
Holly Ellis, Associate Professor of Chemistry and Biochemistry
Michael Squillacote, Associate Professor of Chemistry and Biochemistry
Orlando Acevedo, Associate Professor of Chemistry and Biochemistry

Abstract

Glutaminase interacting protein (GIP) is a 124 amino acid long protein containing a single PDZ domain. This protein intersects a number of important biological pathways. In many of these pathways, the mechanism of function of this domain is still unknown. Its involvement in cancer pathways makes it a good target for drug development. We resolved the solution structures of both free GIP and GIP in complex with the C-terminal peptide analog of Glutaminase L to shed light on the mechanism of binding with the goal of future development of a potential inhibitor for GIP. To understand more of GIP's function, interactions with two target peptides were investigated using different biophysical methods. One of the peptides was homologous to the C-terminus of brain-specific angiogenesis inhibitor 2 (BAI2) and the other one used had a consensus PDZ class I binding motif. Both of the peptides showed moderate binding affinity toward GIP with the BAI2 peptide having comparatively higher affinity. Elucidating the mechanism of interactions for different target partners would help to lay out the network of function for GIP. In a separate project, to understand the mechanism of electron bifurcation in methanogenic archaea, efforts were made to purify either heterodisulfide reductase (Hdr) or the subunit A of Hdr (HdrA) from *Methanothermobacter marburgensis* or *Methanococcus maripaludis*. We were able to purify HdrA with limited purity and showed the presence of [4Fe-4S] clusters in HdrA through EPR studies. However, efforts to purify Hdr from both organisms were with limited success. It is important to continue the efforts to obtain pure

Hdr/HdrA to investigate the mechanism by which electron bifurcation takes place within this enzyme complex.

Acknowledgments

First of all, I would like to show my deepest gratitude to Dr. Smita Mohanty, major professor and co-chair, for her guidance, support, encouragement, and help throughout my graduate study. I would like to convey my heartfelt gratitude to Dr. Eduardus Duin, co-chair of my dissertation committee, for his relentless support, guidance and unbiased help and encouragement for the completion of my graduate study. I would like to thank Dr. Holly Ellis, Dr. Michael Squillacote and Dr. Orlando Acevedo for their help, support and constructive suggestion to my dissertation. I also want to thank my former colleague Dr. David Zoetewey, a postdoc with Dr. Smita Mohanty, with whom I had the privilege to work on several projects and learned a lot about NMR techniques. I would like to thank my friend Suman Mazumder for his continued support. I would like to thank former lab member Dr. Monimoy Banerjee and a very dear colleague Dr. Rajagopalan Bhaskaran who is the former NMR facilities manager of AU Department of Chemistry. I would like to thank Selamawit Ghebreamlak, Dr. Shigeki Saito, Dr. Chengdong Huang, Dr. Janarthanan Krishnamoorthy, Xiao Xiao, Divya Prakash, and Dr. Uma Katre for their meaningful discussions and help. I would like to thank NMR director Dr. Michael Meadows for his help. I would like to thank the chairman of the Department of Chemistry and Biochemistry Dr. J. V. Ortiz for his consistent support and help. I would also like to thank graduate program officer Dr. Rik Blumenthal, Dean of the Graduate School Dr. George Flowers and COSAM interim dean Dr. Charles Savrda for their commitment to the successful completion of my Ph.D. dissertation. I would like to thank all professors and the administrative staff at the

Department of Chemistry and Biochemistry, and all my friends at Auburn who directly or indirectly contributed to the success of my study. I would like to thank U.S. Department of Agriculture PECASE Presidential Early Career Award for Scientists and Engineers Award 2003-35302-12930, National Science Foundation Grant IBN-0628064, and National Institutes of Health Grant DK082397 for their funding to Dr. Smita Mohanty, which financially supported my research in her lab. I would like to thank my parents, all my family members, and relatives for their support and help. Last but not the least; I would like to thank my wife for her continuous encouragement, love and moral support.

Table of contents

Abstract iii

Acknowledgements v

List of Tables xv

List of Figures xvi

Chapter 1 Introduction 1

1.1 NMR 1

1.1.1 Principles of NMR 1

1.1.2 Relaxation processes 4

1.1.2.1 Spin-lattice relaxation 5

1.1.2.2 Spin-spin relaxation 5

1.1.3 Chemical shift 7

1.1.4 Spin-spin coupling 8

1.1.4.1 Scalar coupling (J-coupling) 8

1.1.4.2 Dipolar coupling 9

1.1.5 Nuclear Overhauser effect (NOE) 9

1.1.6 Multidimensional NMR spectroscopy 12

1.1.6.1 Two-dimensional NMR (2D NMR)....12
1.1.6.1.1 COSY....13
1.1.6.1.2 TOCSY....13
1.1.6.1.3 HSQC....13
1.1.6.1.4 NOESY....15
1.1.6.2 Three-dimensional NMR (3D NMR)....15
1.1.7 Protein NMR....18
1.1.8 Study of protein-ligand interaction by NMR....21
1.2 PDZ domain and Glutaminase Interacting Protein (GIP)....24
1.3 Glutaminase Interacting Protein (GIP)/Tax Interacting Protein-1 (TIP-1)....26
1.4 References....29
Chapter 2 Characterization of Glutaminase Interacting Protein (GIP): a PDZ domain....36
2.1 Background....36
2.1.1 Protein-protein interaction network....36
2.1.2 PDZ domain and its classes....38
2.1.3 Glutaminase interacting protein as a class I PDZ domain....39
2.1.4 Objective of the study....41
2.2 Materials and Methods....42
2.2.1 Cloning, over-expression and purification of ^{15}N, ^{13}C-labeled GIP....42
2.2.2 NMR data collection....42

2.2.3 Analysis of dynamics data 43

2.2.4 Structure calculation and refinement 43

2.3 Results 45

2.3.1 NMR structure determination of free GIP 45

2.3.1.1 Introduction 45

2.3.1.2 Backbone assignments of free GIP 47

2.3.1.3 Side-chain assignments of free GIP 53

2.3.1.4 NOE assignments of free GIP 65

2.3.1.5 Structure calculation of free GIP 68

2.3.1.6 Refinement of structures by ARIA 72

2.3.1.7 NMR structure of free GIP 73

2.3.1.8 Accession codes 77

2.3.2 Dynamics of free GIP from ^{15}N relaxation measurements 77

2.4 Conclusion 81

2.5 References 82

Chapter 3 Study of the mechanism of interaction between GIP and the Glutaminase L peptide ..90

3.1 Introduction 90

3.2 Objective of the study 91

3.3 Materials and Methods 92

3.3.1 Cloning, over-expression and purification of ^{15}N, ^{13}C-labeled GIP 92

3.3.1.1 Transformation of *E. coli* BL21DE3pLysS cells with the recombinant plasmid pET-3c/GIP 92
3.3.1.2 Preparation of overnight culture 93
3.3.1.3 Expression of the protein in batch culture 94
3.3.1.4 Cell harvest and lysis 94
3.3.1.5 Protein purification 95
3.3.1.6 NMR sample preparation 95
3.3.2 NMR data collection 95
3.3.3 Analysis of dynamics data 97
3.3.4 Structure calculation and refinement 97
3.4 Results 99
3.4.1 Protein expression 99
3.4.2 Protein purification 104
3.4.3 NMR structure determination of GIP-Glutaminase L peptide complex 104
3.4.3.1 Effect of peptide binding to the resonances of GIP protein 104
3.4.3.2 Backbone and side-chain assignments 111
3.4.3.2.1 For protein 111
3.4.3.2.1 For peptide 115
3.4.3.3 NOE assignments 121
3.4.3.4 Structure calculation 126
3.4.4 Comparison of the structure of free GIP with that of the GIP-Glutaminase L peptide complex 130

3.4.5 Binding and specificity of the Glutaminase L peptide.....132

3.4.6 Dynamics of the GIP-Glutaminase L peptide complex from ^{15}N relaxation measurements.....136

3.4.7 Intermediate chemical exchange within GIP due to the binding of the Glutaminase L peptide.....140

3.5 Discussion.....141

3.5.1 Specificity in the binding interaction between GIP and Glutaminase L peptide.....141

3.5.2 The effects of the Glutaminase L peptide binding on the dynamics of GIP.....144

3.5.3 Comparison to other GIP-peptide complex structures.....145

3.5.4 Comparison between NMR and crystal structures.....146

3.5.5 Potential for drug design.....147

3.6 Accession codes.....148

3.7 References.....149

Chapter 4 Determination of the mode of interaction of Glutaminase Interacting Protein (GIP) with two different interacting partners.....155

4.1 Introduction.....155

4.1.1 PDZ domain and its functions.....155

4.1.2 Binding pocket of PDZ domain.....155

4.1.3 GIP as a PDZ domain.....156

4.1.4 GIP in the brain.....157

4.1.5 Identification of interacting partners in brain.....157

4.2 Materials and Methods.....158

4.2.1 Expression and purification of ^{15}N- and unlabeled GIP.....158

4.2.2 Fluorescence 159

4.2.3 Circular Dichroism (CD) 159

4.2.4 Nuclear Magnetic Resonance (NMR).....160

4.3 Results and Discussion.....161

4.3.1 Protein expression.....161

4.3.2 Protein purification 161

4.3.3 Interaction of BAI2 peptide with GIP.....164

4.3.3.1 Characterization by Fluorescence spectroscopy 164

4.3.3.2 Characterization by CD spectroscopy.....167

4.3.3.3 Characterization by $^{1}H,^{15}N$-HSQC NMR 169

4.3.3.4 Chemical shift perturbations of GIP upon binding to the BAI2 peptide .175

4.3.4 Interaction of the control peptide with GIP 180

4.3.4.1 Characterization by Fluorescence spectroscopy 180

4.3.4.2 Characterization by CD spectroscopy.....183

4.3.4.3 Characterization by $^{1}H,^{15}N$-HSQC NMR 185

4.3.4.4 Chemical shift perturbations of GIP upon binding to the control peptide.....190

4.3.5 Comparison of interaction between GIP and BAI2 peptide with interaction between GIP and control peptide 193

4.4 References 196

Chapter 5 Characterization of subunit A of Heterodisulfide Reductase (HdrA) from *Methanothermobacter marburgensis*.....202

5.1 Introduction 202

5.1.1 Electron bifurcation 202

5.1.2 History of electron bifurcation 202

5.1.3 Mechanism of electron bifurcation 205

5.1.4 Electron bifurcation in other systems 206

5.1.5 Models for electron bifurcation 207

5.1.5.1 Model I 207

5.1.5.2 Model II 207

5.1.5.3 Model III 208

5.2 Objective of this study 210

5.3 Materials and methods 212

5.3.1 Purification of hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC) from *M. marburgensis* 212

5.3.1.1 Growth of *M. marburgensis* cells 212

5.3.1.2 Harvest and sonication of *M. marburgensis* cells 213

5.3.1.3 Purification of hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC) 213

5.3.2 Purification of heterodisulfide reductase (Hdr) from *M. maripaludis* cells 214

5.3.2.1 *M. maripaludis* cells 214

5.3.2.2 Purification of heterodisulfide reductase 214

5.3.3 Purification of HdrA from *M. maripaludis* $HdrA_{marburgensis}$ cells 215

5.3.3.1 *M. maripaludis* $HdrA_{marburgensis}$ cells 215

5.3.3.2 Purification of HdrA 215

5.3.4 Iron determination 216

5.3.5 UV-vis absorption analysis 217

5.3.6 EPR measurements 217

5.4 Results 218

5.4.1 Purification of the hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC) from *M. marburgensis* 218

5.4.2 Purification of heterodisulfide reductase (Hdr) from *M. maripaludis* cells 223

5.4.3 Purification of HdrA from *M. maripaludis* $HdrA_{marburgensis}$ cells 229

5.4.3.1 Purification of HdrA 229

5.4.3.2 Protein and iron determination 232

5.4.3.3 UV-vis absorption of the protein sample 232

5.4.3.4 EPR measurement of the HdrA protein sample 234

5.5 Conclusions and future direction 236

5.6 References 238

Appendix Tables 241

List of Tables

Table 1.1 Relationship of *I* to atomic number and mass number 2

Table 1.2 Correlation observed for some of the most commonly used 3D NMR experiments .20

Table 1.3 Classification of PDZ domains 28

Table 2.1 Sequential alignment of C-terminal binding partners of GIP 40

Table 2.2 Statistics of side-chain assignments of free GIP 61

Table 2.3 NMR structural statistics for the 20 selected lowest energy structures of free GIP ... 75

Table 3.1 Statistics of side-chain assignments of the GIP-Glutaminase L peptide complex ... 114

Table 3.2 Statistics of available proton assignments of the Glutaminase L peptide 116

Table 3.3 NMR structural statistics for the 20 selected lowest energy structures of the GIP-Glutaminase L Peptide Complex 128

Table 4.1 Dissociation constants of various residues of GIP upon binding with the BAI2 peptide by NMR 174

Table 4.2 Dissociation constants of various residues of GIP upon binding with the control peptide by NMR 189

List of Figures

Figure 1.1 Nuclear spin states in a magnetic field 3

Figure 1.2 “Precession” of nucleus in a magnetic field 4

Figure 1.3 The two components of a spinning nucleus in an applied magnetic field 6

Figure 1.4 Energy diagram for a two-spin system 11

Figure 1.5 General scheme for a 2D experiment.....12

Figure 1.6 A) COSY and B) TOCSY spectra 14

Figure 1.7 3D experiment as a combination of two sets of 2D experiments 16

Figure 1.8 Schematic presentation of how with the addition of another evolution time to the 2D experiment (Figure 1.7) can result in another frequency dimension for a 3D experiment 17

Figure 1.9 Effect of chemical exchange on NMR spectra 22

Figure 1.10 Examples of PDZ domain containing proteins 25

Figure 2.1 Different experiment tools to define protein interactions 37

Figure 2.2 Flowchart of structure determination by NMR.....46

Figure 2.3 Sequential assignments of V13-K20 from (^{1}H, ^{13}C)-strips of HNCACB experiment.....49

Figure 2.4 Sequential assignments of V13-K20 from (^{1}H, ^{13}C)-strips of HNCA experiment ...50

Figure 2.5 HNCACB and CBCA(CO)NH strips of I18 residue 51

Figure 2.6 ^{1}H, ^{15}N-HSQC spectrum of free GIP 52

Figure 2.7 Proline cis/trans isomerization 53

Figure 2.8 HC(CO)NH and HSQC-TOCSY spectra showing side-chain assignments of V60 and R59 residues 55

Figure 2.9 HCCH-COSY spectrum of Hα proton of N26 residue showing non-degeneracy of Hβ protons 56

Figure 2.10 HCCH-TOCSY spectrum of Hβ3 proton of P45 residue 57

Figure 2.11 HNHA spectrum of non-degenerate Hα protons of G70 residue 58

Figure 2.12 CC(CO)NH spectrum of Q23 residue 59

Figure 2.13 Aliphatic region of the ^{1}H, ^{13}C-HSQC spectrum of free GIP 62

Figure 2.13 A Part of the aliphatic region of the ^{1}H, ^{13}C-HSQC spectrum with assignments 63

Figure 2.14 Aromatic region of the ^{1}H, ^{13}C-HSQC spectrum with assignments 64

Figure 2.15 ^{15}N-edited HSQC-NOESY spectrum of I33 residue 66

Figure 2.16 ^{13}C-edited HSQC-NOESY spectrum of I33 QD1 proton 67

Figure 2.17 Input and output files for CYANA 1.0.6 70

Figure 2.18 Iterative cycle of CYANA 1.0.6 run 70

Figure 2.19 Input and output files for CYANA 2.1 71

Figure 2.20 Iterative cycle of CYANA 2.1 run 71

Figure 2.21 Ribbon diagrams of the ensemble of the 20 superimposed lowest energy structures of free GIP 74

Figure 2.22 The S^2 values derived using the modelfree analysis from the steady state ^{1}H-^{15}N NOE, R_1 and R_2 relaxation times of free GIP for each non-overlapping well defined residue 79

Figure 2.23 Residues with S^2 values below the threshold of 0.85 are mapped in red onto the structure of free GIP colored blue 80

Figure 3.1 Expression of GIP analyzed by SDS-PAGE 101

Figure 3.2 1D NMR spectrum of non-homogeneously labeled GIP sample 102

Figure 3.3 1D NMR spectrum of homogeneously labeled GIP sample.....103

Figure 3.4 Combined ^{1}H and ^{15}N backbone amide chemical shift perturbations (ΔHN) are plotted as a function of residue number in GIP.....106

Figure 3.5 The magnitudes of ΔHN presented in Figure 3.4 are represented as different colors on a ribbon diagram of free GIP 107

Figure 3.6 Combined HA and CA backbone chemical shift perturbations (ΔHC) are plotted as a function of residue number in GIP.....108

Figure 3.7 The magnitudes of ΔHC presented in Figure 3.6 are represented as different colors on a ribbon diagram of free GIP 109

Figure 3.8 An overlay of free GIP is shown in red and GIP-Glutaminase L peptide at a ratio of 1:3 in blue 110

Figure 3.9 Sequential assignments of V13-K20 in the GIP-Glutaminase L peptide complex from (^{1}H, ^{13}C)-strips of HNCACB experiment.....112

Figure 3.10 ^{1}H, ^{15}N-HSQC spectrum of the GIP-Glutaminase L peptide complex.....113

Figure 3.11 ^{1}H, ^{15}N-HSQC spectrum of the Glutaminase L peptide 117

Figure 3.12 ^{1}H, ^{13}C-HMQC spectrum of the Glutaminase L peptide 118

Figure 3.13 Homonuclear 2D TOCSY spectrum of the Glutaminase L peptide 119

Figure 3.14 2D selectively filtered NOESY spectrum of the Glutaminase L peptide 120

Figure 3.15 ^{13}C-edited HSQC-NOESY spectrum of I33 QD1 proton of GIP in its bound form.....123

Figure 3.16 ^{1}H, ^{13}C-HSQC spectrum of GIP in the bound form.....124

Figure 3.17 Three different HSQC-NOESY spectra of I33 residue of the GIP.....125

Figure 3.18 Ribbon diagrams of the ensemble of the 20 superimposed lowest energy structures of complexed GIP in blue with the Glutaminase L peptide in red 127

Figure 3.19 An overlay of free GIP is shown in green with the complexed GIP protein in blue and the Glutaminase L peptide in red.....131

Figure 3.20 Heavy atom details from the binding site of GIP with the Glutaminase L peptide.....135

Figure 3.21 A plot of ΔS^2 as a function of residue number where ΔS^2 refers to S^2 of the GIP-Glutaminase L peptide complex minus that of free GIP 138

Figure 3.22 The magnitude of ΔS^2 upon binding to the Glutaminase L peptide was mapped onto the structure of free GIP.....139

Figure 4.1 Expression of GIP analyzed by SDS-PAGE.....162

Figure 4.2 Purification of GIP analyzed by SDS-PAGE 163

Figure 4.3 Fluorescence emission spectrum of GIP with the BAI2 peptide 165

Figure 4.4 Non-linear curve fitting assuming 1:1 binding between GIP and the BAI2 peptide where $(F_0 - F_C)/(F_0 - F_{min})$ was plotted against peptide concentration.....166

Figure 4.5 Changes in the CD spectra of GIP upon binding with increasing concentrations of the BAI2 peptide for the wavelength range of 194 nm to 250 nm 168

Figure 4.6 Changes of 2D $^{1}H,^{15}N$-HSQC spectra upon addition of the BAI2 peptide to 100 μM of ^{15}N-labeled GIP 171

Figure 4.7 Expanded region of the spectra demonstrating the chemical shift perturbations of residue E17 upon titration of GIP with the BAI2 peptide 172

Figure 4.8 The NMR titration binding curve for the titration of GIP with the BAI2 peptide..173

Figure 4.9 Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the BAI2 peptide.....178

Figure 4.10 Fluorescence emission spectrum of GIP with the control peptide.....181

Figure 4.11 Non-linear curve fitting assuming 1:1 binding between GIP and the control peptide where $(F_0 - F_C)/(F_0 - F_{min})$ was plotted against peptide concentration.182

Figure 4.12 Changes in the CD spectra of GIP upon binding with increasing concentrations of the control peptide for the wavelength range of 194 nm to 250 nm.....184

Figure 4.13 Changes of 2D $^{1}H,^{15}N$-HSQC spectra upon addition of the control peptide to 100 μM of ^{15}N-labeled GIP 186

Figure 4.14 Expanded region of the spectra demonstrating the chemical shift perturbations of residue N81 upon titration of GIP with the control peptide.....187

Figure 4.15 The NMR titration binding curve for the titration of GIP with the control peptide.....188

Figure 4.16 Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the control peptide 192

Figure 4.17 The Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the BAI2 (red) and the control (black) peptide.....195

Figure 5.1 Model of the structure of the hydrogenase:heterodisulfide reductase complex from *Methanothermobacter marburgensis* 204

Figure 5.2 Models for electron bifurcation 209

Figure 5.3 Chromatography profile of DEAE-Sepharose column for purification of MvhADG/HdrABC complex 219

Figure 5.4 Chromatography profile of Q-Sepharose column for purification of MvhADG/HdrABC complex 220

Figure 5.5 Chromatography profile of Superdex 200 column for purification of MvhADG/HdrABC complex 221

Figure 5.6 15%SDS-PAGE analysis of the fractions from Superdex 200 column for MvhADG/HdrABC complex purification 222

Figure 5.7 EPR spectra of iron-sulfur clusters 224

Figure 5.8 EPR spectrum of Hdr complex from *M. maripaludis* 225

Figure 5.9 Chromatography profile of Superdex 200 column for purification of Hdr from *M. maripaludis* 226

Figure 5.10 17% SDS-PAGE analysis of the Hdr complex from *M. maripaludis* cells 227

Figure 5.11 8% native PAGE analysis of the Hdr complex from *M. maripaludis* cells 228

Figure 5.12 Chromatography profile of nickel column for purification of HdrA from *M. maripaludis* $HdrA_{marburgensis}$ 230

Figure 5.13 12% SDS-PAGE analysis of the HdrA sample from *M. maripaludis* $HdrA_{marburgensis}$ cells 231

Figure 5.14 UV-vis absorption of the HdrA protein sample 233

Figure 5.15 EPR spectra of HdrA protein sample at the temperature of 8 K and at different microwave frequencies 235

Chapter 1

Introduction

1.1 NMR

1.1.1 Principles of NMR

Nuclear magnetic resonance (NMR) was first introduced in 1938. From then on, NMR has become one of the most powerful analytical techniques, widely used in many different fields. NMR is based on the fact that certain nuclei possess spin angular momentum and a resulting magnetic moment. Since a nucleus is positively charged, it would act as a spinning charged particle like a current flowing in a circle. If the nucleus has an angular momentum, P, then such spinning would produce a magnetic field parallel to the spinning axis, and the nucleus would have a magnetic moment, μ. From quantum mechanics, it is known that angular momentum is quantized in half-integral or integral multiples of $h/2\pi$, where h is Planck‟s constant. If I denotes the nuclear spin quantum number, the maximum observable component of angular momentum can be given as:

$$P = Ih/2\pi \qquad (1.1)$$

The spin quantum number can be different for different nuclei such as 0, $\frac{1}{2}$, 1, $\frac{3}{2}$ etc. If I is zero, there will not be any angular momentum for the nucleus, examples are ^{12}C and ^{16}O. The spin quantum number I is related to the atomic number and mass number (T**able 1.1**).

Atomic number	Mass number	Spin quantum number (I)
Even	Even	0
Even or odd	Odd	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots\ldots$
Even	Odd	1, 2, 3, ...

Table 1.1: Relationship of I to atomic number and mass number. Adapted from reference 1.

Since both protons and neutrons spin in the nucleus, they will pair with other protons and neutrons in the same nucleus but with opposite spin and, thus, such a relationship between I and atomic number and mass number can be established.

The angular momentum of the nucleus will follow the $(2I + 1)$ rule to acquire the orientation with respect to the external magnetic field when placed in a uniform magnetic field. A nucleus that has half spin angular momentum ($I = \frac{1}{2}$), such as ^{1}H or ^{13}C, will have two orientations, i.e., a lower energy and a higher energy orientation (**Figure 1.1**). In the lower energy orientation the magnetic moment of the nucleus will be aligned along the external magnetic field whereas in the higher energy orientation it will be aligned against the magnetic field. The potential energy of the nucleus in each orientation equals to $\mu B_0 \cos\theta$, where B_0 is the strength of the external magnetic field and θ is the angle between the axis of the spin and the direction of the magnetic field. The energy difference, ΔE, between the two energy states is proportional to the external magnetic field.

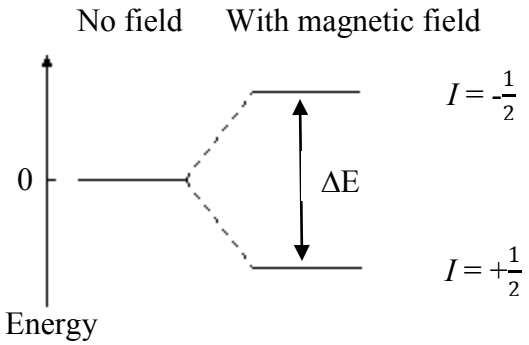


Figure 1.1: Nuclear spin states in a magnetic field. Adapted from reference 1.

Due to the influence of the external magnetic field the spinning nucleus "precesses", i.e., the two ends of the spinning axis follow a circular path but opposite in direction to each other (**Figure 1.2**). For the transition of the nucleus from the lower energy state to the higher energy state, a radio frequency wave that has the exactly equal frequency to that of the "precession" needs to be applied perpendicular to the external magnetic field. The Larmor equation states the relationship between the frequency of this electromagnetic radiation ν and the strength of the magnetic field B_0

$$\nu = \gamma B_0/2\pi \tag{1.2}$$

where γ is the gyromagnetic ratio. Each nucleus has its own characteristic gyromagnetic ratio, for example, ^{1}H has a gyromagnetic ratio (42.576 $MHzT^{-1}$) that is approximately 10 times that of ^{15}N (4.316 $MHzT^{-1}$) and 4 times that of ^{13}C (10.705 $MHzT^{-1}$) (*1*).

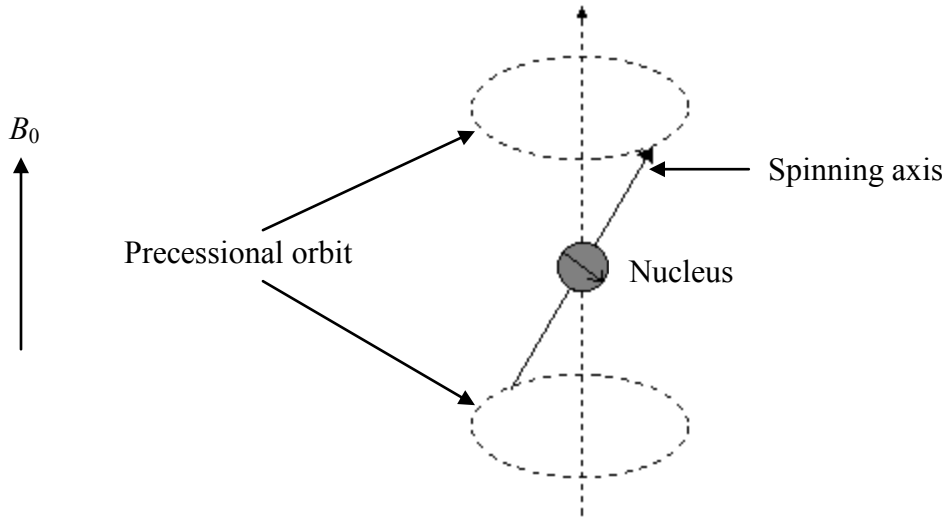


Figure 1.2: "Precession" of nucleus in a magnetic field. Adapted from reference 1.

1.1.2 Relaxation processes

According to the Boltzmann distribution, there is a slight population difference between the two energy states since the nuclei are slightly in excess of number in the lower energy state than in the higher energy state. When the radio frequency wave is applied, it causes the transition of these excess nuclei from the lower energy state to the higher energy state until the population difference becomes zero, as the populations at both energy levels become equal. Such a state is referred to as "saturation" state. To regain the Boltzmann distribution of the higher number of nuclei in the ground state, various relaxation processes take place that allow the nuclei from the higher energy state to come back to the lower energy state. This results in an equilibrium state at an intermediate level between restorations of the initial Boltzmann distribution and complete elimination of that distribution. Such a state can continue to produce an NMR signal.

There are mainly two relaxation processes:

1. Spin-Lattice relaxation (T_1)
2. Spin-Spin relaxation (T_2)

1.1.2.1 Spin-lattice relaxation (T_1)

The precessing nucleus under the influence of an external magnetic field will also face the fluctuating fields generated by the lattice. If the orientation of the field of the lattice is correct and its frequency equals the precession frequency of the nuclei of the higher energy level, then the energy of the nuclei can release energy to the lattice in the form of thermal energy and the nuclei can relax back to the lower energy state along the Z-axis. T_1 depends both on molecular motion of the lattice and the gyromagnetic ratio of the nucleus. The external magnetic field has a very strong influence on T_1 and the higher the magnetic field the slower the T_1 value (which means more efficient relaxation).

1.1.2.2 Spin-spin relaxation (T_2)

This relaxation process is also known as transverse relaxation. In this relaxation process the excited magnetization vector decays in the direction of X-Y plane which is perpendicular to the external magnetic field.

The magnetic field of a precessing nucleus has two components; one that is aligned with the external magnetic field and another one is spinning at processional frequency in the X-Y plane. The component parallel with the applied field is its static component and the other one is rotating component (**Figure 1.3**).

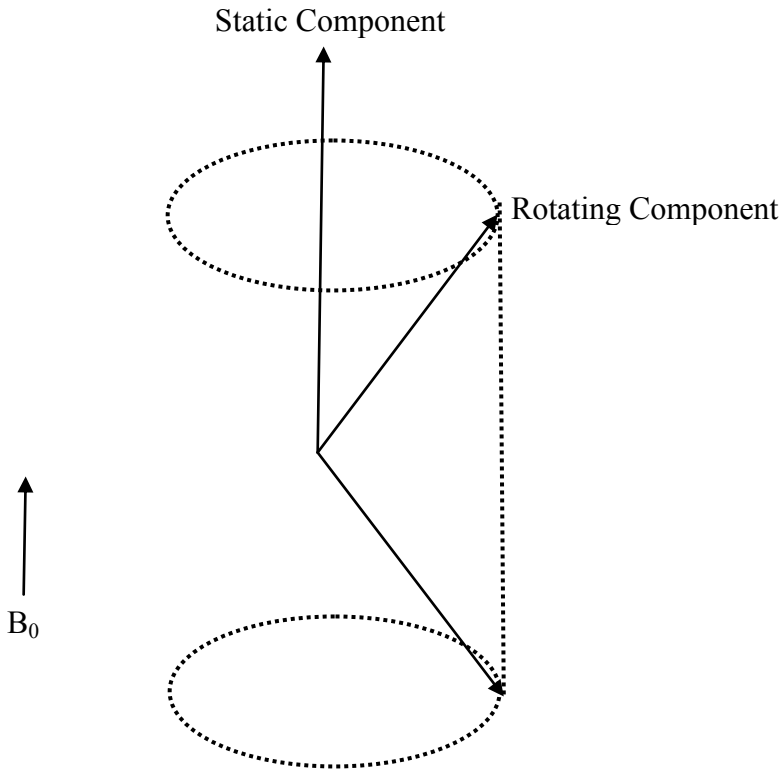


Figure 1.3: The two components of a spinning nucleus in an applied magnetic field. The rotating component present in flipping orientation is also shown here. Adapted from reference 1.

The static component of the magnetic field of the nucleus will add up to the main field as experienced by any neighboring nucleus resulting in the broadening of the resonance signals. If the neighboring nucleus is precessing at the same frequency as that of the rotating component and at the correct orientation, then it would cause a mutual exchange of energy between the two nuclei. Such exchange of energy would cause the spin to relax back to its original state. The time needed for such relaxation is known as spin-spin relaxation or T_2. This would also cause the broadening of the resonance signal. T_2 is shorter than T_1 and can be determined by NMR.

1.1.3 Chemical shift

Depending upon the magnitude of the external magnetic field, gyromagnetic ratio of the nucleus and the molecular environment of the nucleus, a nucleus comes to resonance at a certain frequency. It is the third factor which gives rise to the notion of "chemical shift". Chemical shift of a particular nucleus can be defined by the following equation:

$$\delta = (\nu_s - \nu_{Standard})/Z \qquad (1.3)$$

where ν_s is the resonance frequency of the nucleus in Hz, $\nu_{Standard}$ is the resonance frequency, in Hz, of an internal standard (that usually gives a sharp signal at a high value of the magnetic field) while recording NMR spectra, and Z is the frequency of the instrument in MHz (megahertz=10^6 Hz). Thus the unit for chemical shift (δ) is parts per million or ppm.

Nuclei with different molecular environment show different chemical shifts. This is very useful for structure determination by NMR.

1.1.4 Spin-spin coupling

1.1.4.1 Scalar coupling (J-coupling)

Interaction between nuclei connected through the network of chemical bonds results in scalar coupling. This results in splitting of NMR peaks. This happens due to the two possible spin states for any given nucleus (**Figure 1.1**). Two chemically bonded nuclei influence each other‟s magnetic field by their different spin states. If the nucleus remains aligned parallel to the external magnetic field in the lower energy state, the bonded nucleus will need a slightly lower magnetic strength to come to resonance. Whereas, in the other case, it will experience a lower total magnetic field and will require a little higher value of magnetic field to come to resonance.

However, the reason behind the transmission of the influence of the magnetic state to a nucleus to another lies in the fact of the changed electronic spin states due to the existing nuclear spin states. If another nucleus overlaps with the same affected bonding orbital then the changed electronic spin states affect the nuclear spin states of the second one. It results in a slight change in resonance frequency for the second nucleus. These two nuclei are called J-coupled.

Scalar coupling has many uses in NMR including the three-bond J-coupling for the measurement of dihedral angle, understanding the structural make-ups of atoms in a molecule and, very importantly, coherence transfer or magnetization transfer through scalar couplings. Another important feature of scalar coupling is that it is always constant for a certain set of bonds in a certain molecular structure independent of the external magnetic field. This J-coupling constant property can be very useful in the investigation of various small molecules including drugs.

1.1.4.2 Dipolar coupling

Dipolar coupling is through space interaction between nuclear spins. Dipolar coupling is involved in most spin-spin relaxation. The Nuclear Overhauser effect (NOE) is also an important outcome from dipolar coupling, which results from the change in the intensity of the resonance signal of a nucleus when the signal of dipolar coupled another nucleus is changed. NOEs are very important for the investigation of the structure of various bio-macromolecules (such as protein, DNA, and RNA) and large organic compounds and also the interaction between different molecules.

1.1.5 Nuclear Overhauser effect (NOE)

When a nucleus is irradiated with radio frequency, relaxation processes only through scalar coupling are not enough for that nucleus to reach the equilibrium state. Then, through a dipolar coupling relaxation mechanism, this nucleus can transfer some of its energy to another nucleus that is close enough in space. The second nucleus behaves as if it had been irradiated and relaxes back to the ground state. It causes the population of the ground state to increase and, thus, the intensity of the second nucleus is enhanced. This phenomenon is called nuclear Overhauser effect (NOE).

It can be illustrated by considering two nuclei A and B that are close enough in space for the relaxation process to affect each other. Both of the nuclei can exist in two different spin states, α or β, where α is the lower energy state. Thus, these two nuclei can be represented by four energy states: $\alpha\alpha$, $\alpha\beta$, $\beta\alpha$, and $\beta\beta$ (**Figure 1.4**). The allowed transitions here are between

adjacent levels, such as from αα to αβ or from βα to ββ (W1). When a radiation frequency is applied to irradiate one of the two nuclei, then the populations between sates ββ and βα or between states αβ and αα become equal. However, there is still a population difference that remains between spin states ββ and αα. Dipolar coupling relaxation process (W2) allows restoration of this difference to some extent resulting in the increase of intensity of the NMR resonance line for the second nucleus. This results in a positive NOE which is prevalent for small molecules that tumble in solution fast. For larger molecules, which slowly tumble in the solution, another type of relaxation process operates between βα and αβ (W0). As a result, a decrease in the population difference occurs between αβ and ββ (or αα and βα). This produces lesser intensity in the lines known as negative NOE.

NOE difference measurements can be used to determine the distance between two nuclei. Its intensity is inversely proportional to r^6, where r is the distance between two nuclei. Thus, with the increase in the distance of the two nuclei, there will be a proportional decrease of NOE intensity.

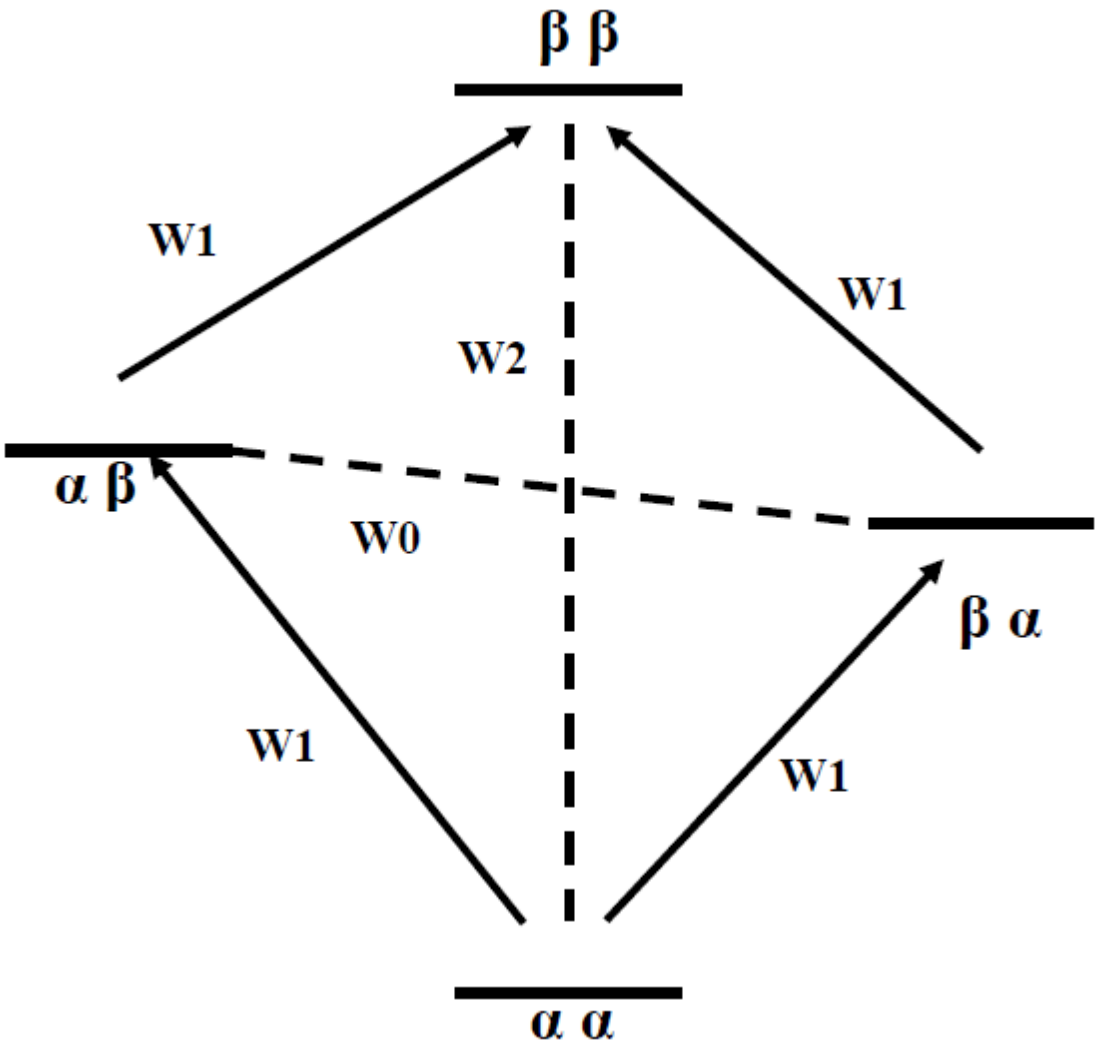


Figure 1.4: Energy diagram for a two-spin system. Adapted from reference (*2*).

1.1.6 Multidimensional NMR spectroscopy

1.1.6.1 Two-dimensional NMR (2D NMR)

With the development of various multidimensional NMR spectroscopic methods in the last few decades it was possible to observe the growth in the successful application of NMR to biological studies. It started off with the introduction of a time period known as the evolution period between preparation and detection periods by Jeener in 1971 which formed the basis for the two-dimensional (2D) NMR spectroscopy. Thus, the time-axis of any 2D experiment can be divided into three (or four) segments (**Figure 1.5**). These are the preparation period, the evolution period and the detection period. The preparation period allows the nuclei to reach thermal equilibrium. Also, it helps to produce the same starting condition each time. The evolution period t_1 is gradually increased. After each t_1, the magnetization is detected in the form of a FID during the detection period t_2. As a result, a series of FIDs are obtained. Fourier transformation of the t_2 dimension yields a set of 1D spectra with the varying intensities of the lines due to the changes in the t_1 duration. A desired 2D spectrum is possible to obtain with a subsequent Fourier transformation of t_1 dimension.

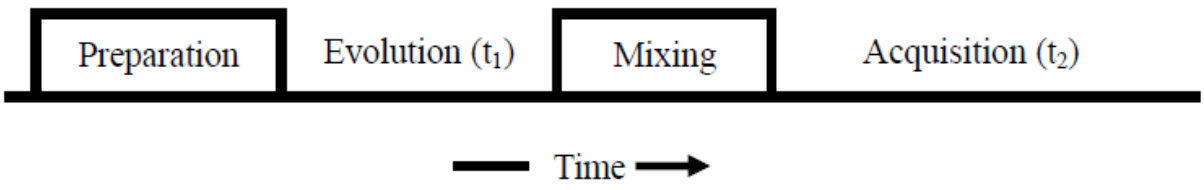


Figure 1.5: General scheme for a 2D experiment. Adapted from reference (*3*).

Some of the very important 2D experiments are discussed below.

1.1.6.1.1 COSY

In this experiment, the magnetization is transferred between protons that are chemically bonded (up to 3 bonds) on adjacent nuclei (**Figure 1.6A**). Thus, it provides the information on the protons that are ^{3}J-coupled. This is one of the first and simplest multi-dimensional experiments (*4*).

1.1.6.1.2 TOCSY

In this experiment, information on all the protons attached to nuclei within a given spin system (**Figure 1.6 B**) is obtained. This includes protons that are beyond ^{3}J chemical bonds. In this experiment, following the evolution period, during the mixing period, the spin is locked in the transverse plane for some time. Scalar coupling results in the transfer of coherence during this mixing period.

1.1.6.1.3 HSQC

In NMR, proton is more sensitive (has higher gyromagnetic ratio) than any other heteronuclei. To get a good signal of the heteronuclei, an HSQC (Heteronuclear Single-Quantum Coherence) (*5*) experiment utilizes the INEPT (Insensitive Nuclei Enhancement by Polarization Transfer) sequence to transfer the magnetization of the proton to its bonded heteronuclei (^{13}C or ^{15}N). This is then transferred back to the magnetization of the proton by a second INEPT sequence for the detection. An HSQC spectrum has two axes; one is for the proton chemical shift and another one is for the heteronuclear chemical shift.

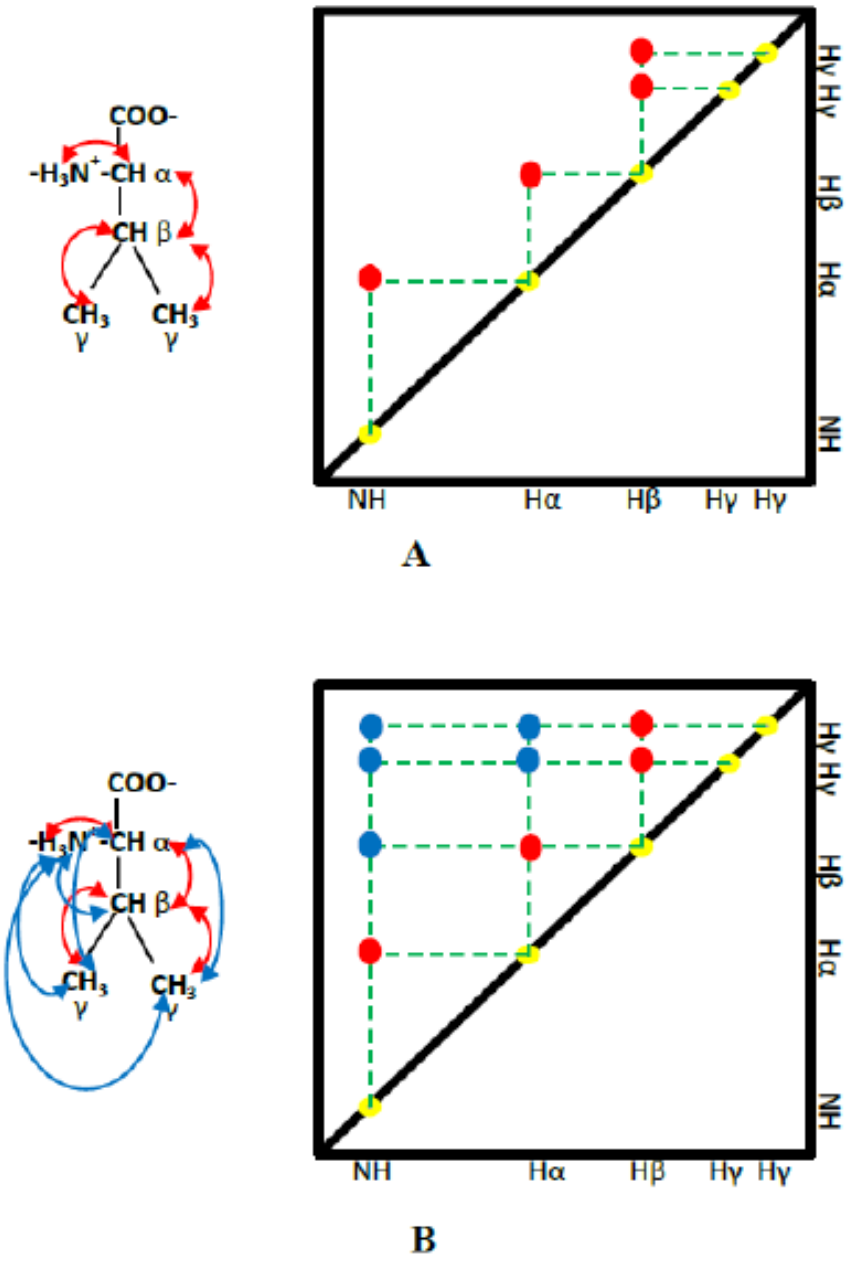


Figure 1.6: A) COSY and B) TOCSY spectra. Adapted from reference (*6*).

1.1.6.1.4 NOESY

In this experiment, dipolar coupling is involved between two spatially close (nearer than 5 angstrom) nuclei. Magnetization is transferred through the J-coupling during the mixing time. For structure calculations, NOESY is one of the most useful information since it connects the nuclei through space. The distance information comes as a function of the intensity of the peaks.

1.1.6.2 Three-dimensional NMR (3D NMR)

For the determination of the structure of small proteins, 2D experiments have been used quite successfully over time (*7*). However, with the increase in the size (more than 100 residues) of the proteins, 2D experiments alone were not enough anymore to get the structure. There are two basic reasons for this limitation of 2D experiments:

i. For the larger protein, due to the large volume of information for the high number of residues of the protein, only the space of two-dimension becomes insufficient. As a result, too much overlap of the peaks within the spectrum makes it impossible to interpret the data.
ii. As the size of the protein increases, the rotational correlation time increases. This results in slower movement of the protein in the solution leading to the broadening of the line-width of the resonance which can become larger than the J-coupling constant (*7*).

In order to improve the chance of determination of the structure of larger proteins using NMR, the dimension needs to be increased to get rid of the overlap and also heteronuclear coupling is essentially utilized to make use of the scalar coupling which is larger than the line-

widths. That is why, larger proteins are routinely over-expressed by growing heterologous expression systems in minimal media containing ^{15}N or ^{13}C- labeled component as their sole source of nitrogen or carbon (*7*). 3D NMR, principally, can be easily constructed by combining two sets of 2D NMR experiments (**Figure 1.7**). As illustrated in **Figure 1.7**, by removing the detection period of the first set of 2D experiment and preparation period for the second, a 3D experiment combining two evolution periods (t_1 and t_2), independent of each other, is originated.

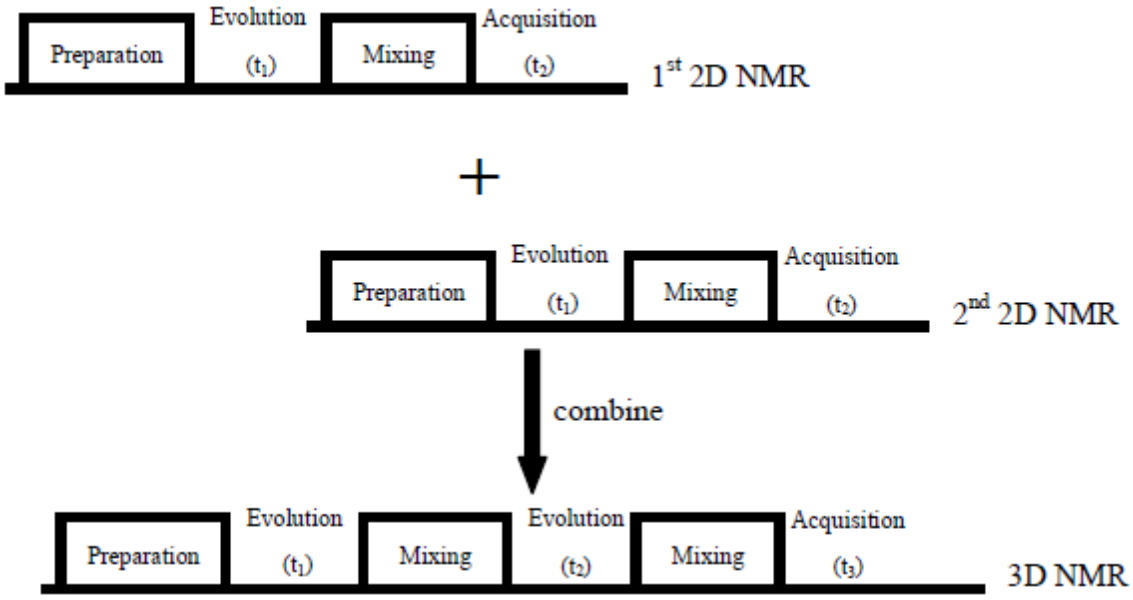


Figure 1.7: 3D experiment as a combination of two sets of 2D experiments. Adapted from reference (*3*).

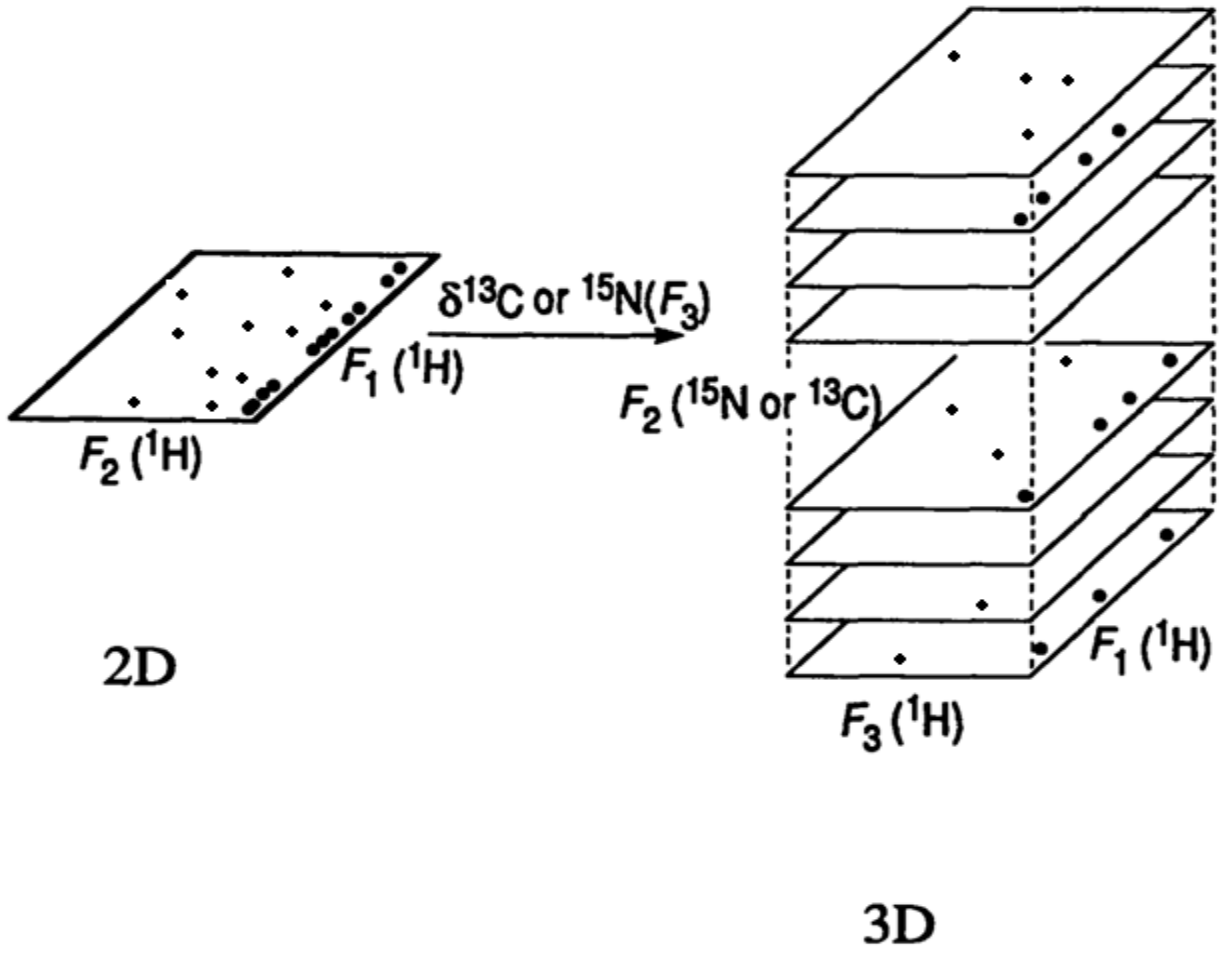


Figure 1.8: Schematic presentation of how with the addition of another evolution time to the 2D experiment (**Figure 1.7**) can result in another frequency dimension for a 3D experiment. The black dots representing NOE cross peaks in 2D spectrum are hard to attribute to the correct proton destination in the F_2 dimension. But, in the 3D spectrum, the expansion of another frequency dimension arising from the heteronucleus allows the determination of NOEs involving protons that lie in three different planes. Now, in the 3D spectrum, each plane corresponds to the specific chemical shift of the heteronucleus whereby NOE peaks arising from the interaction between protons attached to the heteronucleus and the other protons on F_3 dimension can be detected. Some of the peaks cannot be seen because of the overlapped planes in the presentation used in this figure. Adapted from reference (*8*).

1.1.7 Protein NMR

There are two principal methods to determine the structure of a protein; one is X-ray crystallography and another one is NMR. Both of these techniques have their own pros and cons. The focus on NMR as an alternative tool to X-ray crystallography to determine the structure of proteins has grown over the years for several reasons. Not all proteins can be crystallized and even, if it is crystallized, it might not produce good enough diffraction data to get the structure. Also, with proteins in crystals one could be missing some important dynamic information that the proteins in solution might possess, something that can be detected by NMR. However, NMR has an intrinsic disadvantage of larger line-width attributed to longer tumbling time with increasing size of the protein. Also, it needs a very high concentration of protein as a sample (~300-600 μL protein of 0.1-3mM) and concentrated protein tends to aggregate.

To determine a 3D structure of a protein by NMR, the first step is to assign the back-bone of the protein. Various heteronuclear 3D experiments are employed for this purpose. Among these, the most common ones are HNCA (*9*), HN(CO)CA (*9, 10*), HNCACB (*11*), CBCA(CO)NH (*11*), HNCO (*9*), and HN(CA)CO (*12*). All these experiments are composed of a 2D HSQC plane of ^{15}N and ^{1}H in X and Y axes while in the Z-dimension ^{13}C chemical shifts are placed. In the HNCA experiment, the amide proton is correlated with the Cα atom of its own residue (residue i) and of the residue preceding it (residue i-1). On the other hand, the HN(CO)CA experiment allows the correlation between the amide proton (residue i) and Cα atom of its preceding residue (residue i-1). Assignments from these two experiments can be accomplished in parallel to match the chemical shifts. Similarly, HNCO and HN(CA)CO spectra can be examined together to determine the correlation between amide proton and carbonyl

carbons and HNCACB and CBCA(CO)NH spectra for both Cα and Cβ atoms‟ correlation with amide protons. All these experiments are simultaneously assigned to correctly obtain the chemical shifts of each possible nucleus without any ambiguity (**Table 1.2**).

Once the backbone assignment is done, the next step will be to assign the side chains of the protein using various 3D heteronuclear experiments such as ^{15}N-edited HSQC-TOCSY (*13, 14*) and HCC(CO)NH (*9*).

The basic principle of any NMR structure determination is to assign a specific resonance to each proton and then to identify the NOE interactions between a pair of protons. A number of experiments are used to determine these NOE interactions such as ^{15}N/^{13}C-edited HSQC-NOESY (*13-15*). Initially, the sequential and short-range inter-residual NOEs are assigned as they are comparatively easier to assign. Then, the long-range NOEs are dealt with which are much harder to assign but are the most important for determining the global fold of the protein structure.

The structure calculation of the protein can then be initiated by using the obtained NOE restraints and dihedral angle restraints which are entered into computer programs like CYANA (*16*) or XPLOR-NIH (*17, 18*). For soluble proteins, the energy function of the structure is further lowered by using a water refinement module in ARIA (*19*). The output of these computational calculations consists of a set of best structures from all the probable calculated structures characterized by good convergence of the well defined parts of the protein. A good set of

structures should have low RMSD, a low energy function and few angle and distance violations. A software program named PROCHECK (*20*) is used to assess these attributes of the structure.

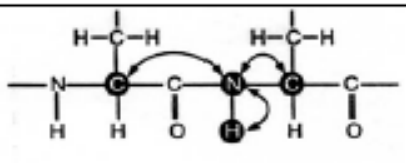
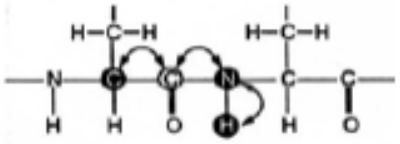
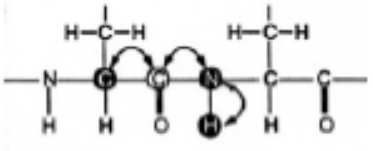
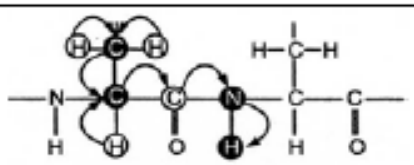
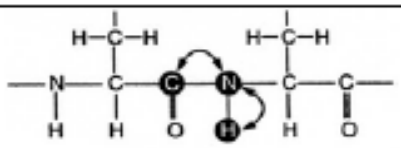
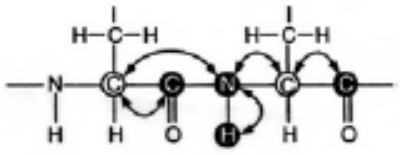
Experiment	Correlation	Magnetization Transfer
HNCA	$H^N(i)$, $^HN(i)$, Cα(i) and Cα(i-1)	
HN(CO)CA	$H^N(i)$, $^HN(i)$, Cα(i-1)	
CBCANH (HNCACB)	Cβ(i-1), Cα(i-1), Cβ(i), Cα(i), $H^N(i)$, $^HN(i)$	
CBCA(CO)NH	Cβ(i-1), Cα(i-1), $H^N(i)$, $^HN(i)$	
HNCO	CO(i-1), $H^N(i)$, $^HN(i)$	
HN(CA)CO	CO(i), $H^N(i)$, $^HN(i)$	

Table 1.2: Correlation observed for some of the most commonly used 3D NMR experiments. Adapted from reference (*6*).

1.1.8 Study of protein-ligand interaction by NMR

The study of the interaction between proteins and other molecules (such as DNA, RNA, sugars, or even another protein/peptide) in solution has become more common (*21*). Such studies hold key to understanding various biological processes, for example, the interaction between enzyme and its substrate/inhibitor or the binding of various transcription factors to DNA. NMR is a very useful and powerful tool to investigate such interactions.

One important aspect of the NMR study of protein-ligand interactions is to determine the effect of chemical exchange on NMR spectra, that is, to determine whether the bound and free form of the protein coexist in the fast or slow exchange regime on the NMR time scale. In fast exchange, a single average resonance peak is observed, whereas in slow exchange, two different resonance peaks are observed for a single nucleus (**Figure 1.8**). For intermediate exchange the two resonance peaks will appear to coalesce together into one, and if there is no exchange, then there would hardly be any line broadening (**Figure 1.9**). As previously discussed, line-width of the resonance peaks is inversely proportional to spin-spin relaxation (T_2). This phenomenon can easily be correlated with the strength of the interaction for the protein-complex. The faster the chemical exchange indicates, the looser the interaction between protein and its interacting partner. ^{15}N-HSQC experiment is routinely used to observe the effect of chemical exchange on the protein for its specific residues. These observations can sometimes even lead to a basic idea on what part of the protein is actually involved in binding. Also, the determination of dissociation constant (K_D) values for that specific interaction is possible through these experiments. However, though, such a method for determination of K_D values is not very accurate (*21*).

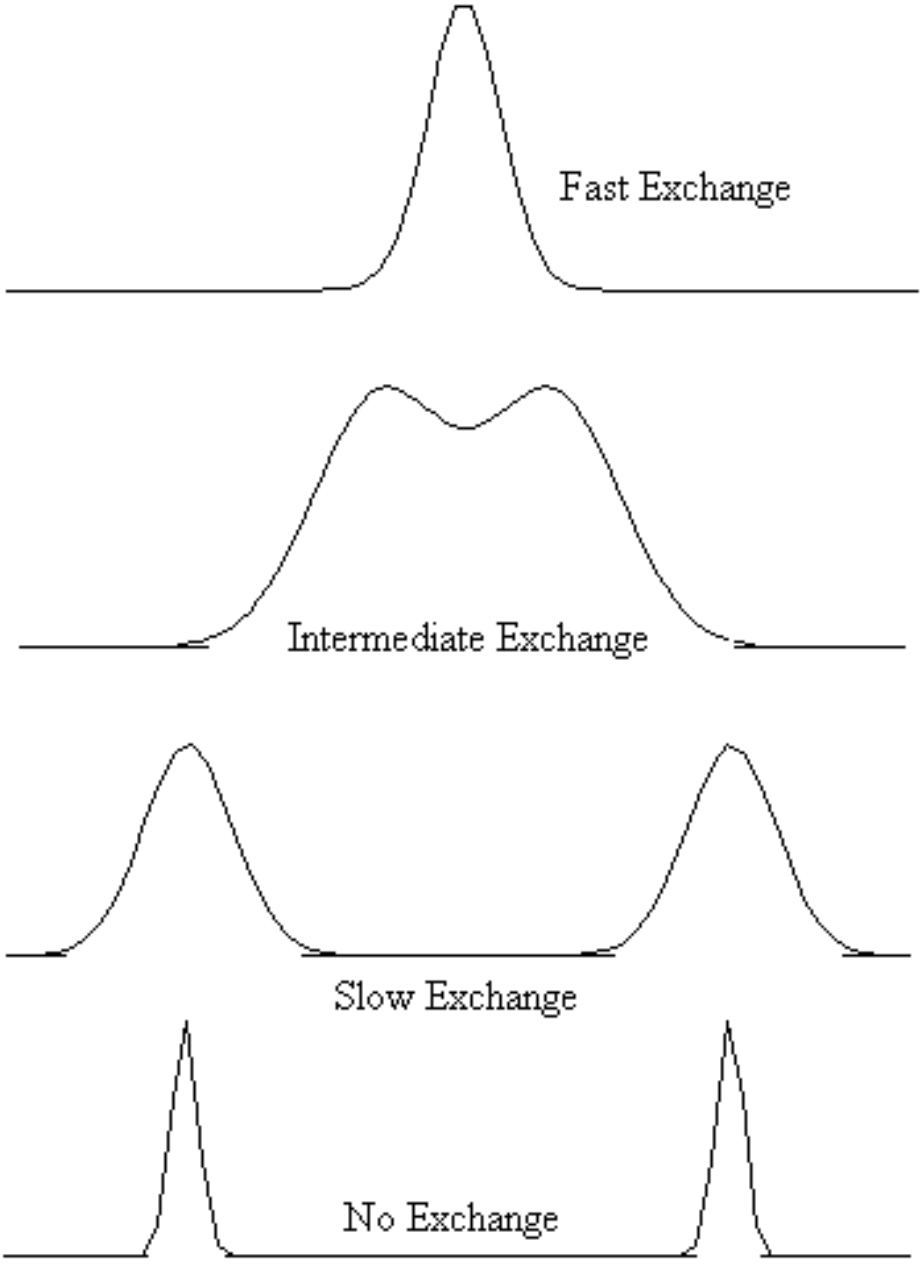


Figure 1.9: Effect of chemical exchange on NMR spectra. Drawn according to http://web.nmsu.edu/~snsm/classes/chem435/Lab8/.

To determine the structure of the protein-interacting partner complex, it is important to assign the resonances of protons for both the protein and interacting partner individually. For the protein, as for its free form, the same different types of 3D heteronuclear experiments are done to obtain the resonances of the protons of the protein in its complex form. If the interacting partner is a protein and isotopically labeled, then the same experimental procedures can be followed for the assignment of protons of the partner as well. However, if the interacting partner is not isotopically labeled (such as unlabeled peptide), then along with other conventional experiments, a unique experiment, known as filtered NOESY experiment, is used. This experiment can be designed in such a way that any resonance that arises from labeled nuclei should be eliminated. Thus, only those resonances that originate from the unlabeled peptide are detected and assigned. In this way, successful assignment of the nuclei (mainly protons) present in the unlabeled peptide is possible. Additionally, filtered NOESY experiments can also be employed to determine the intermolecular NOEs between the protein and the peptide. This allows building up NOEs necessary to dock the peptide onto the protein in NMR calculation.

1.2 PDZ domain and Glutaminase Interacting Protein (GIP)

To maintain an efficient and active cellular physiology, it is important for the cells to maintain effective signaling systems and protein-protein interaction is at the root of these signaling systems. There are several motifs/domains/modules that are involved in such interactions. PDZ domains (Post Synaptic Density 95 (PSD-95), Discs Large (Dlg) and Zonula Occludentes (ZO-1)) (*22-24*) are one of the most ubiquitous and well known domains involved in protein-protein interactions. The PDZ domain is widespread in the nature. It is involved in multiple processes and possibly numerous others are yet to be discovered. Protein scaffolding (*25, 26*), maintaining cell polarities (*27*), localizing and clustering of ion-channels are to name only a few of myriad of the processes it plays a role in. It is an 80-100 amino acid long motif. Usually, PDZ-domain containing proteins, having more than one PDZ domains, are involved in the formation of multimeric protein complexes (**Figure 1.10**). This domain is primarily found in eukaryotic organisms (*28*), but can be found in a slightly different form in prokaryotes and plants as well (*29-31*). Based on its specificity toward the sequence of its binding partner, PDZ domain can be classified into three major broad classes:

a. Class I PDZ domain (binding motif S/T-X-Φ-COOH, where Φ is a hydrophobic amino acid and X is any amino acid)
b. Class II PDZ domain (binding motif Φ-X-Φ-COOH)
c. Class III PDZ domain (binding motif X-X-C-COOH)

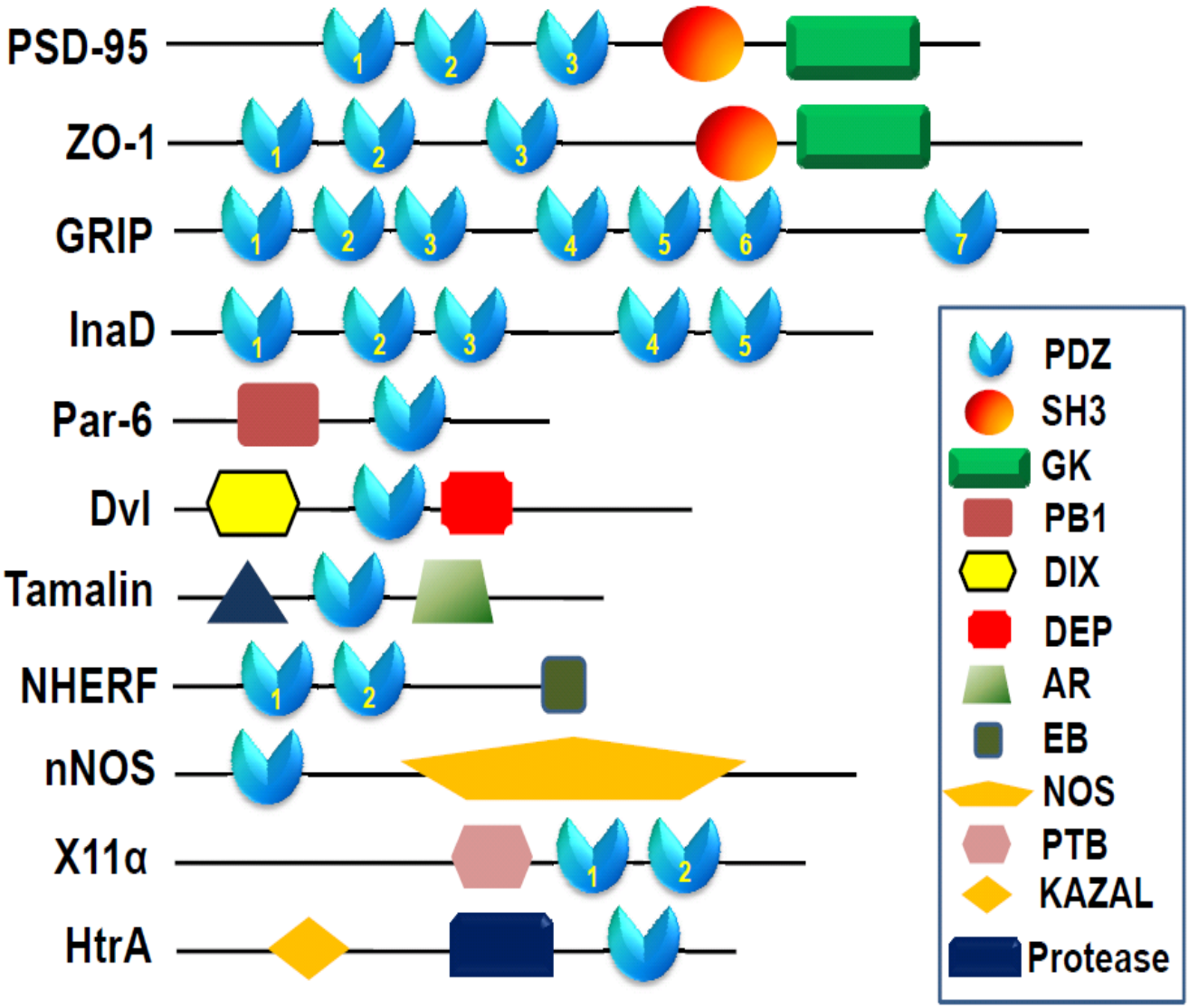


Figure 1.10: Examples of PDZ domain containing proteins. Adapted from reference (*28*).

However, some PDZ domains cannot be categorized into any of the above classes (**Table 1.3**) (*32*).

PDZ domains are involved in various cancer pathways (*33-40*). G-protein coupled receptors and ion channels are very important candidates for drug development. It was found that GIP interacts with these proteins for its proper function. Thus, developing a drug molecule that will compete with GIP for binding with these targets could prove promising (*33*). Because of its diverse functions and implications in various diseases including cancer, it is very important to investigate the interacting partners of PDZ domains and gain structural insight into the mode of binding of PDZ domains with their partner protein, which are critical for the development of drug candidates (*41*).

1.3 Glutaminase Interacting Protein (GIP)/Tax Interacting Protein-1 (TIP-1)

GIP, also known as Tax Interacting Protein 1 (TIP-1) is a small PDZ domain containing protein. This protein is 124 amino acid long with a molecular mass of 13.7 kDa. GIP contains a single PDZ domain which is unique among PDZ containing proteins. GIP is also an excellent protein for structure-function studies by solution NMR, since, it is a small globular protein having good solubility properties and stability (NMR sample can be stored even up to several months without any aggregation). Additionally, research methods have been developed in our laboratory to over-express and purify the recombinant protein in milligram quantities in a single step (*42*). Additionally, GIP is implicated in many cancer pathways due to its interactions with a growing list of partner proteins all with different roles in the cell. The role of GIP in many of

these processes is not yet understood at the molecular level. Thus, to understand the functions of GIP, it is important to characterize the interaction between GIP with different binding partner proteins to gain an insight into its mechanism of interaction and mode of recognition.

In this dissertation work, we have solved the solution structures of GIP both in the free state and also bound to a substrate, the C-terminal octa peptide of Glutaminase L (KENLESMV-COOH) using solution NMR. This is the first NMR structure of a complex of GIP. With this structural information, essential knowledge can be obtained on the mechanism of interactions and mode of recognition between GIP and those interacting partners that contain C-terminal recognition motif. This knowledge will be essential for structure-based drug design with either GIP as target or its partner proteins. Further, we also characterized the interaction between GIP and a peptide mimic of a human Brain-specific Angiogenesis Inhibitor 2 (BAI2) using various biophysical techniques. Discovery of the complete network of interacting partners for GIP is necessary to comprehend fully the function of GIP in the human brain and other parts of the body.

PDZ Domain	Consensus binding sequence	Ligand protein
Class I	P_{-3} -P_{-2} - P_{-1} -P_0 S/T-X-φ-COOH	
Syntrophin (*43*)	E-S-T-V-COOH	Voltage-gated Na+ channel
PSD-95 (*26*)	E-T-D-V-COOH	Shaker-type K+ channel
GIP (*44*)	E-S-M-V-COOH	Glutaminase-L
Class II	φ-X-φ-COOH	
hCASK (*45*)	E-Y-Y-V-COOH	Neurexin
Erythrocyte p55 (*46*)	E-Y-F-I-COOH	Glycophorin C
Class III	X-X-C-COOH	
Mint-1 (*47*)	D-H-W-C-COOH	N-type Ca +2 Channel
SITAC (*48*)	Y-X-C-COOH	L6 antigen
Other		
nNOS (*49*)	G-D-X-V-COOH	MelR
MAGI PDZ2 (*50*)	S/T-W-V-COOH	Phage display
Engineered from SF6 (*51*)	K/R-Y-V-COOH	Synthesized peptide

Table 1.3: Classification of PDZ domains. Adapted from reference (*32*).

1.4 References

1. Rahman, A. U. (1986) *Nuclear Magnetic Resonance: Basic Principles*, Springer-Verlag, New York.
2. Prestegard, J. H., Bougault, C. M., and Kishore, A. I. (2004) Residual dipolar couplings in structure determination of biomolecules, *Chemical reviews 104*, 3519-3540.
3. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J., and Rance, M. (2006) *Protein NMR spectroscopy: principles and practice.*, second ed., Academic Press.
4. Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976) Two dimensional spectroscopy. Application to nuclear magnetic resonance., *J. Chem. Phys. 64*, 2229-2246.
5. Bodenhausen, G., and Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chemical Physics Letters 69*, 185-189.
6. Huang, C. (2010) Production, characterization and structure determination of the C-terminal domain of Stt3p: the catalytic subunit of yeast oligosaccharyl transferase, In *Department of Chemistry and Biochemistry*, p 253, Auburn University, USA.
7. Clore, G. M., and Gronenborn, A. M. (1993) *NMR of proteins*, CRC Press Inc., USA.
8. Clore, G. M., and Gronenborn, A. M. (1991) Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy, *Science 252*, 1390-1399.
9. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *Journal of the American Chemical Society 114*, 6291-6293.
10. Bax, A., and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins, *Journal of biomolecular NMR 1*, 99-104.

11. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity, *Journal of Magnetic Resonance, Series B 103*, 203-216.

12. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C′ chemical shifts in 15N/13C-labelled proteins, *Journal of Magnetic Resonance (1969) 97*, 213-217.

13. Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., and Campbell, I. D. (1990) Comparison of techniques for 1H-detected heteronuclear 1H/15N Spectroscopy, *Journal of Magnetic Resonance (1969) 87*, 488-501.

14. Palmer Iii, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *Journal of Magnetic Resonance (1969) 93*, 151-170.

15. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J Biomol NMR 4*, 845-858.

16. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol Biol 278*, 353-378.

17. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package, *J Magn Reson 160*, 65-73.

18. Schwieters, C. D., Kuszewski, J. J., and Marius Clore, G. (2006) Using Xplor–NIH for NMR molecular structure determination, *Progress in Nuclear Magnetic Resonance Spectroscopy 48*, 47-62.

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

20. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J Biomol NMR 8*, 477-486.

21. Roberts, G. C. K. (1993) *NMR of macromolecules : a practical approach*, IRL Press at Oxford University Press, Oxford ; New York.

22. Ponting, C. P., and Phillips, C. (1995) DHR domains in syntrophins, neuronal NO synthases and other intracellular proteins, *Trends in biochemical sciences 20*, 102-103.

23. Kennedy, M. B. (1995) Origin of PDZ (DHR, GLGF) domains, *Trends in biochemical sciences 20*, 350.

24. Sheng, M., and Sala, C. (2001) PDZ domains and the organization of supramolecular complexes, *Annual review of neuroscience 24*, 1-29.

25. 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.

26. 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.

27. Bhat, M. A., Izaddoost, 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*, 833-845.

28. Lee, H. J., and Zheng, J. J. (2010) PDZ domains and their binding partners: structure, specificity, and modification, *Cell communication and signaling : CCS 8*, 8.

29. Pallen, M. J., and Ponting, C. P. (1997) PDZ domains in bacterial proteins, *Molecular microbiology 26*, 411-413.

30. Ponting, C. P. (1997) Evidence for PDZ domains in bacteria, yeast, and plants, *Protein science : a publication of the Protein Society 6*, 464-468.

31. 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, *Nature structural biology 7*, 749-753.

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

33. Wang, N. X., Lee, H. J., and Zheng, J. J. (2008) Therapeutic use of PDZ protein-protein interaction antagonism, *Drug news & perspectives 21*, 137-141.

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

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

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

37. Dev, K. K. (2004) Making protein interactions druggable: targeting PDZ domains, *Nature reviews. Drug discovery 3*, 1047-1056.

38. 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., and Madden, D. R. (2007) Targeting CAL as a negative regulator of DeltaF508-CFTR cell-surface expression: an RNA interference and structure-based mutagenetic approach, *The Journal of biological chemistry 282*, 8099-8109.

39. Georgescu, M. M., Morales, F. C., Molina, J. R., and Hayashi, Y. (2008) Roles of NHERF1/EBP50 in cancer, *Current molecular medicine 8*, 459-468.

40. Georgescu, M. M. (2008) NHERF1: molecular brake on the PI3K pathway in breast cancer, *Breast cancer research : BCR 10*, 106.

41. Banerjee, M. (2011) Human Glutaminase Interacting Protein (GIP): a Potential Candidate for Anti-Cancer Drug Design, In *Department of Chemistry and Biochemistry*, p 172, Auburn University, USA.

42. 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-Us 47*, 9208-9219.

43. 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, *Nature structural biology 5*, 19-24.

44. 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.

45. 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.

46. Marfatia, S. M., Byron, O., Campbell, G., Liu, S. C., and Chishti, A. H. (2000) Human homologue of the Drosophila discs large tumor suppressor protein forms an oligomer in solution. Identification of the self-association site, *The Journal of biological chemistry 275*, 13759-13770.

47. Maximov, A., Sudhof, T. C., and Bezprozvanny, I. (1999) Association of neuronal calcium channels with modular adaptor proteins, *The Journal of biological chemistry 274*, 24453-24456.

48. Borrell-Pages, M., Fernandez-Larrea, J., Borroto, A., Rojo, F., Baselga, J., and Arribas, J. (2000) The carboxy-terminal cysteine of the tetraspanin L6 antigen is required for its interaction with SITAC, a novel PDZ protein, *Molecular biology of the cell 11*, 4217-4225.

49. Stricker, N. L., Christopherson, K. S., Yi, B. A., Schatz, P. J., Raab, R. W., Dawes, G., Bassett, D. E., Jr., Bredt, D. S., and Li, M. (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences, *Nature biotechnology 15*, 336-342.

50. 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, *The Journal of biological chemistry 275*, 21486-21491.

51. Schneider, S., Buchert, M., Georgiev, O., Catimel, B., Halford, M., Stacker, S. A., Baechi, T., Moelling, K., and Hovens, C. M. (1999) Mutagenesis and selection of PDZ domains that bind new protein targets, *Nature biotechnology 17*, 170-175.

Chapter 2

Characterization of Glutaminase Interacting Protein (GIP): a PDZ domain

2.1 Background

2.1.1 Protein-protein interaction network

One of the challenging tasks to understand cells and diseases is to know how within the cells a network of proteins is connected. This knowledge of protein networks will help to shed light on the inner machinery of the cells. It would also allow scientists to specifically target a protein within that network to treat a disease and, thus, help narrowing down the potential targets to combat diseases. Proteins can interact with other proteins, metabolites and DNA or RNA in a cell. Several experimental tools have been employed to determine the protein-protein interaction either as a direct approach such as yeast two-hybrid screening, mass spectrometry (MS) and immunoprecipitation or on a genome-wide level such as chromatin immunoprecipitation (ChIP) on chip assays and double knockout assays in yeast (**Figure 2.1**) (*1*).

However, constructing a comprehensive protein-protein interaction network is well beyond the scope of this thesis. Besides establishing an interaction network between different proteins, it is also very important to investigate the mechanisms by which the proteins interact with each other. With an insight into the structure, binding mechanisms and mode of interactions between different proteins, it is possible to design the most effective and selective drug

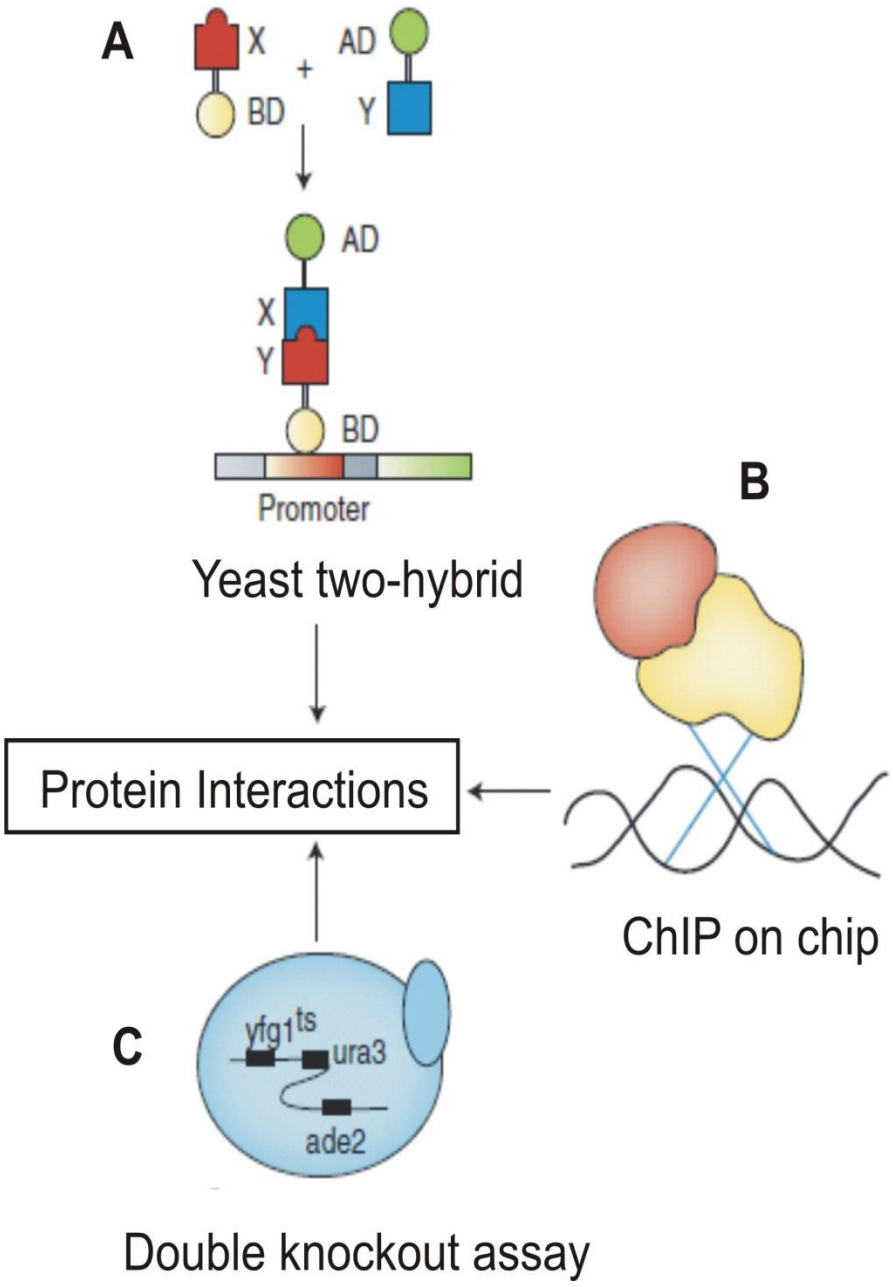


Figure 2.1: Different experiment tools to define protein interactions. A. Yeast two-hybrid screening B. ChIP on chip assay C. Double knockout assay in yeast. Adapted from reference (*1*).

compounds that may have an expected therapeutic effect as a result of their interactions with the targeted proteins.

These protein-protein interaction networks are important for maintaining a continuous and ordered communication within cells. The interactions between proteins are facilitated via a number of different interacting domains such as SH2 (Src Homology 2) (*2*), SH3 (Src Homology 3) (*3*), PH (Pleckstrin-homology) (*4*), PDZ (Post synaptic density 95, Discs large and Zonula occludentes) (*5, 6*) and others (*7*) which can be present either as a single domain or multiple domains in a single protein.

2.1.2 PDZ domain and its classes

In nature, there are many protein-protein interaction modules present. One of the most important of these interaction modules is the PDZ domain (*8*). These domains are small and contain 80-100 amino acid residues. 1-2 α-helices and 5-6 β-strands comprise these domains. Animals contain many PDZ domain/s containing proteins. However, in yeast and plants, "PDZ-like" domains that are structurally similar, but not exactly same, have been found (*9, 10*). PDZ-like domain consists of 5 β-strands (β1-β5) capped by 2 α-helices (α2 and α3) and also two short β-strands at the N and C termini (βN and βC). Also, a well-defined α-helix (α1) is formed between the β1 and β2 loop (*11*). PDZ domains are involved in various important cellular functions, including signaling pathways and acting as scaffolds to organize multimeric complexes often with the help of other protein-protein interaction modules (*7*). PDZ domains usually recognize the unstructured C-terminal end of their interacting partner proteins (*12*). But,

in rare cases, proteins with internal motifs that structurally mimic the C-terminus can bind to PDZ domains (*13, 14*). PDZ domains can be categorized mainly into three classes according to the sequence specificity of their binding partners (*15*). They are class I **(**X-S/T-X-Φ-COOH) (*16*), class II (X-Φ-X-Φ-COOH) (*7*), class III (X-E/D-X-Φ-COOH) (*17*) and, also, various other minor classes (*18*) where Φ is any hydrophobic residue and X is any residue (*19*).

2.1.3 Glutaminase interacting protein as a class I PDZ domain

Glutaminase Interacting Protein (GIP) is a PDZ domain containing protein that has a number of important functions (*20*). It is also known as Tax Interacting Protein-1 (TIP-1) (*21*). GIP is a very small protein containing only 124 amino acid residue. Also, it is unique among PDZ containing proteins since the whole protein is composed solely of a single PDZ domain without any other additional domain/s. All other PDZ domain containing proteins usually have either more than one PDZ domains and/or contain other domains such as SH2, SH3 etc. (*7*). Over the last several years, there has been an increasing number interacting partner proteins reported for GIP including Glutaminase L (*20*), β-Catenin (*22, 23*), Fas (*24, 25*), HTLV Tax (Human T-lymphotropic virus Tax) (*21*), HPV E6 (Human papillomavirus E6) (*26*), Rhotekin (*27*) and Kir 2.3 (*28, 29*). All these interacting partners contain the PDZ class I (X-S/T-X-I/L/V-COOH) binding motif. To get an insight into the mechanism of GIP‟s recognition and mode of interaction with such a wide range of proteins, it is critical to investigate these binding events to understand the molecular basis of the functions that these proteins carry out in the cells. For example, β-Catenin and Rhotekin are important in the Wnt and Rho signaling pathways, respectively. Fas is a member of the Tumor Necrosis Factor (TNF) family of receptors, while

HTLV Tax and HPV E6 are both viral proteins from oncogenic viruses. Lastly, GIP regulates the inward rectifier potassium channel Kir 2.3 in renal epithelial cells. GIP has been shown to be involved in a variety of different cancer and cell signaling pathways with its numerous binding-partner proteins. Also, GIP is involved in the regulation of Glutaminase L, which has been shown to be up-regulated in various cancers (*30-32*). By doing sequence alignment of all of these discovered interacting partners; it is possible to identify the optimal consensus sequence for GIP binding as to be E-S-X-V-COOH (**Table 2.1**) (*19*).

Binding Partner	Position							
	-7	-6	-5	-4	-3	-2	-1	0
Glutaminase L	K	E	N	L	E	S	M	V
Kir 2.3	S	Y	R	R	E	S	A	I
HTLV Tax	K	H	F	R	E	T	E	V
HPV E6	R	Q	A	T	E	S	T	V
Rhotekin	R	T	W	L	Q	S	P	V
β-Catenin	L	A	W	F	D	T	D	L
FAS	R	N	E	I	Q	S	L	V
Consensus	X	X	X	X	E	S	X	V

Table 2.1: Sequential alignment of C-terminal binding partners of GIP. Adapted from reference (*19*).

2.1.4 Objective of the study

As a first step towards understanding the mechanism of recognition and mode of interaction of GIP with its various partner proteins, it is imperative to solve the high resolution structure of GIP at atomic level. The atomic structures of proteins both in the free-state and bound with their substrate provide snapshots of many complex features of the biological event including residues involved in the binding, site of interaction etc. Solution-state NMR enables us to investigate the protein under biological condition and also allows examining the dynamics of these processes in the timescale of picoseconds to seconds. In this chapter, the NMR experiments and analysis method are described that were used to determine the atomic structure and the dynamics of free GIP in solution. We determined the NMR structure of free GIP in solution with a backbone RMSD of 0.45 Å. We also investigated the dynamics of the free GIP. Comparison of this structural and dynamic information of free GIP with those of GIP bound with a surrogate peptide that mimics the C-terminus of Glutaminase L (Chapter 3 of this dissertation paper) yielded insight into the mechanism of interaction of GIP with its binding partners.

2.2 Materials and Methods

The research work described here was carried out in the laboratory of Dr. Smita Mohanty.

2.2.1 Cloning, over-expression and purification of ^{15}N, ^{13}C-labeled GIP

According to the protocol developed in Dr. Smita Mohanty's laboratory, the double-labeled free GIP protein was prepared by Dr. Smita Mohanty and other group members (23).

2.2.2 NMR Data collection

All NMR data were collected on a Bruker Avance 600 MHz spectrometer with a triple resonance $^{1}H/^{13}C/^{15}N$ TCI cryoprobe equipped with z-axis pulsed field gradients at either the Department of Chemistry and Biochemistry, Auburn University, Auburn, AL, or the New York Structural Biology Center (NYSBC), New York, NY. The data were processed using NMRPipe (*33*) and analyzed using Sparky (*34*). For structure determination of free GIP, samples between 500 μM and 1 mM of uniformly $^{15}N/^{13}C$-labeled GIP in 50 mM phosphate buffer containing 5% D_2O pH 6.5, 1 mM EDTA and 0.01% (w/v) NaN_3 were prepared. All NMR experiments were performed at 298 K. To determine the ^{15}N T_1 values, NMR spectra were recorded with relaxation delays of 10, 600, 50, 500, 100, 400, 200, 300 and 10 ms. To determine ^{15}N T_2 values, NMR spectra were recorded with delays of 17, 153, 34, 17, 136, 51, 119, 68, 102, 85 and 34 ms. The relaxation times were randomized and some points repeated in order to avoid any systematic errors that may arise when the data are collected sequentially. The relaxation rates were

calculated by least squares fitting of peak heights versus relaxation delay to a single exponential decay. Steady state ^{1}H-^{15}N NOE values were calculated from the ratio of peak heights in a pair of NMR spectra acquired with and without proton saturation. These dynamics data were analyzed in collaboration with Dr. David Zoetewey in Dr. Smita Mohanty‟s research group. For backbone and side-chain assignments of free GIP, the following NMR experiments were recorded at 298 K: 2D ^{1}H,^{15}N-HSQC (*35*), 3D HNCACB (*36*), 3D CC(CO)NH (*37*), 3D CBCA(CO)NH (*36*), 3D ^{15}N-edited HSQC-TOCSY (*38, 39*) with an 80 ms mixing time, 3D HC(CO)NH (*37*), 3D HNHA (*40*), 3D HNCO (*37*) and 3D HN(CA)CO (*41*) at NYSBC by Dr. Smita Mohanty. NOE distance restraints were obtained from 3D ^{15}N-edited HSQC-NOESY (*38, 39, 42*) and 3D ^{13}C-edited HSQC-NOESY (*38, 39, 42*) spectra collected both at NYSBC and also again at AU with the ^{13}C carrier frequency in the aliphatic (44 ppm) and aromatic (125 ppm) regions and mixing times of 140 for ^{15}N and 110 ms for ^{13}C, respectively (*19*).

2.2.3 Analysis of dynamics data

Measured relaxation parameters R_1, R_2 and the steady-state ^{1}H-^{15}N NOE for each residue were used as inputs in the Modelfree 4.15 program developed by Palmer et al (*43, 44*) to analyze ^{15}N-backbone dynamics. The τ_c value for free GIP was calculated using the program Tensor2 for the core region A11-Q112 (*45, 46*). Of five different models, the best one was chosen according to the selection criteria (*43*) to get the order parameter (S^2) that represents the degree of spatial restriction within the ^{1}H-^{15}N bond vector. These values range from zero for completely isotropic internal motions to unity for totally restricted motion and represent dynamics in the picosecond to nanosecond time scale (*19*).

2.2.4 Structure calculation and refinement

A total of 4303 cross peaks were assigned manually using Sparky (*34*) for free GIP. The assignments were corrected or confirmed with both the CANDID module of CYANA 1.0.6 and NOEASSIGN module of CYANA 2.1 (*47*), using the standard protocol of eight iterative cycles of NOE assignment and structure calculation. A total of 118 dihedral angles restrains were derived from the TALOS (*48*) program based on the chemical shift index (CSI) and primary sequence of GIP for free protein calculations. Additionally, a total of 64 hydrogen bond distance restraints (two restraints per bond) for the free protein were derived from the CSI by TALOS. During the iterative NOE assignments, a total of 1134 assignments for free GIP were removed due to overlap, redundancy, or unresolved ambiguity that resulted from low stringency in the initial peak picking phase and high stringency in the final assignments. The final assignments averaged over 25 NOEs per residue for free protein. Final refinement of the 100 lowest energy structures of the 200 total calculated structures was performed with the water refinement protocol implemented in ARIA (*49*). The 20 structures with the lowest potential energy and best Ramachandran statistics as assessed by PROCHECK (*50*) were selected for analysis. The structures were visualized with VMD and figures were created using Pymol (*51, 52*). (**Table 2.3** shows the complete structural statistics for structure of GIP alone (*19*).)

2.3 Results

2.3.1 NMR Structure determination of free GIP

2.3.1.1 Introduction

De Novo structure determination of small (MW <25 kDa) water soluble molecules such as proteins by NMR spectroscopy is very useful to understand the mechanisms of function of the protein under study. Several steps need to be followed to determine the structure of a protein by NMR. These steps can be summarized in a flowchart (**Figure 2.2**). Up to this point, the first two steps of the flowchart have been discussed both in the materials and methods section. The next step is to assign the resonance for each individual spin-active nucleus to ultimately utilize those resonances to establish a spatial relationship between these spin-active nuclei through NOE assignments. In the series of steps for the atomic structure determination by NMR, the sequential assignment is the initial step whereby the resonances of the backbone nuclei (^{15}N, ^{1}HN, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO) of the protein chain are assigned. Once this crucial step is accomplished, then the side-chain nuclei attached to these backbone nuclei can be assigned comparatively more easily. When most of the resonances of the nuclei within the protein are assigned, then these resonances would allow assigning NOE resonances which is one of the distance constraints for initial structure calculation by computational method. Hydrogen bonds and dihedral angles (both derived from CSI analysis by TALOS program) are finally used in the three-dimensional structure calculation process.

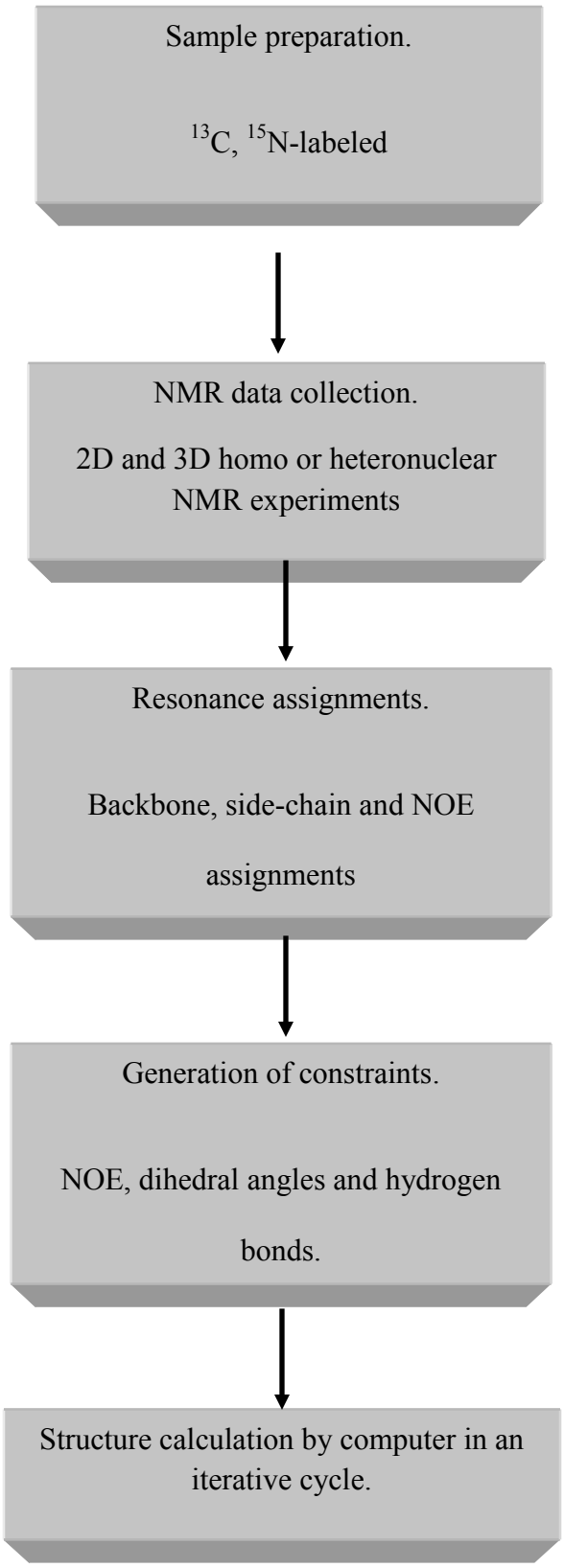


Figure 2.2: Flowchart of structure determination by NMR.

2.3.1.2 Backbone assignments of free GIP

Backbone assignment of free GIP was carried out previously in our laboratory (*24*). However, I carried out the backbone assignment as described below again with previously collected NMR data and with some new data to proceed further with the side chain and 3D NOESY data assignments of free GIP.

Sequential assignments were accomplished by HNCAB and HNCA experiments. HNCACB experiment allowed the sequential assignments of both C_α and C_β atoms of both i and (i-1) residues of the peptide chain (**Figure 2.3**). In HNCA experiment, only C_α atoms of both i and (i-1) residues were assigned (**Figure 2.4**). This sequential assignment was continuous as long as there was no ambiguity or absence of the peaks occurred. To further resolve any ambiguity, HNCO and CBCA(CO)NH experiments were helpful. In the CBCA(CO)NH experiment, only C_α and C_β atoms of the (i-1) residues were assigned. This helped to reconfirm the assignments of the HNCACB experiment (**Figure 2.5**). Although, in the figure (**Figure 2.5**) the peak intensity of the (i-1) residue of HNCACB spectrum is almost same as that of the (i-1) residue of CBCA(CO)NH spectrum, more often than not, the peak intensity of the (i-1) residue of the HNCACB experiments is less than that of the (i-1) residue of CBCA(CO)NH experiment due to the difference in the transfer of magnetization in these two different experiments. This feature also gives an added advantage during the assignments of C_α and C_β atoms of i and (i-1) residues of HNCACB. HNCO experiment also helped to remove ambiguities and reconfirm the assignments. All these assignments were done by continuously referring to the table of Statistics Calculated for All Chemical Shifts from Atoms of the 20 Common Amino Acids (Biological

Magnetic Resonance Data Bank, BMRB, http://www.bmrb.wisc.edu/) which is always updated. Microsoft excel was used to facilitate the sequential assignment. Initially, some unknown numbers were given to each residue. Later on, with the help of excel sheet and the above mentioned table, specific amino acid types and their sequence were obtained.

Once sequential assignment was done, it was quite easy to assign the (^{1}H, ^{15}N)-HSQC spectrum (**Figure 2.6**). N-terminus (M1 residue) of the protein was absent from the (^{1}H, ^{15}N)-HSQC spectrum due to the exchange of the free amide protons with the deuterated solvent (5-10% D_2O). Also, five proline residues were absent from the spectrum due to its unique cyclic structure. However, due to the cis- to trans-isomerization of the proline residues (**Figure 2.7**), the neighboring residues experienced two different chemical environments, consequently appearing at two different chemical shift positions. Glycine at position 6 and Valine at position 9 were affected by the cis- to trans-isomerization of P5 and P8 and were assigned as G6A and V9A (**Figure 2.6**).

Another noticeable thing in the ^{1}H, ^{15}N-HSQC spectrum was that, the disordered regions within the protein such as N- and C-termini have higher peak intensity with a corresponding higher data height in the Sparky program than the regions that are ordered such as α-helix and β-sheet. This happened due to the fast exchange of the flexible regions within the NMR timescale.

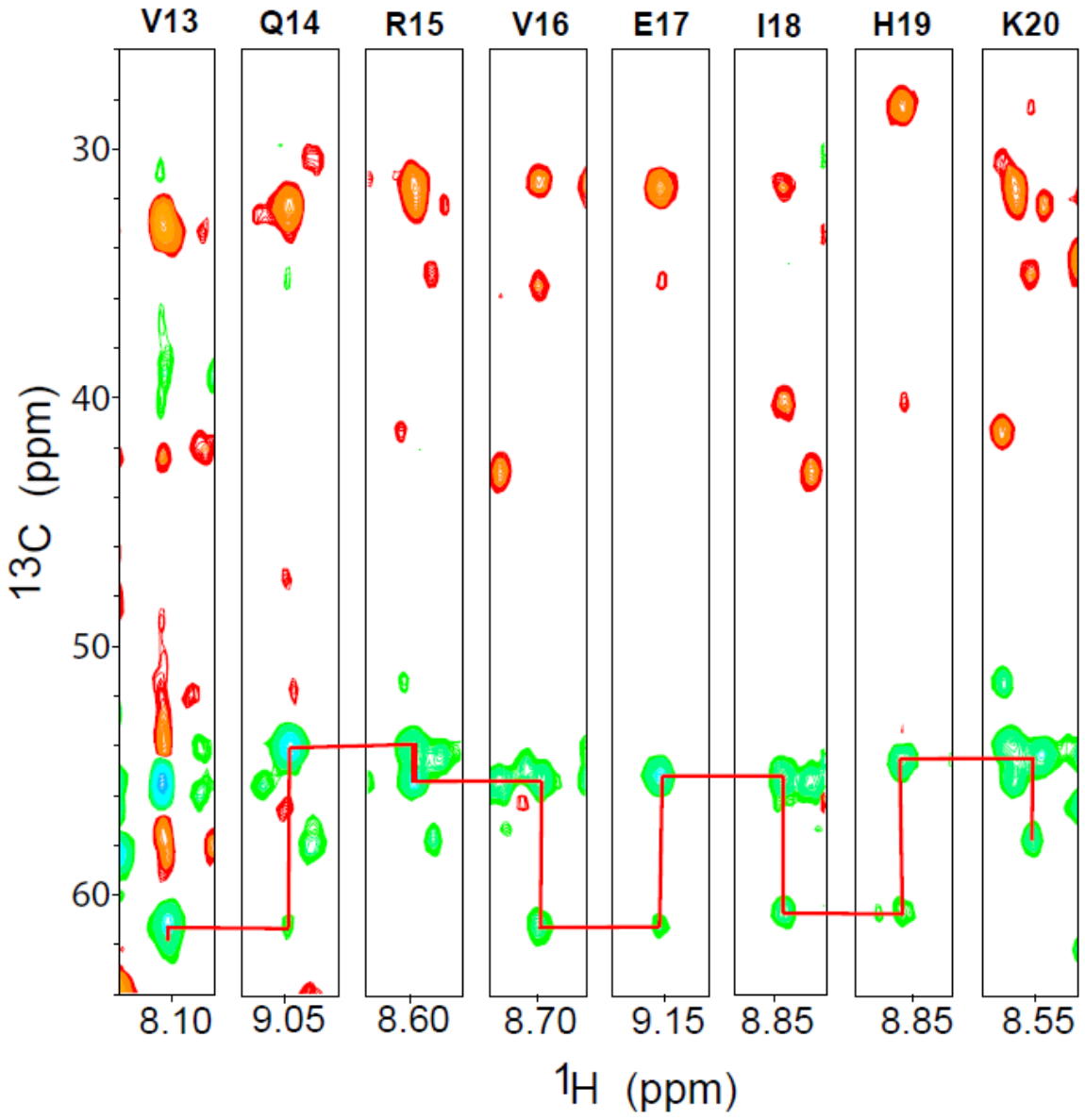


Figure 2.3: Sequential assignments of V13-K20 from (^{1}H, ^{13}C)-strips of HNCACB experiment (*19, 24*). Only the C_α atoms of the residues were connected with red lines to show the sequential assignment. Positive signals are green and negative signals are red. C_α appeared as positive signal and C_β appeared as negative signal.

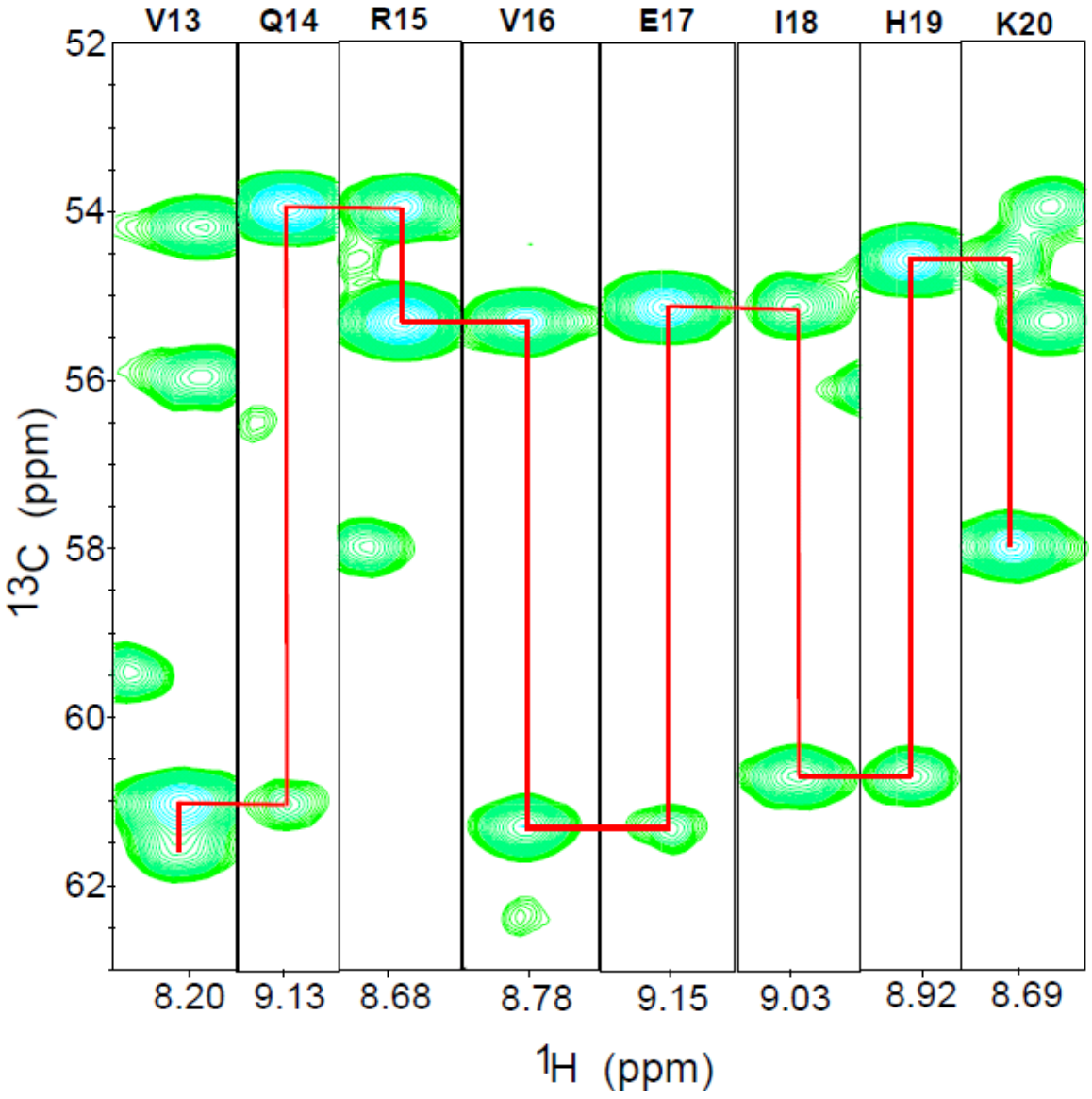


Figure 2.4: Sequential assignments of V13-K20 from (^{1}H, ^{13}C)-strips of HNCA experiment (*19, 24*). C_{α} atoms of the residues were connected with red lines to show the sequential assignment.

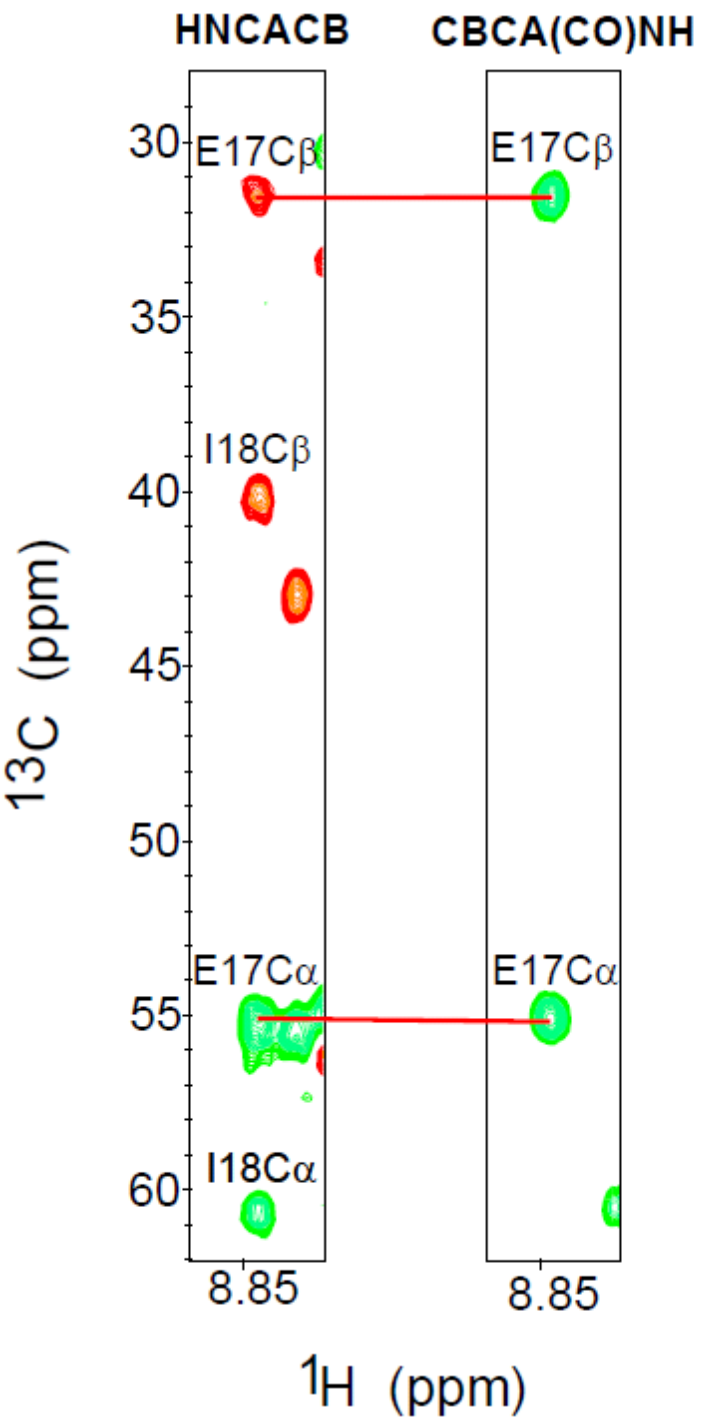


Figure 2.5: HNCACB and CBCA(CO)NH strips of I18 residue (*19, 24*). Red lines were used to connect the C_α and C_β atoms of the (i-1) residue on both spectra.

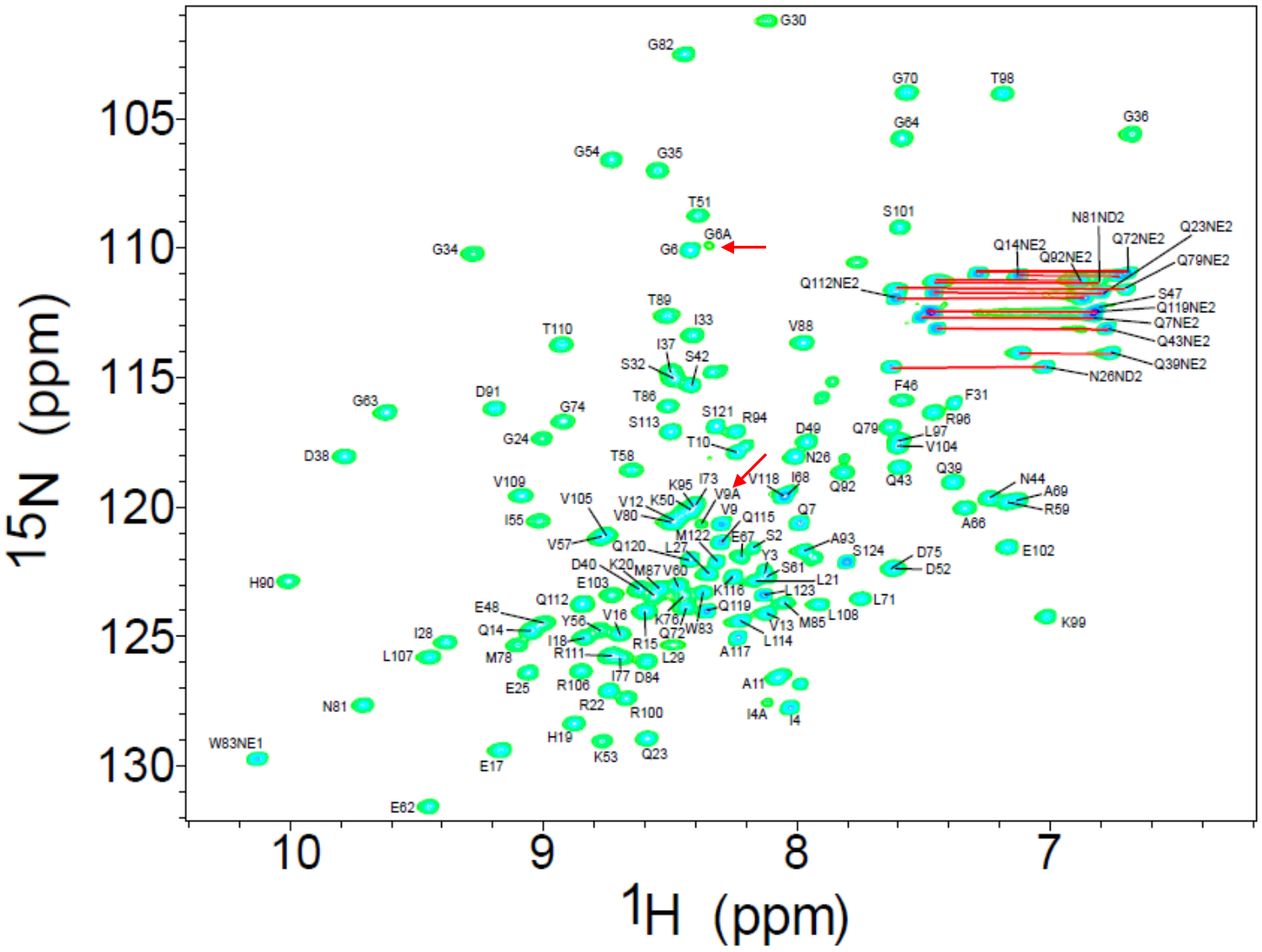


Figure 2.6: ^{1}H, ^{15}N-HSQC spectrum of free GIP (*19, 24*). Red lines connected the non-degenerate protons of the side-chain amide groups of Asparagine and Glutamine residues. Two red arrows located G6A and V9A.

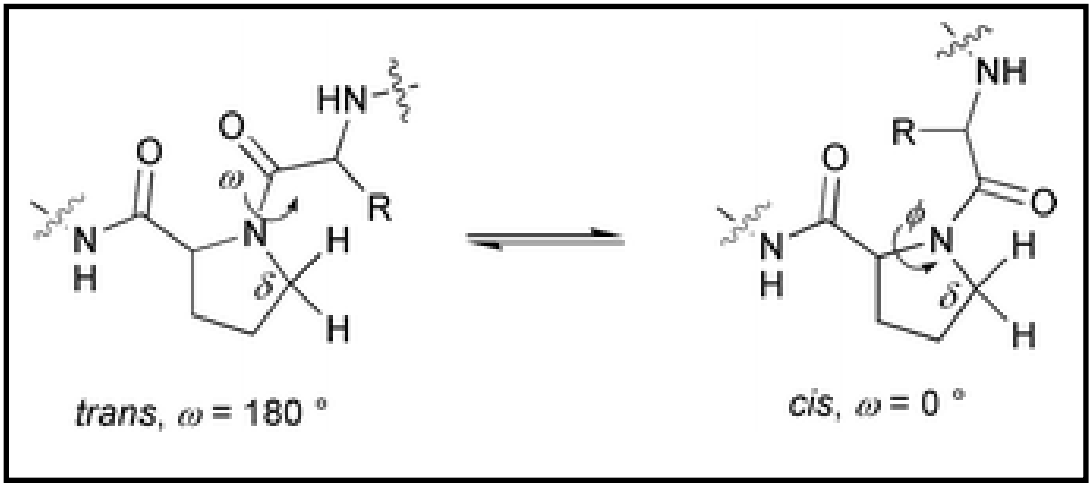


Figure 2.7: Proline cis/trans isomerization.

2.3.1.3 Side-chain assignments of free GIP

With the completion of the backbone assignments (*24*), the next step is to assign the side-chain nuclei of the amino acid residues of the protein. This step is relatively straight-forward. Having the assigned resonances of the backbone atoms, to assign side-chains attached to these backbone atoms (e.g. amide protons), one has to start with a specific amide proton and attached nitrogen resonance of a specific amino acid residue to find the resonances of the side-chain atoms of that residue or the one preceding it from the different spectra.

In HC(CO)NH experiment, assignments of side-chain protons of (i-1) residues were accomplished. Side-chain protons of i-residues were assigned in an HSQC-TOCSY experiment. In the later experiment, side-chain protons of the (i-1) residue could also appear as a negative signal (**Figure 2.8**). Thus, it is a good practice to use these two spectra side-by-side for assigning side-chain protons as a tool for reconfirmation of the assignments. Other spectra used to assign side-chain protons include HCCH-COSY (**Figure 2.9**) and HCCH-TOCSY (**Figure 2.10**). These spectra helped to assign the non-degenerate protons of the side-chains. For the assignment of non-degenerate protons of Glycine C_α, an HNHA experiment was useful (**Figure 2.11**). Sometimes, it was also possible to determine non-degenerate protons in ^{15}N-edited HSQC-NOESY or in ^{13}C-edited HSQC-NOESY. For the unambiguous assignments of the NOEs, detection of non-degenerate protons was very important. To assign side-chain carbons, a CC(CO)NH experiment was used which gives resonances of the side-chain carbon atoms of the (i-1) residues (**Figure 2.12**).

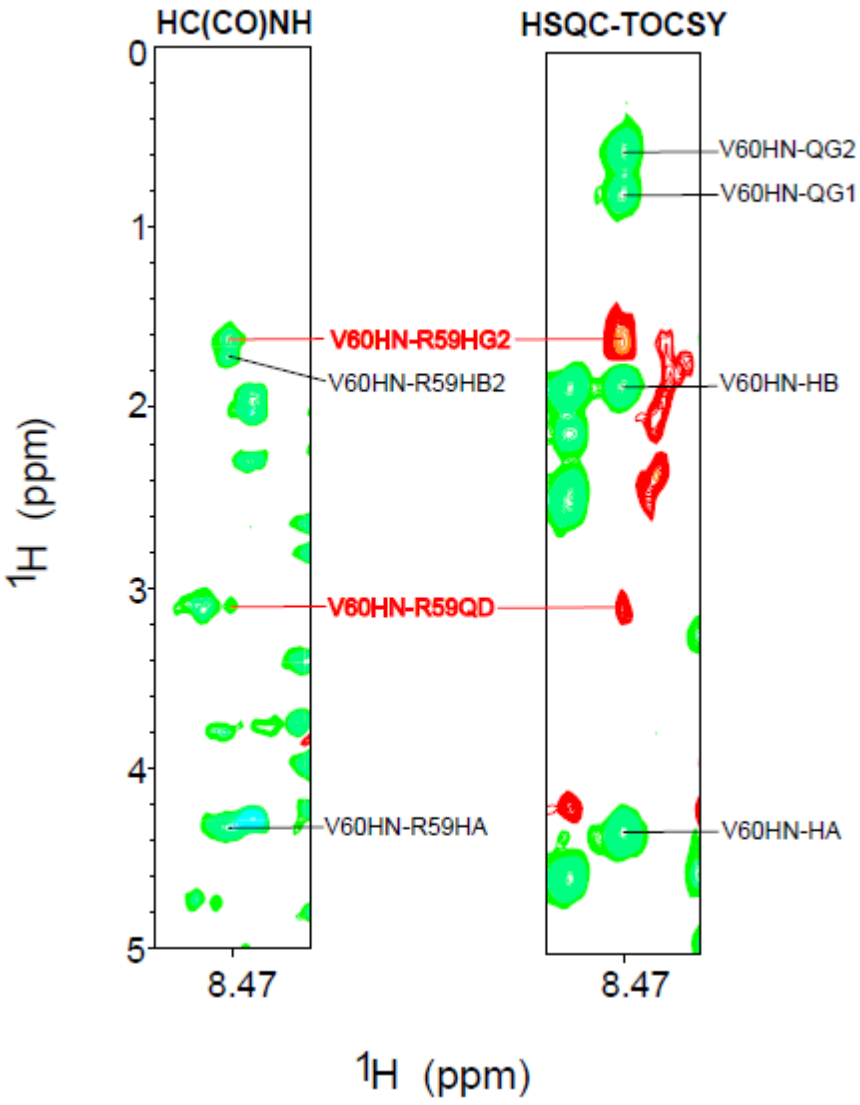


Figure 2.8: HC(CO)NH and HSQC-TOCSY spectra showing side-chain assignments of V60 and R59 residues.

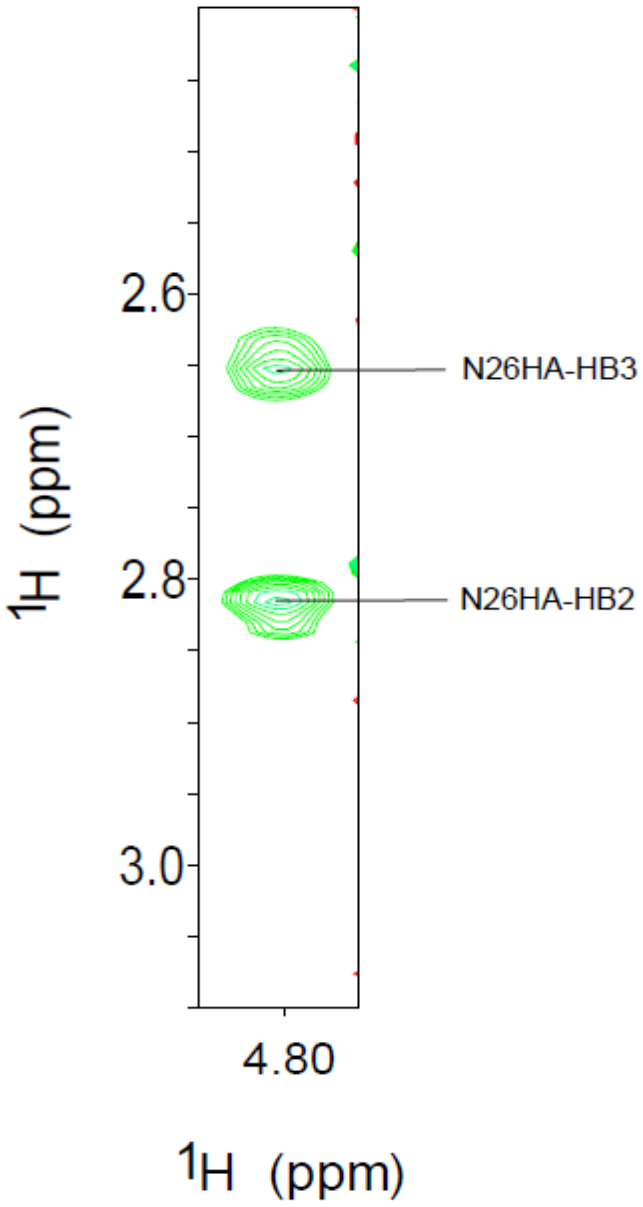


Figure 2.9: HCCH-COSY spectrum of Hα proton of N26 residue showing non-degeneracy of Hβ protons.

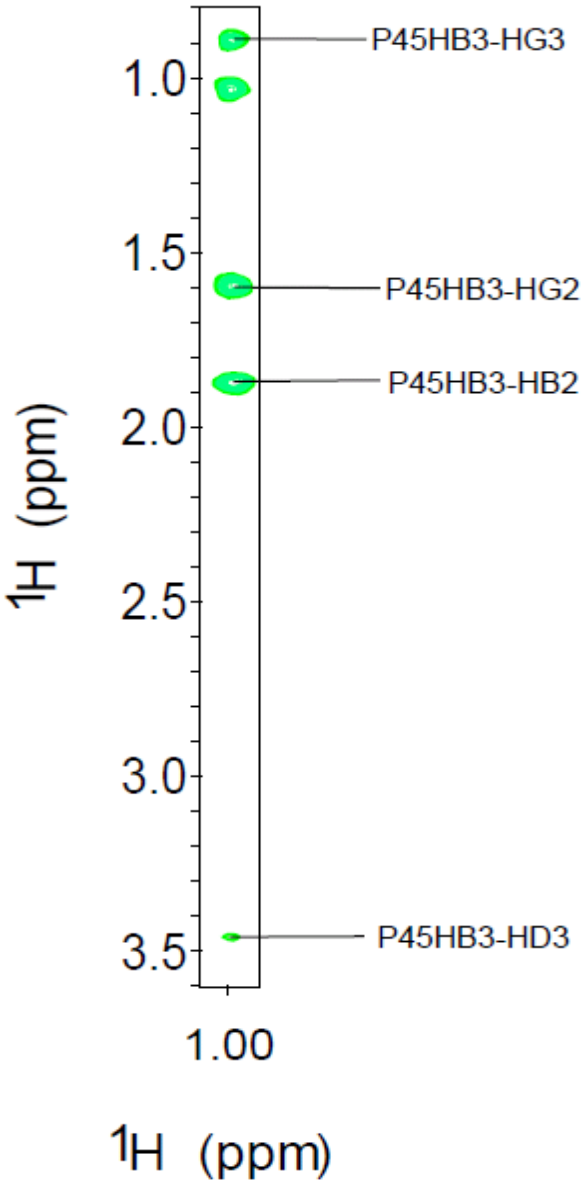


Figure 2.10: HCCH-TOCSY spectrum of Hβ3 proton of P45 residue.

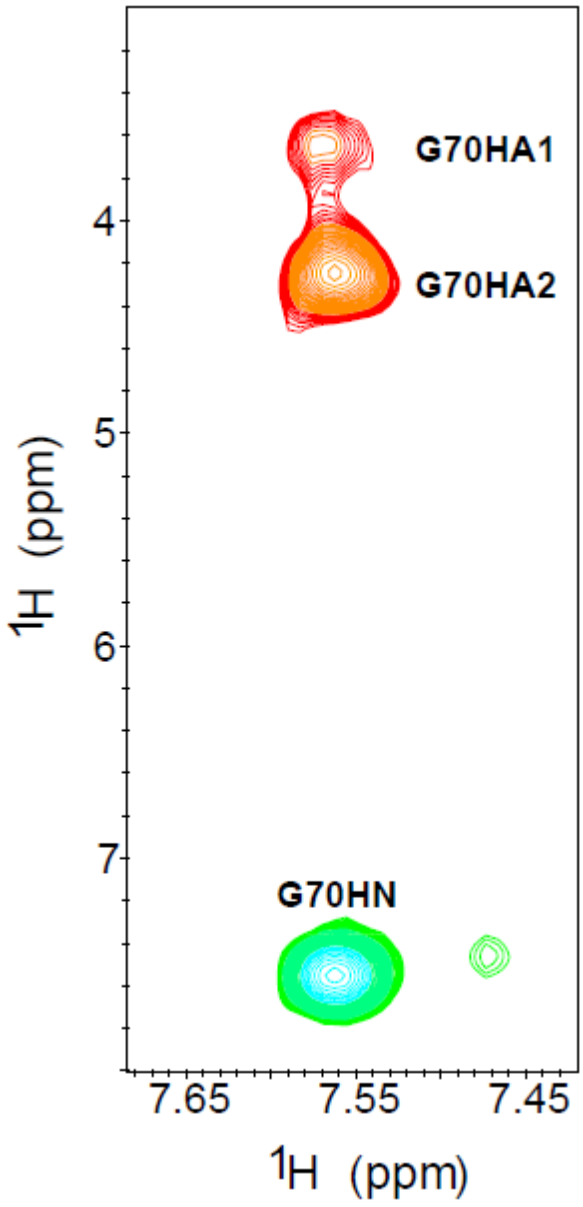


Figure 2.11: HNHA spectrum of non-degenerate Hα protons of G70 residue.

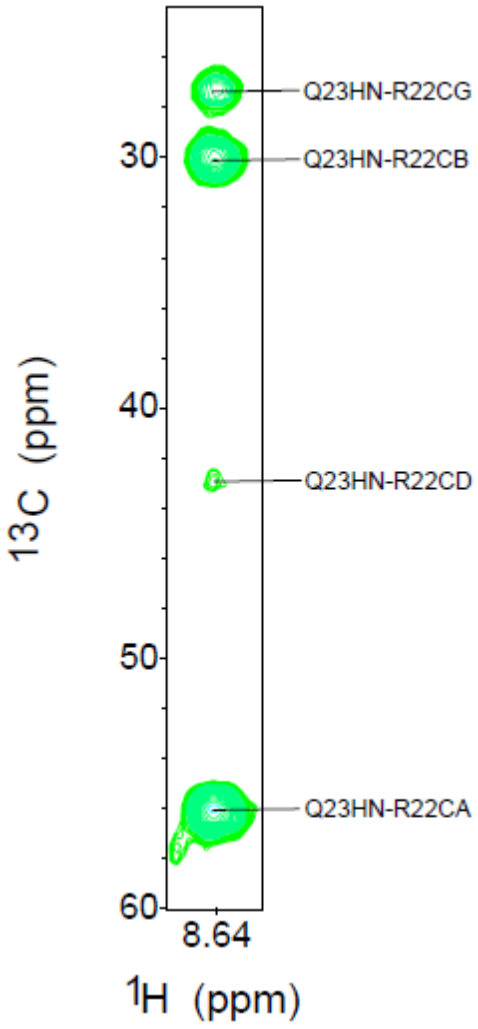


Figure 2.12: CC(CO)NH spectrum of Q23 residue.

The statistics of the assignments of the side-chains are summarized in the **Table 2.2**. In summary, around 92, 95 and 90 percent of all carbon, hydrogen and nitrogen nuclei, respectively, were unambiguously assigned. Although, the peaks in the ^{1}H, ^{13}C-HSQC spectrum are overlapping more than those in the ^{1}H, ^{15}N-HSQC spectrum, this amount of assignments was sufficient to assign most of the peaks of the aliphatic region of the ^{1}H, ^{13}C-HSQC spectrum (**Figure 2.13** and **Figure 2.13A**) and all of the peaks of the aromatic region of the ^{1}H, ^{13}C-HSQC spectrum (**Figure 2.14)**. Assignment of the full ^{1}H, ^{13}C-HSQC spectrum was instrumental in the assignments of the cross-peaks in the ^{13}C-edited HSQC-NOESY spectrum later on.

Atom	C(CO)	Cα	Cβ	Cγ	Cδ	Cε	Cζ	Cη	Total C
% of Assignment	100	100	99	81	73	58	50	100	92.3
Found vs. Expected	124/124	124/124	111/112	74/91	33/45	11/19	2/4	1/1	466/520
Atom	HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Total H
% of Assignment	99	100	99	87	92	79	100	100	95.2
Found vs. Expected	118/119	124/124	111/112	79/91	44/48	31/39	4/4	1/1	512/538
Atom	N				Nδ	Nε			Total N
% of Assignment	99				40	56			90
Found vs. Expected	118/119				2/5	13/23			133/147

Table 2.2: Statistics of side-chain assignments of free GIP.

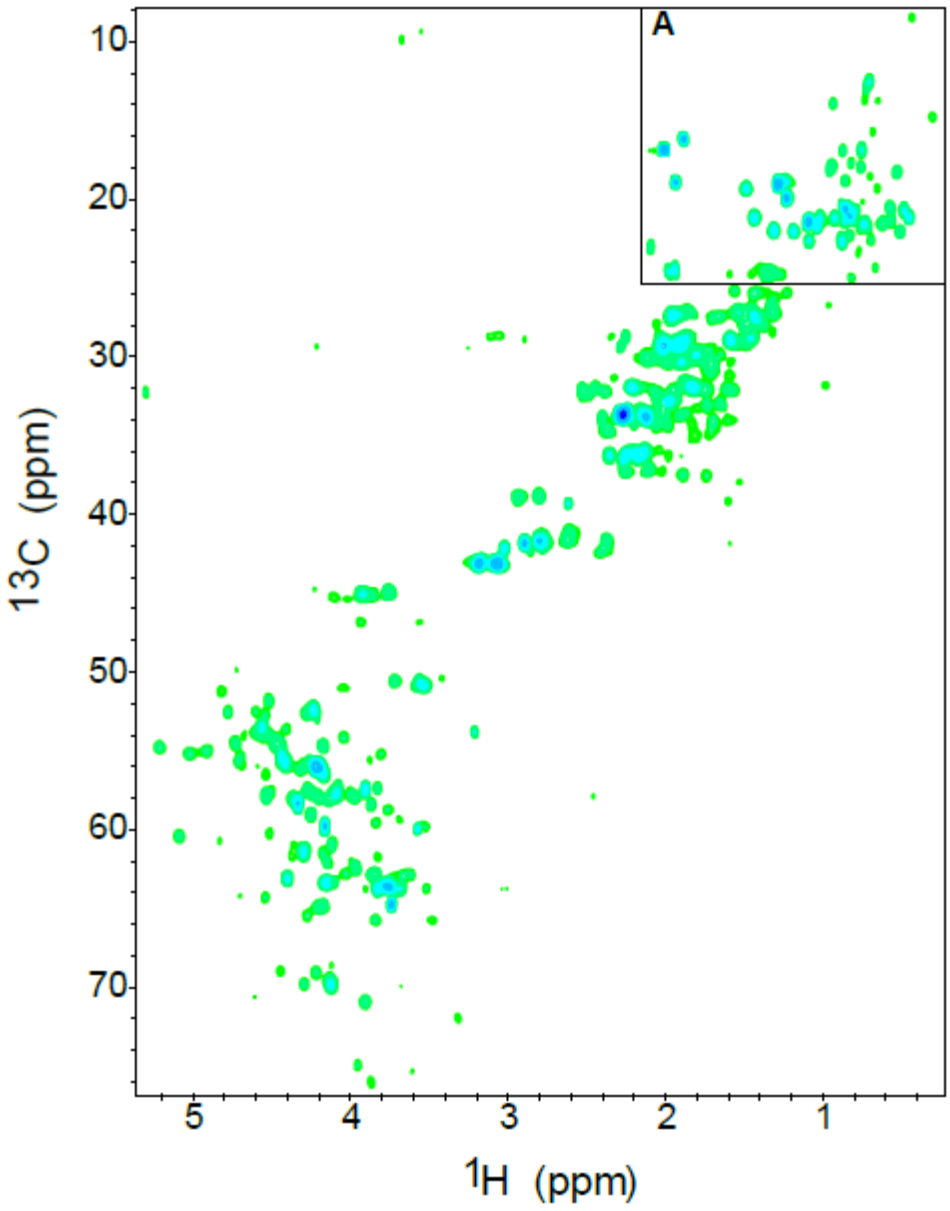


Figure 2.13: Aliphatic region of the ^{1}H, ^{13}C-HSQC spectrum of free GIP. Inset A contains methyl groups. This inset is blown up in **Figure 2.13A**.

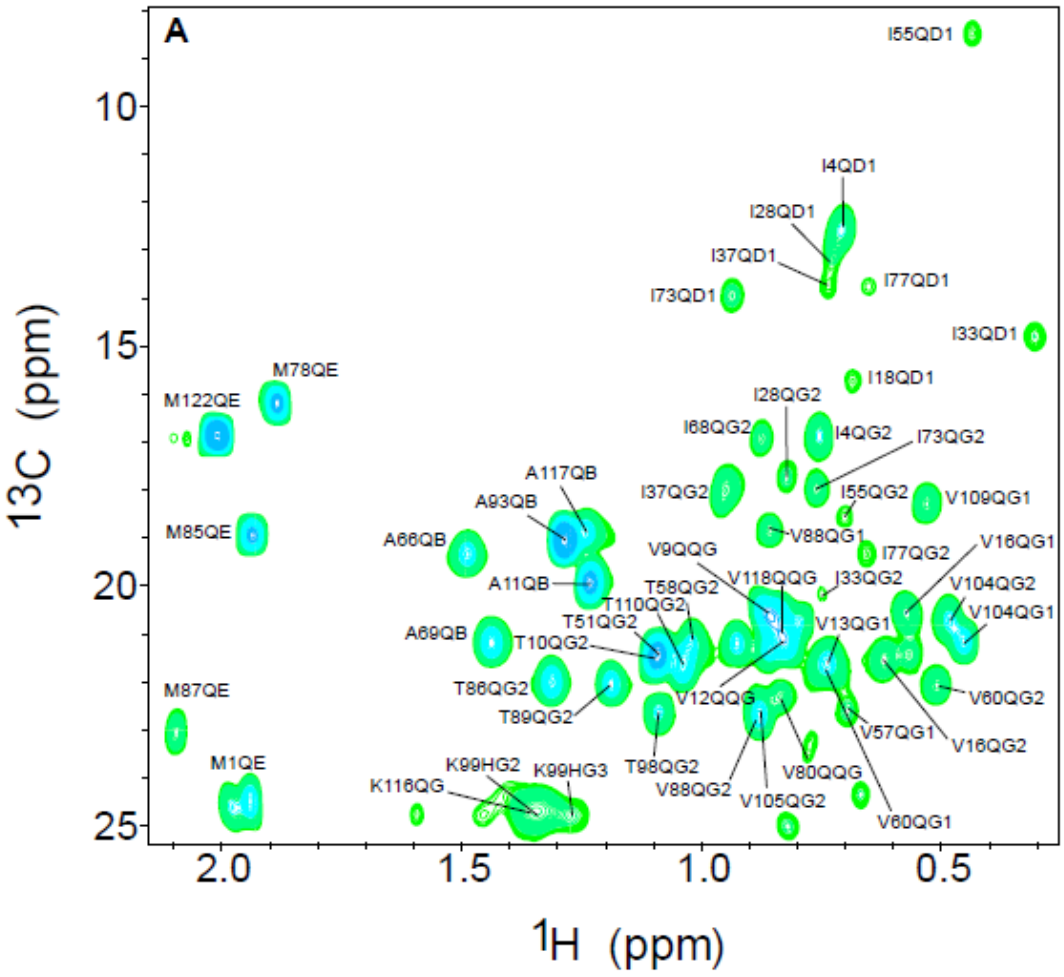


Figure 2.13A: Part of the aliphatic region of the ^{1}H, ^{13}C-HSQC spectrum with assignments.

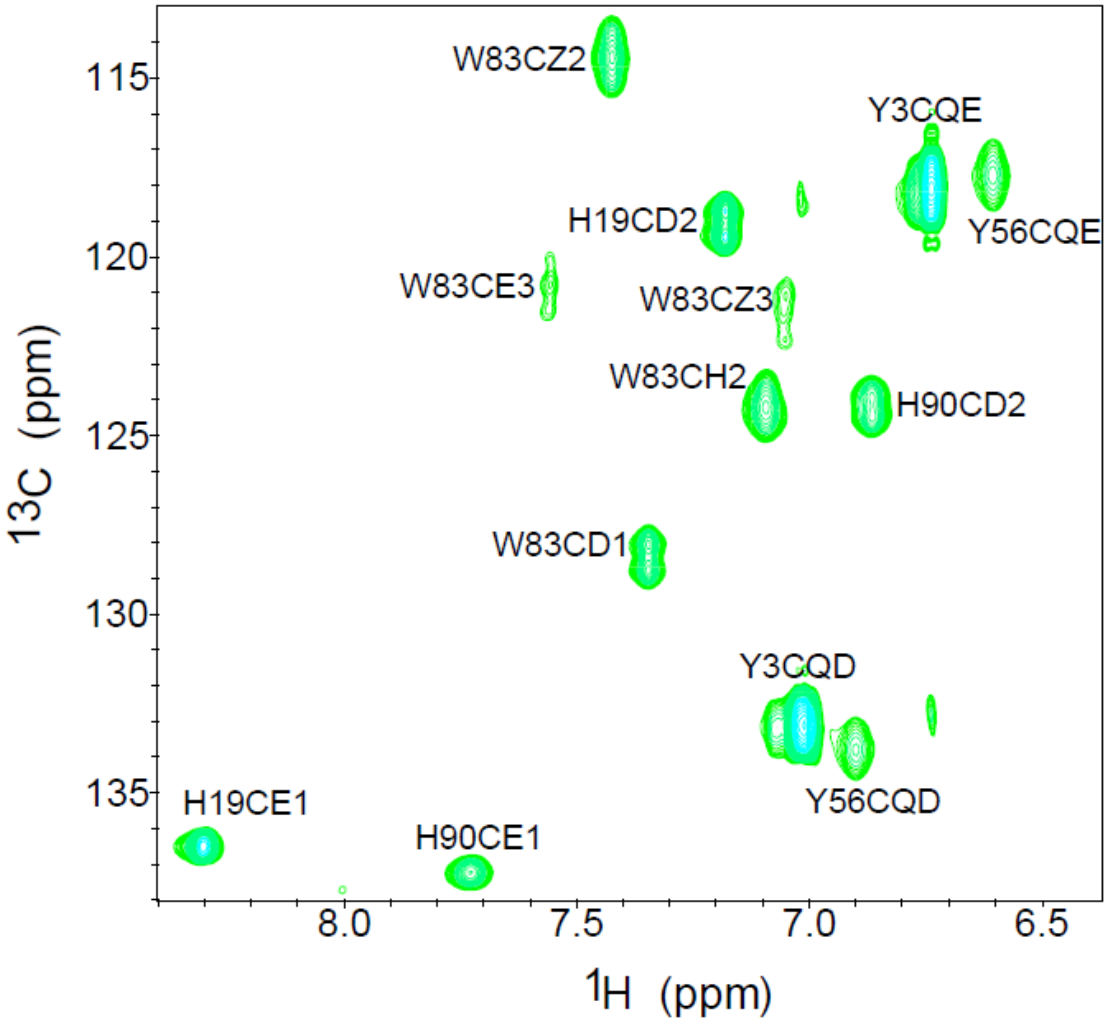


Figure 2.14: Aromatic region of the ^{1}H, ^{13}C-HSQC spectrum with assignments.

2.3.1.4 NOE assignments of free GIP

NOE assignment is the last and most crucial stage of the resonance assignment step. The sole purpose of all the previous resonance assignments (both backbone and side-chain) was to utilize those resonances in this step so that appropriate and unambiguous NOEs between different protons could be determined. For the structure calculation, NOE constraints are the most important distance constraints. Both ^{15}N-edited HSQC-NOESY (**Figure 2.15**) and ^{13}C-edited HSQC-NOESY (**Figure 2.16**) were used to find out NOEs between the protons. Though it sounds quite straight-forward for NOE assignment having most of the protons assigned, this stage is quite challenging due to the fact of the overlap between peaks or absence of expected peaks. That is why proper care was taken during NOE assignment to maintain a balance between not picking up a useful NOE and assigning NOEs of data heights which are not actually representative of those NOEs. To find the global fold of the protein, enough long-range NOEs are needed for the calculation. But, unfortunately, long-range NOEs are usually weak and can easily be shadowed by the intra-residue, short- and medium-range NOEs. Due to these factors, NOE assignment requires an iterative process of manual assignment and correction with the concomitant run and check of structure calculation. Initially, using Sparky a total of 4303 NOE cross peaks were assigned manually. These assignments were then either corrected or confirmed by the process of structure calculation using both CYANA1.0.6 and CYANA 2.1. A total of 1134 of those NOE cross peaks were removed and the rest of the peaks were used in the final structure calculation. Among the finally used cross peaks, a total of 1824 were either sequential, medium or long-range and the rest were just intra-residue NOEs (**Table 2.3**) (*19*).

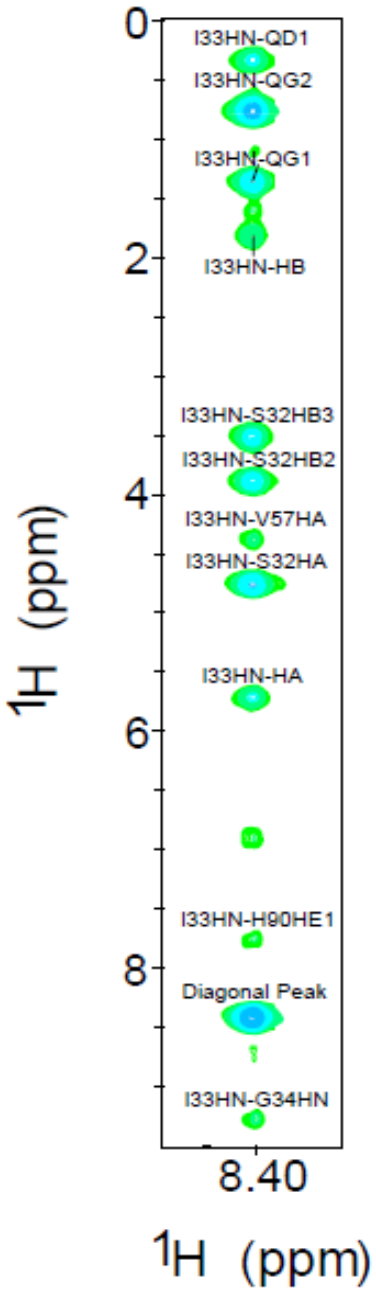


Figure 2.15: ^{15}N-edited HSQC-NOESY spectrum of I33 residue. The assignments shown here were manually picked in Sparky which were later confirmed, removed or corrected in the iterative process.

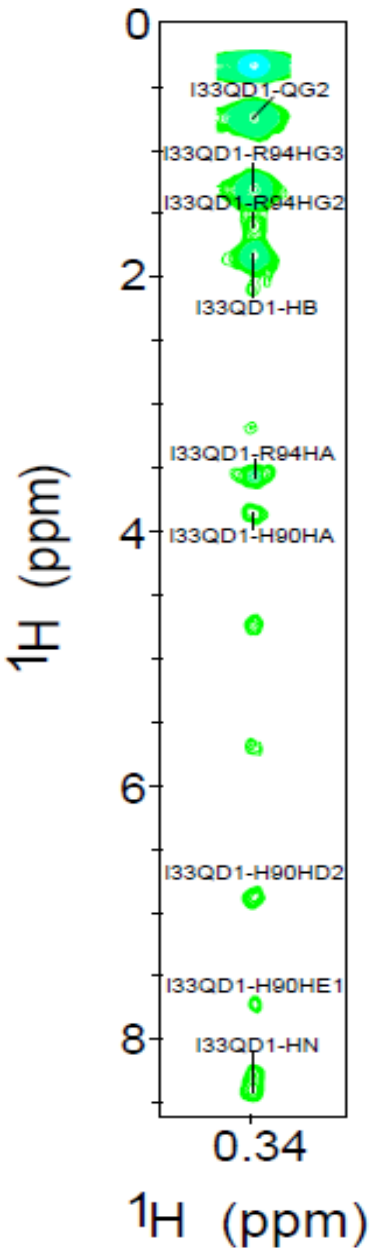


Figure 2.16: ^{13}C-edited HSQC-NOESY spectrum of I33 QD1 proton. The assignments shown here were manually picked in Sparky which were later confirmed, removed or corrected in the iterative process.

2.3.1.5 Structure calculation of free GIP

To calculate the structure of free GIP, both the CANDID module of CYANA 1.0.6 and NOEASSIGN module of CYANA 2.1 were used. The reason for using both of the versions of CYANA is that somehow only the older version of CYANA allowed the use of both defined upper limits of distance constraints (as *.upl files) and undefined (as *.peaks files) distance constraints. The newer version would only allow the usage of *.upl files. The advantage of using both types of files in this case of structure calculation is that it helps to initially determine a structure based on the given upper limits of the distance constraints (*.upl files), then the other undefined peaks can be defined with respect to this initial structure through an iterative process (**Figure 2.18**). In CYANA 1.0.6, besides *.peaks files dihedrals and hydrogen bonds are also given as *.aco and *.upl files. During each of the CYANA run, the output files were checked for the possible indication for the improvement in the next run, for example, by examining listed violations in *.ovw files or suggested peak assignments by the program itself in *.ass files. After several runs of checking and correcting, an enriched *.upl file was constructed which can then be used in the NOEASSIGN module of CYANA 2.1.

In CYANA 2.1, the latest refined *.upl file was used as an input along with the dihedrals and hydrogen bonds. After each run, the violations and energy functions were checked from *.ovw files along with the close examination of output *.pdb files (**Figure 2.20**). Suspicious upper limits of distance constraints from the *.upl file were removed to achieve lesser violations (distance and angle) and lower energy functions. An initial structure was also used as an input at a later stage of the CYANA 2.1 run. The addition of an initial structure as an input file in the run

helped lowering root mean square deviation (RMSD) of the output structures. Final water refinement was done to get 100 lowest energy structures from 200 calculated structures. Of these, 20 structures of lowest potential energy and best Ramachandran statistics found from PROCHECK were used for analysis. Their structural statistics were summarized in the **Table 2.3 (*19*)**. The ensemble of these 20 structures is shown in **Figure 2.21 (*19*)**.

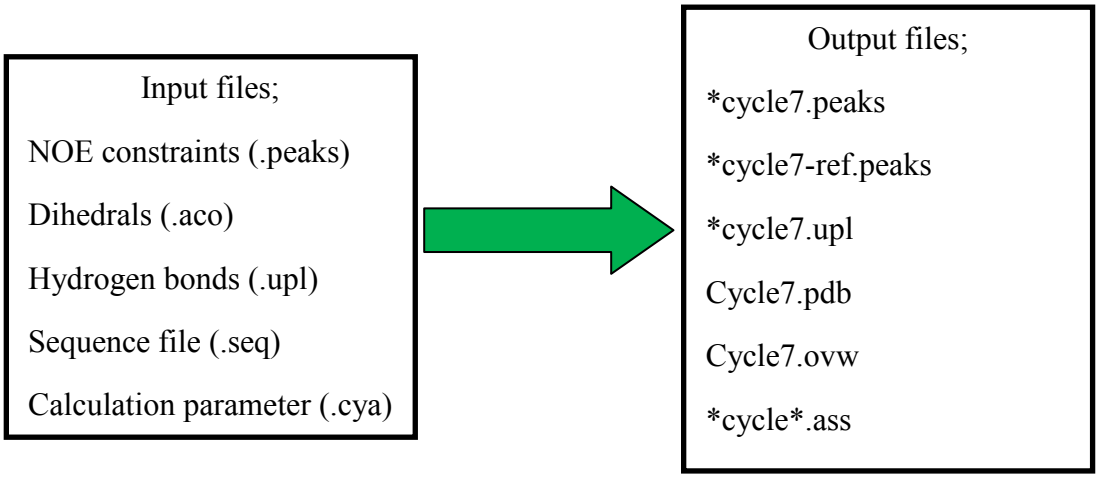


Figure 2.17: Input and output files for CYANA 1.0.6. * denotes possible preceding or following letters. A structure can also be used as an input file as *.cor/*.pdb file. In all cases, calculation parameter file (*.cya) needs to be changed accordingly.

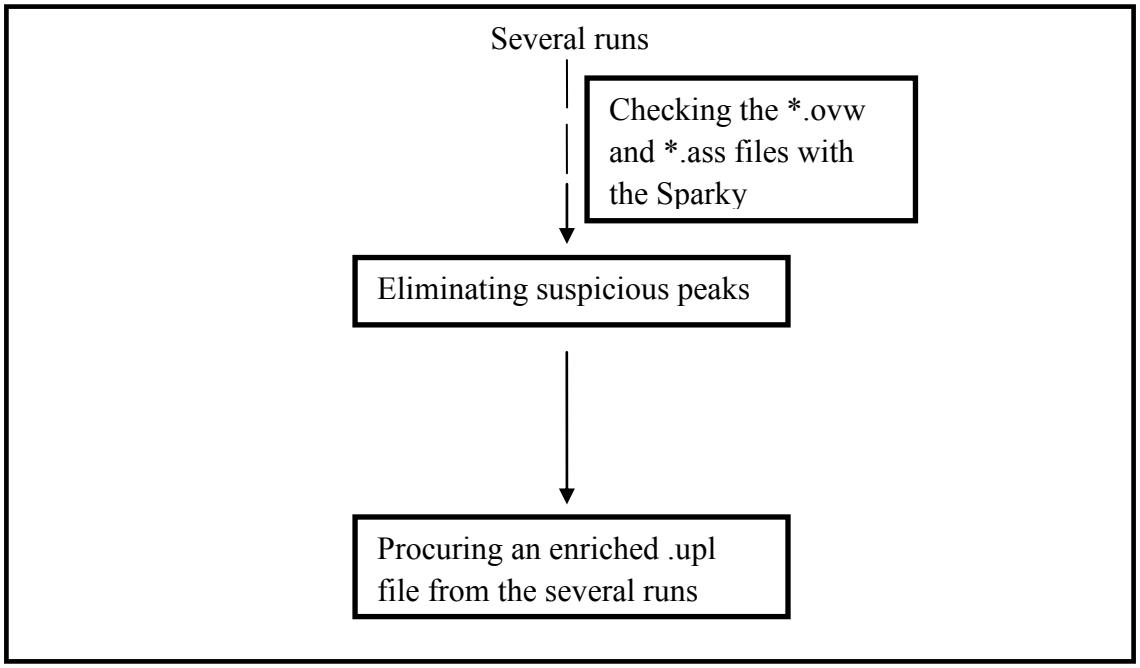


Figure 2.18: Iterative cycle of CYANA 1.0.6 run.

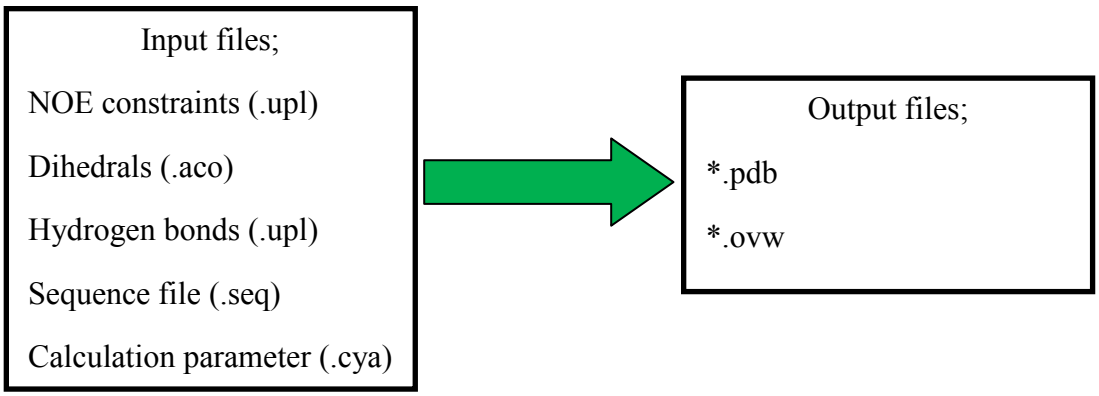


Figure 2.19: Input and output files for CYANA 2.1. A structure can also be used as an input file as *.cor/*.pdb file. In all cases, calculation parameter file (*.cya) needs to be changed accordingly.

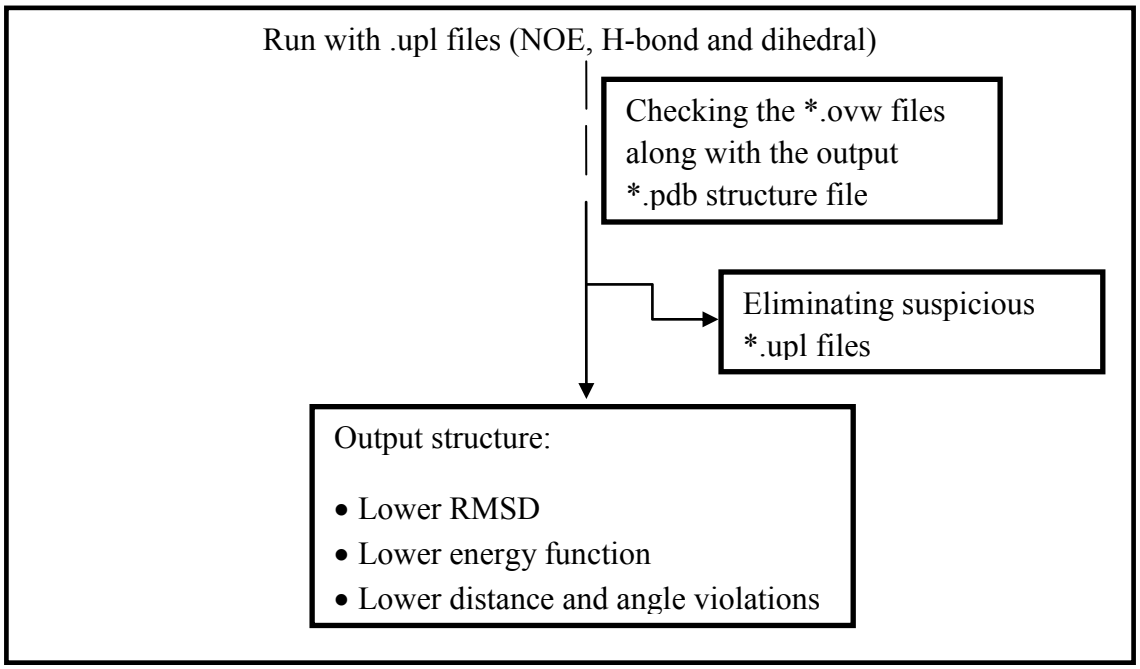


Figure 2.20: Iterative cycle of CYANA 2.1 run.

2.3.1.6 Refinement of structures by ARIA

It has been reported that NMR structures can be significantly improved by using the refinement protocol in explicit solvent (*53-55*). ARIA is computer software that allows the refinement of NMR structures using a Water refinement protocol. ARIA utilizes slightly modified OPLS (Optimized Potentials for Liquid Simulations) force field to involve Lennard-Jones van der Waals and electrostatic interactions during the water refinement. For the water refinement, the structures are immersed in a 7.0 Å shell of water molecules while keeping the distance between a heavy atom of the protein and oxygen atom of water at 4.0 Å (*56*). Using the same distance and angle restraints files, as those used for the final structural calculation by CYANA, ARIA 1.2 employs seven cycles of simulated annealing (SA) protocol for structure calculation followed by a final cycle of water refinement protocol. SA protocol is composed of four phases (*55*):

i. 1100 steps of torsion angle simulated annealing at 10,000 K
ii. 550 steps of first torsion angle dynamics cooling phase from 10,000 K to 2000 K
iii. 5000 steps of second Cartesian dynamics cooling phase from 2000 K to 1000 K
iv. 2000 steps of third Cartesian dynamics cooling phase from 1000 K to 0 K

In the simulated annealing protocol, the 200 best structures are calculated and arranged according to their total energy. Among these 200 structures, only the 100 best structures are then used in the final cycle of water refinement.

2.3.1.7 NMR structure of free GIP

A PDZ domain is usually composed of six β-strands (β1-β6) and two α-helices (α1 and α2) (*16*). Being composed solely of a single PDZ domain, the NMR structure of GIP resembles the characteristic PDZ domain with six β-strands (β1-β6) and two α-helices (α1 and α2) (**Figure 2.21**). However, GIP contains two additional β strands (βa and βb), between β1 and β2, anti-parallel to each other connected by a turn (*19*). The C- and N-termini of the protein are very disordered signifying their free movement in the solution. Another region of the protein, quite unstructured and flexible, is the loop region between the β2 and the β3 strand (β2- β3 loop). Apart from these regions, the free GIP protein appears quite structured in the core and illustrated by the convergence of those parts of the structures in the ensemble of 20 superimposed lowest energy structures (**Figure 2.21**).

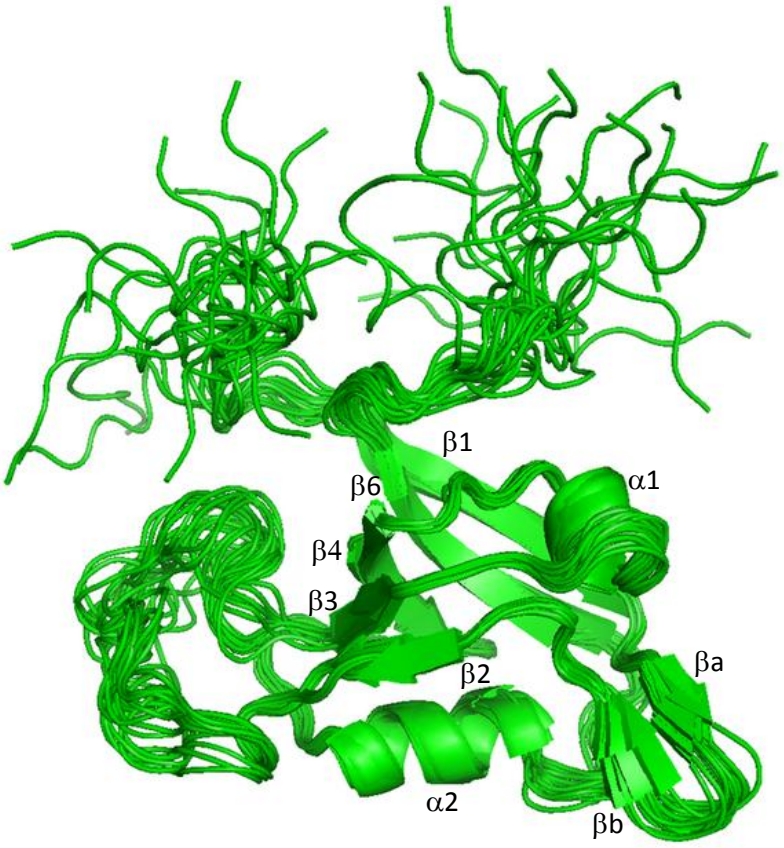


Figure 2.21: Ribbon diagrams of the ensemble of the 20 superimposed lowest energy structures of free GIP. Adapted from reference (*19*).

Assignments	Free GIP
Sequential |i-j|=1	871
Medium 2≤|i-j|≤4	331
Long |i-j|>4	622
Intermolecular	0
Hydrogen Bonds [a]	64
Dihedral Constraints [b]	118
Ensemble Average [c]	
Total energy	-3625 ± 125
NOE energy	1131 ± 189
VDW energy	-937 ± 75
Bonds energy	85 ± 5
Dihedral energy	657 ± 10
Angle energy	318 ± 22
Improper energy	963 ± 78
Electrostatic energy	-4712 ± 67
Ramachandran Plot [d]	
Favorable	68.6
Additionally Allowed	26.6
Generously Allowed	3.4
Disallowed	1.5
RMSD (Å) [e]	
Well-ordered Backbone	0.45
Well-ordered Sidechain	0.92

Table 2.3: NMR structural statistics for the 20 selected lowest energy structures of free GIP. Adapted from reference (*19*).

[a] Hydrogen bonds were defined by a set of two distance restraints per bond for residues of predicted secondary structure based on TALOS (*48*) predictions from CSI.

[b] Dihedral constraints were derived from TALOS (*48*) predictions from CSI.

[c] Energy terms were calculated by the water refinement module of ARIA 1.2 (*49*).

[d] Ramachandran plot statistics were calculated by PROCHECK (*50*).

[e] Well ordered regions included residues 11-19, 29-36 and 54-112.

2.3.1.8 Accession codes

The accession codes for free GIP in the BioMagnetic Resonance Bank (BMRB) and the Protein Data Bank (PDB) are 17254 and 2L4S, respectively. In BMRB, the chemical shifts of the resonances and, in PDB, the atomic coordinates for free GIP have been deposited (*19*).

2.3.2 Dynamics of free GIP from ^{15}N relaxation measurements

Using the Lipari-Szabo formalism based model-free analysis (*57*), the order parameters (S^2) for free GIP were calculated with the data collected by another member in our research group, using steady-state ^{1}H-^{15}N NOE intensities, R_1 and R_2 relaxation rates. Those residues that could not be analyzed as a result of low intensity or absence from the HSQC spectra due to the overlapping were excluded from the data analysis. Excluded residues include M1, P5, P8, V12, K20, L21, L29, G30, P41, P45, K50, D52, V57, R59, P65, I68, A69, I73, D75, V80, M87, K95, V105 and V118. Aside the N-terminus and five proline residues, S^2 values for other residues could not be measured mainly for two reasons: spectral overlap and line broadening. In total, 100 of 118 residues (excluding the N-terminus and 5 prolines) were analyzed to determine the S^2 values. It is important to remember here that, the higher the S^2 value, the lesser mobile it is. Well-defined secondary structure of the protein should be more ordered and less mobile. Analysis of the dynamics data reveals the same pattern of mobility for free GIP protein. The defined secondary structure of free GIP showed relatively restricted mobility of 0.85 or above (**Figure 2.22**), whereas, C-and N-termini of the protein and various loops including the βa-βb hairpin, the β2-β3 loop and a few other short loops between secondary structural elements exhibited greater flexibility (**Figure 2.22** & **Figure 2.23**). When the RMSD values for individual

residues obtained from structural calculation were plotted on the same graph containing information on S^2 values, a correlation was found between the order parameters and the overall RMSD values (**Figure 2.22**). Higher RMSD values corresponded to lower S^2 values. An average high S^2 value of 0.89 for the core region (A11-Q112) of free GIP was calculated from the modelfree analysis. This high value corresponds to the restricted backbone mobility of a well folded protein. However, as we go toward the termini of the protein this value drops low very suddenly (**Figure 2.22**) (*19*).

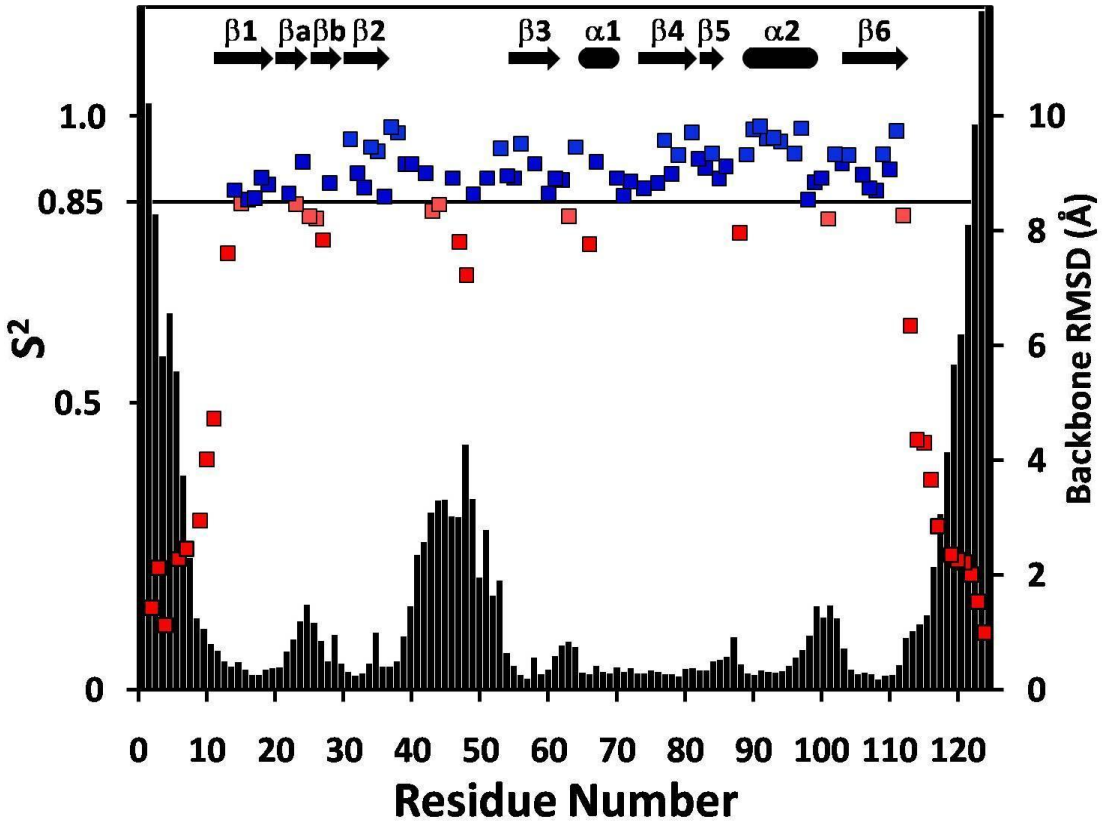


Figure 2.22: The S^2 values derived using the modelfree analysis from the steady state ^{1}H-^{15}N NOE, R_1 and R_2 relaxation times of free GIP for each non-overlapping well defined residue. Residues with order parameters above the threshold 0.85 were colored in blue while those below were colored in red. The backbone RMSD of free GIP for each residue was overlaid on this plot in black. Adapted from reference (*19*).

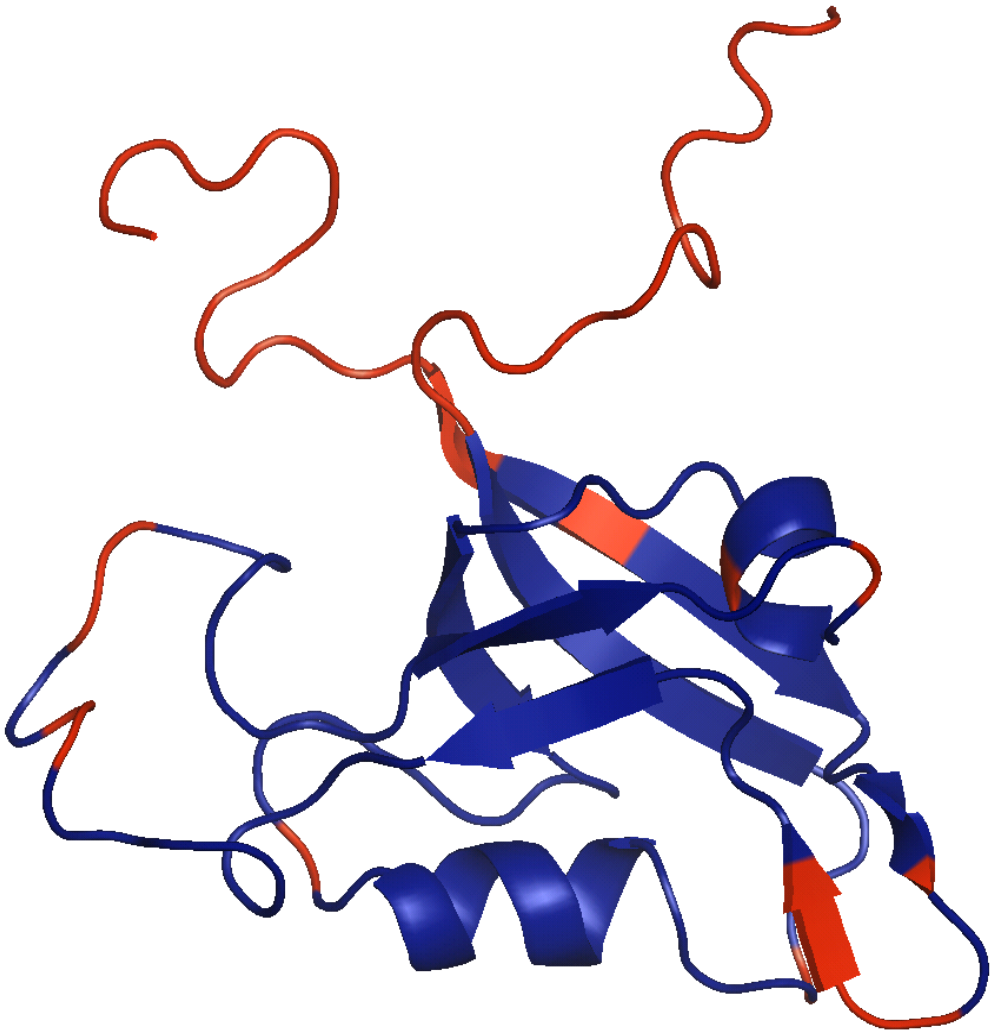


Figure 2.23: Residues with S^2 values below the threshold of 0.85 are mapped in red onto the structure of free GIP colored blue. Adapted from reference (*19*).

2.4 Conclusion

We solved the solution structure of free GIP using NMR and determined the dynamics of free GIP protein. The global structure of GIP is consistent with that of the canonical PDZ domain although there are small differences. The dynamics corresponds coherently to the structure of GIP. The more structured the region of the protein is, the lower is its mobility and randomness. This structural and dynamics study of free GIP would allow us to compare and contrast with those of bound GIP that forms complex with a substrate (Chapter 3). Such a comparative study should shed light on the mechanism of interaction between GIP and its binding partner.

2.5 References

1. Pellegrini, M., Haynor, D., and Johnson, J. M. (2004) Protein interaction networks, *Expert Rev Proteomics 1*, 239-249.
2. Neel, B. G. (1993) Structure and function of SH2-domain containing tyrosine phosphatases, *Semin Cell Biol 4*, 419-432.
3. Kardinal, C., Posern, G., Zheng, J., Knudsen, B. S., Moarefi, I., and Feller, S. M. (1999) Rational development of cell-penetrating high affinity SH3 domain binding peptides that selectively disrupt the signal transduction of Crk family adapters. Amgen Peptide Technology Group, *Ann N Y Acad Sci 886*, 289-292.
4. Falke, J. J. (2007) Membrane Recruitment as a Cancer Mechanism: A Case Study of Akt PH Domain, *Cellscience 4*, 25-30.
5. Fanning, A. S., and Anderson, J. M. (1996) Protein-protein interactions: PDZ domain networks, *Curr Biol 6*, 1385-1388.
6. Woods, D. F., and Bryant, P. J. (1993) ZO-1, DlgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions, *Mechanisms of development 44*, 85-89.
7. Garcia-Mata, R., and Burridge, K. (2007) Catching a GEF by its tail, *Trends Cell Biol 17*, 36-43.
8. Fanning, A. S., and Anderson, J. M. (1996) Protein-protein interactions: PDZ domain networks, *Current Biology 6*, 1385-1388.
9. Jelen, F., Oleksy, A., Smietana, K., and Otlewski, J. (2003) PDZ domains - common players in the cell signaling, *Acta Biochim Pol 50*, 985-1017.

10. Spaller, M. R. (2006) Act globally, think locally: systems biology addresses the PDZ domain, *ACS Chem Biol 1*, 207-210.

11. Lee, H. J., and Zheng, J. J. (2010) PDZ domains and their binding partners: structure, specificity, and modification, *Cell communication and signaling : CCS 8*, 8.

12. 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.

13. Christopherson, K. S., Hillier, B. J., Lim, W. A., and Bredt, D. S. (1999) PSD-95 assembles a ternary complex with the N-methyl-D-aspartic acid receptor and a bivalent neuronal NO synthase PDZ domain, *J Biol Chem 274*, 27467-27473.

14. Lemaire, J. F., and McPherson, P. S. (2006) Binding of Vac14 to neuronal nitric oxide synthase: Characterisation of a new internal PDZ-recognition motif, *FEBS Lett 580*, 6948-6954.

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

16. 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.

17. Wiedemann, U., Boisguerin, P., Leben, R., Leitner, D., Krause, G., Moelling, K., Volkmer-Engert, R., and Oschkinat, H. (2004) Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides, *J Mol Biol 343*, 703-718.

18. 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.

19. 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-Us 50*, 3528-3539.

20. 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.

21. 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.

22. 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, *Journal of Biological Chemistry 278*, 38758-38764.

23. Zhang, J., Yan, X., Shi, C., Yang, X., Guo, Y., Tian, C., Long, J., and Shen, Y. (2008) Structural basis of beta-catenin recognition by Tax-interacting protein-1, *J Mol Biol 384*, 255-263.

24. 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-Us 47*, 9208-9219.

25. 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.

26. 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.

27. 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.

28. Le Maout, S., Welling, P. A., Brejon, M., Olsen, O., and Merot, J. (2001) Basolateral membrane expression of a K+ channel, Kir 2.3, is directed by a cytoplasmic COOH-terminal domain, *Proc Natl Acad Sci U S A 98*, 10475-10480.

29. 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.

30. Perez-Gomez, C., Campos-Sandoval, J. A., Alonso, F. J., Segura, J. A., Manzanares, E., Ruiz-Sanchez, P., Gonzalez, M. E., Marquez, J., and Mates, J. M. (2005) Co-expression of glutaminase K and L isoenzymes in human tumour cells, *Biochem J 386*, 535-542.

31. Gao, P., Tchernyshyov, I., Chang, T. C., Lee, Y. S., Kita, K., Ochi, T., Zeller, K. I., De Marzo, A. M., Van Eyk, J. E., Mendell, J. T., and Dang, C. V. (2009) c-Myc suppression

of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism, *Nature 458*, 762-765.

32. Gallagher, F. A., Kettunen, M. I., Day, S. E., Lerche, M., and Brindle, K. M. (2008) 13C MR spectroscopy measurements of glutaminase activity in human hepatocellular carcinoma cells using hyperpolarized 13C-labeled glutamine, *Magn Reson Med 60*, 253-257.

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

34. Goddard, T. D., and Kneller, D. G. SPARKY 3, University of California, San Francisco.

35. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *Journal of the American Chemical Society 114*, 10663-10665.

36. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity, *Journal of Magnetic Resonance, Series B 103*, 203-216.

37. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *Journal of the American Chemical Society 114*, 6291-6293.

38. Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., and Campbell, I. D. (1990) Comparison of techniques for 1H-detected heteronuclear 1H---15N Spectroscopy, *Journal of Magnetic Resonance (1969) 87*, 488-501.

39. Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *Journal of Magnetic Resonance (1969) 93*, 151-170.

40. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins, *Journal of the American Chemical Society 115*, 7772-7777.

41. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C' chemical shifts in 15N---13C-labelled proteins, *Journal of Magnetic Resonance (1969) 97*, 213-217.

42. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J Biomol NMR 4*, 845-858.

43. Mandel, A. M., Akke, M., and Palmer, A. G., 3rd. (1995) Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme, *J Mol Biol 246*, 144-163.

44. Palmer, 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 carbon-13 heteronuclear NMR spectroscopy, *Journal of the American Chemical Society 113*, 4371-4380.

45. Dosset, P., Hus, J.-C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data, *Journal of Biomolecular NMR 16*, 23-28.

46. Tsan, P., Hus, J.-C., Caffrey, M., Marion, D., and Blackledge, M. (2000) Rotational Diffusion Anisotropy and Local Backbone Dynamics of Carbon Monoxide-Bound Rhodobacter capsulatus Cytochrome c,,, *Journal of the American Chemical Society 122*, 5603-5612.

47. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol Biol 278*, 353-378.

48. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR 13*, 289-302.

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

50. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J Biomol NMR 8*, 477-486.

51. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J Mol Graph 14*, 33-38, 27-38.

52. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

53. Linge, J. P., and Nilges, M. (1999) Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation, *J Biomol NMR 13*, 51-59.

54. Spronk, C. A., Linge, J. P., Hilbers, C. W., and Vuister, G. W. (2002) Improving the quality of protein structures derived by NMR spectroscopy, *J Biomol NMR 22*, 281-289.

55. Xia, B., Tsui, V., Case, D. A., Dyson, H. J., and Wright, P. E. (2002) Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water, *J Biomol NMR 22*, 317-331.

56. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent, *Proteins 50*, 496-506.

57. Lipari, G., and Szabo, A. (1980) Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes, *Biophysical Journal 30*, 489-506.

Chapter 3

Study of the mechanism of interaction between GIP and the Glutaminase L peptide

3.1 Introduction

GIP has been shown to be important as a scaffolding protein in the mammalian brain by demonstrating its association with Glutaminase L in astrocytes and neurons (*1*). Activated Glutaminase catalyzes the production of glutamate and ammonia from the substrate glutamine, which is an important energy generation reaction in mammalian tissues (*2*). Various other functions of Glutaminase have been reported including involvement in synaptic transmission, hepatic ureagenesis, renal ammoniagenesis and regulation of cerebral concentrations of glutamine and glutamate (*3, 4*). Two different gene loci in two different chromosomes encode two different isoforms of the enzyme. They are kidney-type (K) isozyme (encoded by a gene located in chromosome 2) and liver-type (L) isozyme (encoded by a gene located in chromosome 12) (*5*). Localization of these two isozymes has been demonstrated by immunostaining (*6*). For Glutaminase L, the compartment is neuronal nuclei and Glutaminase K has been found in mitochondria. This suggests that GIP plays a role in the determination of the subcellular distribution of Glutaminase L and, also, in possible interactions with other nuclear proteins (*6*). The presence of the class I binding motif (ESMV-COOH) at the C-terminal end of Glutaminase L, but not in Glutaminase K, allows these two isozymes to be differentially regulated and spatially localized, even when they are present in the same tissue (*5*). Glutamine catabolism is a key pathway in the energy generation processes of both tumor cells and normally dividing cells

(*7-9*). Several PDZ domain-containing proteins such as alpha-1-syntrophin (SNT) and GIP has been reported to interact with the C-terminus of Glutaminase L (*10*).

3.2 Objective of the study

To understand the mechanism by which Glutaminase L interacts with GIP, it is important to determine the structure of GIP in complex with Glutaminase L. PDZ domains interact with the C-terminus of the interacting partner, and it has been reported that peptides representing the C-terminal end of the binding partner can act as surrogates for the corresponding partner proteins in vitro (11). Thus, the study of the binding of GIP was carried out with a peptide mimic of the C-terminus of Glutaminase L that would essentially reflect the real binding between GIP and Glutaminase L. In this chapter, we determined the first solution NMR structure of GIP bound to a C-terminal peptide used as a surrogate for Glutaminase L. The C-terminal Glutaminase L peptide, hereinafter referred to as the Glutaminase L peptide has the KENLESMV sequence. Also, to understand how the addition of Glutaminase L peptide affects the dynamics of the protein, the dynamics of the GIP-Glutaminase L peptide complex has been investigated and compared with that of the free GIP. Important insights into the binding mechanism have been gained by demonstration of perturbation of both NMR chemical shifts and backbone dynamics within GIP through ligand binding. Comparison of the structural analysis between the free and bound states of GIP enables to learn the mechanism of interaction between GIP and Glutaminase L peptide. With this information, it is possible to design a small molecule inhibitor for GIP as a potential drug candidate for the treatment of cancer. In addition, because of its promiscuity for

having many binding partners, such an inhibitor could prove to be effective against a number of class I PDZ domains with the possibility of treatment of other diseases (*10*).

3.3 Materials and Methods

The research work described here was carried out in the laboratory of Dr. Smita Mohanty.

3.3.1 Cloning, over-expression and purification of ^{15}N, ^{13}C-labeled GIP

Following the method developed previously in Dr. Smita Mohanty's laboratory (*11*), transformation, over-expression and purification of ^{15}N, ^{13}C-labeled GIP described below was carried out.

3.3.1.1 Transformation of *E. coli* BL21DE3pLysS cells with the recombinant plasmid pET-3c/GIP

SOC (Super Optimal broth with Catabolite repression) medium and LB (Lysogeny Broth) agar medium was incubated at 37 ^{0}C. Both the competent cells (*E. coli* BL21DE3pLysS) and the plasmid (pET-3c/GIP) were thawed on ice ~30 min. 1 µL of plasmid was added to the cells and mixed gently with the pipette tip. The competent cells with the added plasmid were kept on ice for ~20 minutes. The cells were then heat shocked by putting them in the water bath set exactly at 42 ^{0}C for 45 seconds. To reduce the shocks to the cells, they were transferred to water mixed

with ice and kept in there for an additional 20 minutes. After that, 200 µL of SOC medium was added to the cells under sterile conditions. The cells were then incubated in the 37 ^{0}C shaker for 15 min. 50 µL of the cells containing SOC medium were spread-plated on the LB-agar plate. The plate was incubated overnight at 37 ^{0}C. After completion of ~ 16 hours, when the colonies grew visibly, the plate was sealed with parafilm and kept at 4 ^{0}C.

3.3.1.2 Preparation of overnight culture

500 mL of the M9 minimal medium was prepared having the following composition:

KH_2PO_4	6.5 g
K_2HPO_4	5 g
Na_2HPO_4 (anhydrous)	4.5 g
K_2SO_4	1.2 g
$^{15}NH_4Cl$	0.6 g

The volume was adjusted to 500 mL and the medium was sterilized by autoclaving. To the M9 minimal medium, the following nutrients and antibiotics were added aseptically:

20% ^{13}C glucose	10 mL
5mg/mL Thiamine	2.5 mL
1M $MgSO_4$	1 mL
Yeast extract	1 mL
0.1 M $CaCl_2$	250 µL

Trace elements	2.5 mL
100 mg/mL ampicillin	520 µL

Three flasks of 15 mL of sterile M9 minimal media were prepared. To each of the flask, one single colony of the transformed cells from LB-agar plate was added using sterile tips. All the three flasks were transferred immediately to the 37 ^{0}C shaker for overnight incubation.

3.3.1.3 Expression of the protein in batch culture

After measuring the OD_{600} of all three overnight cultures, the culture with the maximum OD_{600} was used to inoculate the M9 medium to a final OD_{600} of ~0.1 in 500 mL medium. The flask was immediately transferred to the 37 ^{0}C shaker. The OD_{600} of the undiluted culture was checked intermittently for every 1.5-2 hours. Once the OD_{600} reached 0.4-0.5, the culture was induced by adding 500 µl of 1M IPTG. The flask was transferred immediately to the shaker to continue incubation at 30 ^{0}C for ~15 hrs.

3.3.1.4 Cell harvest and lysis

The cell culture was kept on ice for chilling for ~30 min. Centrifugation of the cell culture was done at 8000 rpm for 30 min at 4 ^{0}C. The supernatant was discarded. The cells were frozen in liquid nitrogen for 5 minutes. They were subsequently incubated on ice to thaw for ~1.5-2 hours. Once the cell mass becomes fluid, the sample was again frozen in liquid nitrogen for 5 minutes. This procedure was repeated 5 to 6 times. A lysis mixture was prepared with a crushed half-tablet of cocktail protease inhibitor, 10 mL lysis buffer (50 mM phosphate buffer at

pH 8, 4 mM Ethylene Diamine Tetra Aceticacid (EDTA), 200 mM NaCl and 4 % glycerol) and 150 µL of 0.1 M phenylmethylsulfonyl fluoride (PMSF). The cells were mixed with this mixture and sonicated with a 10 second pulse for 8-10 times. The cell lysate was then centrifuged at 12,000 rpm for 25 min at 4 ^{0}C. The supernatant was carefully stored at 4 ^{0}C.

3.3.1.5 Protein purification

The protein was purified from the supernatant by a single-step FPLC method of purification using a Sephacryl S-100 column (*11*). The buffer used for this size-exclusion chromatography was 20 mM sodium phosphate buffer at pH 6.5 containing 150 mM NaCl, 1 mM EDTA and 0.01% (w/v) sodium azide. The protein was collected from the fraction no. 41-44. These fractions were pooled together.

3.3.1.6 NMR sample preparation

The pooled fraction was concentrated down to ~1 mL, to which, then 10 mL of NMR buffer (50 mM phosphate buffer at pH 6.5 containing 5% D_2O, 1 mM EDTA and 0.01% (w/v) NaN_3) was added. This was again concentrated down to 1 mL and another 10 mL of NMR buffer was added. It was then concentrated down to ~ 1 ml. The OD_{280} of the sample was checked to determine the protein concentration and it was stored at 4 ^{0}C. Finally, 50 µL of D_2O was added to the sample to make a final concentration of 5 % D_2O for NMR experiments.

3.3.2 NMR Data collection

All NMR data were collected on a Bruker Avance 600 MHz spectrometer with a triple resonance $^{1}H/^{13}C/^{15}N$ TCI cryoprobe equipped with z-axis pulsed field gradients at either the Department of Chemistry and Biochemistry, Auburn University, Auburn, AL, Bruker BioSpin Corporation**, Billerica, MA,** or the New York Structural Biology Center, New York, NY. The data were processed using NMRPipe (*12*) and Sparky (*13*). For structure determination, samples between 500 μM and 1 mM of uniformly $^{15}N/^{13}C$-labeled GIP in 50 mM phosphate buffer containing 5% D_2O pH 6.5, 1 mM EDTA and 0.01% (w/v) NaN_3 were prepared with addition of the Glutaminase L peptide (Chi Scientific, Maynard, MA, USA) at a 1:3 protein to peptide ratio. All NMR experiments were performed at 298 K. Dynamics data were collected by Mohiuddin Ovee and David Zoetewey. To determine the ^{15}N T_1 values, NMR spectra were recorded with relaxation delays of 10, 600, 50, 500, 100, 400, 200, 300 and 10 ms. To determine ^{15}N T_2 values, NMR spectra were recorded with delays of 17, 153, 34, 17, 136, 51, 119, 68, 102, 85 and 34 ms. The relaxation times were randomized and some points repeated in order to avoid any systematic errors that may arise when the data are collected sequentially. The relaxation rates were calculated by least squares fitting of peak heights versus relaxation delay to a single exponential decay. Steady state ^{1}H-^{15}N NOE values were calculated from the ratio of peak heights in a pair of NMR spectra acquired with and without proton saturation. For backbone and side-chain assignments of the GIP-Glutaminase L peptide complex the following spectra were recorded at 298 K: 2D $^{1}H,^{15}N$-HSQC (*14*), 3D HNCACB (*15*), 3D CC(CO)NH (*16*), 3D CBCA(CO)NH (*15*), 3D ^{15}N-edited HSQC-TOCSY (*17, 18*) with an 80 ms mixing time, 3D HC(CO)NH (*16*), 3D HNHA (*19*), 3D HNCO (*16*) and 3D HN(CA)CO (*20*). NOE distance restraints were collected from 3D ^{15}N-edited HSQC-NOESY (*17, 18, 21*) and 3D ^{13}C-edited HSQC-NOESY (*17, 18, 21*) with the ^{13}C carrier frequency in the aliphatic (44 ppm) and aromatic (125 ppm)

regions and mixing times of 140 for ^{15}N and 110 ms for ^{13}C, respectively. For complex structure determination of GIP with the Glutaminase L peptide, selectively filtered 2D NOESY(*22*) with a mixing time of 100 ms, 3D ^{15}N-filtered and 3D ^{13}C-filtered NOESY experiments, each with mixing times of 120 ms, were performed (*23*). The backbone and side-chain assignments of the Glutaminase L peptide were obtained with an unlabeled peptide sample (~4mM) from the following spectra: 2D $^{1}H,^{15}N$-HSQC, 2D ^{1}H-^{13}C-HMQC, homonuclear 2D TOCSY (*24*) and ROESY (*25*) each with a mixing time of 60 ms (*10*).

3.3.3 Analysis of dynamics data

Dynamics data were analyzed by Mohiuddin Ovee along with Dr. David Zoetewey in Dr. Smita Mohanty's laboratory. Measured relaxation parameters R_1, R_2 and the steady-state ^{1}H-^{15}N NOE for each residue were used as inputs in the Modelfree 4.15 program developed by Palmer et al (*26, 27*) to analyze ^{15}N-backbone dynamics. The τ_c value for GIP-Glutaminase L peptide complex was calculated using the program Tensor2 for the core region A11-Q112 (*28, 29*). Of five different models, the best one was chosen according to the selection criteria (*26*) to get the order parameter (S^2) that represents the degree of spatial restriction within the ^{1}H-^{15}N bond vector. These values range from zero for completely isotropic internal motions to unity for totally restricted motion and represent dynamics on the picosecond to nanosecond time scale (*10*).

3.3.4 Structure calculation and refinement

A total of 2866 NOE cross peaks were assigned manually using Sparky (*13*) for the GIP-Glutaminase L peptide complex. The assignments were corrected or confirmed with both the CANDID module of CYANA 1.0.6 and NOEASSIGN module of CYANA 2.1 (*30*), using the standard protocol of eight iterative cycles of NOE assignment and structure calculation. The CANDID module of CYANA 1.0.6 was used on the complex to initially fit the Glutaminase L peptide into the binding pocket of GIP because it allowed the intermolecular assignments to be fixed separately from the intramolecular assignments. To calculate the complex structure, 36 glycine residues were added as a flexible linker between the protein and the peptide. A total of 118 dihedral angles restrains were derived from the TALOS (*31*) program based on the chemical shift index (CSI) and primary sequence of GIP for protein-peptide complex calculations. Additionally, a total of 66 hydrogen bond distance restraints (two restraints per bond) for the protein-peptide complex were derived from the CSI by TALOS. During the iterative NOE assignments, a total of 490 assignments for the GIP-Glutaminase L peptide complex were removed due to overlap, redundancy, or unresolved ambiguity that resulted from low stringency in the initial peak picking phase and high stringency in the final assignments. The final assignments averaged over 18 and 12 NOEs per residue for protein in the complex, and for the peptide in the complex, respectively. Final refinement of the 100 lowest energy structures of the 200 total calculated structures was performed with the water refinement protocol implemented in ARIA (*32*). The 20 structures with the lowest potential energy and best Ramachandran statistics as assessed by PROCHECK (*33*) were selected for analysis. The structures were visualized with VMD and figures were created using Pymol (*34, 35*). **Table 3.3** shows the complete structural statistics for structure of GIP in complex with the Glutaminase L peptide (*10*).

3.4 Results

3.4.1 Protein Expression

As described above, ^{15}N, ^{13}C-labeled GIP was expressed in *E. coli* cells grown in minimal media supplemented with ^{15}N-labeled ammonium chloride and ^{13}C-labeled glucose (**Figure 3.1**). As seen in NMR studies later on, the isotope labeling of the protein GIP was successful since both the isotope labeled nitrogen and carbon nuclei provided good signals in NMR. However, the initial efforts to isotopically label the protein using the Lysogeny Broth (LB) medium as a growth medium for overnight cell cultures produced inhomogeneous labeling of the nucleus (carbon or nitrogen) even though the starter culture was diluted 25 times in the minimal media. Still use of such a small percentage of LB media was sufficient for the dilution of the isotope labeling; making it impossible to carry out isotope filtered experiments. A simple 1D NMR experiment (*36*) was carried out to check the homogeneity of the isotope labeling (**Figure 3.2**). Methyl ($-CH_3$) protons of Leucine 108 of GIP appears in a 1D NMR spectrum at a value of less than zero in the ppm scale which is completely separate from any other peaks of the spin-active nuclei. Thus, observing the splitting patterns of the methyl ($-CH_3$) protons of that residue would help to determine the homogeneity of the isotope labeling of the protein. If there is non-homogeneous isotope labeling, then there would be still spin-inactive ^{12}C present in the protein which would cause no splitting of the ($-CH_3$) protons resulting in a single proton peak. However, the available ^{13}C present in such case, would still split ($-CH_3$) protons, thus, the resulting 1D spectrum should have three peaks for the ($-CH_3$) protons (**Figure 3.2**). But, if the isotope labeling is homogeneous, then there should only be two peaks for the ($-CH_3$) protons resulting from the splitting by ^{13}C (**Figure 3.3**). Protein samples prepared from the earlier

protocol (using LB medium) gave three peaks for the (-CH_3) protons of leucine 108 of the protein. Of which, the intermediate peak is due to the contribution from the (-CH_3) protons attached to ^{12}C, whereas, two adjacent peaks on both sides of the middle peak is produced by the splitting of (-CH_3) protons attached to ^{13}C (**Figure 3.2**). The comparatively higher intensity of the middle peak compared to the two shoulder peaks indicated presence of a higher percentage of unlabeled spin-inactive ^{12}C nucleus in the protein (**Figure 3.2**). This suggested that the GIP proteins prepared using LB medium as the growth medium for the overnight culture was not homogeneously isotope labeled. To achieve homogeneous isotope labeling of GIP protein, the protocol was changed. The growth medium used for the starter culture was changed to M9 minimal medium as well. The purified GIP protein from such expression was checked for the homogeneity of the isotope labeling and this protein sample showed almost 100% isotope labeling (**Figure 3.3**). For the purpose of the structure determination of GIP-Glutaminase L peptide complex, all the ^{13}C, ^{15}N-labeled GIP protein was produced following the latest protocol. The production of homogeneously labeled protein was a prerequisite for the successful operation of the filtered NOESY experiments. Thus, confirmation of an available homogeneously isotope labeled GIP protein was a very important step in the determination of complex structure. This also wonderfully shows a practical application of the spin-spin coupling having an impact on the research.

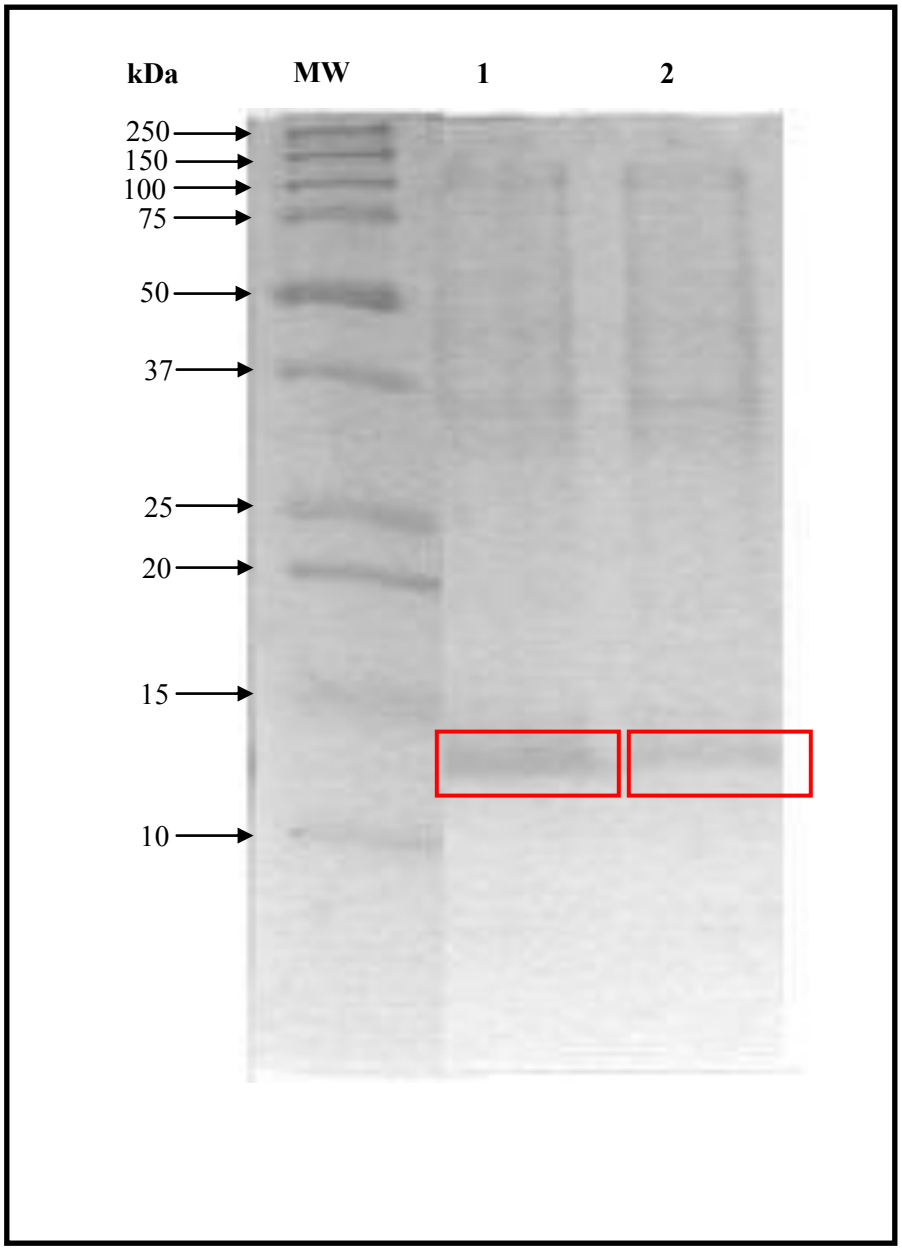


Figure 3.1: Expression of GIP analyzed by SDS-PAGE. Both lane 1 and 2 show the expression of GIP prior to purification (the red rectangle spots the protein of expected size). The lane MW is for protein marker.

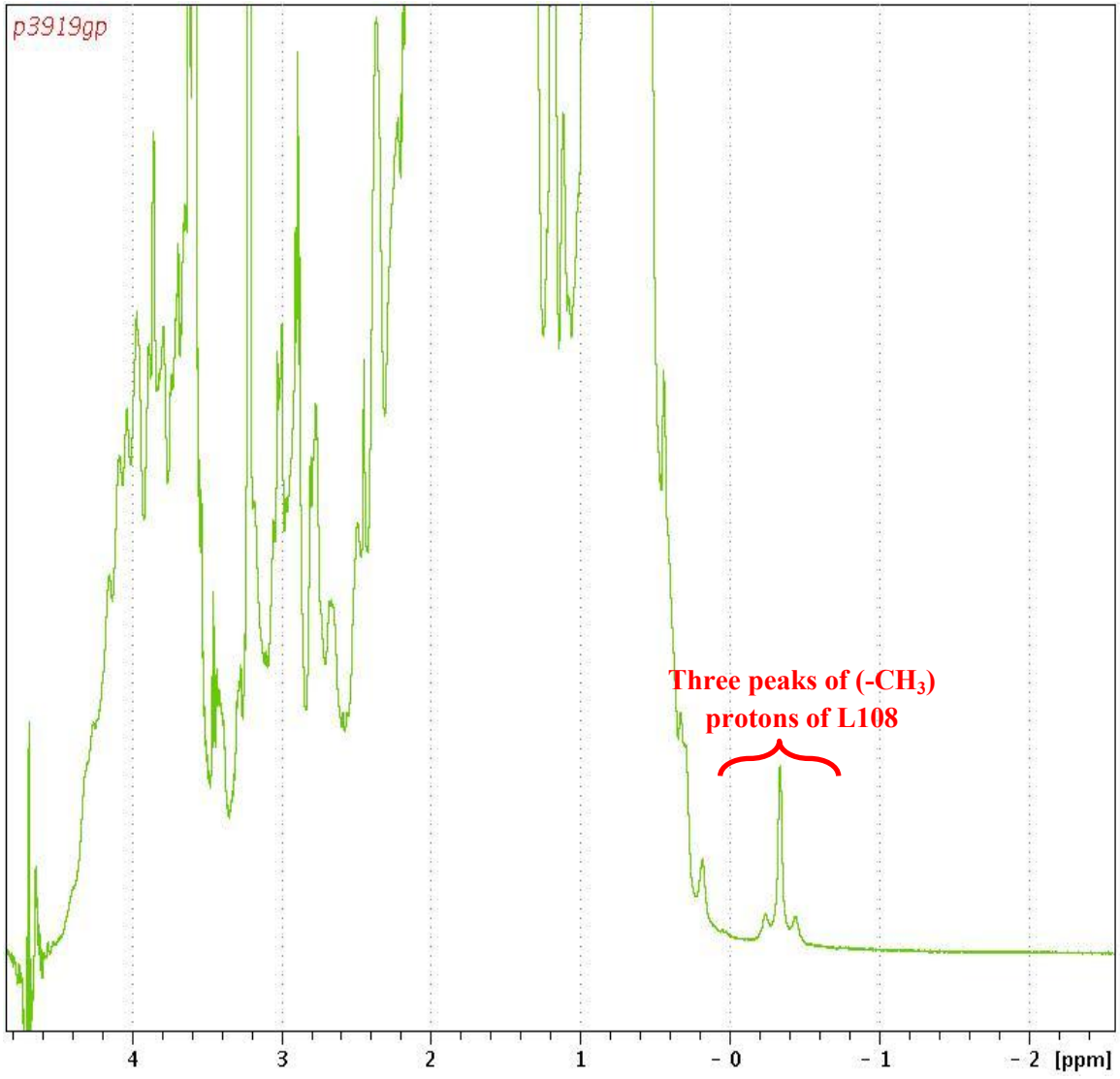


Figure 3.2: 1D NMR spectrum of non-homogeneously labeled GIP sample. p3919gp was the name for the pulse program used for this NMR experiment.

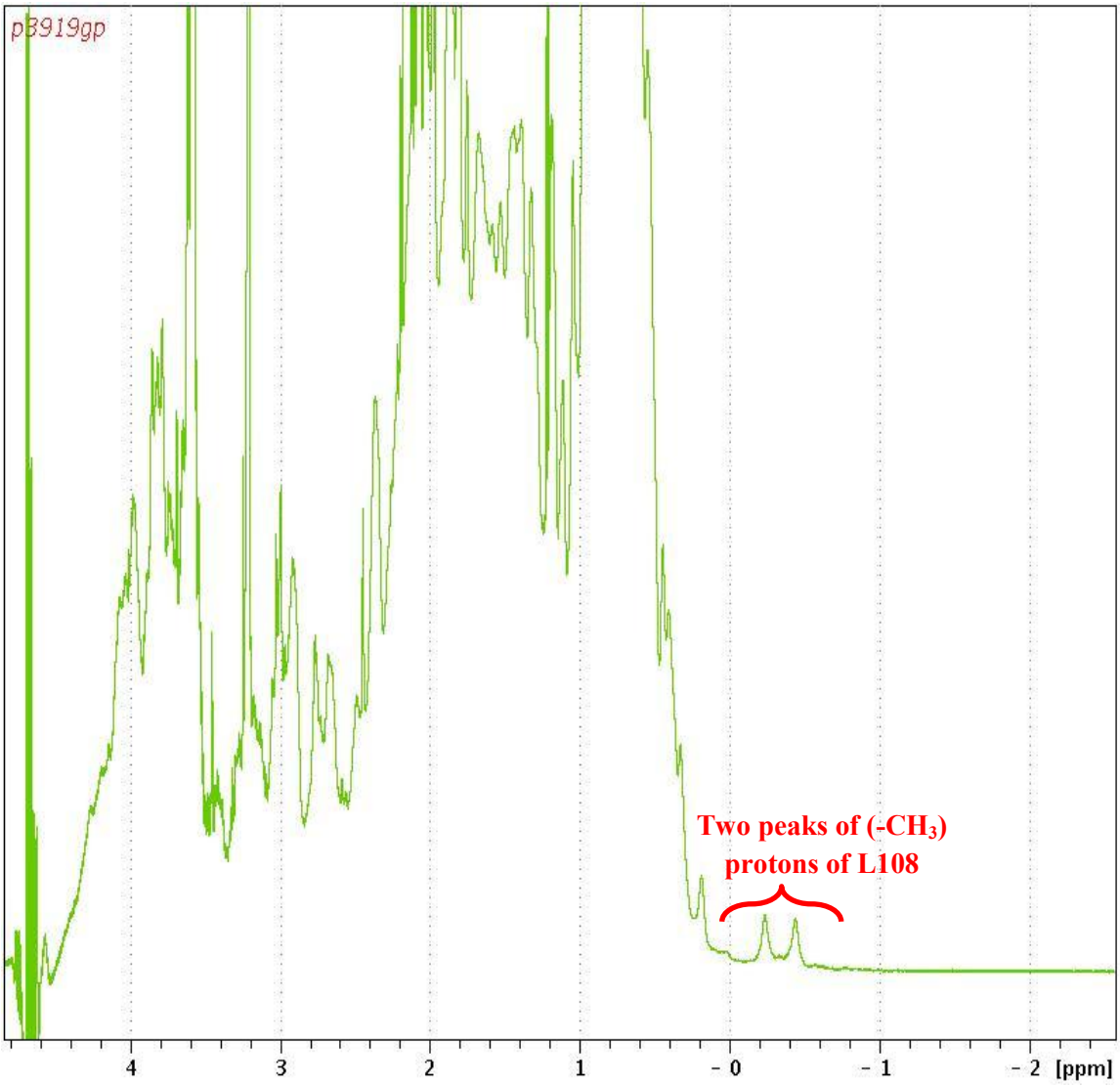


Figure 3.3: 1D NMR spectrum of homogeneously labeled GIP sample. p3919gp was the name for the pulse program used for this NMR experiment.

3.4.2 Protein Purification

Using size-exclusion chromatography as a single step, GIP was purified on a Sephacryl S-100 column (GE Healthcare). The production of the ^{15}N, ^{13}C-labeled recombinant GIP is around 15.2 mg per liter of bacterial culture.

3.4.3 NMR Structure determination of GIP-Glutaminase L peptide complex

3.4.3.1 Effect of peptide binding to the resonances of GIP protein

As GIP was titrated against Glutaminase L peptide, it was possible to track the movement of the resonance peaks in the ^{1}H, ^{15}N-HSQC spectra, because most of the resonances of the protein residues were in the fast chemical exchange on the NMR time scale. However, there were some exceptions. The amino acid residues I18, L21, I28-G35, Q39, D40, Q43, N44, E48, I55, E62, A66, E67, A69 and R96 had peak intensities that either hugely decreased or were below the level of the noise threshold, presumably due to intermediate to slow exchange on the NMR time scale. But, as GIP reaches saturation and the predominant state becomes the bound state, then these undetectable resonances reappeared often in remote regions of the HSQC spectrum relative to their initial positions. For the assignments of the residues that were assumed to be critical to complex formation, such phenomenon produced considerable uncertainty. Such residues include I28-E48 and R96, which are located within the β2 strand, the β2-β3 loop and the α2 helix. This is evidence that GIP interacts with the Glutaminase L peptide primarily through the β-strand addition mechanism (*37*) instead of a direct interaction with the α2 helix. Residues that were not predicted to be part of the binding region also underwent intermediate to slow chemical

exchange, such as I18, I55, and E62-A69, which belong to the β1 and β3 strands and the α1 helix, respectively. This observation points to the fact that, due to the binding interaction between the protein and peptide, there are some long range allosteric interactions within the protein.

Like ^{1}H, ^{15}N-HSQC spectra, the ^{1}H, ^{13}C-HSQC spectra were also significantly different when compared between free GIP and the GIP- Glutaminase L peptide complex. It was quite impossible to assign a number of key protein side-chain nuclei purely based on free GIP assignments, since there were a couple of factors that created this uncertainty. Firstly, there was severe overlap of carbon and proton chemical shifts and, secondly, the protein-peptide interaction resulted in large chemical shift perturbations. Therefore, it was necessary to reassign the whole protein in its complexed state using the following 3D experiments: HNCACB, CBCA(CO)NH, HCC(CO)NH, CC(CO)NH, HSQC-TOCSY and HCCH-TOCSY. This helped to reassign even the residues, such as L27 – G35, which initially disappeared but reappeared in distant locations with the course of the titration of GIP with Glutaminase L peptide. This re-assignment of the protein in the complex was very essential considering the amount of chemical shift perturbations for all of the resonances, both backbone and side-chain. Such significant changes in chemical shifts are nicely illustrated from **Figure 3.4** to **Figure 3.8**. Thus, to proceed with structure calculations, each resonance must be reassigned with accuracy (*10*).

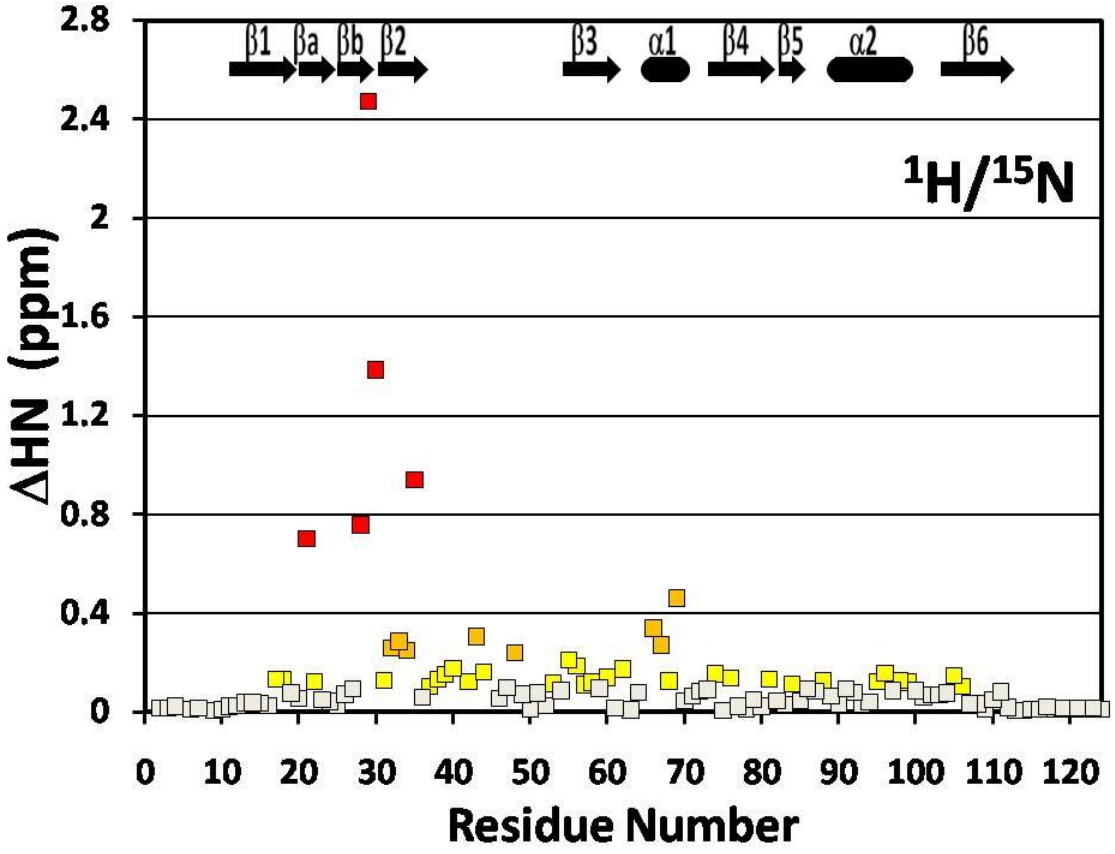


Figure 3.4: Combined ^1H and ^{15}N backbone amide chemical shift perturbations (ΔHN) are plotted as a function of residue number in GIP by the equation $\Delta HN=\{(H_f-H_b)^2+((N_f-N_b)/10)^2\}^{½}$, with 10 as a scaling factor. H_f, H_b, N_f and N_b are the chemical shifts of each residue's amide ^1H and ^{15}N in the free (GIP alone) and bound (GIP-Glutaminase L peptide complex) states, respectively. Adapted from reference (*10*).

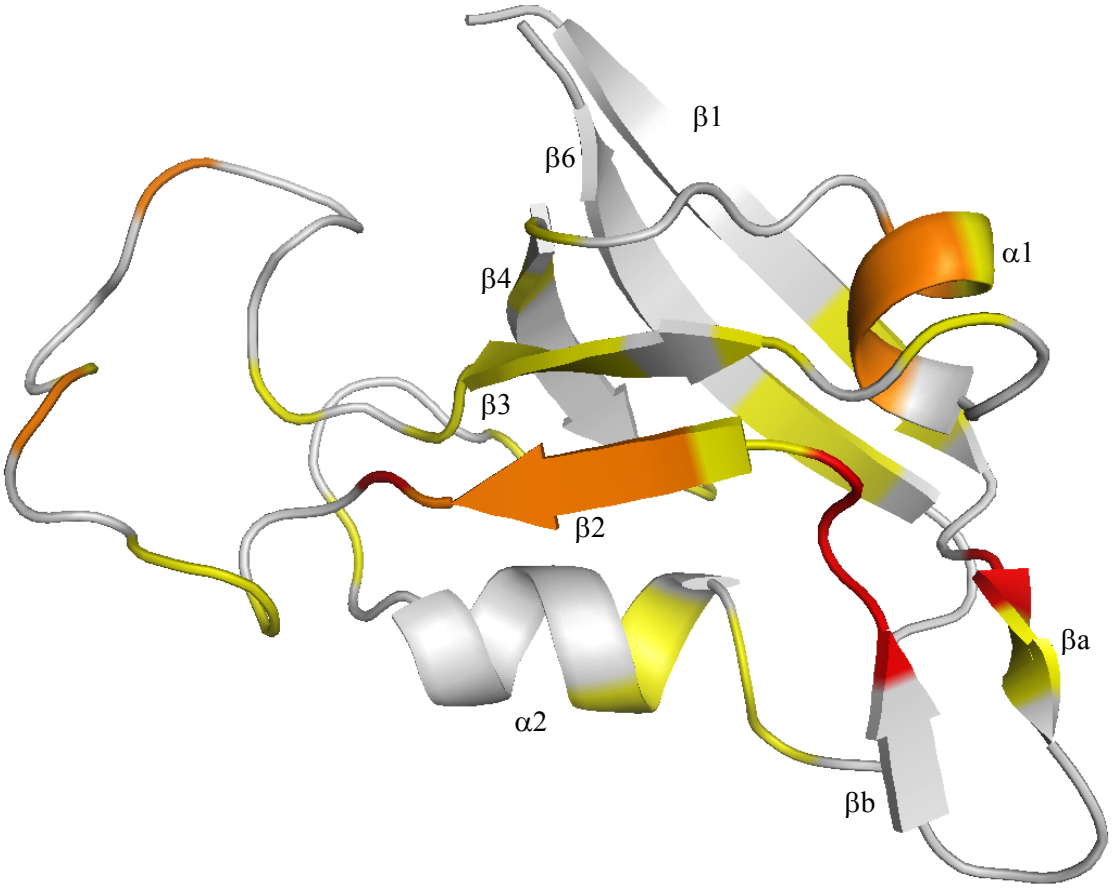


Figure 3.5: The magnitudes of ΔHN presented in **Figure 3.4** are represented as different colors on a ribbon diagram of free GIP. White is < 0.1 ppm, yellow is < 0.2 ppm, orange is < 0.5 ppm and red is > 0.5 ppm. Only residues A11-Q112 are shown as residues M1-T10 and A113-S124 are highly disordered and have chemical shifts perturbations of < 0.05 ppm. Adapted from reference (*10*).

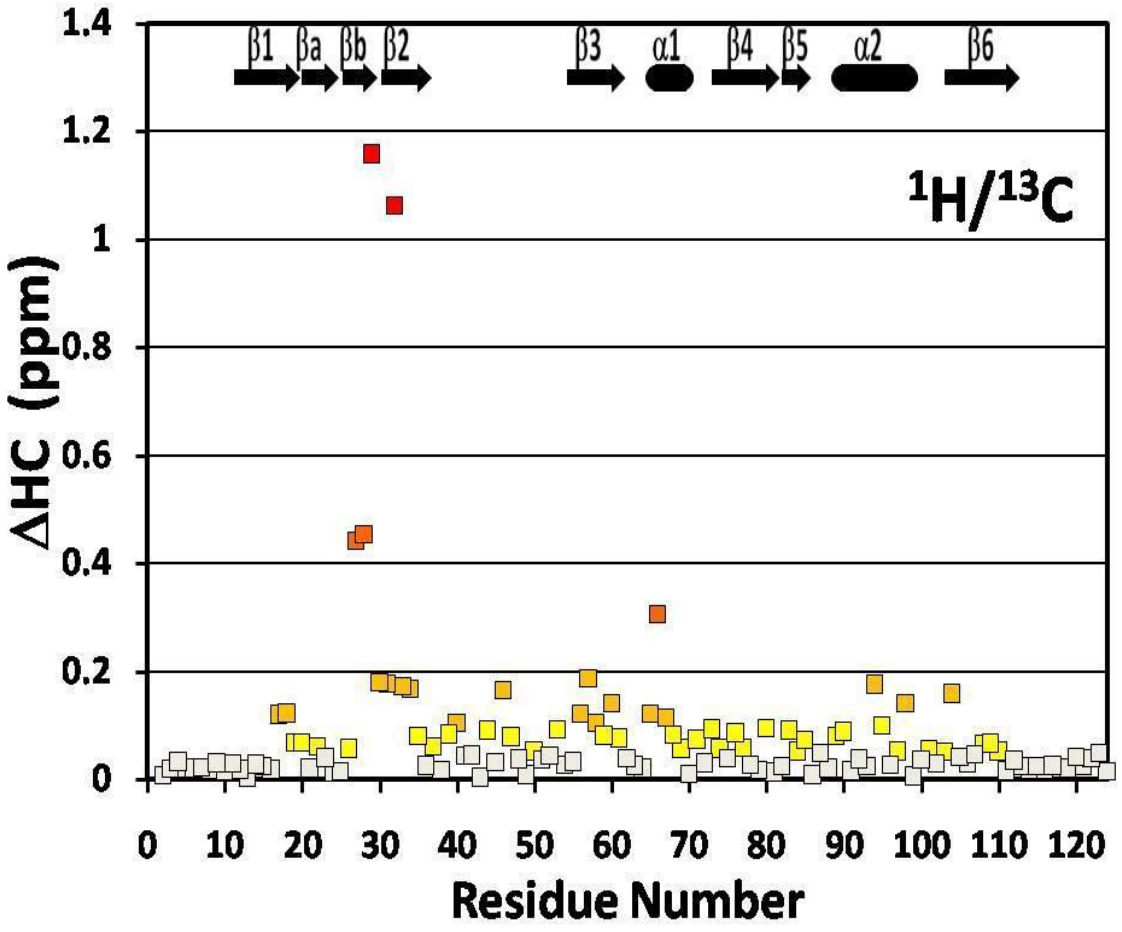


Figure 3.6: Combined HA and CA backbone chemical shift perturbations (ΔHC) are plotted as a function of residue number in GIP by the equation $\Delta HC=\{(H_f-H_b)^2+((C_f-C_b)/4)^2\}^{½}$, with 4 as a scaling factor. H_f, H_b, C_f and C_b are the chemical shifts of each residue's alpha ^1H and ^{13}C in the free (GIP alone) and bound (GIP-Glutaminase L peptide complex) states, respectively. Adapted from reference (*10*).

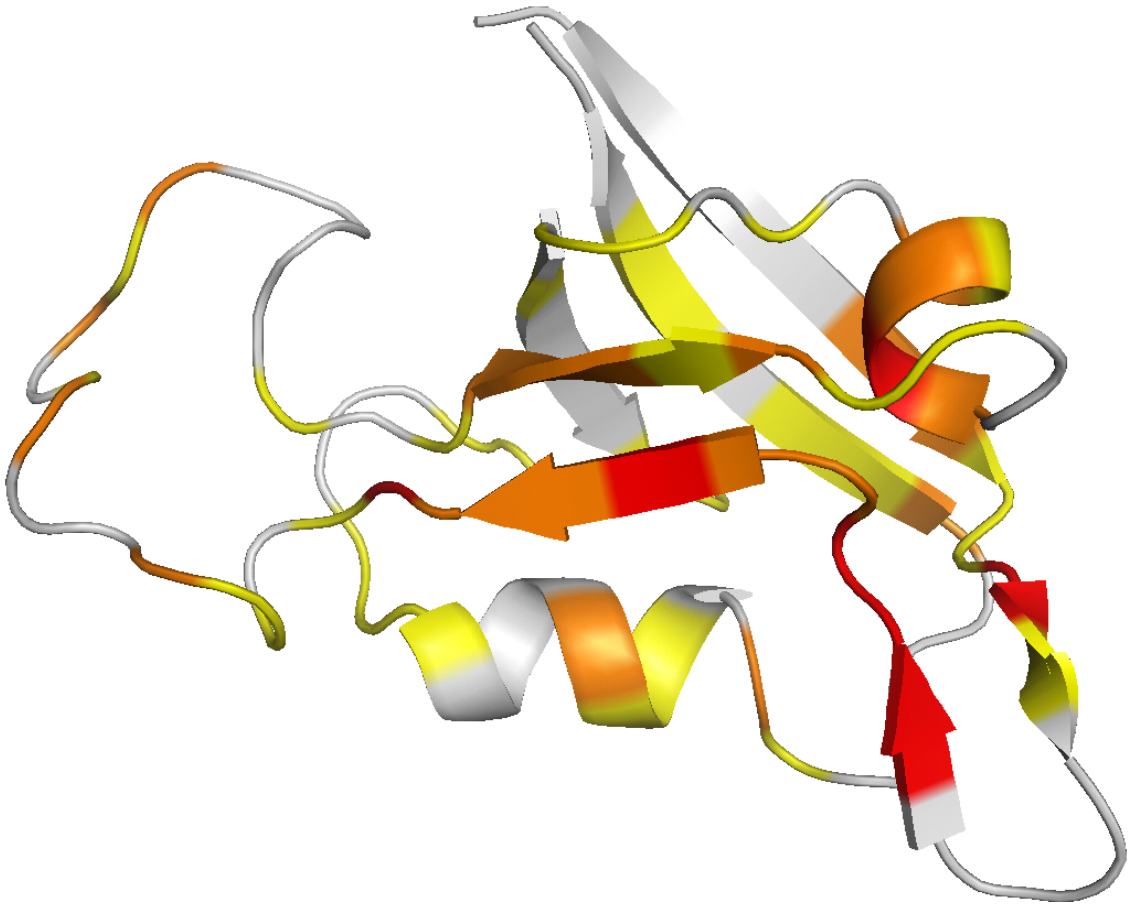


Figure 3.7: The magnitudes of ΔHC presented in **Figure 3.6** are represented as different colors on a ribbon diagram of free GIP. White is < 0.05 ppm, yellow is < 0.1 ppm, orange is < 0.2 ppm, red-orange is < 0.5 ppm and red is > 0.5 ppm. Only residues A11-Q112 are shown as residues M1-T10 and A113-S124 are highly disordered and have chemical shifts perturbations of < 0.05 ppm. Adapted from reference (*10*).

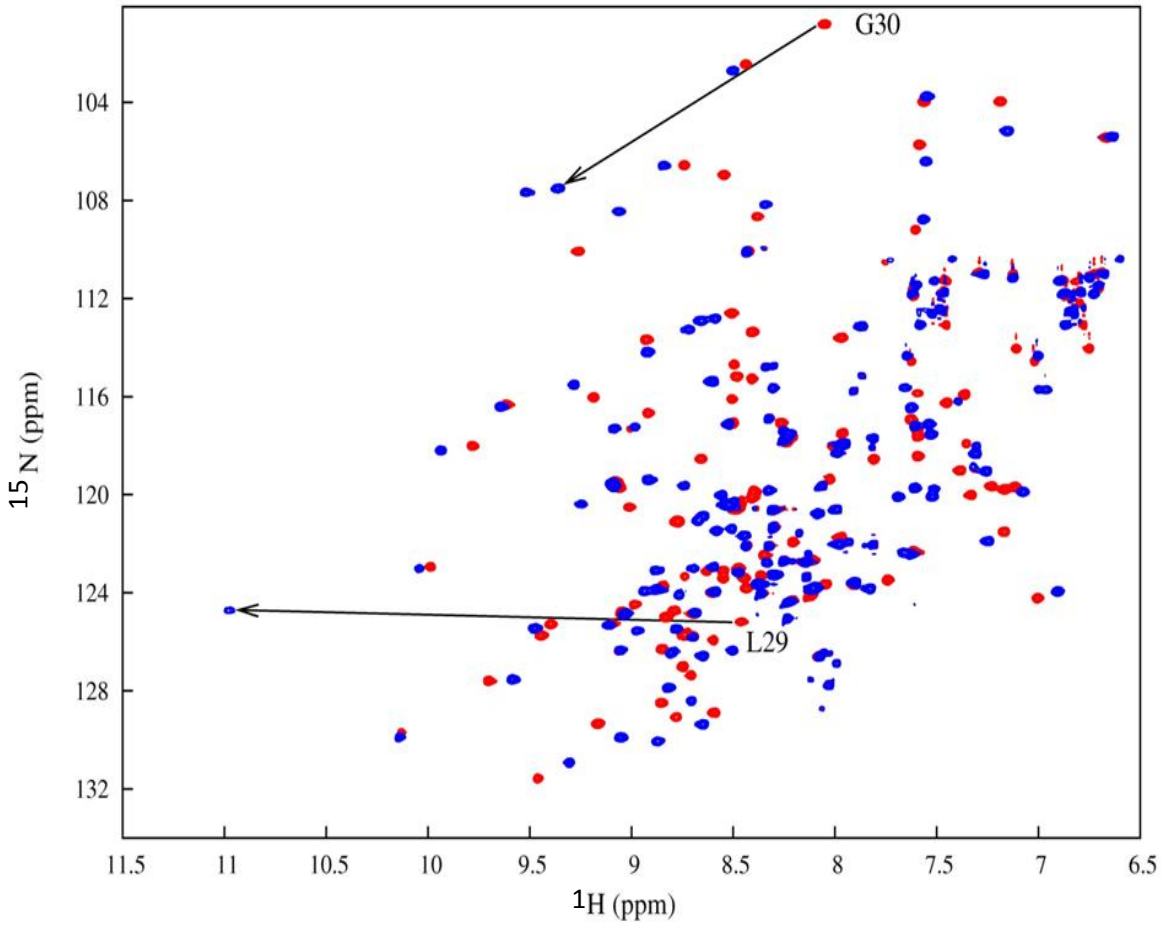


Figure 3.8: An overlay of free GIP is shown in red and GIP-Glutaminase L peptide at a ratio of 1:3 in blue, but at a lower contour threshold to highlight L29. Arrows indicate the dramatic chemical shift perturbations of L29 and G30. Adapted from reference (*10*).

3.4.3.2 Backbone and side-chain assignments

3.4.3.2.1 For protein

With the available sequential assignments for free GIP, assigning the backbone for GIP in its bound form was not hard. A 3D HNCACB experiment was used to perform sequential assignments of the GIP in complexed state (**Figure 3.9**). A CBCA(CO)NH spectrum was quite useful in the confirmation of the assignments of the HNCACB spectrum. However, there were certain peaks in the ^{1}H, ^{15}N-HSQC spectrum **(Figure 3.10)** which required some efforts to identify them for the purpose of acquiring a complete sequential assignment such as L29. This peak goes into an intermediate exchange from fast exchange as the protein goes from the free to the bound state. At a higher concentration of Glutaminase L peptide (1:3 protein to peptide ratio), it re-appears barely at a high contour level in a completely different location (**Figure 3.8**). After assigning certain side-chain and NOESY experiments, assignments of peaks like this one were confirmed.

Several experiments were used to assign side-chains of the protein in its bound form such as HCC(CO)NH, CC(CO)NH, HSQC-TOCSY and HCCH-TOCSY. The statistics of the assignments of the side-chains were summarized in the **Table 3.1**. In summary, around 93, 95 and 92 percent of all carbon, hydrogen and nitrogen nuclei, respectively, were unambiguously assigned.

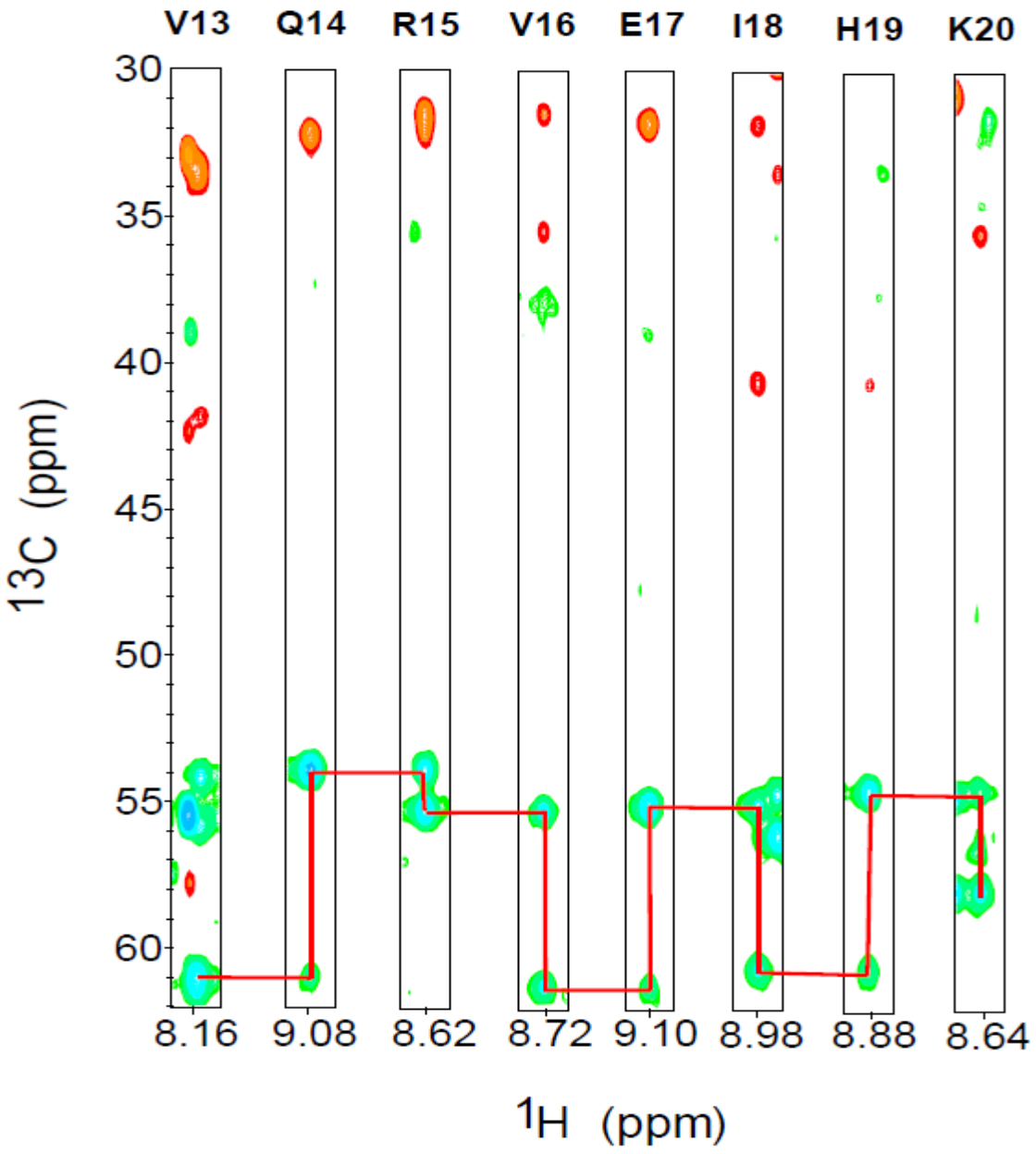


Figure 3.9: Sequential assignments of V13-K20 in the GIP-Glutaminase L peptide complex from (^{1}H, ^{13}C)-strips of the HNCACB experiment. Only the C_α atoms of the residues were connected with red lines to show the sequential assignment. Positive signals are green and negative signals are red. C_α appears as positive signal and C_β appears as negative signal.

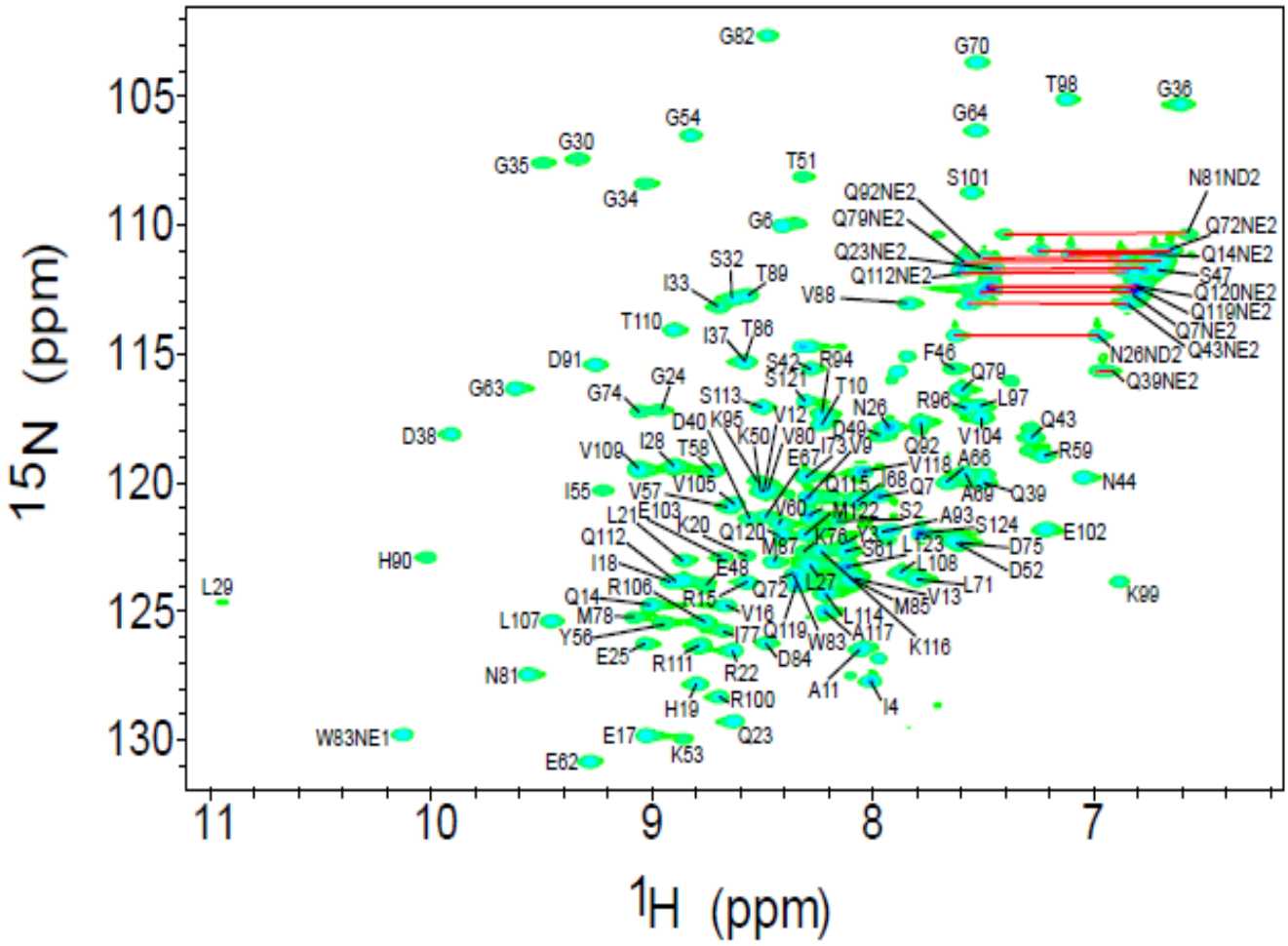


Figure 3.10: ^{1}H, ^{15}N-HSQC spectrum of the GIP-Glutaminase L peptide complex. Red lines connected the non-degenerate protons of the side-chain amide groups of Asparagine and Glutamine residues.

Atom	C (CO)	Cα	Cβ	Cγ	Cδ	Cε	Cζ	Cη	Total C
% of Assignment	0	99	99	87	82	84	50	100	93.2
Found vs. Expected	0/124	123/124	111/112	79/91	37/45	16/19	2/4	1/1	369/396
Atom	HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Total H
% of Assignment	98	99	99	86	98	85	100	100	95.5
Found vs. Expected	117/119	123/124	111/112	78/91	47/48	33/39	4/4	1/1	514/538
Atom	N				Nδ	Nε			Total N
% of Assignment	98				40	70			92
Found vs. Expected	117/119				2/5	16/23			135/147

Table 3.1: Statistics of side-chain assignments of the GIP-Glutaminase L peptide complex.

3.4.3.2.2 For peptide

To assign the resonances for the residues of the Glutaminase L peptide in its bound form, initially, resonances of the peptide residues were assigned from its free from. Subsequently, those resonances were used as a guiding reference for the assignment of the residues of the peptide in the bound form. Assignment of free Glutaminase L peptide was done by ^{1}H, ^{15}N-HSQC, homonuclear 2D TOCSY, ^{1}H, ^{13}C-HMQC and ROESY experiments. The ^{1}H, ^{15}N-HSQC experiment (**Figure 3.11**) was used to assign the amide protons of the Glutaminase L peptide based on the usual chemical shifts for the amide protons of respective amino acids (Biological Magnetic Resonance Data Bank, BMRB, http://www.bmrb.wisc.edu/) and on the assignments from other experiments such as homonuclear 2D TOCSY (**Figure 3.12**), ^{1}H, ^{13}C-HMQC (**Figure 3.13**) and ROESY. Among the eight residues of the peptide, in addition to K301, two (E302 and M307) did not give any peaks in the spectrum for some unknown reasons. Amide protons of K301 and E302 were never assigned. However, amide proton of M307 was assigned from other spectra. 2D TOCSY experiment helped to assign non-degenerate protons of the side-chains of the peptide. Most of the assignments of the side-chains of the free peptide were done in this experiment.

To assign the resonances of the residues of the peptide in its bound state, a special 2D selectively filtered NOESY experiment that results into four different 2D NOESY spectra (*22*), was used. In this experiment, NOEs that arise from protons attached to either $^{12}C/^{14}N$ (peptide) or $^{13}C/^{15}N$ (protein) can be selectively filtered. Thus, there should be one spectrum among the resulting four spectra, which would allow only NOEs that originate from protons attached to ^{12}C

or ^{14}N. Through the comparison with the resonances of the residues of the peptide in its free from, assignment of those of the peptide in its bound form, from such spectrum, was achieved (**Figure 3.14**). Moreover, such an assignment process also helped to determine the structure of the Glutaminase L peptide in its bound form (*10*).

The statistics of the side-chain assignments for Glutaminase L peptide are summarized in **Table 3.2**. Assignment of the available protons of the peptide was solely considered here since only these protons would produce any possible NOE relationship with the protons of the protein. In summary, about 95% of all the possible protons of the peptide were assigned unambiguously.

Atom	HN	Hα	Hβ	Hγ	Hδ	Hε	Total H
% of Assignment	75	100	100	100	100	100	95.2
Found vs. Expected	6/8	8/8	12/12	7/7	5/5	2/2	40/42

Table 3.2: Statistics of available proton assignments of the Glutaminase L peptide.

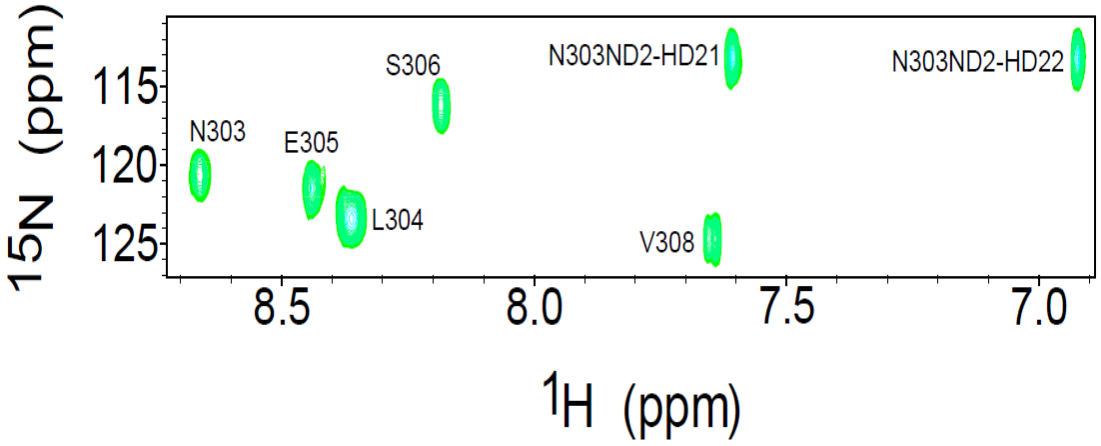


Figure 3.11: ^{1}H, ^{15}N-HSQC spectrum of the Glutaminase L peptide.

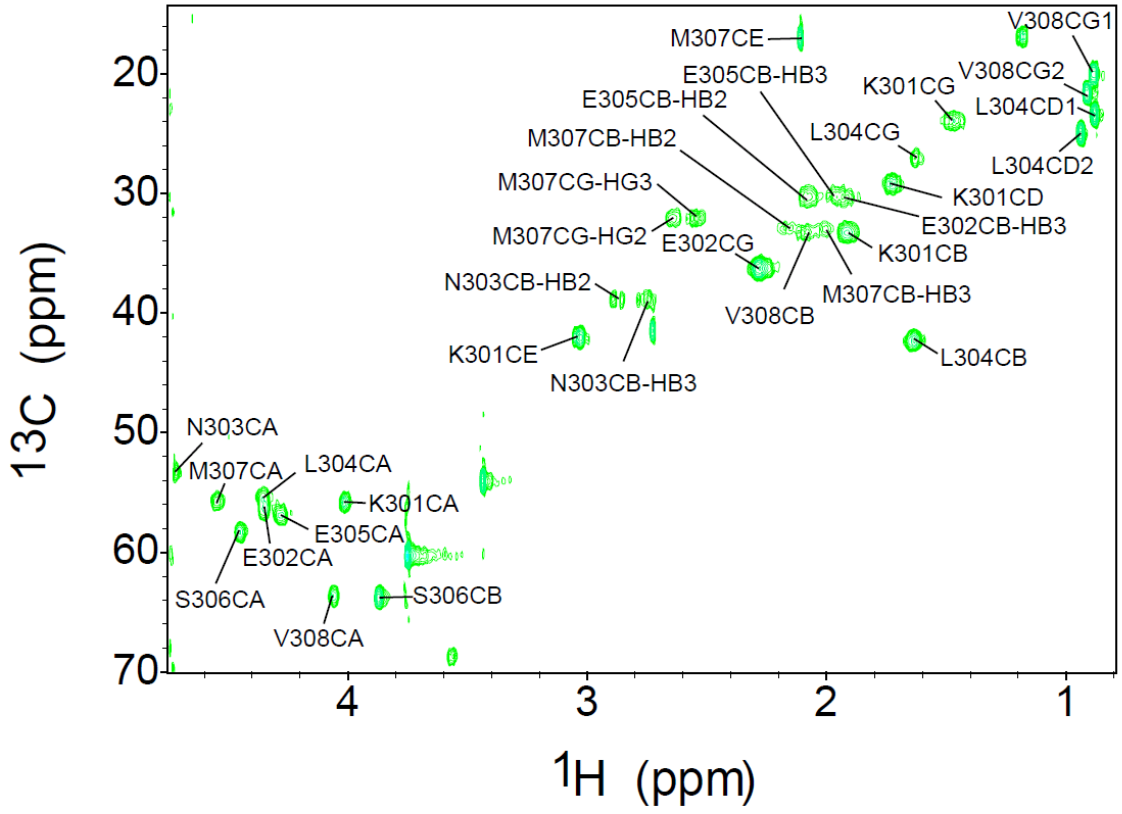


Figure 3.12: ^{1}H, ^{13}C-HMQC spectrum of the Glutaminase L peptide.

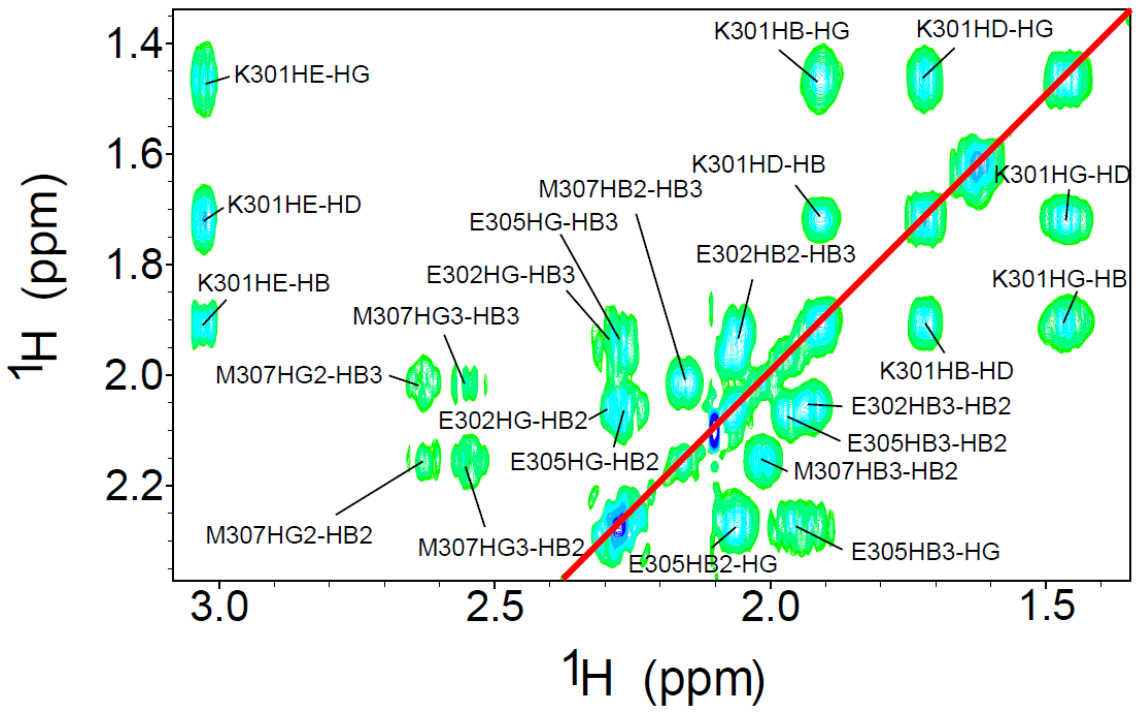


Figure 3.13: Homonuclear 2D TOCSY spectrum of the Glutaminase L peptide. Red line crosses the diagonal peaks. Notice the duplicate peaks on either side of the red line.

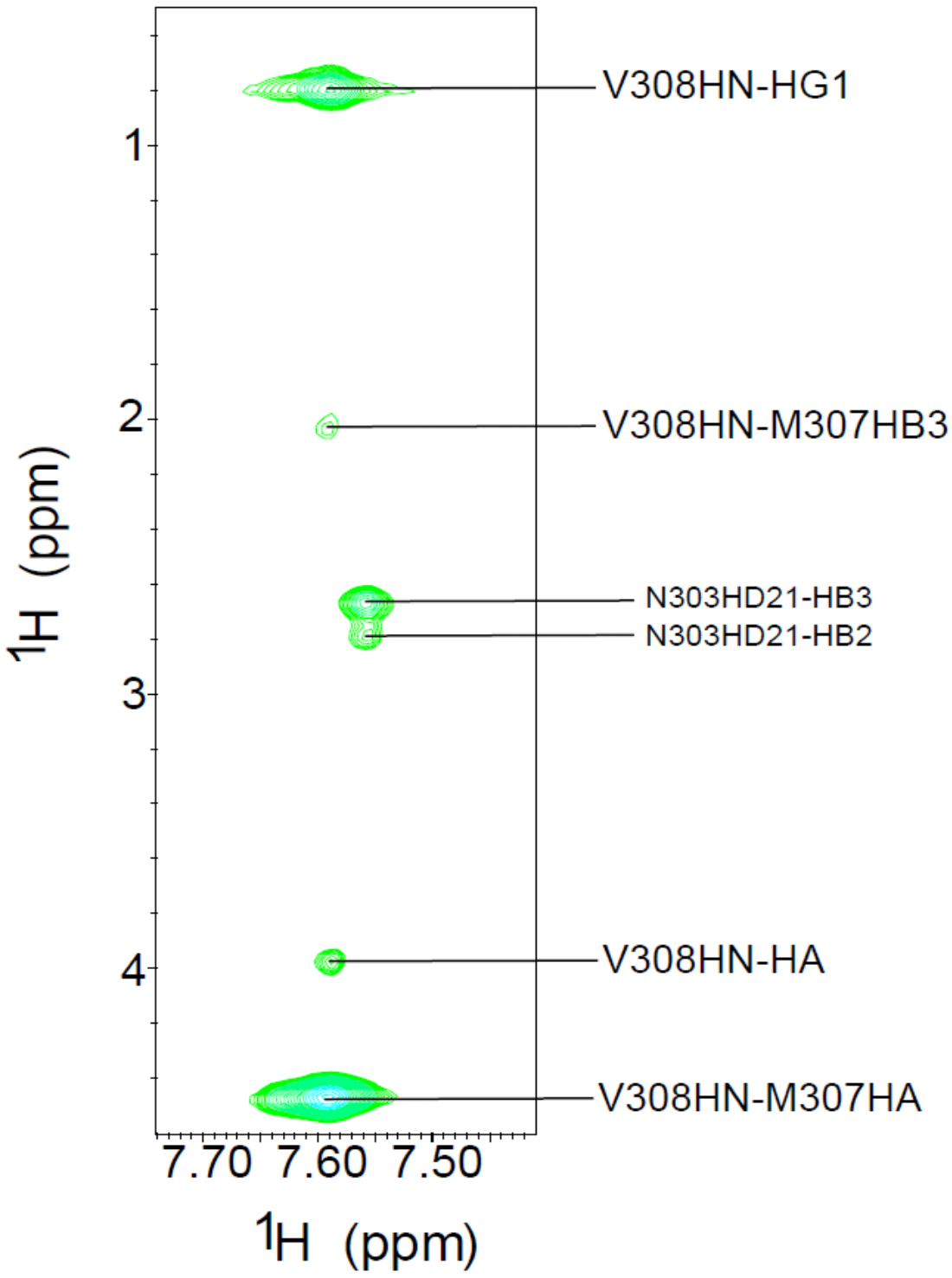


Figure 3.14: 2D selectively filtered NOESY spectrum of the Glutaminase L peptide.

3.4.3.3 NOE assignments

Traditional 3D ^{15}N- and ^{13}C-edited HSQC-NOESY experiments were used to assign the NOEs for GIP in the bound form. Interestingly, several intermolecular NOEs between GIP and Glutaminase L peptide were also assigned in these two experiments. To assign the ^{13}C-edited HSQC-NOESY spectrum (**Figure 3.15**), a 2D ^{1}H, ^{13}C-HSQC spectrum was constructed from the ^{13}C-edited HSQC-NOESY spectrum itself by compressing all the data from the proton z-dimension into a single plane. Although this resulted in a much overlapped spectrum (**Figure 3.16**), the presence of such a base spectrum was extremely helpful in the assignment of the ^{13}C-edited HSQC-NOESY spectrum.

To find intermolecular NOEs between the unlabeled peptide and the ^{13}C, ^{15}N-labeled protein in the complex, F1-filtered/F3-selected NOESY experiments with both ^{15}N/^{14}N and ^{13}C/^{12}C filtering methods were used. Although these filtered experiments were supposed to have only NOEs from unlabeled peptide, it appeared that the experiment was not that stringent and a lot of intramolecular NOEs "bleed through" to add up the ambiguities. To remove ambiguities in the assignments of ^{15}N-filtered HSQC-NOESY experiment, one approach was to do a control experiment with the same pulse sequence on a free GIP sample (**Figure 3.17**). Theoretically, such a spectrum should not have any NOEs. But, since there were "bleeding through", this spectrum was helpful to establish NOEs only from the unlabeled peptide in the filtered NOESY spectrum with the simultaneous comparison to the controlled spectrum (**Figure 3.17**). This way, a good number of possible intermolecular NOEs were manually assigned in both traditional 3D ^{15}N-edited HSQC-NOESY and ^{15}N-filtered HSQC-NOESY spectra.

Usually, standard ^{13}C-filtered NOESY is the experiment that is most often used for the determination of the structure of a complex. When compared with most of the other complexes of PDZ domains, the GIP-Glutaminase L peptide complex appears to have much fewer observable NOEs in the ^{13}C-filtered NOESY spectrum. The reason behind the lack of observable NOEs is due to line broadening resulting from intermediate to slow exchange of residues in the entire β2 strand. Thus, only the strongest NOEs were seen which are very important in the ligand binding. Initially, the assignments of intermolecular NOEs done on the traditional 3D (unfiltered) NOESY spectrum were ambiguous. However, with the establishment of the peptide's relative position in the binding site, those ambiguities could be sorted out. These additional unambiguous assignments were very instrumental for the final structure calculation as they added up to the total number of intermolecular NOEs (*10*).

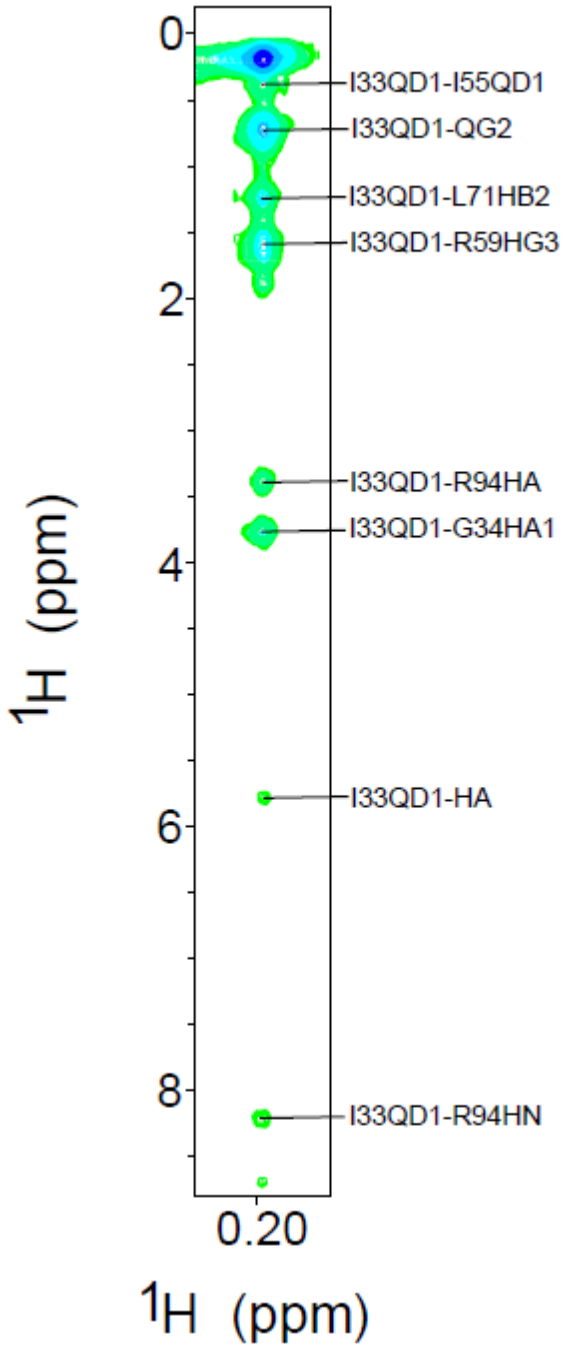


Figure 3.15: ^{13}C-edited HSQC-NOESY spectrum of the I33QD1 proton of GIP in its bound form. The assignments shown here were manually picked in Sparky which were later confirmed, removed or corrected in the iterative process.

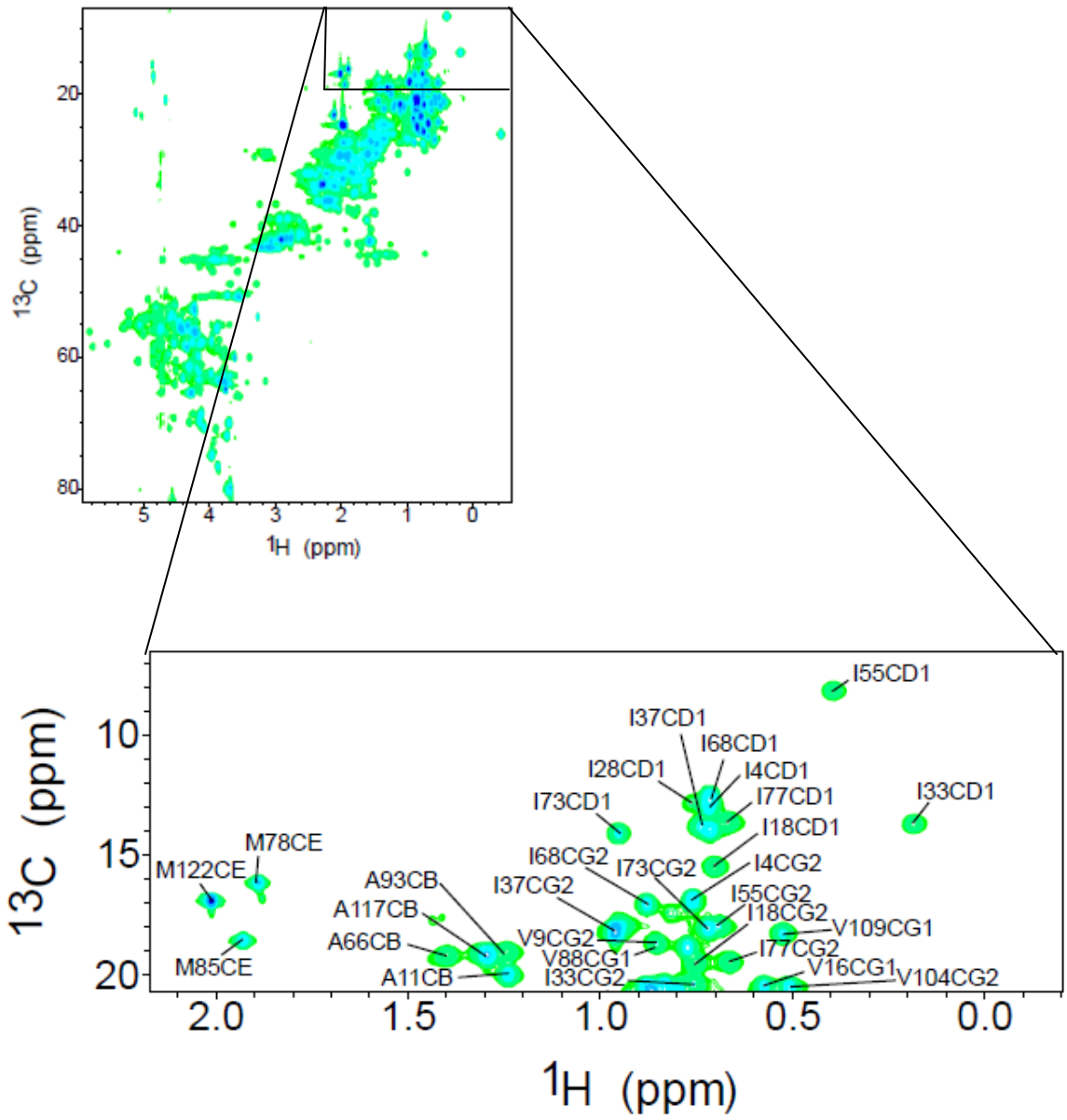


Figure 3.16: ^{1}H, ^{13}C-HSQC spectrum of GIP in the bound form. *Top*- Full spectrum, *Bottom*- Part of the spectrum was blown up and shown with assignments.

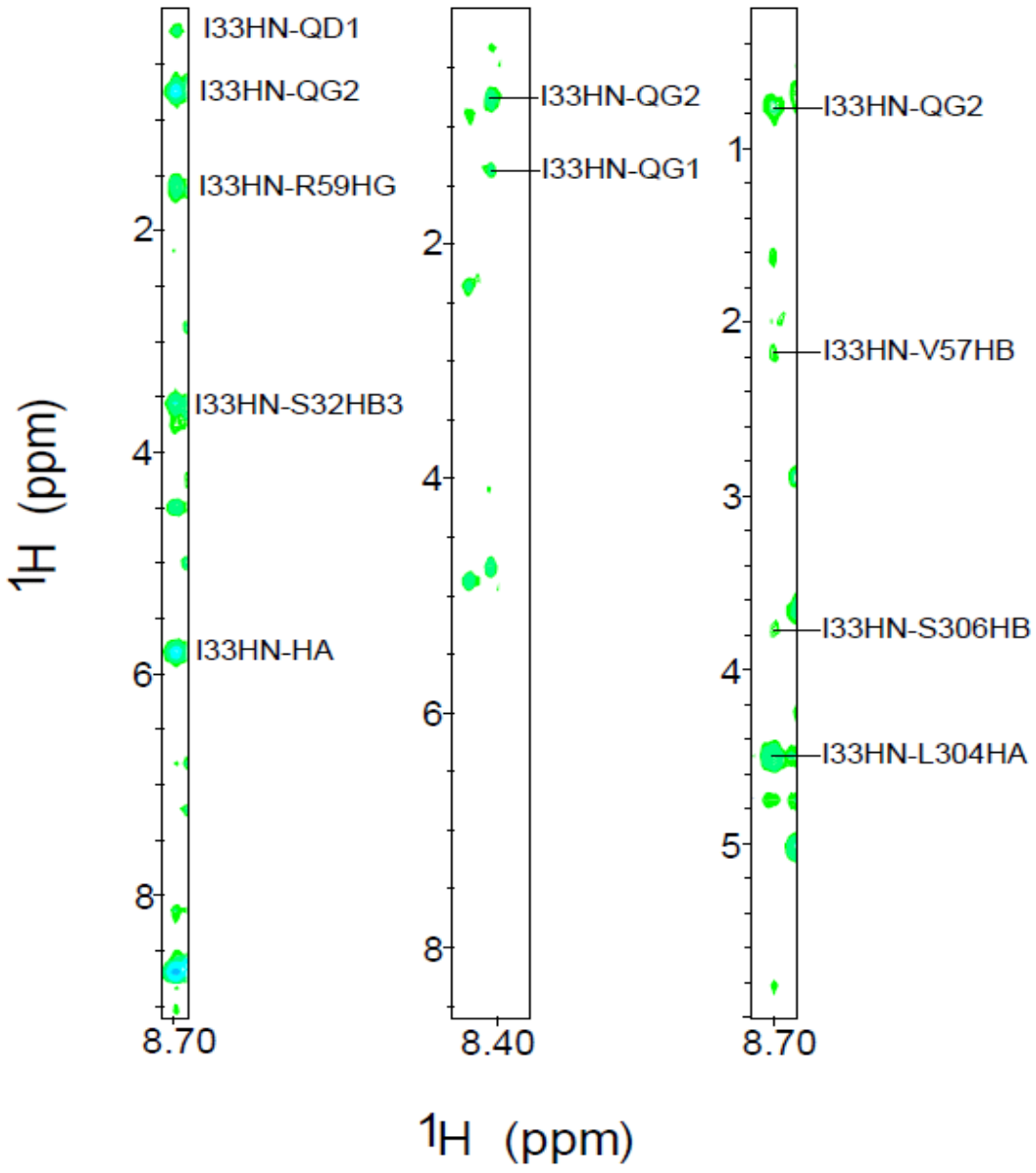


Figure 3.17: Three different HSQC-NOESY spectra of I33 residue of the GIP. *Left*- Traditional 3D ^{15}N-edited HSQC-NOESY spectrum; *middle*- ^{15}N-filtered HSQC-NOESY control spectrum; *right*- ^{15}N-filtered HSQC-NOESY spectrum. The assignments shown here were manually picked in Sparky which were later confirmed, removed or corrected in the iterative process.

3.4.3.4 Structure calculation

Initially, a total of 2866 NOE cross peaks were assigned manually for the GIP-Glutaminase L peptide complex. But, as with the free GIP structure calculation, during the iterative process of GIP-Glutaminase L peptide complex structure calculation, a total of 490 assignments were removed. The selective formation of specific hydrogen bonds between the negatively charged C-terminal Val carboxyl oxygens from the Glutaminase L peptide to the amide protons of L29 and G30 from GIP could be directly identified from their very large induced chemical shift perturbations (**Figure 3.8**) (*38*). These hydrogen bonds greatly enhanced the iterative assignment process in fitting the Glutaminase L peptide into the structure of GIP. Final water refinement was done to get the 100 lowest energy structures from 200 calculated structures. Of these, 20 structures of lowest potential energy and best Ramachandran statistics found with PROCHECK were used for analysis. Their structural statistics were summarized in the **Table 3.3 (*10*)**. The ensemble of these 20 structures is shown in **Figure 3.18 (*10*)**.

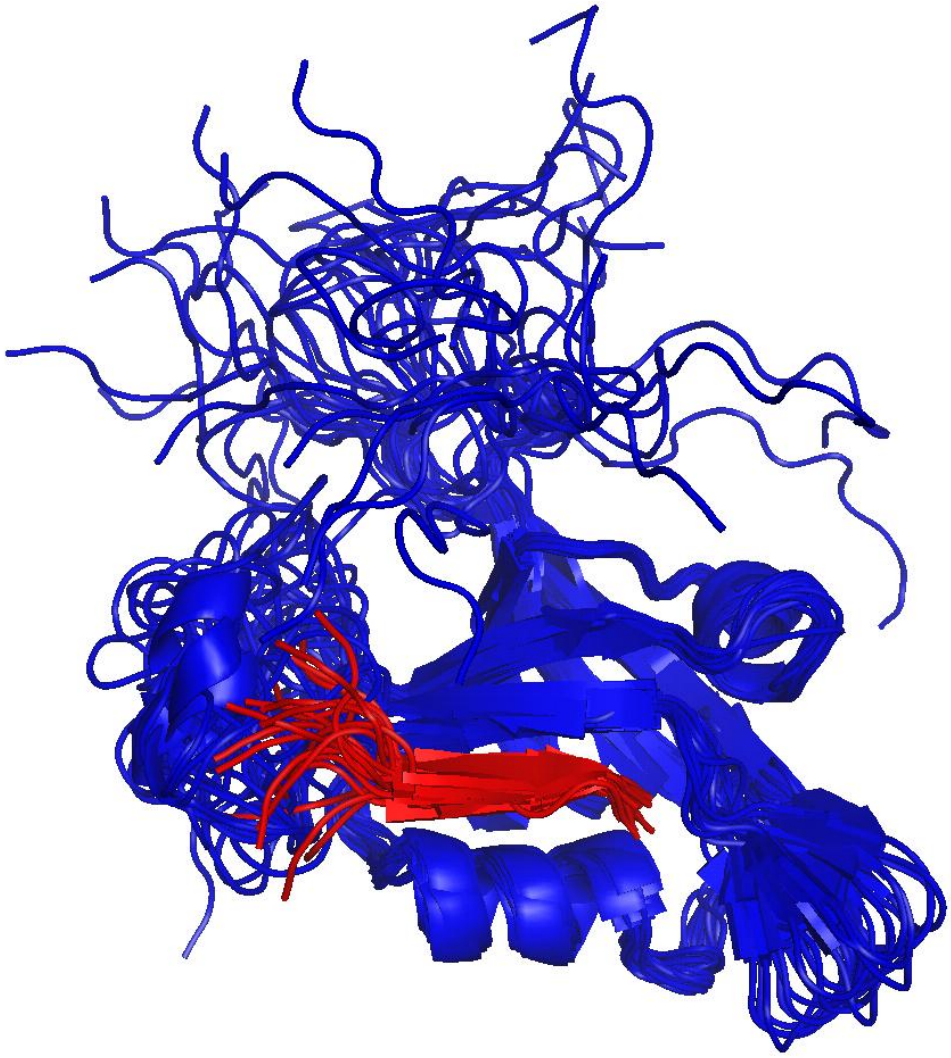


Figure 3.18: Ribbon diagrams of the ensemble of the 20 superimposed lowest energy structures of complexed GIP in blue with the Glutaminase L peptide in red. Adapted from reference (*10*).

Assignments	GIP-Glutaminase L complex
Sequential |i-j|=1	718
Medium 2≤|i-j|≤4	241
Long |i-j|>4	360
Intermolecular	37
Hydrogen Bonds [a]	66
Dihedral Constraints [b]	118
Ensemble Average [c]	
Total energy	-4816 ± 175
NOE energy	1586 ± 302
VDW energy	-1096 ± 67
Bonds energy	170 ± 8
Dihedral energy	749 ± 13
Angle energy	434 ± 26
Improper energy	1009 ± 89
Electrostatic energy	-6082 ± 123
Ramachandran Plot [d]	
Favorable	71.2
Additionally Allowed	24.3
Generously Allowed	2.7
Disallowed	1.8
RMSD (Å) [e]	
Well-ordered Backbone	0.67
Well-ordered Sidechain	1.28

Table 3.3: NMR structural statistics for the 20 selected lowest energy structures of the GIP-Glutaminase L Peptide Complex. Adapted from reference (*10*).

[a] Hydrogen bonds were defined by a set of two distance restraints per bond for residues of predicted secondary structure based on TALOS (*31*) predictions from CSI.

[b] Dihedral constraints were derived from TALOS (*31*) predictions from CSI.

[c] Energy terms were calculated by the water refinement module of ARIA 1.2 (*32*).

[d] Ramachandran plot statistics were calculated by PROCHECK (*33*).

[e] Well ordered regions included residues 11-19, 29-36 and 54-112.

3.4.4 Comparison of the structure of free GIP with that of the GIP-Glutaminase L peptide complex

Overall, the structures of both free GIP and the GIP-Glutaminase L peptide complex were somewhat similar, containing the same fold. However, to accommodate the additional β-strand of the Glutaminase-L peptide, the protein underwent changes in an allosteric manner in the complex. Binding with the peptide made the α2 helix of GIP move away from β2 by 0.95 Å to accommodate the additional β-strand (**Figure 3.19**). In both free GIP and the complex, the β2-β3 loop was largely unstructured. However, this loop appeared to have a few NOEs with the Glutaminase L peptide in the complex. This observation is in accordance with the report that GIP interacts with the C-terminal β-catenin peptide through its PFS loop (residues 45-47) (*39*). This suggests specificity in the nature of the interaction of GIP with different binding partners. Due to the closeness of the α1 helix to the binding site, significant chemical shift perturbations were observed in that region (**Figure 3.4** to **3.7**). But, such changes in chemical shifts were not reflected on the three-dimensional structure of the complex (**Figure 3.19**). Without complete structure determination, it could be misleading to infer any direct protein-ligand interactions simply based on the chemical shift perturbation map. This fact is illustrated by our observation of significant changes in chemical shifts of the residues that are not part of the binding pocket. Also, through structural comparison of free and bound protein, it was not easy to determine the specific interactions that caused the relatively large changes in the chemical shifts for residues that are located away from the binding site (*10*).

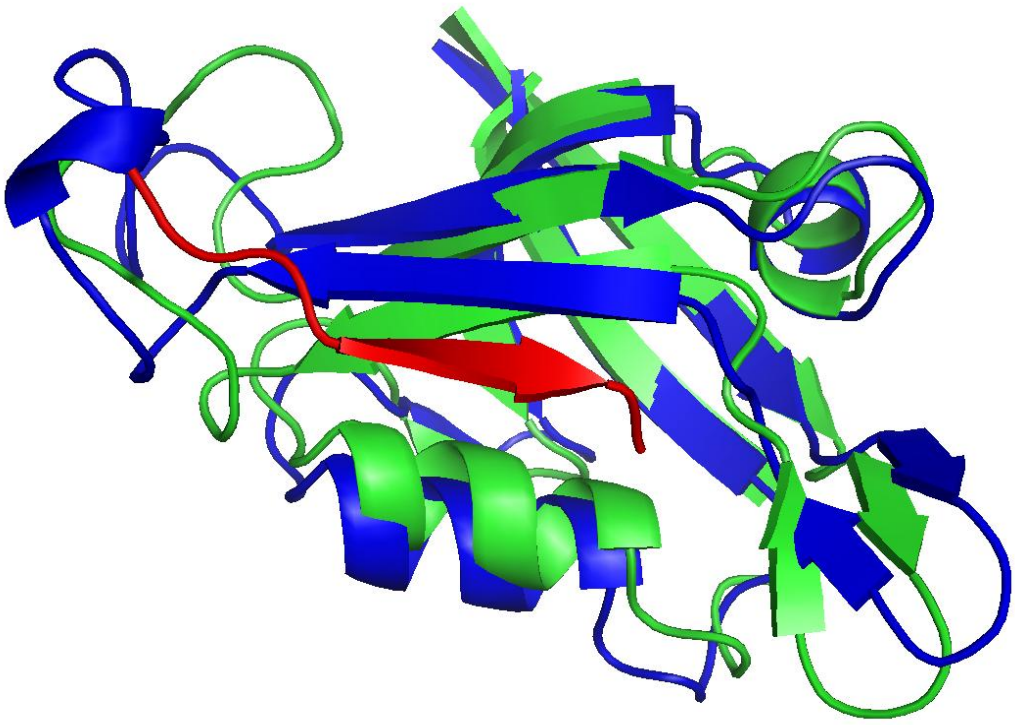


Figure 3.19: An overlay of free GIP is shown in green with the complexed GIP protein in blue and the Glutaminase L peptide in red. Adapted from reference (*10*).

3.4.5 Binding and specificity of the Glutaminase L peptide

The C-terminus of a binding partner binds in the binding pocket of the PDZ domain, in a process called β-strand addition, as an additional antiparallel β-sheet to the β2 strand of the protein (*37*). The binding pocket is created by the groove formed between the α2 helix and β2 strand of the protein. Specificity of this binding interaction comes from the sequences of the C-terminus of the interacting protein. Traditionally, the last four residues of the C-terminus of the peptide/ligand are numbered as positions -3, -2, -1 and 0 starting with the C-terminal residue as P_0 (*10*). There is a consensus GLGF loop located at the beginning of the β2 strand of PDZ domain that forms a series of hydrogen bonds between the backbone amides of the protein and the COO- of the C-terminal peptide. In addition, a hydrophobic interaction is facilitated by this loop to allow the sequence selectivity for the C-terminal residue of the substrate peptide.

A more detailed picture of the peptide bound to GIP is shown in **Figure 3.20**. In GIP, the canonical GLGF motif of PDZ domain is replaced by $I^{28}LGF^{31}$ motif, suggesting that while G28 is the consensus amino acid in the binding motif of PDZ domains, the mutation to Ile is tolerated perhaps due to the structural role it plays in forming the βa-βb hairpin. Whereas, G30 of this motif could be deemed as an absolute requirement, since it is the only amino acid that can accommodate the geometry needed for the formation of hydrogen bonds from L29 and G30 of GIP to the COO- at position P_0 of the C-terminal peptide. The charged carboxyl group from the C-terminal Val (P_0) of the Glutaminase L peptide formed two hydrogen bonds to the backbone amide protons of L29 and G30 of GIP. The hydrophobic side-chain of Val (P_0) of the peptide ligand buries itself in the hydrophobic pocket formed by L29, F31, L97 and I33 as well as T98 at

the periphery (**Figure 3.20**). As the protein binds to the peptide, the above mentioned two hydrogen bonds formed between the ligand and the protein caused unusually large chemical shift changes of up to 2.5 ppm for the amides of L29 and G30 in the ^{1}H, ^{15}N-HSQC spectra (**Figure 3.8**). When chemical shift perturbations of both HN/N and HA/CA pairs were mapped onto the structure of GIP (**Figure 3.4** to **Figure 3.7**), we observed that the regions near the binding site, including the β2, α2 and the β2-β3 loop were generally the most perturbed, however, α1, which did not appear to be directly involved in the binding was also significantly affected. This clearly demonstrated the allosteric mode of binding for GIP with Glutaminase L peptide. The residue H90 at the beginning of α2 (α2:1 in PDZ nomenclature) was oriented into the binding pocket and made a specific hydrogen bond with the Ser at P_{-2} of the peptide (**Figure 3.20**). This is a general feature of class I PDZ domains as the residue at position α2:1 provides the sequence selectivity that distinguishes between different classes (*40*). Generally, there is no specificity at P_{-1} (**Table 2.1**). The Glutaminase L peptide has Met at $P_{-1,}$ which was oriented away from the binding pocket toward the solvent. Some class I PDZ domains have specificity towards E/D or a small amino acid at P_{-3} (*40*). This interaction comes from hydrogen bonds between E at P_{-3} from the Glutaminase L peptide with Y56 and T58 of GIP. Alternately, a transient salt-bridge could potentially exist, but did not appear to be formed with R59 (**Figure 3.20**) of GIP. This particular salt-bridge has been observed in the crystal structures of GIP with β-catenin (*39*) and Kir 2.3 (*41*). However, no NOEs were observed to support the formation of a salt bridge between E at P_{-3} of the Glutaminase L peptide with R59 of GIP. In contrast to the static nature of a crystal environment, the dynamic flexibility of the protein side chains in solution contributed to the above observation. It is possible that the flexibility of these side chains would allow them to come close enough to form a transient salt-bridge. However, these results demonstrated that both

E at P_{-3} and R59 were solvent-exposed, thus decreasing the strength of such an interaction in solution. Thus, the salt-bridges observed in the two crystal structures could be due to packing artifacts of crystallization, while the true nature of the salt-bridge in solution is more dynamic (*10*).

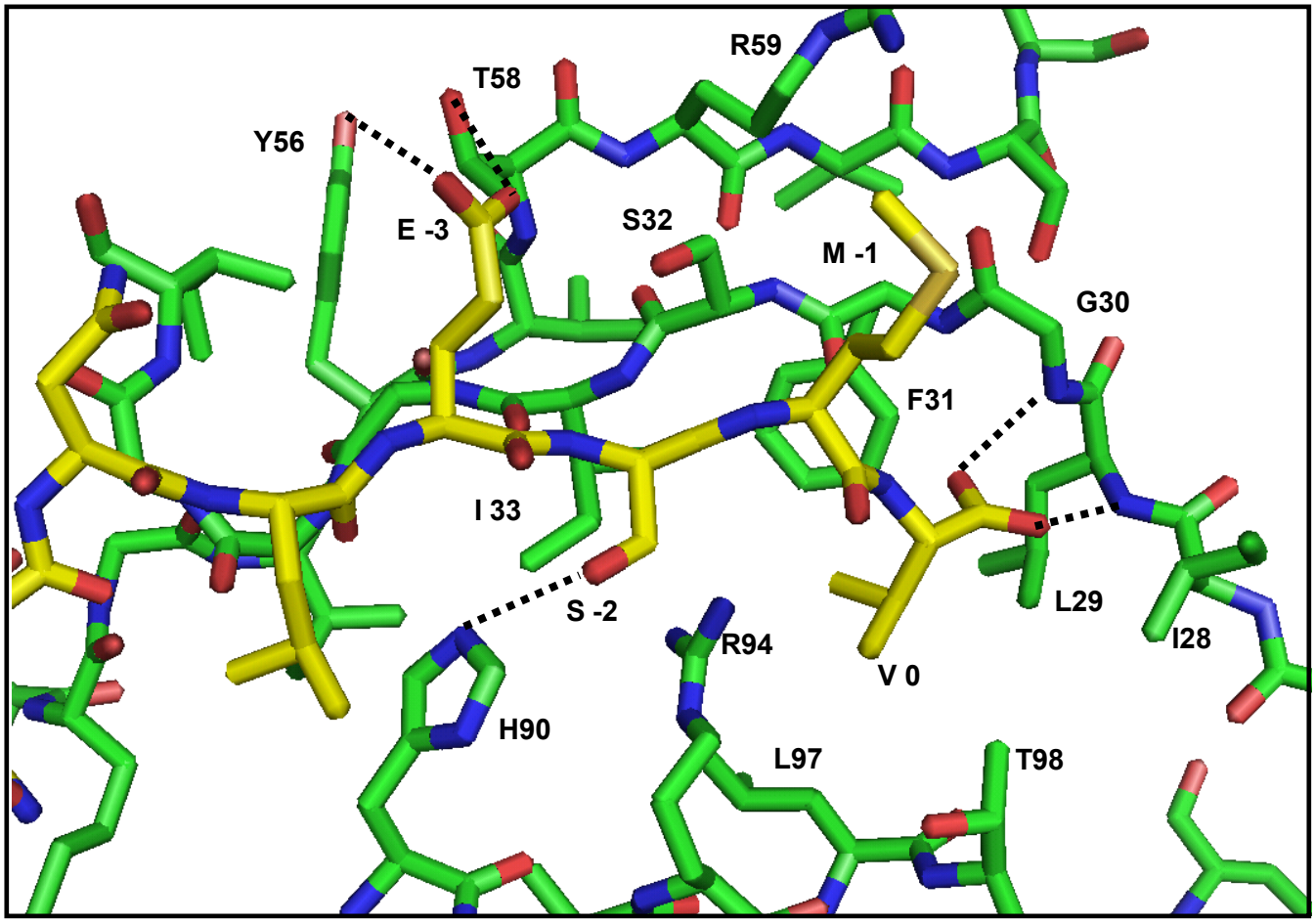


Figure 3.20: Heavy atom details from the binding site of GIP with the Glutaminase L peptide. The Glutaminase L peptide was colored in yellow and GIP in green. Potential hydrogen bonds (marked as dashed lines) could be seen from H90 with S at P_{-2}, the COO- from V at P_0 with the L29 & G30 amide nitrogens, and E at P_{-3} with Y56 and T58. V at P_0 buries its side chain into a hydrophobic pocket created by L29, F31, I33, L97 and partially T98. Adapted from reference (*10*).

3.4.6 Dynamics of the GIP-Glutaminase L peptide complex from ^{15}N relaxation measurements

Study on the dynamics of the GIP-Glutaminase L peptide complex was carried out to elucidate the binding mechanism of the Glutaminase L peptide to GIP. Using the Lipari-Szabo formalism-based model-free analysis (*42*), the order parameters (S^2) for GIP-Glutaminase L were calculated using steady-state ^1H-^{15}N NOE intensities, R_1 and R_2 relaxation rates. Those residues that could not be analyzed as a result of low intensity or absence from the HSQC spectra due to the overlapping were excluded from the data analysis. Excluded residues include M1, P5, P8, V12, V13, L21, N26, F31, G35, I37, D40, P41, Q43, P45, E48, D49, K50, D52, Y56, S61, P65, Q72, D75, V80, W83, M85, T86 and A93. Of these, L29, G30, F31, G35, D40, Q43, E48, D49 were from residues that form part of the binding pocket including the ILGF motif (canonical GLGF) and the β2-β3 loop, and they could not be measured as a result of being too close to the intermediate exchange regime to provide sufficient intensity required for observation in the NMR dynamics data. Aside the N-terminus and five proline residues, S^2 values for rest of the excluded residues could not be measured mainly for two reasons: spectral overlap and line broadening. In total, 96 of the 118 residues (excluding the N-terminus and 5 prolines) were analyzed to determine the S^2 values. Additionally, ΔS^2 values between bound and free states were determined for 84 residues. The generalized order parameters, S^2, were broadly similar for both the free and complexed states, but exhibited certain differences as explained below. The core region (A11-Q112) of the GIP-Glutaminase L peptide complex had an average S^2 value of 0.87 (0.89 for free GIP) as calculated based on the model-free analysis. Although, in general, the core of the protein maintained its structure and flexibility upon binding to the Glutaminase L peptide, however, specific residues exhibited either an increase or decrease in flexibility. Among the

residues for which ΔS^2 could be calculated, G36, G54, A66 and T98 showed a substantial ($\Delta S^2 > 0.06$) decrease in flexibility. Furthermore, residues I4, T51, G74, R96 and K99 showed smaller but still significant increases in S^2 ($0.03 < \Delta S^2 < 0.06$) where the average variance in ΔS^2 was ±0.015 for all measured residues. Twelve other residues showed positive, but statistically insignificant increases in S^2. Likewise twenty-four residues showed statistically insignificant decreases in S^2 upon binding. However, residues Q14, H19, I28, D38, N44, F46, T58, G63, G70, D91 and V109 showed a small but statistically significant ($-0.03 > \Delta S^2 > -0.06$) increase in flexibility. Additionally, residues R15, I18, G24, E25, L27, G34, K76, I77, H90, Q92, E103, R106, L107, R111 and many of the measured residues in the unstructured termini (M1-T10, S113-S124) showed a substantial increase in flexibility ($\Delta S^2 < -0.06$) as shown in **Figure 3.21**. When these residues were mapped onto the structure of free GIP (**Figure 3.22**), the biggest decreases in flexibility were displayed by the residues at the C-terminal end of the α2 helix near the binding site and at the hinge points of the β2-β3 loop. However, residues, located either on the β4 and β6 strands that are distal to the binding site or in the flexible loops such as the βa-βb hairpin and the β2-β3 loop, showed the biggest increases in backbone flexibility (*10*).

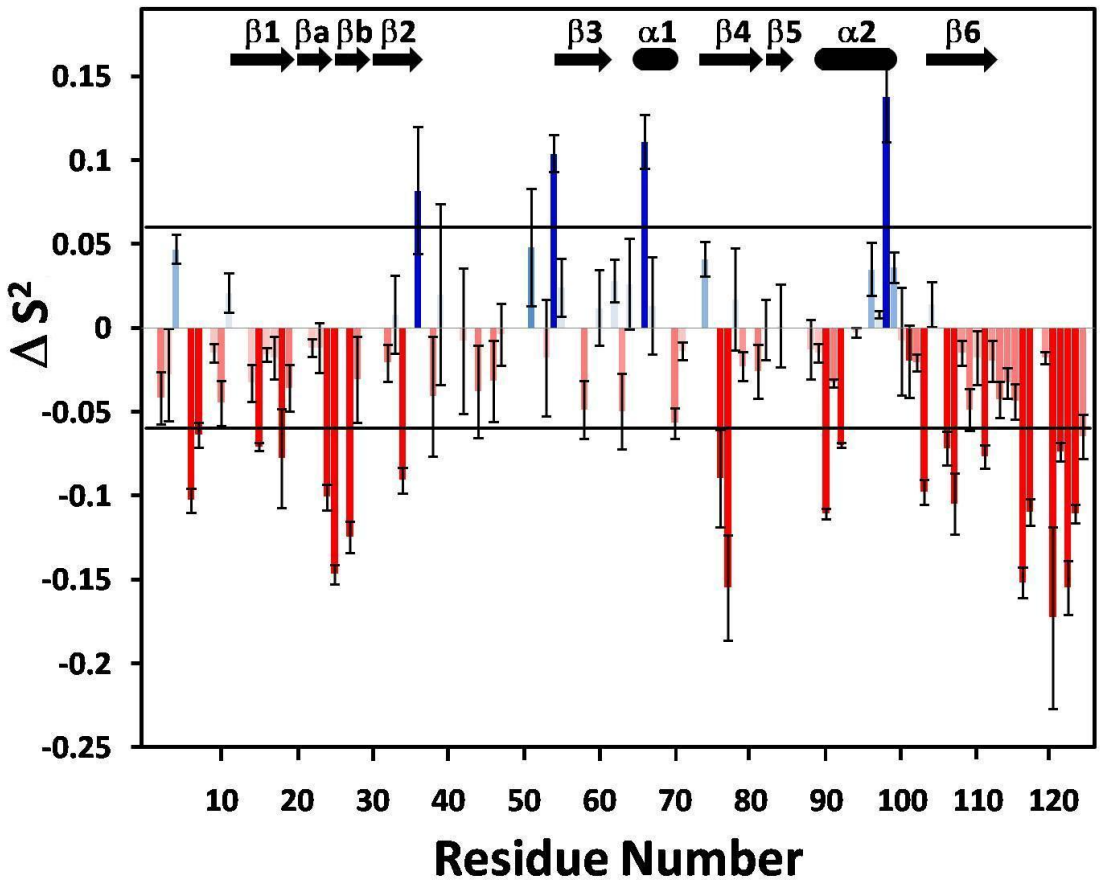


Figure 3.21: A plot of ΔS^2 as a function of residue number where ΔS^2 refers to S^2 of the GIP-Glutaminase L peptide complex minus that of free GIP. Positive values are indicated with increasing blue intensity while negative values are indicated with increasing red intensity Adapted from reference (*10*).

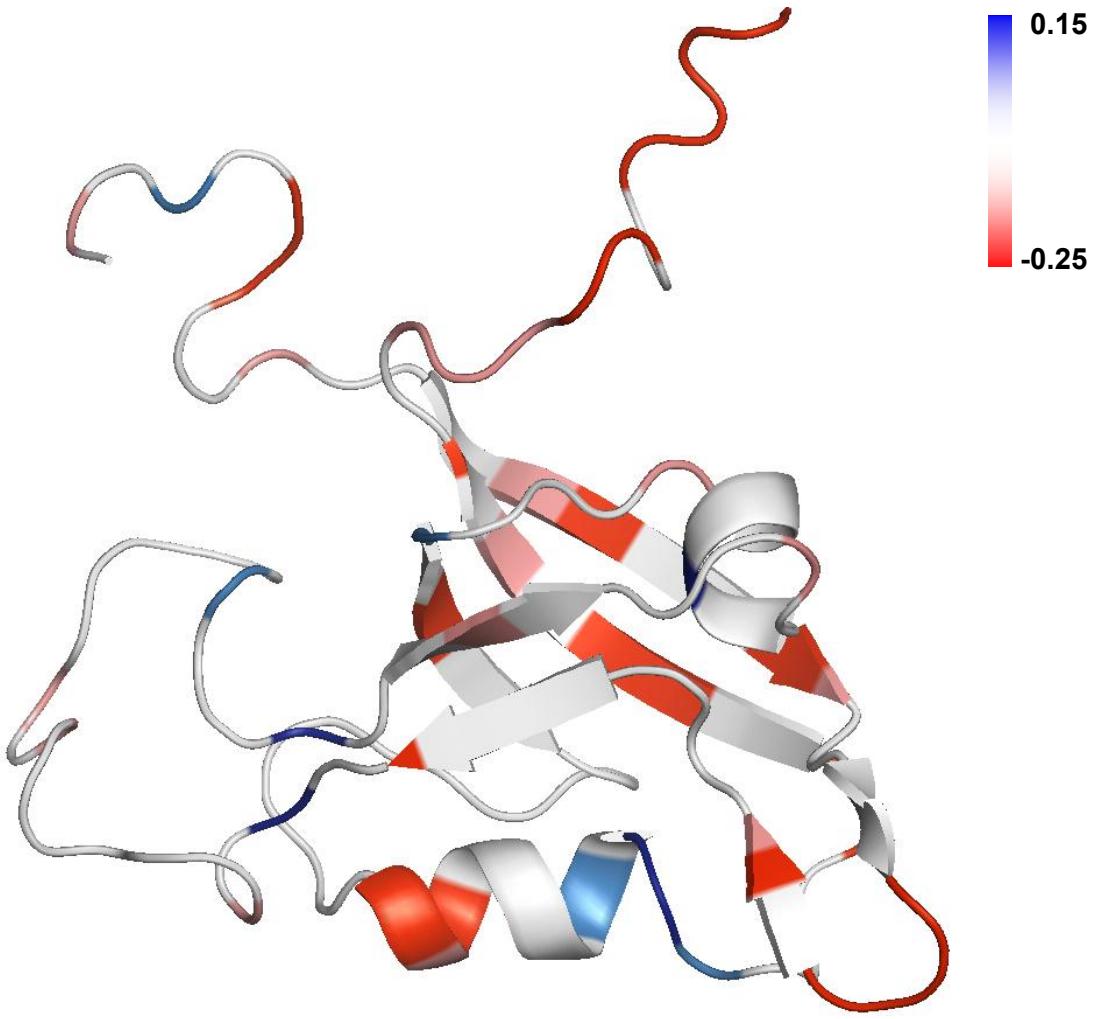


Figure 3.22: The magnitude of ΔS^2 upon binding to the Glutaminase L peptide was mapped onto the structure of free GIP and was indicated by darker intensity for red (increased flexibility) or blue (decreased flexibility). Residues were colored white for one of the following reasons: they could not be measured in both structures due to overlap, they had ΔS^2 values between the threshold values 0.06 and -0.06, or the residue was a proline. Adapted from reference (*10*).

3.4.7 Intermediate chemical exchange within GIP due to the binding of the Glutaminase L peptide

Most of the residues of GIP were in fast exchange regime while being titrated with the Glutaminase L peptide. But, the residues that are located within the binding pocket appeared to be in intermediate exchange. Residues L27, I28, L29, G30, F31, S32, I33, G34, and G35 had disappeared or greatly diminished in intensity due to intermediate to slow exchange at a low protein to peptide ratio. But, as the protein approached the saturation point, these residues reappeared in new locations in the ^{1}H, ^{15}N-HSQC spectrum at a higher concentration of Glutaminase L peptide. In addition, residues L29 and G30 had lower intensity in the 2D ^{1}H, ^{15}N-HSQC spectrum compared to all other residues of the protein due to line broadening caused by intermediate exchange, both in the free and complexed states of GIP. Based on the dynamics and chemical shift perturbations studies, we observed that, residues lining the binding pocket showed significant chemical shift perturbations along with substantial changes in the measurable order parameters. It appeared that both ends of the binding pocket experienced opposite effects in S^2 values. One end of the binding pocket that is near the C-terminus of the peptide is composed of the ILGF loop and the other C-terminal half of the α2 helix (K95-R100). The residues L29, G30 and F31 of the IGLF loop were in intermediate exchange, which precluded the measurement of ΔS^2. The residues R96, T98 and K99 from the α2 helix) experienced a decreased flexibility upon binding the C-terminal end of the peptide. On the other hand, at the opposite end of the binding pocket, residues from both β2 and α2 (G34, H90 and Q92) experienced an increase in flexibility. This observation was consistent with the relatively high RMSD for the N-terminal end of the Glutaminase L peptide. This increased or decreased flexibility in the binding pocket of the

protein and that of the peptide suggests that the substrate specificity is limited to the C-terminal four residues of Glutaminase L (*10*).

3.5 Discussion

3.5.1 Specificity in the binding interaction between GIP and Glutaminase L peptide

A number of interactions between GIP and the Glutaminase L peptide i.e. E-S/T-X-I/L/V-COOH (**Table 2.1**) provides specificity for the recognition process. The amide protons of residues L29 and G30 in the ILGF loop are uniquely positioned in such a way that allows them to form a pair of hydrogen bonds to both carboxyl oxygen atoms of V at P_0 from the Glutaminase L peptide (**Figure 3.20**). Very large chemical shift perturbations observed for these two residues reflect on the nature of these interactions (**Figure 3.4** & **3.8**). The proximity of the negatively charged carboxyl oxygens at P_0 position of the C-terminus of the Glutaminase L peptide to L29 and G30 of the protein caused dramatically different chemical environment at the binding site with large chemical shift perturbations although the protein structure is not significantly affected globally. When compared to the dynamics of free GIP, the effect of peptide binding on the dynamics of the protein appears to be dramatic. The disappearance of residues L27-G35 during the course of the titration due to the intermediate to slow exchange and their reappearance at saturation of the binding site, illustrates the dramatic effect of ligand binding on protein dynamics.

The specificity for a hydrophobic residue at P_0 of the ligand comes from the hydrophobic pocket created by L29, F31, I33 and L97 of the protein. Val seems to be preferred at P_0 more than Leu or Ile possibly due to the steric hindrance on this hydrophobic interaction with the longer side chain of these amino acids. This phenomenon was also observed and discussed in chapter 4 of this dissertation, where binding affinity of the interactions of two ligands (RDGDFQTEV-COOH and RGGSRL-COOH) with GIP was compared. The observance of a high affinity and stronger interaction (almost 7 times) for the ligand with V at its P_0 position than the one with L at P_0 position, is very likely due to the steric hindrance caused by the long side chain of L. The steric nature of these hydrophobic interactions could be confirmed through point mutation of one or more of the following residues in the binding pocket of GIP: L29V, L97V or T98A. Residue L97, located at position α2:8, is highly conserved across class I PDZ domains and is known to confer specificity at P_0 (*40*). The side-chains of L29 and L97 interact to form the majority of the surface area of this hydrophobic pocket. These mutations would likely change the selectivity at P_0 from Val to Ile, Leu or potentially a larger hydrophobic amino acid currently not allowed such as Phe or Trp.

Specificity for S/T at P_{-2} is due to H90 at position α2:1 of GIP. However, there is no specificity at the P_{-1} position. The likely reasons for the lack of specificity could be steric in nature. Firstly, the geometry of G30 is a prerequisite to sterically allow the binding of the C-terminus of the target protein to a PDZ protein. Thus, it could be an evolutionary trade-off between specificity for the C-terminus and sequence specificity at P_{-1}. Secondly, because the

binding occurs through β-strand addition, alternating amino acids are oriented away from the binding site.

To identify and distinguish between common and unique features of binding for each ligand, the first NMR structure of the GIP-Glutaminase L peptide complex was compared to the crystal structures of GIP bound to other target proteins. The mode of binding between GIP and each of its ligands is unique and specific. For example, unlike the specific interactions seen between the PFS loop of GIP with β-catenin (*39*), there are only a few interactions that occur between the β2-β3 loop of GIP and the Glutaminase L peptide. Additionally, the E at P_{-3} of the Glutaminase L peptide makes specific hydrogen bonds to Y56 and T58 of GIP rather than the salt-bridge observed between the D or E at P_{-3} of β-catenin or Kir 2.3 respectively with R59 of GIP (*39, 41*). Thus, it is necessary to experimentally determine the structure of GIP in complex with each of its known ligands to understand the mechanism of interactions for each binding partner. By maximizing the common features and taking advantage of the unique features of ligand binding, we should be able to efficiently design a competitive inhibitor with higher affinity than any of the natural ligands. Specificity for E at P_{-3} of the peptide is due to the formation of a hydrogen bond with Y56 and/or T58 of GIP. Since Y56 and T58 can each act both as hydrogen bond donors or acceptors, this explains why P_{-3} can also accommodate multiple side-chains. Furthermore, the lack of side-chains in three glycines in a row: G34, G35 and G36 of GIP render the ability to the protein to bind multiple partners. The lack of side-chains in such a stretch of three residues allows enough space for the different side-chains of the residues of the interacting partners that are located close to this region of the β2 sheet of the GIP but some

residues apart from those involved in the binding interactions. Finally, at positions beyond P_{-3}, GIP shows some specificity such as those observed in the interaction with β-catenin. During molecular recognition of β-catenin by GIP, a hydrogen bond is formed between the main-chain oxygen atom of tryptophan residue at P_{-5} and NE2 atom of Q43 (*39*). Thus, for any future drug design effort, an aromatic residue at P_{-5} or P_{-6} (**Table 2.1**) could provide additional specificity to GIP (*10*).

3.5.2 The effects of the Glutaminase L peptide binding on the dynamics of GIP

When the dynamics of free GIP was compared to that of the GIP-Glutaminase L complex, 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. One such exception is residue G34, which is part of the β2 strand that forms an antiparallel β-sheet with the Glutaminase L peptide. Although, the dynamic nature of the residue is expected to be more stabilized, yet it actually becomes more disordered. While it is part of the binding site, it is located on the opposite end of the β2-strand from the ILGF binding loop and is near the hinge-point between the β2 strand and the β2-β3 loop (residues G36-G54). Additionally, H90, D91 and Q92 show increased flexibility. While H90 makes a direct H-bond to the S at P_{-2}, (**Figure 3.20**) the specificity of the Glutaminase L peptide is limited to the four C-terminal residues, while the N-terminal four residues are disordered with higher RMSD values. However, overall, the region of GIP, where the peptide directly interacts, becomes more rigid. But, this decrease in flexibility in those regions is apparently offset by an increase in flexibility that is distributed throughout the rest of the protein including core regions of the protein that are distal to the binding site such as

β1, β4 and β6 strands and flexible regions of the protein such as the βa-βb hairpin and β2-β3 loops as well as both termini (*10*).

3.5.3 Comparison to other GIP-peptide complex structures

Both the N-terminal (M1-T10) and C-terminal (S113-S124) regions of GIP are completely unstructured both in the free form and in the bound form with very few observed NOEs and correspondingly high RMSDs in our structural ensembles (**Figure 2.21** & **3.18**). The dynamics study further supports this observation, indicating that these regions are completely unstructured (**Figure 2.22**). Previously, it has been reported that the C-terminal truncation of GIP leads to a decreased affinity for full length β-catenin *in vivo* (*43*). However, the binding modes of the β-catenin and Glutaminase L peptides to GIP were found generally to be similar (*11*). Therefore, it is unlikely that the reported decrease in full length β-catenin affinity to a C-terminally truncated GIP is due to an interaction between the canonical C-terminal binding motif of β-catenin and the C-terminus (113-124) of GIP. Moreover, upon binding with Glutaminase L, β-catenin or FAS peptide, the C-terminal region of GIP showed very little change in the chemical shifts (*11*). Therefore, a possible explanation for the above observation is the decrease in the affinity for the full length β-catenin upon C-terminal truncation of GIP could be the interaction of the C-terminus of GIP with either a different region of full length β-catenin or another interacting partner protein *in vivo*. An *in vivo* 2-hybrid interaction studies between various deletion mutants for both GIP and β-catenin supports this hypothesis (*43*). From these studies, it was observed that a central core region of β-catenin (173-483) lacking the class I C-terminus still

maintained some affinity for GIP (*43*). In light of our structural and dynamics characterization, the best plausible explanation is that the central core region of β-catenin interacts directly and specifically with the C-terminus of GIP. Thus, apparently, β-catenin and GIP each bind to the other protein's C-terminus (*10*).

3.5.4 Comparison between NMR and crystal structures

To comprehend the dynamic nature of a protein in solution, NMR is the technique of choice for structure determination. While there is good agreement between NMR and crystal structures of free GIP, there are a few key differences. First, in both the free and bound state NMR structures of GIP, both the N- and C-termini (regions 1-10 and 113-124) are highly dynamic and unstructured. Whereas, in the crystal structure of free GIP, the C-terminus forms a helix. This is very likely an artifact of crystallization. Second, in the NMR structures, the β2-β3 loop from G36-G54 is considerably more flexible in comparison to the crystal structures where this region has a defined structure (*39, 41*). Flexibility in this loop is also supported by the dynamics data, where significantly lower order parameters compared to the rest of the central core region were observed. Also, relatively few NOEs were observed compared to other regions of the protein. Moreover, all of the observed NOEs were medium range ($|i-j|<5$) or shorter, but there were no unambiguously defined long-range NOEs ($|i-j|>5$). This was the case for both free GIP as well as the GIP-Glutaminase L peptide complex. However, for the complex, there were some intermolecular NOEs between the loop and the peptide, indicating a conformational change in this flexible loop upon binding. This conformational change is observed from the decrease in flexibility of G36 and G54 near the hinge-point of the β2-β3 loop while flexibility increases on

either side of the hinge point. Although, a distinct conformational change is observed (**Figure 3.19**), the loop still remains relatively unstructured compared to the rest of the core protein in both free and bound states. Third, in comparison to crystal structures, the non-canonical β-hairpin formed by residues L21-I28 has a higher relative backbone RMSD of around 0.85 Å in the free form of GIP compared to the rest of the core structured portion of the protein at 0.45 Å. In the GIP-Glutaminase L complex the corresponding RMSD values are 2.73 Å and 0.67 Å. Like the β2-β3 loop, this β-hairpin structure also has mostly medium or short-range NOEs. Since it is exposed to the solvent, it does not make as many contacts with the rest of the protein. This results in very few long-range NOEs for this region and, therefore, this hairpin structure remains relatively unconstrained during the structural calculation. Comparatively, there were more long-range NOEs for this hairpin loop in free GIP than in the complex. That is why; there is an increase in RMSD for this structure within the complex compared to free GIP. The above observation is further supported by the dynamics study as increases in flexibility is observed for residues G24, E25, L27 and I28 in the complex (*10*).

3.5.5 Potential for drug design

Because GIP is very specific for certain types of molecular interactions, designing a drug that would target this protein is a promising endeavor. Since, cells contain literally hundreds of PDZ domains, if a drug is intended to target only the PDZ domain within GIP or broadly other PDZ domains that may share the same specificity as GIP, it is essential that the design of the drug molecule be very specific toward its desired target. Thus, the structural insights gained in

this chapter could prove very useful for the future design of a very specific drug molecule. Also, targeting GIP could lead to promising anticancer therapeutics.

3.6 Accession codes

The accession codes for GIP-Glutaminase L peptide complex in the BioMagnetic Resonance Bank (BMRB) and the Protein Data Bank (PDB) are 17255 and 2L4T, respectively. In BMRB, the chemical shifts of the resonances and, in PDB, the atomic coordinates for GIP-Glutaminase L peptide complex have been deposited (*10*).

3.7 References

1. Olalla, L., Gutierrez, A., Jimenez, A. J., Lopez-Tellez, J. F., Khan, Z. U., Perez, J., Alonso, F. J., de la Rosa, V., Campos-Sandoval, J. A., Segura, J. A., Aledo, J. C., and Marquez, J. (2008) Expression of the scaffolding PDZ protein glutaminase-interacting protein in mammalian brain, *J Neurosci Res 86*, 281-292.
2. Krebs, H. A. (1935) Metabolism of amino-acids: The synthesis of glutamine from glutamic acid and ammonia, and the enzymic hydrolysis of glutamine in animal tissues, *Biochem J 29*, 1951-1969.
3. 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.
4. Curthoys, N. P., and Watford, M. (1995) Regulation of glutaminase activity and glutamine metabolism, *Annu Rev Nutr 15*, 133-159.
5. Aledo, J. C., Gomez-Fabre, P. M., Olalla, L., and Marquez, J. (2000) Identification of two human glutaminase loci and tissue-specific expression of the two related genes, *Mamm Genome 11*, 1107-1110.
6. Olalla, L., Gutierrez, A., Campos, J. A., Khan, Z. U., Alonso, F. J., Segura, J. A., Marquez, J., and Aledo, J. C. (2002) Nuclear localization of L-type glutaminase in mammalian brain, *J Biol Chem 277*, 38939-38944.
7. Medina, M. A., Sanchez-Jimenez, F., Marquez, J., Rodriguez Quesada, A., and Nunez de Castro, I. (1992) Relevance of glutamine metabolism to tumor cell growth, *Mol Cell Biochem 113*, 1-15.

8. Brand, K. (1985) Glutamine and glucose metabolism during thymocyte proliferation. Pathways of glutamine and glutamate metabolism, *Biochem J 228*, 353-361.

9. Kovacevic, Z., and McGivan, J. D. (1983) Mitochondrial metabolism of glutamine and glutamate and its physiological significance, *Physiol Rev 63*, 547-605.

10. 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-Us 50*, 3528-3539.

11. 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-Us 47*, 9208-9219.

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

13. Goddard, T. D., and Kneller, D. G. SPARKY 3, University of California, San Francisco.

14. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *Journal of the American Chemical Society 114*, 10663-10665.

15. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity, *Journal of Magnetic Resonance, Series B 103*, 203-216.

16. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *Journal of the American Chemical Society 114*, 6291-6293.

17. Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., and Campbell, I. D. (1990) Comparison of techniques for 1H-detected heteronuclear 1H---15N Spectroscopy, *Journal of Magnetic Resonance (1969) 87*, 488-501.

18. Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *Journal of Magnetic Resonance (1969) 93*, 151-170.

19. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins, *Journal of the American Chemical Society 115*, 7772-7777.

20. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C' chemical shifts in 15N---13C-labelled proteins, *Journal of Magnetic Resonance (1969) 97*, 213-217.

21. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J Biomol NMR 4*, 845-858.

22. 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.

23. 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.

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

25. Bax, A., and Davis, D. G. (1985) Practical aspects of two-dimensional transverse NOE spectroscopy, *Journal of Magnetic Resonance (1969) 63*, 207-213.

26. Mandel, A. M., Akke, M., and Palmer, A. G., 3rd. (1995) Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme, *J Mol Biol 246*, 144-163.

27. Palmer, 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 carbon-13 heteronuclear NMR spectroscopy, *Journal of the American Chemical Society 113*, 4371-4380.

28. Dosset, P., Hus, J.-C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data, *Journal of Biomolecular NMR 16*, 23-28.

29. Tsan, P., Hus, J.-C., Caffrey, M., Marion, D., and Blackledge, M. (2000) Rotational Diffusion Anisotropy and Local Backbone Dynamics of Carbon Monoxide-Bound Rhodobacter capsulatus Cytochrome c$_{2}$, *Journal of the American Chemical Society 122*, 5603-5612.

30. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol Biol 278*, 353-378.

31. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR 13*, 289-302.

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

33. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J Biomol NMR 8*, 477-486.

34. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J Mol Graph 14*, 33-38, 27-38.

35. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

36. Hwang, T. L., and Shaka, A. J. (1995) Water Suppression That Works - Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients, *J Magn Reson Ser A 112*, 275-279.

37. Jelen, F., Oleksy, A., Smietana, K., and Otlewski, J. (2003) PDZ domains - common players in the cell signaling, *Acta Biochim Pol 50*, 985-1017.

38. 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.

39. Zhang, J., Yan, X., Shi, C., Yang, X., Guo, Y., Tian, C., Long, J., and Shen, Y. (2008) Structural basis of beta-catenin recognition by Tax-interacting protein-1, *J Mol Biol 384*, 255-263.

40. 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.

41. Yan, X., Zhou, H., Zhang, J., Shi, C., Xie, X., Wu, Y., Tian, C., Shen, Y., and Long, J. (2009) Molecular mechanism of inward rectifier potassium channel 2.3 regulation by tax-interacting protein-1, *J Mol Biol 392*, 967-976.

42. Lipari, G., and Szabo, A. (1980) Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes, *Biophysical Journal 30*, 489-506.

43. 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, *Journal of Biological Chemistry 278*, 38758-38764.

Chapter 4

Determination of the mode of interaction of Glutaminase Interacting Protein (GIP) with two different interacting partners

4.1 Introduction

4.1.1 PDZ domain and its functions (*1*)

Glutaminase interacting protein (GIP) (*2*), also known as tax interacting protein-1 (TIP-1) (*3*), is a 13.7 kDa PDZ domain-containing protein. PDZ domains are one of the most important protein-protein interaction modules found in nature (*4*). PDZ domain-mediated interactions contribute to cell signaling pathways, adhesion and receptor and ion transporter function (*5*). PDZ domains often act as scaffolds, specifying protein interactions required for the formation of multimeric complexes (*6*). The diversity of PDZ domain-protein interactions and their involvement in maintenance of normal physiological functions of the body are significant in the context of clinical disorders. Several human diseases are known to occur as a result of inappropriate protein-protein interactions, which in turn affect gene expression and regulation, transport of biomolecules across the membranes, cell adhesion, antigen recognition and signal transduction (*7*).

4.1.2 Binding pocket of PDZ domain (*1*)

The binding pocket of PDZ domains and the mode of binding to the interacting partner proteins are each well characterized (*5, 8-10*). The GLGF motif present in the binding pocket of

PDZ domains plays a major role in the binding interactions with the target protein. PDZ domains were therefore previously referred to as GLGF repeat domains (*11*). PDZ domains exhibit sequence specificity towards the unstructured C-terminal ends of their interacting protein partners. Peptides representing these C-terminal recognition motifs have been shown to act as surrogates for their corresponding partner proteins *in vitro* (*12*). Several classes of PDZ domains have been reported based on this specificity: class I {X-S/T-X-Φ-COOH}, class II {X-Φ-X-Φ-COOH}(*6*), class III {X-E/D-X-Φ-COOH}(*13*) and other minor classes (*14*) where Φ is any hydrophobic residue and X is any residue. The interacting peptide forms an additional anti-parallel β-strand between the β2 strand and the α2 helix of PDZ domain (*5*).

4.1.3 GIP as a PDZ domain (*1*)

GIP is an unusual class I PDZ domain protein in the sense that it is solely composed of a single PDZ domain (*6*). Structurally, GIP is made up of two α-helices (α1 and α2) and six β-strands (β1, βa, βb, β2, β3, β4, β5 and β6) (*10, 15*). GIP is also striking for the promiscuity of its binding profile. A number of different binding partners have been identified with roles in diverse cellular processes. Some of the reported interacting proteins include Glutaminase L, β-Catenin, Fas, HTLV (Human T-lymphotropic virus) Tax and HPV (Human papillomavirus) E6, which are involved in signaling pathways, energy generation pathways or oncogenic processes (*2, 3, 8, 10, 12, 16-24*).

4.1.4 GIP in the brain (*1*)

GIP is known to function as a key scaffolding protein in the mammalian brain (*25*), contributing to the bioenergetics of both normal and cancer cells through its interaction with Glutaminase L (*2, 16-18*). GIP may also mediate normal brain cellular functions through interactions with other as yet unidentified partner proteins. To fully understand the mechanism of function of GIP in the brain, it is necessary to identify the proteins that interact with GIP in brain cells.

4.1.5 Identification of interacting partners in brain (*1*)

Among the various methods available for the investigation of novel protein-protein interactions, the yeast two-hybrid genetic selection system (Y2H) is a powerful technique with several advantages over traditional biochemical approaches (*7*). This method was developed by Song and Fields in *Saccharomyces cerevisiae* (baker‟s yeast) and involves the expression within the yeast cell nucleus of two proteins being assessed for interaction (*26*). Each protein is expressed as a chimera, fused to one domain of the yeast Gal4 transcription factor. Interaction of the two fusion proteins brings the two domains of Gal4 into close enough proximity to restore transcription factor function, detected by activation of Gal4-responsive reporter genes. In this study, our collaborators in Ege University, Izmir, Turkey used the yeast two-hybrid system to screen a human fetal brain cDNA library for GIP-interacting proteins. From that screening, Brain-specific angiogenesis inhibitor 2 (BAI2) was identified as a novel interacting partner of GIP. Here, CD, fluorescence and NMR techniques were used to further confirm BAI2 as an interacting partner of GIP by using a peptide RDGDFQTEV-COOH representing the BAI2 C-

terminus. To compare the interaction between GIP and RDGDFQTEV-COOH, another arbitrary peptide RGGSRL-COOH hereinafter termed as control peptide was designed based on the peptide sequence specificity for PDZ domain (**Table 2.1**) and used to determine the comparative strength of interaction by CD, fluorescence and NMR techniques.

4.2 Materials and Methods (*1*)

The research work described here was carried out in the laboratory of Dr. Smita Mohanty.

4.2.1 Expression and purification of ^{15}N- and unlabeled GIP

GIP protein was expressed in *E. coli* and purified according to our lab protocol (*12*), *E coli* (strain BL21DE3pLysS) was transformed with plasmid pET-3c/GIP and cells were cultured in M9 minimal media containing ^{15}N-labeled ammonium chloride for ^{15}N-labeled GIP and in LB-ampicillin media for unlabeled GIP. An overnight culture was diluted 1:25, {v/v} in minimal media (or LB-ampicillin media for unlabeled protein) and grown at 37 °C to an OD_{600} of 0.4-0.5. Expression was induced with 1 mM IPTG at 30° C, and after 12 h. incubation (for unlabeled GIP, after 4 hours), cells were harvested by centrifugation. The harvested cells were lysed by sonication using lysis buffer containing 50 mM phosphate buffer at pH 8, 200 mM NaCl, 4 mM EDTA, 4% glycerol, and 1 mM PMSF. After centrifugation of the lysed cells, the supernatant was retained for further purification. ^{15}N- and unlabeled GIP were each purified in a single-step using size exclusion chromatography with a Sephacryl S-100 column {GE Healthcare} according to our lab protocol (*12*). Pooled fractions of pure protein were concentrated.

4.2.2 Fluorescence

All fluorescence spectra were recorded on a PerkinElmer Precisely LS 55 Luminescence spectrofluorometer at 25 °C (λ_{ex} 280 nm). Emission spectra were recorded over a range of 300-500 nm with 1 nm steps. All experiments were carried out in 20 mM phosphate buffer, pH 6.5, 150 mM NaCl, 0.1 mM EDTA and 0.01% NaN_3. Stock solutions of the synthetic peptide sequence RDGDFQTEV-COOH, hereafter known as BAI2 peptide and control peptide were prepared in water at a concentration of 10 mM. The target peptides were obtained with >95% purity from Chi Scientific (MA). The stock solutions were then diluted to 1 mM. Aliquots of the 1 mM peptide solutions were directly added to a cuvette containing 2 mL of 1 μM unlabeled GIP. All titration experiments were corrected to take the dilution effect into account. Emission from the control was corrected by recording subtraction spectra between sample and control probes.

4.2.3 Circular Dichroism (CD)

All circular dichroism (CD) experiments were performed on a Jasco J-810 automatic recording spectropolarimeter. Far-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 30 μM. Data were averaged over 100 scans for each protein sample and over 50 scans for each control sample. Response time was 1 s, and scan speed was 100 nm min^{-1}.

4.2.4 Nuclear Magnetic Resonance (NMR)

All NMR data were collected at 298 K on a Bruker Avance 600 MHz spectrometer equipped with a triple resonance H/C/N TCI cryoprobe at the Department of Chemistry and Biochemistry, Auburn University, Auburn, AL. The data were processed using NMRPipe (*27*) and analyzed using Sparky (*28*). The ligand titration experiments were performed and monitored by a series of 2D ^{15}N-edited HSQC experiments. The interaction study was carried out by titration of 100 μM ^{15}N-labeled GIP with the BAI2 peptide and control peptide. The amide chemical shift perturbations ($\Delta\delta$) were calculated as $\Delta\delta = \sqrt{[\{|\Delta\delta^{15}\mathrm{N}|/10\}^2 + \{|\Delta\delta^{1}\mathrm{H}|\}^2]}$. In the equation, $\Delta\delta^{15}$N was divided by 10 to account for the difference in the gyromagnetic ratio of the ^{15}N and ^{1}H nuclei to give roughly equal weighting for both types of chemical shift changes. The program ModelTitr (*29*) was used to calculate the dissociation constant values for various residues of GIP.

4.3 Results and Discussion

4.3.1 Protein expression

As described above, unlabeled GIP was expressed in bacterial cells growing in LB media and ^{15}N-labeled GIP was expressed in M9 minimal media containing ^{15}N-labeled ammonium chloride. The SDS-PAGE analysis of expression of both unlabeled and ^{15}N-labeled GIP upon induction is given in the **Figure 4.1**. GIP as a 13.7 kDa size protein appeared as a prominent band in both of the lanes for labeled and unlabeled protein at its due place in the gel (**Figure 4.1**).

4.3.2 Protein purification

Using size-exclusion chromatography as a single step, GIP was purified in a Sephacryl S-100 column (GE Healthcare). The production of the unlabeled and ^{15}N-labeled recombinant GIP is around 46 mg and 12 mg per liter of bacterial culture (**Figure 4.2**). Comparing the expression profile for unlabeled and ^{15}N-labeled GIP in the **Figure 4.1**, higher amount of production for unlabeled GIP than ^{15}N-labeled GIP is observed. Recently, Turck *et al.* at Max Planck Institute of Psychiatry has demonstrated that, when *E. coli* cells were grown in ^{15}N-labeled media, consistent lower level of protein expression and alteration of growth rates and metabolite levels were observed as compared to when cells grown on unlabeled media (*30*).

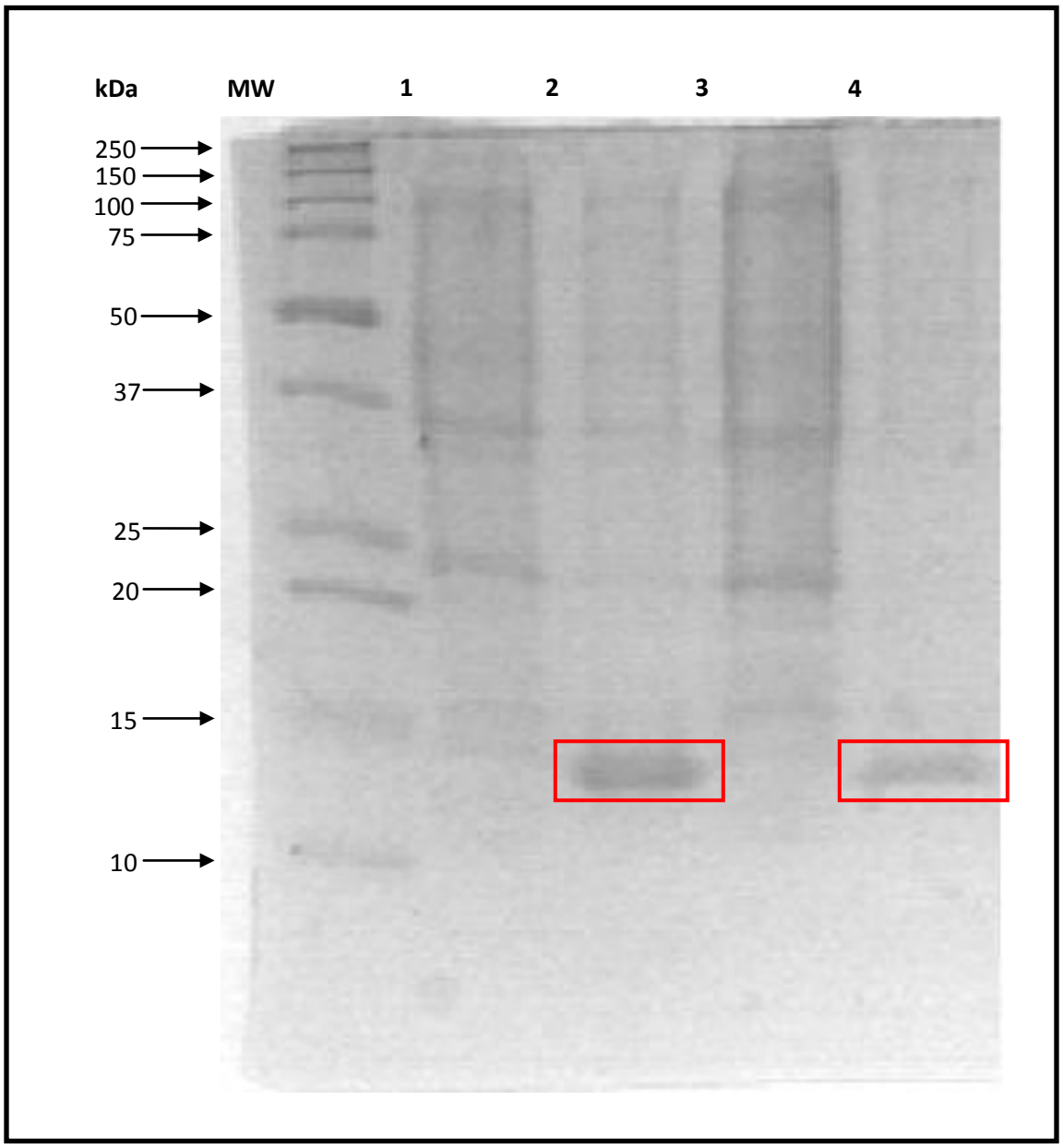


Figure 4.1: Expression of GIP analyzed by SDS-PAGE. *Lane 1*- Unlabeled GIP expression in T_0 cells before induction. *Lane 2*- Unlabeled GIP expression in T_5 cells after complete induction. *Lane 3*- ^{15}N-labeled GIP expression in T_0 cells before induction. *Lane 4*- ^{15}N-labeled GIP expression in T_{12} cells after complete induction. The red rectangle spots the protein of expected size. The lane MW is for protein marker.

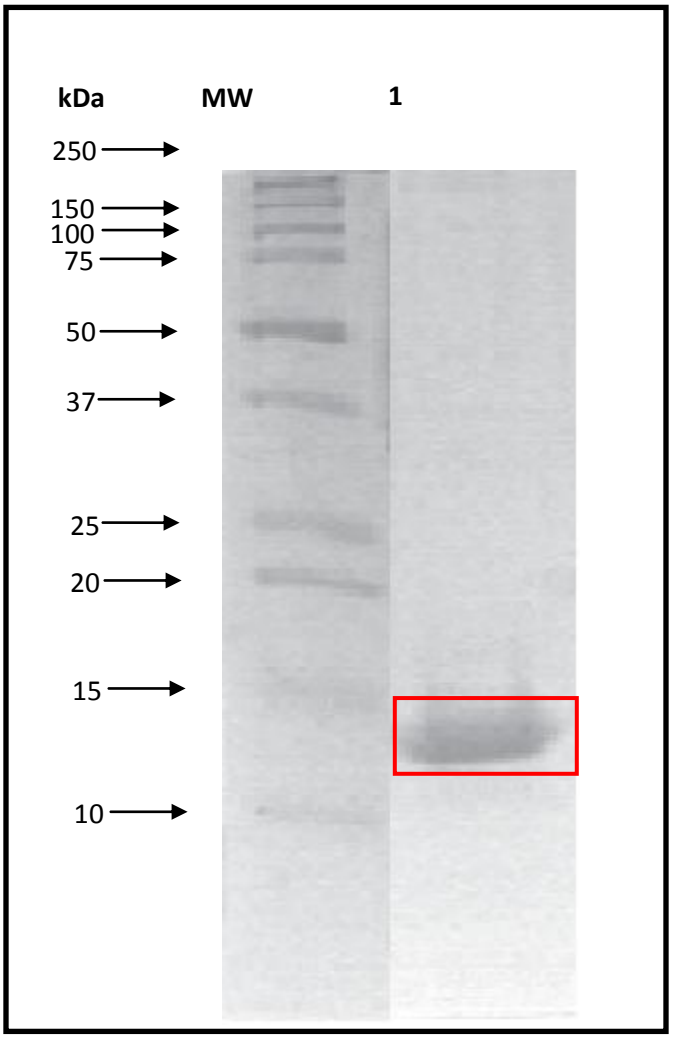


Figure 4.2: Purification of GIP analyzed by SDS-PAGE. Lane 1 shows the purified ^{15}N-labeled GIP without any impurities. The red rectangle spots the protein of expected size. The lane MW is for protein marker.

4.3.3 Interaction of BAI2 Peptide with GIP (*1*)

4.3.3.1 Characterization by Fluorescence spectroscopy

When the peptide was titrated against unlabeled GIP, it showed a small but consistent decrease in fluorescence intensity (**Figure 4.3**). The dissociation constant K_D ($K_D = 1/K_a$) was determined using the OriginPro 6.1 software. The decrease in the fluorescence intensity was calculated as $(F_0 - F_C)/(F_0 - F_{min})$, where F_0 is the initial fluorescence intensity of free GIP; F_C is the corrected fluorescence intensity at a ligand concentration $[C]$, and F_{min} is the fluorescence intensity at the saturating concentration of the peptide. The data were fitted to a nonlinear regression of the plot of $(F_0 - F_C)/(F_0 - F_{min})$ against $[C]$ with the equation corresponding to a single binding site (**Figure 4.4**). The titration of the BAI2 peptide with GIP yielded a dissociation constant of 0.71 μM. To determine the thermodynamic nature of the interaction, the free energy change of the association was calculated using the following equation: $\Delta G = -RT \ln K_a$, where K_a is the association constant, T is the temperature and R is the universal gas constant. By putting the experimentally determined K_a ($K_a = 1/K_D$) value into this equation, the ΔG value for binding of the BAI2 peptide to GIP was calculated to be-35.08 kJ mol^{-1}, which reflects the spontaneous binding of the peptide to GIP.

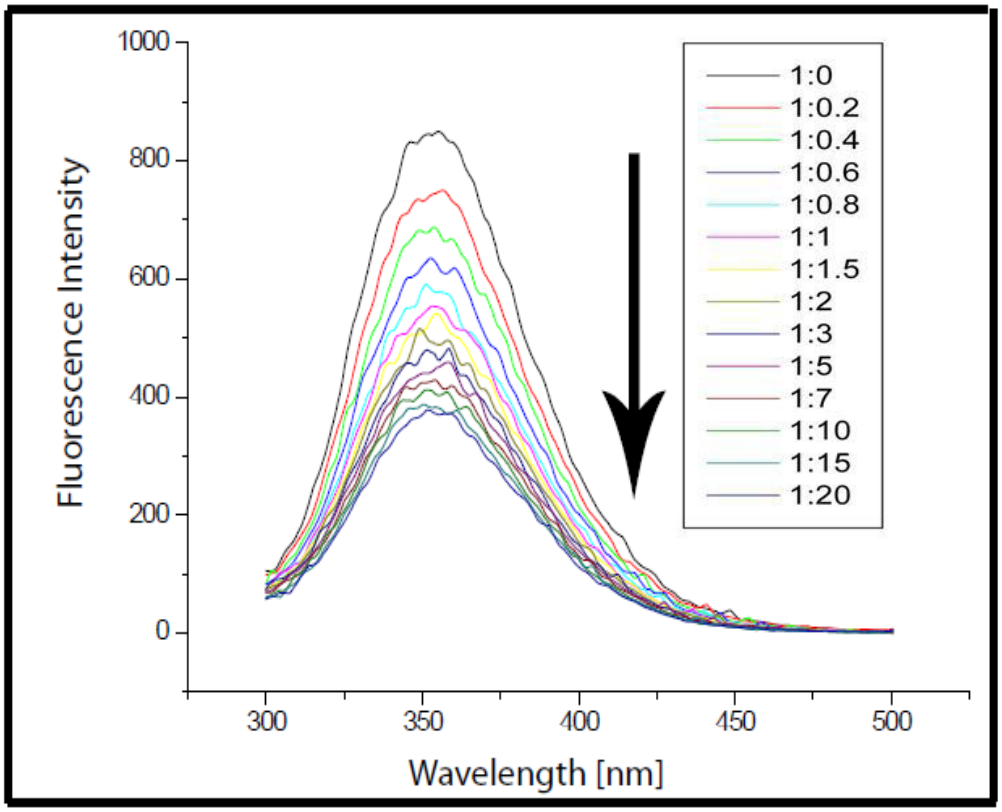


Figure 4.3: Fluorescence emission spectrum of GIP with the BAI2 peptide. Fluorescence emission plots corresponding to (top to bottom) 0 to 20 μM concentrations of the peptide to 1 μM protein sample. In the legend, protein to peptide ratios are indicated with the respective color codes. The black arrow indicates the quenching of fluorescence of GIP upon peptide binding in a downward fashion.

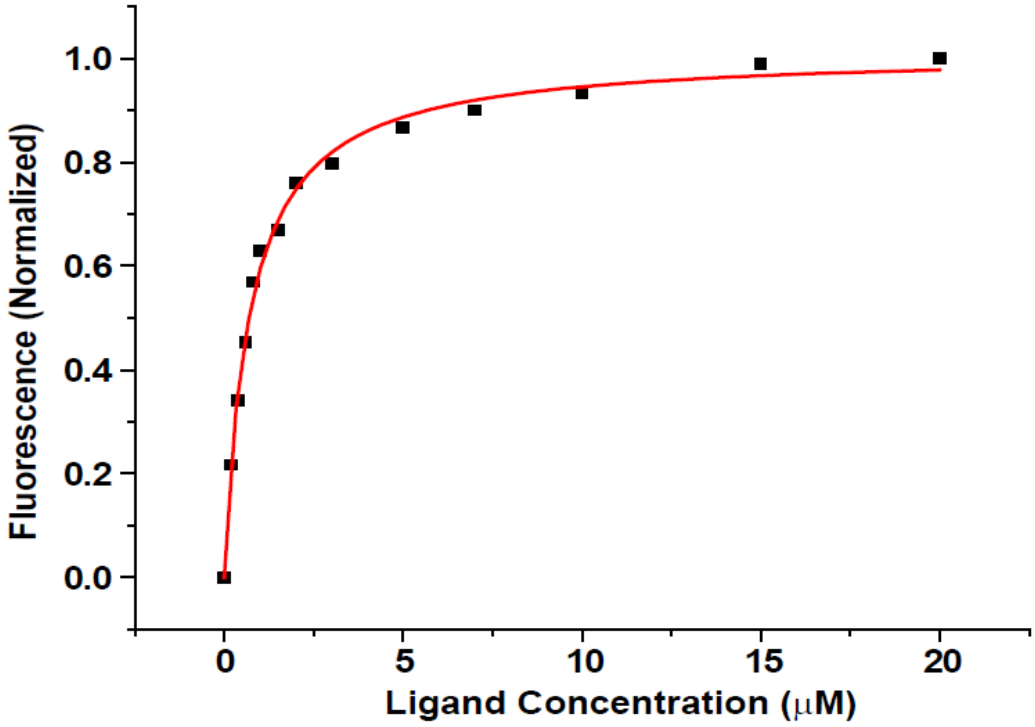


Figure 4.4: Non-linear curve fitting assuming 1:1 binding between GIP and the BAI2 peptide where $(F_0 - F_C)/(F_0 - F_{min})$ was plotted against peptide concentration.

4.3.3.2 Characterization by CD spectroscopy

CD spectroscopy is another powerful tool to investigate the effect of any ligand binding on the secondary structure of the protein. The phosphate buffer, as well as the BAI2 peptide alone, showed minimal signal in the CD measurements. However, any contribution from the peptide and buffer was subtracted from the CD spectrum obtained in subsequent analyses of GIP with peptide. The secondary structure of GIP showed significant changes in the CD spectrum with the titration of different concentrations of the peptide (**Figure 4.5**). CD data of the GIP-peptide complex was deconvoluted using the program CDPro (*31*) and the secondary structure content was calculated. From the deconvolution results, the helix content was found to be reduced by ~ 47%, random coil content by ~ 8% and the β-sheet structure content increased by ~ 29%. The changes in the secondary structure of GIP with the addition of increasing concentration of BAI2 peptide is comparable to that observed with other previously reported binding partners of GIP such as Glutaminase L, FAS and β-catenin (*12*). Although, the increase in β-sheet content in all these cases can be explained by the mode of binding of these peptides to the GIP through β-strand addition, closer examination of the representative complex structure of GIP with its binding partner does not show any change in the helical content but does indicate some displacement of the helical structure in space (*8, 10, 32*). CD spectroscopy is sensitive enough to detect even slight changes in the secondary structure of the protein upon interaction with the ligand but is not always sufficient to get a complete picture of the structural features of protein-peptide interactions.

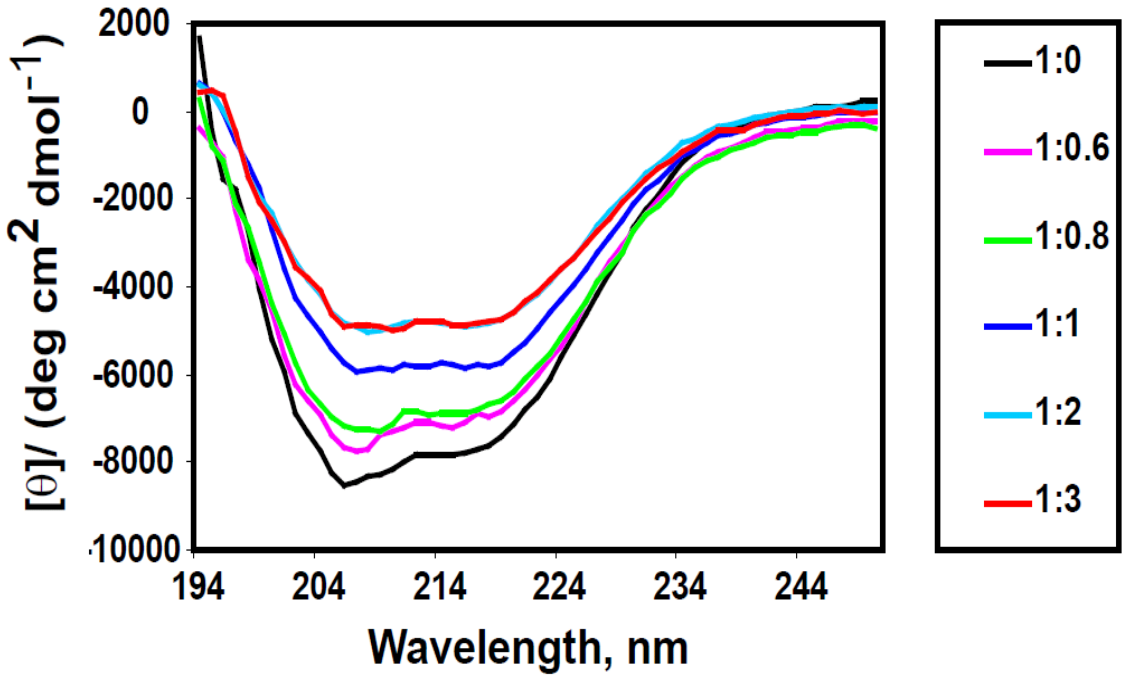


Figure 4.5: Changes in the CD spectra of GIP upon binding with increasing concentrations of the BAI2 peptide for the wavelength range of 194 nm to 250 nm. The protein to peptide ratios for the corresponding color codes are indicated in the legend.

4.3.3.3 Characterization by $^{1}H,^{15}N$-HSQC NMR

To examine the interaction of GIP with the C-terminal BAI2 peptide more thoroughly, an NMR analysis was undertaken. NMR can be employed as a very powerful technique for monitoring structure-activity relationships (SAR) in protein-protein or protein-ligand interactions studies (*33*). The chemical shifts of the backbone amides of a folded protein are extremely sensitive to any changes in their chemical environments, such as temperature, pH, ionic strength, or binding to a ligand. For this reason, the 2D $^{1}H,^{15}N$-HSQC spectrum is often called the fingerprint region of a protein, as the exact pattern is unique to each protein under a specific set of environmental conditions. Upon ligand binding, the chemical shifts of the residues involved in the binding change, which is reflected in a series of 2D $^{1}H,^{15}N$-HSQC spectra (*10, 12*). However, when the binding is allosteric, which affects the protein globally rather than locally, the chemical environments of most of the residues in a protein experience a change. Thus, residues that are not directly part of the binding pocket may also show change in their chemical shifts (*10*). Therefore, any perturbation in the chemical shifts from their original positions may indicate a change in the conformation of the protein upon binding with the ligand (*34*). However, it is important to note that, for GIP, such chemical shift perturbations should not necessarily indicate a drastic conformational change in the protein (*10*). To investigate whether BAI2 peptide binds to the protein, ^{15}N-labeled GIP protein was titrated with the synthetic BAI2 peptide to excess (~60 times that of the protein) until complete saturation was achieved. During the course of the titration, the fingerprint region of the protein in the 2D $^{1}H,^{15}N$-HSQC spectra was monitored. The fingerprint region of the HSQC spectra of GIP was collected in the absence and presence of different concentrations of the peptide and the spectra were overlaid (**Figure 4.6**). From the overlay, it was evident that most of the residues of GIP showed moderate changes in chemical

shifts upon binding with the peptide, while other residues showed more dramatic changes. Using the program ModelTitr (*29*), the dissociation constant (K_D) values for various residues of GIP were calculated (**Table 4.1**) by non-linear least-squares fitting of the chemical shift data against ligand concentration to the Langmuir isotherm that involved the assumption of a stoichiometry of 1:1 between the ligand and the protein (*i.e.* one binding site) (**Figure 4.8**). The dilution effect on the concentration of the protein due to the addition of the ligand was corrected in the program. The calculated dissociation constant (K_D) value from NMR technique (97.77 μM on an average) was different from the value obtained from fluorescence technique. Since the dissociation constant (K_D) value varies depending upon techniques and initial protein concentration used (*35-37*), such a difference in the K_D values obtained from two different techniques is acceptable. From the K_D values of both fluorescence and NMR techniques, the dissociation constant (K_D) value falls in the range of low to mid μM, which indicates a moderate affinity of GIP for the BAI2 peptide.

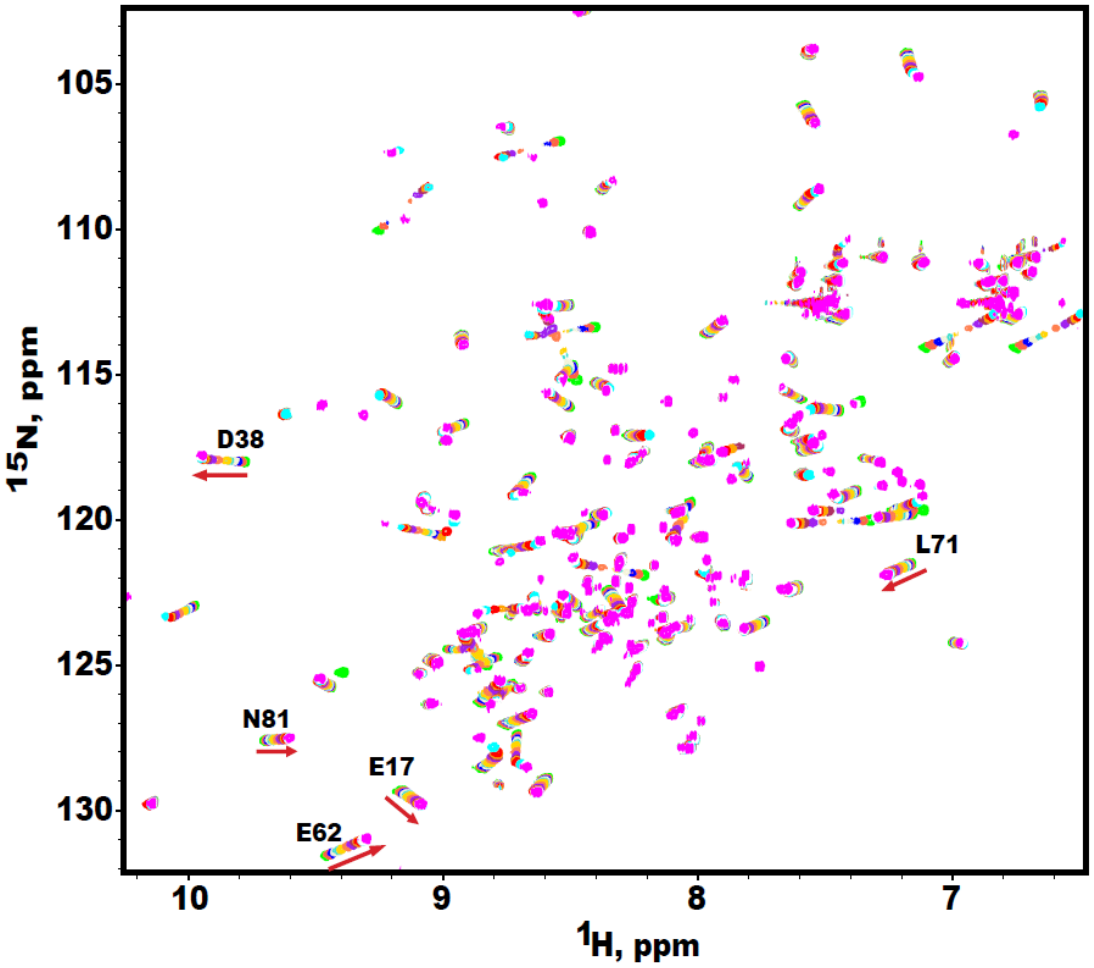


Figure 4.6: Changes of 2D $^{1}H,^{15}N$-HSQC spectra upon addition of the BAI2 peptide to 100 μM of ^{15}N-labeled GIP. The 2D $^{1}H,^{15}N$-HSQC spectra demonstrating chemical shift perturbations of residues upon titration of the peptide to GIP. Ratios of GIP to the peptide range from 1:0 to 1:60.

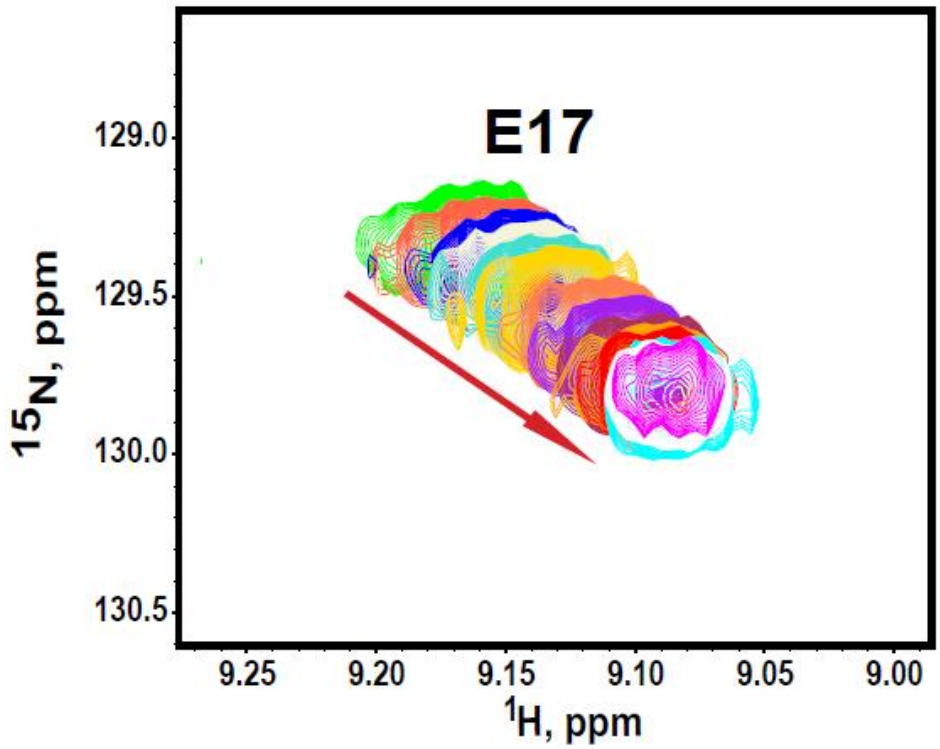


Figure 4.7: Expanded region of the spectra demonstrating the chemical shift perturbations of residue E17 upon titration of GIP with the BAI2 peptide. Ratios of GIP to the peptide are 1:0 (green), 1:0.2 (tomato), 1:0.4 (blue), 1:0.6 (beige), 1:0.8 (turquoise), 1:1 (gold), 1:2 (coral), 1:3 (purple), 1:5 (maroon), 1:7 (orange), 1:10 (red), 1:20 (cyan), 1:40 (white), 1:60 (magenta).

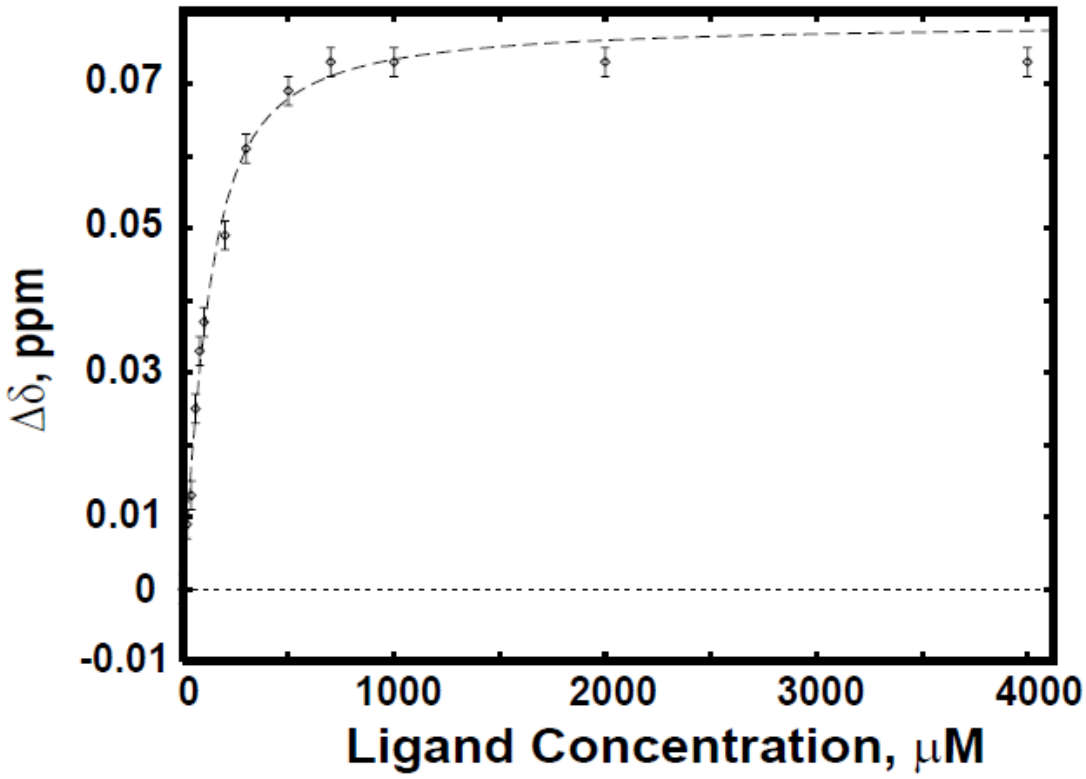


Figure 4.8: The NMR titration binding curve for the titration of GIP with the BAI2 peptide. The plot shows the changes in the chemical shift of E17 induced by the addition of peptide versus the peptide concentration. Dashed line is the titration curve as fit by the program ModelTitr from NMRPipe. The apparent dissociation constant K_D corresponding to residue E17 was determined by fitting the chemical shift change of the residue to increasing concentrations of peptide. The determined K_D value was 64.8 ± 10.6% μM.

Interaction with BAI2 peptide	
Residues of GIP	Dissociation constants, μM
E17	64.78±10.64%
R22	101±5.6%
D38	92.5±3%
F46	137.7±7.57%
E62	102.3±5.27%
A66	86.58±2.20%
L71	85.85±7.82%
N81	84.28±7.29%
T86	104.3±6.46%
E102	118.4±6.26%

Table 4.1: Dissociation constants of various residues of GIP upon binding with the BAI2 peptide by NMR.

4.3.3.4 Chemical shift perturbations of GIP upon binding to the BAI2 peptide (*1*)

Mapping the chemical shift perturbation with respect to residue number for a protein is a way to demonstrate the putative interacting portions of a protein with its interacting partner. For the mapping study of GIP with the BAI2 peptide, a series of the 2D $^{1}H,^{15}N$-HSQC spectra of GIP while titrating with increasing peptide concentrations were analyzed. The chemical shifts of most of the residues of GIP in both free and complex forms were determined. During analysis of the 2D $^{1}H,^{15}N$- HSQC spectra, the amide proton and nitrogen resonances of most residues showed gradual shifts with increasing peptide concentration, indicating that the complex was in the fast exchange regime on the NMR time scale. However, some residues disappeared or decreased in intensity below the noise level threshold with increasing peptide concentrations but reappeared at higher peptide concentrations suggesting that these residues were in intermediate exchange on the NMR time scale. For example, Leu 29 and Gly 30 initially disappeared with increasing peptide concentrations but reappeared at high peptide concentrations. Some of the residues could not be characterized for this mapping study because of the complete absence of the peak from the HSQC spectrum or peak overlap. These residues included Met 1, all five proline residues, Val 12, Leu 21, Phe 31, Glu 48, Lys 50, Val 57, Val 80 and Val 105.

Residues that constitute the β2 strand (residues 31 to 35) and the α2 helix (residues 90 to 97) of the protein showed the most chemical shift perturbations compared to other residues as seen on the 2D HSQC spectrum and mapping of chemical shift perturbations (**Figure 4.6** and **Figure 4.9**). This observation is consistent with that of interaction of GIP with a canonical C-terminal binding motif recognition peptide (*10, 12*). Most of the residues located within this

region showed greater than 0.1 ppm perturbations except residues Gln 92, Ala 93 and Leu 97 (**Figure 4.9**). The large perturbations occurred because the peptide directly interacted with most of these residues of the β2 strand and α2 helix. Residues Leu29 and Gly30 showed very large perturbations (greater than 1.0 ppm) (**Figure 4.9**) probably due to the hydrogen bonding formed between these two residues and the C-terminal end of the peptide (*38*). Such large chemical shift perturbations for Leu29 and Gly30 are reminiscent of our previous work on the interaction of GIP with a C-terminal peptide analog of Glutaminase L that was reported previously (*10*). Also, another cluster of residues showing prominent perturbations were residues 66 to residues 71 that form the α1 helix of the protein (**Figure 4.9**). Within this region, residues Ala 66, Glu 67, Ile 68 and Ala 69 showed greater perturbations (greater than 0.1 ppm). The significant changes in chemical shifts of this region (α1 helix) of the protein were not due to the direct interaction with the peptide but rather due to the change in the surrounding environment of the helix since this helix is in close proximity to the binding pocket of the protein. In the work shown in the previous chapter, several long-range NOEs were observed between Ile 28 and the α1 helix indicating a close spatial proximity between the βa-βb loop and the α1 helix for the free state of the protein but only a very few NOEs were present for that region of the complex form of the protein with Glutaminase L peptide (BMRB entry: 17254 and 17255) (*10*). Thus, the reason for comparatively higher chemical shift perturbation for residue Ile 28 (greater than 0.5 ppm) (**Figure 4.9**) could be twofold. First it is very close to the binding pocket. Second the binding of the BAI2 peptide to the protein probably resulted in the disruption of the interaction (NOEs) between residue Ile 28 and α1 helix. Although there were certain pockets of residues that showed significant chemical shift perturbations, the binding of the peptide to the protein seemed to induce a change in the chemical environment over nearly the entire protein except for the

termini. The N- and C-termini of the protein did not show any significant changes in the chemical shifts (**Figure 4.9**) upon peptide binding. Thus, the mode of BAI2 peptide binding to GIP can be characterized as allosterically driven analogous to the binding of the Glutaminase L peptide to GIP (*10*).

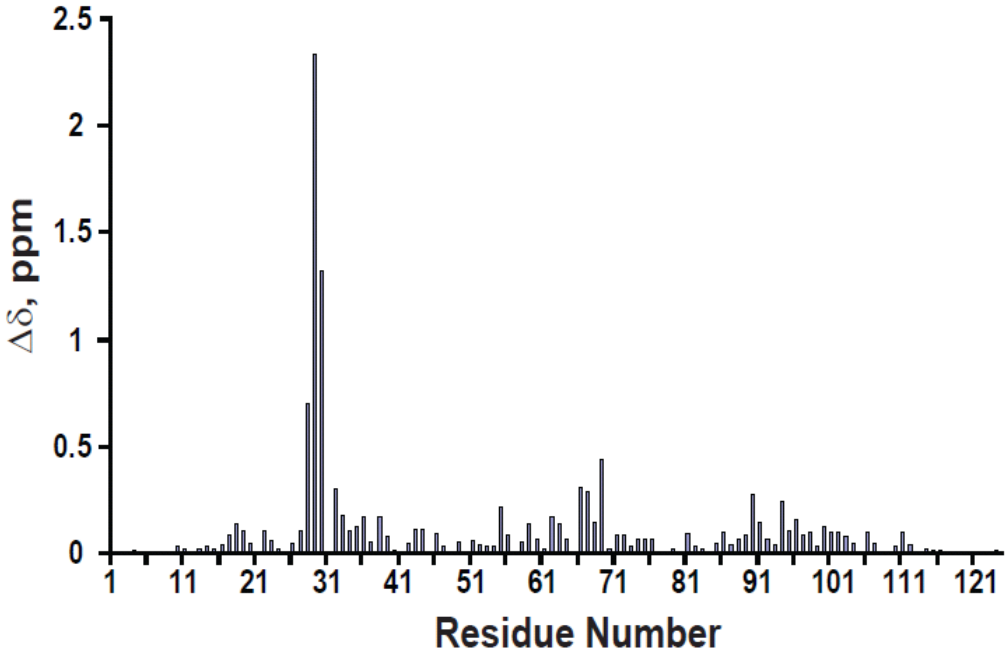


Figure 4.9: Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the BAI2 peptide.

BAI2 is a member of the adhesion-G protein-coupled receptors (GPCRs) (*39, 40*)**.** It is composed of 521-amino acids and mainly expressed in neurons (*41*). BAI2 possesses a Src homology 3 (SH3) domain, composed of 50-60 amino acids that mediates protein-protein interactions and was previously reported as interacting with the C-terminus of Brain-Specific Angiogenesis Inhibitor 1 (BAI1) via its SH3 domain as shown by *in vitro* binding assays (*41*). This was the first study reporting an interaction between BAI2 and GIP with an extensive biophysical characterization of their interaction (*1*).

4.3.4 Interaction of the control peptide with GIP

4.3.4.1 Characterization by Fluorescence spectroscopy

When the control peptide was titrated against unlabeled GIP, it showed a small but consistent decrease in fluorescence intensity (**Figure 4.10**). The dissociation constant K_D ($K_D = 1/K_a$) was determined using the OriginPro 6.1 software. The decrease in the fluorescence intensity was calculated as $(F_0 - F_C)/(F_0 - F_{min})$, where F_0 is the initial fluorescence intensity of free GIP; F_C is the corrected fluorescence intensity at a ligand concentration [C], and F_{min} is the fluorescence intensity at the saturating concentration of the peptide. The data were fitted to a nonlinear regression of the plot of $(F_0 - F_C)/(F_0 - F_{min})$ against [C] with the equation corresponding to a single binding site (**Figure 4.11**). The titration of the control peptide with GIP yielded a dissociation constant of 1.07 μM. To determine the thermodynamic nature of the interaction, the free energy change of the association was calculated using the following equation: $\Delta G = -RT \ln K_a$, where K_a is the association constant, T is temperature and R is universal gas constant. By putting the experimentally determined K_a ($K_a = 1/K_D$) value into this equation, the ΔG value for binding of the BAI2 peptide to GIP was calculated to be -34.06 kJ mol^{-1}, which reflects the spontaneous binding of the peptide to GIP.

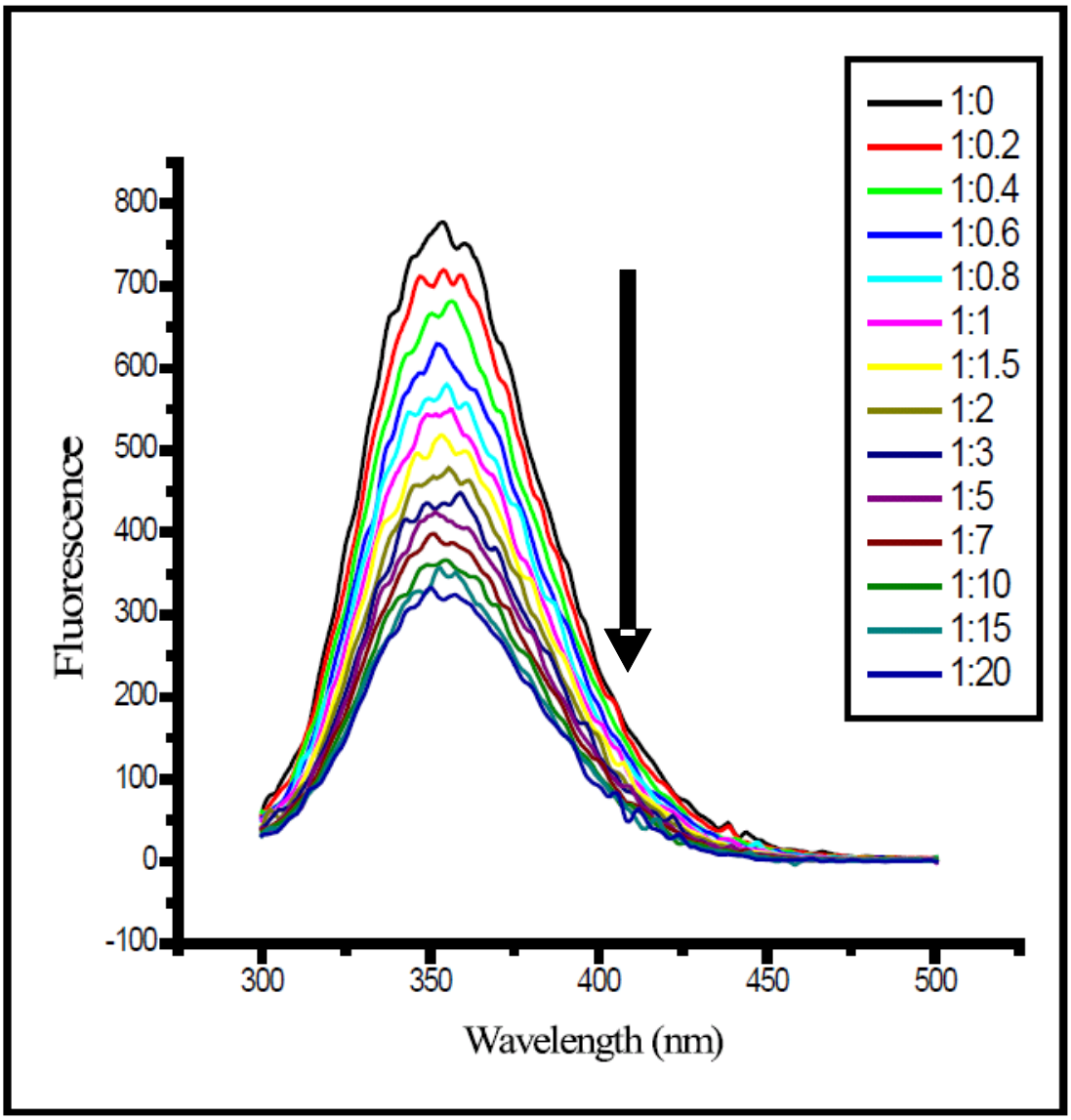


Figure 4.10: Fluorescence emission spectrum of GIP with the control peptide. Fluorescence emission plots corresponding to (top to bottom) 0 to 20 μM concentrations of the peptide to 1 μM protein sample. In the legend, protein to peptide ratios are indicated with the respective color codes. Black arrow indicates the quenching of fluorescence of GIP upon peptide binding in a downward fashion.

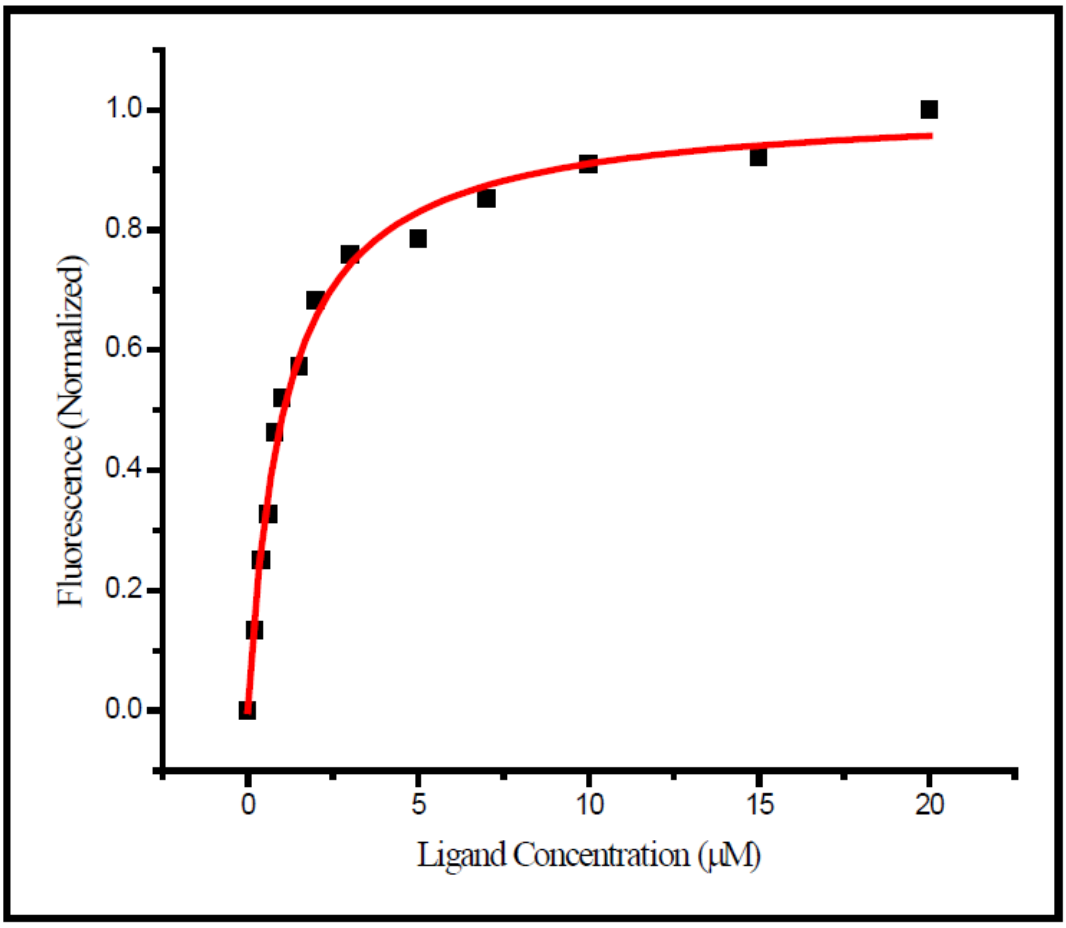


Figure 4.11: Non-linear curve fitting assuming 1:1 binding between GIP and the control peptide where $(F_0 - F_C)/(F_0 - F_{min})$ was plotted against peptide concentration.

4.3.4.2 Characterization by CD spectroscopy

Like the BAI2 peptide, the interaction between the control peptide and GIP was also characterized by CD spectroscopy. The control peptide did not alter the CD spectrum from that obtained with phosphate buffer. However, any contribution from the peptide and buffer was subtracted from the CD spectrum obtained in subsequent analyses of GIP with peptide. The secondary structure of GIP showed significant changes in the CD spectrum with the titration of different concentrations of the peptide (**Figure 4.12**).

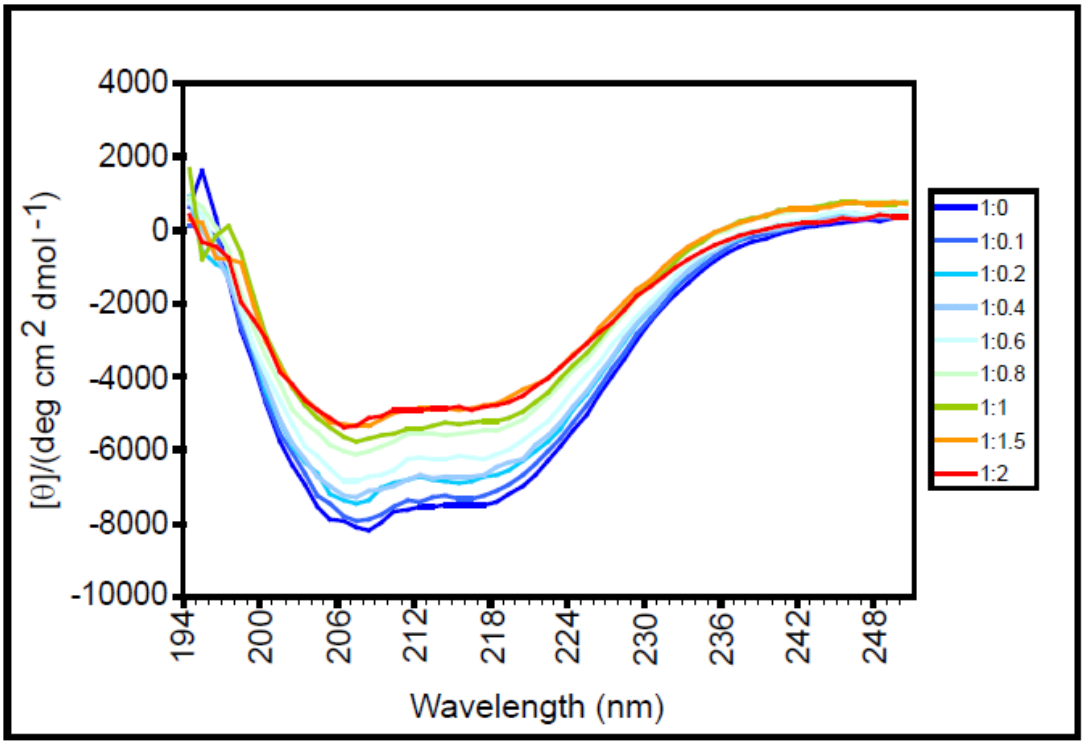


Figure 4.12: Changes in the CD spectra of GIP upon binding with increasing concentrations of the control peptide for the wavelength range of 194 nm to 250 nm. The protein to peptide ratios for the corresponding color codes are indicated in the legend.

4.3.4.3 Characterization by $^{1}H,^{15}N$-HSQC NMR

Like the interaction with the BAI2 peptide, for the investigation of any possible binding and, less importantly, a subsequent conformational change in GIP, ^{15}N-labeled GIP protein was titrated against the control peptide to excess (~60 times that of the protein) until complete saturation was achieved. During the course of the titration, the fingerprint region of the protein in the 2D $^{1}H,^{15}N$-HSQC spectra was monitored. The fingerprint region of the HSQC spectra of GIP was collected in the absence and presence of different concentrations of the peptide and the profiles were overlaid (**Figure 4.13**). From the overlay, it was evident that most of the residues of GIP showed changes in chemical shifts only slightly (if any) upon binding with the peptide. Using the program ModelTitr (*29*), the dissociation constant (K_D) values for various residues of GIP were calculated (**Table 4.2**) by non-linear least-squares fitting of the chemical shift data against ligand concentration to the Langmuir isotherm that involved the assumption of a stoichiometry of 1:1 between the ligand and the protein (*i.e.* one binding site) (**Figure 4.15**). The dilution effect on the concentration of the protein due to the addition of the ligand was corrected in the program. The calculated dissociation constant (K_D) value from NMR technique (717.97 μM on an average) was different from the value obtained from fluorescence technique as was the case with the BAI2 peptide. From the K_D values of both fluorescence and NMR techniques, the dissociation constant (K_D) value falls in the range of low to mid μM, which indicates a moderate affinity of GIP for the control peptide.

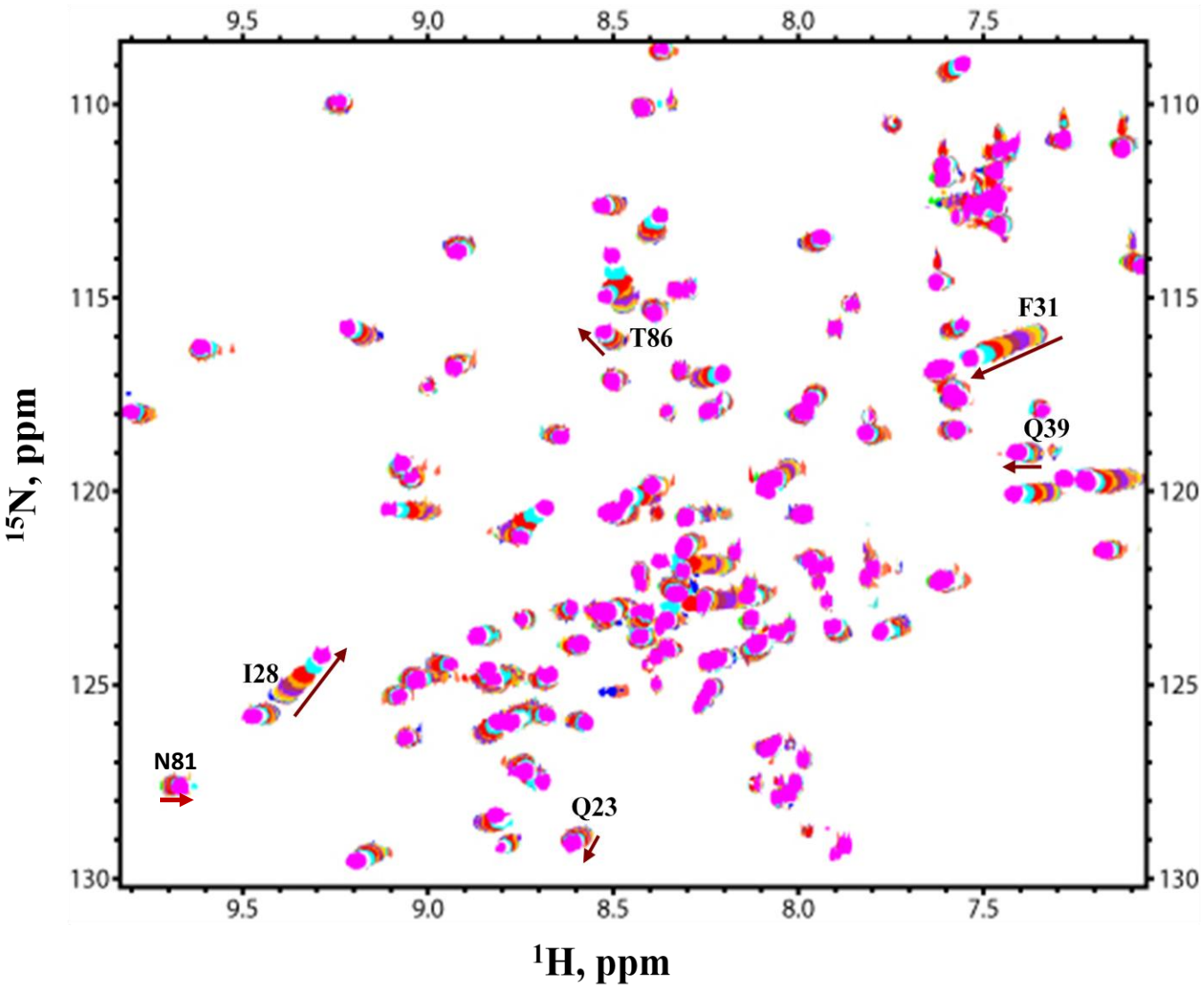


Figure 4.13: Changes of 2D ^{1}H,^{15}N-HSQC spectra upon addition of the control peptide to 100 μM of ^{15}N-labeled GIP. The 2D ^{1}H,^{15}N-HSQC spectra demonstrating chemical shift perturbations of residues upon titration of the peptide to GIP. Ratios of GIP to the peptide ranged from 1:0 to 1:60.

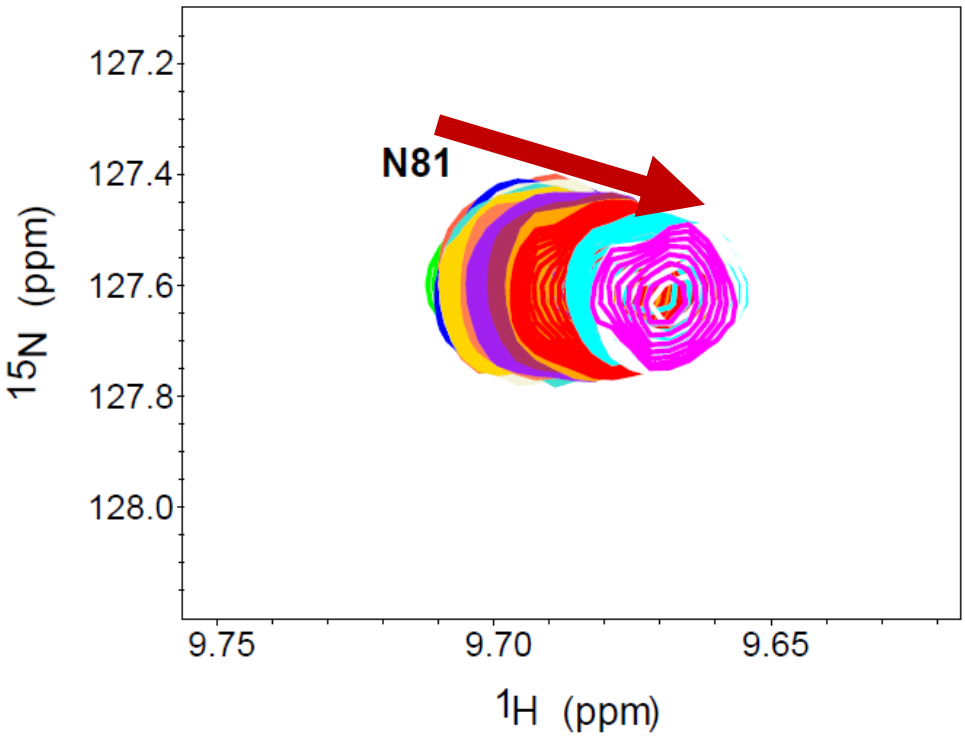


Figure 4.14: Expanded region of the spectra demonstrating the chemical shift perturbations of residue N81 upon titration of GIP with the control peptide. Ratios of GIP to the peptide were 1:0 (green), 1:0.2 (tomato), 1:0.4 (blue), 1:0.6 (beige), 1:0.8 (turquoise), 1:1 (gold), 1:2 (coral), 1:3 (purple), 1:5 (maroon), 1:7 (orange), 1:10 (red), 1:20 (cyan), 1:40 (white), 1:60 (magenta).

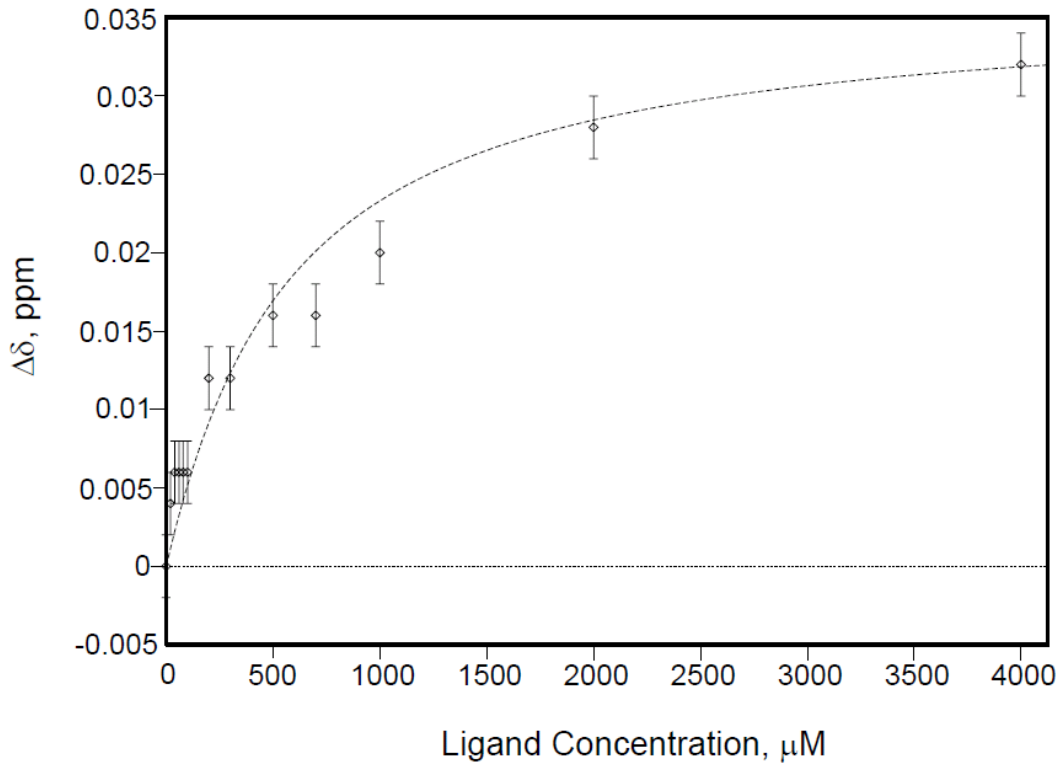


Figure 4.15: The NMR titration binding curve for the titration of GIP with the control peptide. The plot shows the changes in the chemical shift of N81 induced by the addition of peptide versus the peptide concentration. Dashed line is the titration curve as fit by the program ModelTitr from NMRPipe. The apparent dissociation constant K_D corresponding to residue N81 was determined by fitting the chemical shift change of the residue to increasing concentrations of peptide. The determined K_D value was 510.1 ± 24.8% μM.

Interaction with RGGSRL	
Residues of GIP	**Dissociation constants, μM**
Q23	527.4±59.23%
I28	1169±6.70%
F31	734.1±3.10%
G34	485.7±66.98%
Q39	1049±19.21%
I68	458.6±24.61%
N81	510.1±24.79%
D84	953±31.25%
T86	480.6±63.81%
R106	812.2±17.21%

Table 4.2: Dissociation constants of various residues of GIP upon binding with the control peptide by NMR.

4.3.4.4 Chemical shift perturbations of GIP upon binding to the control peptide

For the mapping study of GIP with the control peptide, a series of the 2D $^{1}H,^{15}N$-HSQC spectra of GIP with increasing peptide concentrations were analyzed. The chemical shifts of most of the residues of GIP in both free and complex forms were determined. During analysis of the 2D $^{1}H,^{15}N$- HSQC spectra, the amide proton and nitrogen resonances of most residues showed gradual shifts with increasing peptide concentration, indicating that the complex was mostly in the fast exchange regime in the NMR time scale. Unlike the perturbation study with the BAI2 peptide, Leu 29 and Gly 30 did not reappear at a higher peptide concentration and could not be mapped in the study. Along with these two, some of the residues could not be characterized for this mapping study because of the complete absence of the peak from the HSQC spectrum or peak overlapping. These residues included Met 1, all five proline residues, Val 12, Leu 21, Phe 31, Glu 48, Lys 50, Val 57, Val 80 and Val 105.

Residues that constitute the β2 strand (residues 31 to 35) and the α1 helix (residues 66 to 71) of the protein showed the most chemical shift perturbations compared to other residues as seen on the mapping of chemical shift perturbations (**Figure 4.16**). Most of the residues located within this region showed greater than 0.05 ppm perturbations except residues Gly 34, Gly 70 and Leu 71 (**Figure 4.16**). Residues that form α2 helix (residues 90 to 97) had perturbations greater than 0.01 ppm. Apparently, these three clusters of residues were most perturbed due to the interaction with the control peptide. This feature is consistent with that of the interaction of GIP with a canonical C-terminal binding motif recognition peptide (*10, 12*). As in the interaction with the BAI2 peptide, in this perturbation study with the control peptide residue Ile 28 showed

perturbation of greater than 0.1 ppm (**Figure 4.16**) for the same possible reasons listed in case of interaction with the BAI2 peptide (section **4.3.3.4**). Although there were certain pockets of residues that showed significant chemical shift perturbations, the binding of the peptide to the protein seemed to induce a change in the chemical environment over nearly the entire protein except for the termini. The N- and C-termini of the protein did not show any significant changes in the chemical shifts (**Figure 4.16**) upon peptide binding.

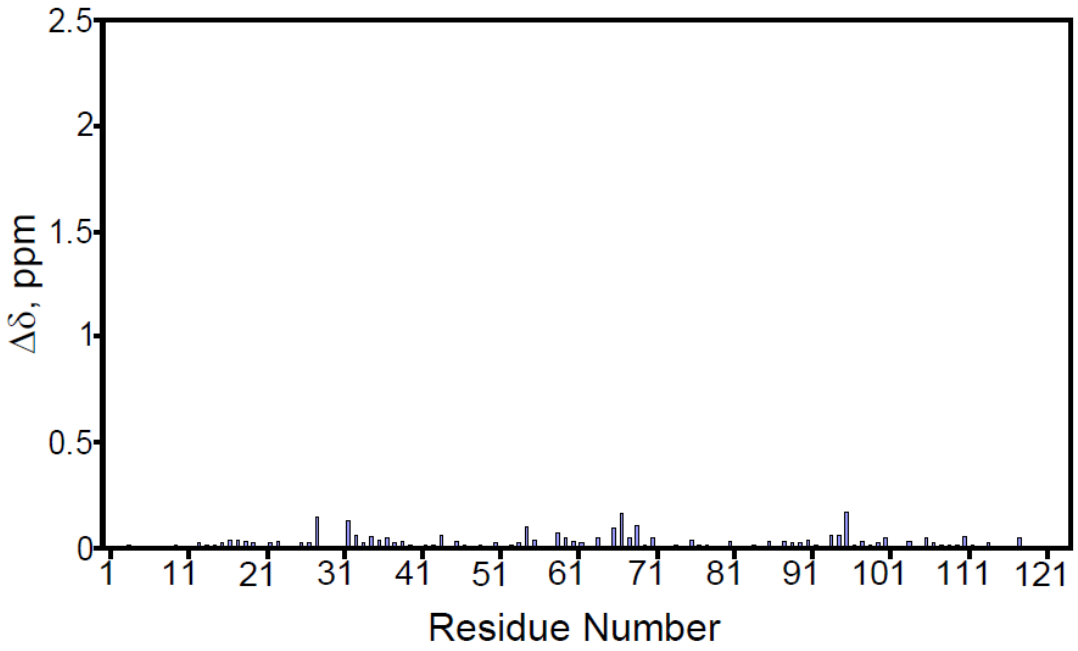


Figure 4.16: Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the control peptide.

4.3.5 Comparison of interaction between GIP and BAI2 peptide with interaction between GIP and control peptide

To determine the comparative strength of the interaction between two interacting partners, NMR would be deemed as the most appropriate experimental tool since it is the most sensitive technique. The interaction between GIP and the two interacting partners was monitored by examining the series of 2D ^{1}H, ^{15}N-HSQC titration spectra corresponding to the increasing concentrations of the interacting peptides. These spectra were then overlaid to reflect the perturbations of the residues upon binding. Comparison of these two overlays for both of the peptides easily revealed that interaction with the BAI2 peptide appeared to cause more perturbations than with the control peptide (**Figure 4.6** and **Figure 4.13**). The dissociation constant (K_D) values determined from NMR for the residues of GIP were on an average about 7 times lower for the BAI2 peptide than with the control peptide (**Table 4.1** and **Table 4.2**). This suggests that the binding of BAI2 peptide to GIP is at least 7 times stronger than that of control peptide to GIP. Moreover, the overlay of chemical shift perturbations map for both of the peptides easily reflected the overall greater chemical shift perturbations for the BAI2 peptide (**Figure 4.17**). The calculated Gibbs‟ free energy (ΔG) from the K_D values determined from the fluorescence technique also showed an amount of 1.02 kJ mol^{-1} extra energy released as a result of binding of the BAI2 peptide to GIP compared to that of the control peptide to GIP. In summary, GIP seems to interact with BAI2 peptide more strongly than the control peptide. Such a preference of interaction might lie in the sequence of the peptide. As discussed in the previous chapter, one of the important interaction between GIP and the canonical C-terminus of the peptide is the hydrophobic interaction formed between the hydrophobic residue at the P_0 position of the peptide and the hydrophobic pocket created by Leu 29, Phe 31, Leu 97 and Ile 33 as well

as Thr 98 at the periphery of GIP (*10*). The BAI2 peptide (RDGDFQTEV) has a Val at its P_0 whereas the control peptide (RGGSRL) has Leu at its P_0. Both are hydrophobic, but Leu is one (-CH_2-) group long than Val. Larger side chain of Leu might cause a steric hindrance in the hydrophobic pocket of GIP leading to the disruption of interaction between GIP and the control peptide. Whereas, the smaller side chain of Val at P_0 of the BAI2 peptide allows a more favorable interaction with GIP. Such a phenomenon could also be observed when a microarray technique was utilized to determine the protein interaction network of mouse PDZ domain with moderate to high affinity ($K_D \leq 10$ μM). Among the 20 interacting peptides used, 16 had Val at their P_0 position (*42*).

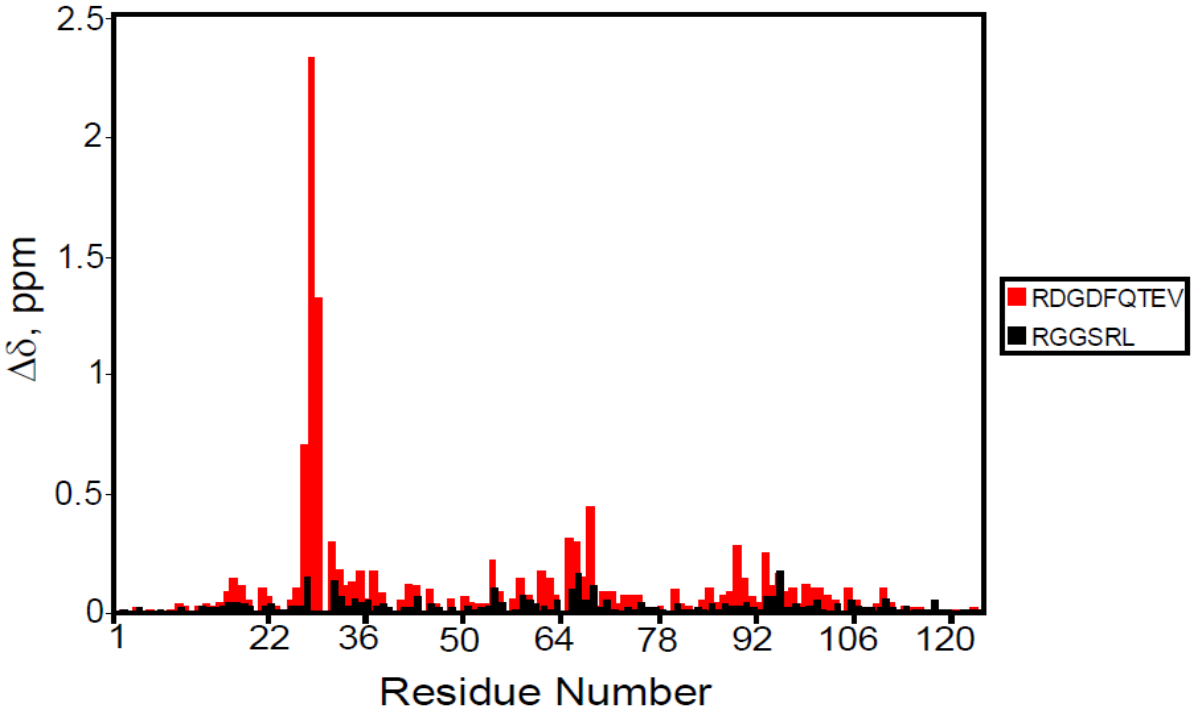


Figure 4.17: Chemical shift perturbations (Δδ) of the GIP backbone amide groups upon binding with the BAI2 (red) and the control (black) peptide.

4.4 References

1. Zencir, S., Ovee, M., Dobson, M. J., Banerjee, M., Topcu, Z., and Mohanty, S. (2011) Identification of brain-specific angiogenesis inhibitor 2 as an interaction partner of glutaminase interacting protein, *Biochem Biophys Res Commun 411*, 792-797.
2. Olalla, L., Aledo, J. C., Bannenberg, C., and Marquez, J. (2001) The C-terminus of human glutaminase L mediates association with PDZ domain-containing proteins, *Febs Letters 488*, 116-122.
3. 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.
4. Fanning, A. S., and Anderson, J. M. (1996) Protein-protein interactions: PDZ domain networks, *Current Biology 6*, 1385-1388.
5. Jelen, F., Oleksy, A., Smietana, K., and Otlewski, J. (2003) PDZ domains - common players in the cell signaling, *Acta Biochim Pol 50*, 985-1017.
6. Garcia-Mata, R., and Burridge, K. (2007) Catching a GEF by its tail, *Trends in Cell Biology 17*, 36-43.
7. Topcu, Z., and Borden, K. L. B. (2000) The yeast two-hybrid system and its pharmaceutical significance, *Pharmaceutical Research 17*, 1049-1055.
8. Zhang, J. X., Yan, X. J., Shi, C. W., Yang, X., Guo, Y., Tian, C. L., Long, J. F., and Shen, Y. Q. (2008) Structural Basis of beta-Catenin Recognition by Tax-interacting Protein-1, *Journal of Molecular Biology 384*, 255-263.
9. Hung, A. Y., and Sheng, M. (2002) PDZ domains: Structural modules for protein complex assembly, *Journal of Biological Chemistry 277*, 5699-5702.

10. 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-Us 50*, 3528-3539.

11. Cho, K. O., Hunt, C. A., and Kennedy, M. B. (1992) The Rat-Brain Postsynaptic Density Fraction Contains a Homolog of the Drosophila Disks-Large Tumor Suppressor Protein, *Neuron 9*, 929-942.

12. 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.

13. Wiedemann, U., Boisguerin, P., Leben, R., Leitner, D., Krause, G., Moelling, K., Volkmer-Engert, R., and Oschkinat, H. (2004) Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of superbinding peptides, *Journal of Molecular Biology 343*, 703-718.

14. Tonikian, R., Zhang, Y. N., Sazinsky, S. L., Currell, B., Yeh, J. H., Reva, B., Held, H. A., Appleton, B. A., Evangelista, M., Wu, Y., Xin, X. F., 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 Biology 6*, 2043-2059.

15. Durney, M. A., Birrane, G., Anklin, C., Soni, A., and Ladias, J. A. A. (2009) Solution structure of the human Tax-interacting protein-1, *Journal of Biomolecular Nmr 45*, 329-334.

16. Medina, M. A., Sanchezjimenez, F., Marquez, J., Quesada, A. R., and Decastro, I. N. (1992) Relevance of Glutamine-Metabolism to Tumor-Cell Growth, *Molecular and Cellular Biochemistry 113*, 1-15.

17. Brand, K. (1985) Glutamine and Glucose-Metabolism during Thymocyte Proliferation - Pathways of Glutamine and Glutamate Metabolism, *Biochemical Journal 228*, 353-361.

18. Kovacevic, Z., and Mcgivan, J. D. (1983) Mitochondrial Metabolism of Glutamine and Glutamate and Its Physiological Significance, *Physiological Reviews 63*, 547-605.

19. 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, *Journal of Biological Chemistry 278*, 38758-38764.

20. 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, *Journal of Biological Chemistry 272*, 20979-20981.

21. Hampson, L., Li, C. G., 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, *International Journal of Oncology 25*, 1249-1256.

22. 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, *Journal of Biological Chemistry 275*, 33962-33968.

23. Le Maout, S., Welling, P. A., Brejon, M., Olsen, O., and Merot, J. (2001) Basolateral membrane expression of a K+ channel, Kir 2.3, is directed by a cytoplasmic COOH-

terminal domain, *Proceedings of the National Academy of Sciences of the United States of America 98*, 10475-10480.

24. Alewine, C., Olsen, O., Wade, J. B., and Welling, P. A. (2006) TIP-1 has PDZ scaffold antagonist activity, *Molecular Biology of the Cell 17*, 4200-4211.

25. Olalla, L., Gutierrez, A., Jimenez, A. J., Lopez-Tellez, J. F., Khan, Z. U., Perez, J., Alonso, F. J., de la Rosa, V., Campos-Sandoval, J. A., Segura, J. A., Aledo, J. C., and Marquez, J. (2008) Expression of the scaffolding PDZ protein glutaminase-interacting protein in mammalian brain, *Journal of Neuroscience Research 86*, 281-292.

26. Fields, S., and Song, O. K. (1989) A Novel Genetic System to Detect Protein Protein Interactions, *Nature 340*, 245-246.

27. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes, *Journal of Biomolecular Nmr 6*, 277-293.

28. Kneller, T. D. G. a. D. G. SPARKY 3, *University of California*, San Francisco.

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. Filiou, M. D., Varadarajulu, J., Teplytska, L., Reckow, S., Maccarrone, G., and Turck, C. W. (2012) The 15N isotope effect in Escherichia coli: a neutron can make the difference, *Proteomics 12*, 3121-3128.

31. 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 Biochem 287*, 252-260.

32. Yan, X. J., Zhou, H., Zhang, J. X., Shi, C. W., Xie, X. Q., Wu, Y. N., Tian, C. L., Shen, Y. Q., and Long, J. F. (2009) Molecular Mechanism of Inward Rectifier Potassium Channel 2.3 Regulation by Tax-Interacting Protein-1, *Journal of Molecular Biology 392*, 967-976.

33. Mao, H. Y., 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, *Journal of the American Chemical Society 123*, 10429-10435.

34. 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, *Journal of Biological Chemistry 284*, 32167-32177.

35. Hu, H. Y., Horton, J. K., Gryk, M. R., Prasad, R., Naron, J. M., Sun, D. A., Hecht, S. M., Wilson, S. H., and Mullen, G. P. (2004) Identification of small molecule synthetic inhibitors of DNA polymerase beta by NMR chemical shift mapping, *Journal of Biological Chemistry 279*, 39736-39744.

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

37. Jemth, P., and Gianni, S. (2007) PDZ domains: Folding and binding, *Biochemistry 46*, 8701-8708.

38. 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, *Nature Structural Biology 5*, 19-24.

39. Jeong, B. C., Kim, M. Y., Lee, J. H., Kee, H. J., Kho, D. H., Han, K. E., Qian, Y. R., Kim, J. K., and Kim, K. K. (2006) Brain-specific angiogenesis inhibitor 2 regulates VEGF through GABP that acts as a transcriptional repressor, *Febs Letters 580*, 669-676.

40. Okajima, D., Kudo, G., and Yokota, H. (2010) Brain-specific angiogenesis inhibitor 2 (BAI2) may be activated by proteolytic processing, *Journal of Receptors and Signal Transduction 30*, 143-153.

41. Oda, K., Shiratsuchi, T., Nishimori, H., Inazawa, J., Yoshikawa, H., Taketani, Y., Nakamura, Y., and Tokino, T. (1999) Identification of BAIAP2 (BAI-associated protein 2), a novel human homologue of hamster IRSp53, whose SH3 domain interacts with the cytoplasmic domain of BAI1, *Cytogenetics and Cell Genetics 84*, 75-82.

42. Stiffler, M. A., Grantcharova, V. P., Sevecka, M., and MacBeath, G. (2006) Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays, *J Am Chem Soc 128*, 5913-5922.

Chapter 5

Characterization of subunit A of Heterodisulfide Reductase (HdrA) from *Methanothermobacter marburgensis*

5.1 Introduction

5.1.1 Electron bifurcation

In a recently discovered process, it has been found that enzymes can produce electrons with a very low redox potential without the involvement of ATP-hydrolysis or radical-SAM enzymes. This process has been termed electron bifurcation. In this process, two electrons enter the bifurcation cycle at a certain redox potential. Of these two electrons, one comes out at a much lower potential and the other comes out at a much higher potential. Although, these electrons are physically separated, there is a tight coupling in this process. To produce one type of electron the other needs to be simultaneously generated while at the same time they continue on separate electron paths in the enzyme complex. One electron stays on a high-potential branch and the other on a low-potential branch.

5.1.2 History of electron bifurcation

Buckel *et al.* first developed the concept of electron bifurcation while trying to explain their observations on enzymes from *Clostridia*(*1, 2*). This concept was then adopted by Thauer *et al.* for methanogens, specifically for *Methanothermobacter marburgensis*(*3, 4*). In this organism, they identified the methylviologen-reducing hydrogenase (Mvh)/heterodisulfide reductase (Hdr)

as the enzyme complex (MvhADG/HdrABC) that performs the electron bifurcation (**Figure 5.1**) (*4*). A similar complex was also discovered in *Methanobacterium thermoautotrophicum* by this group (*5*). Independently, the phenomenon of electron bifurcation was also successfully identified by the Leigh group in *Methanococcus maripaludis*(*6*).

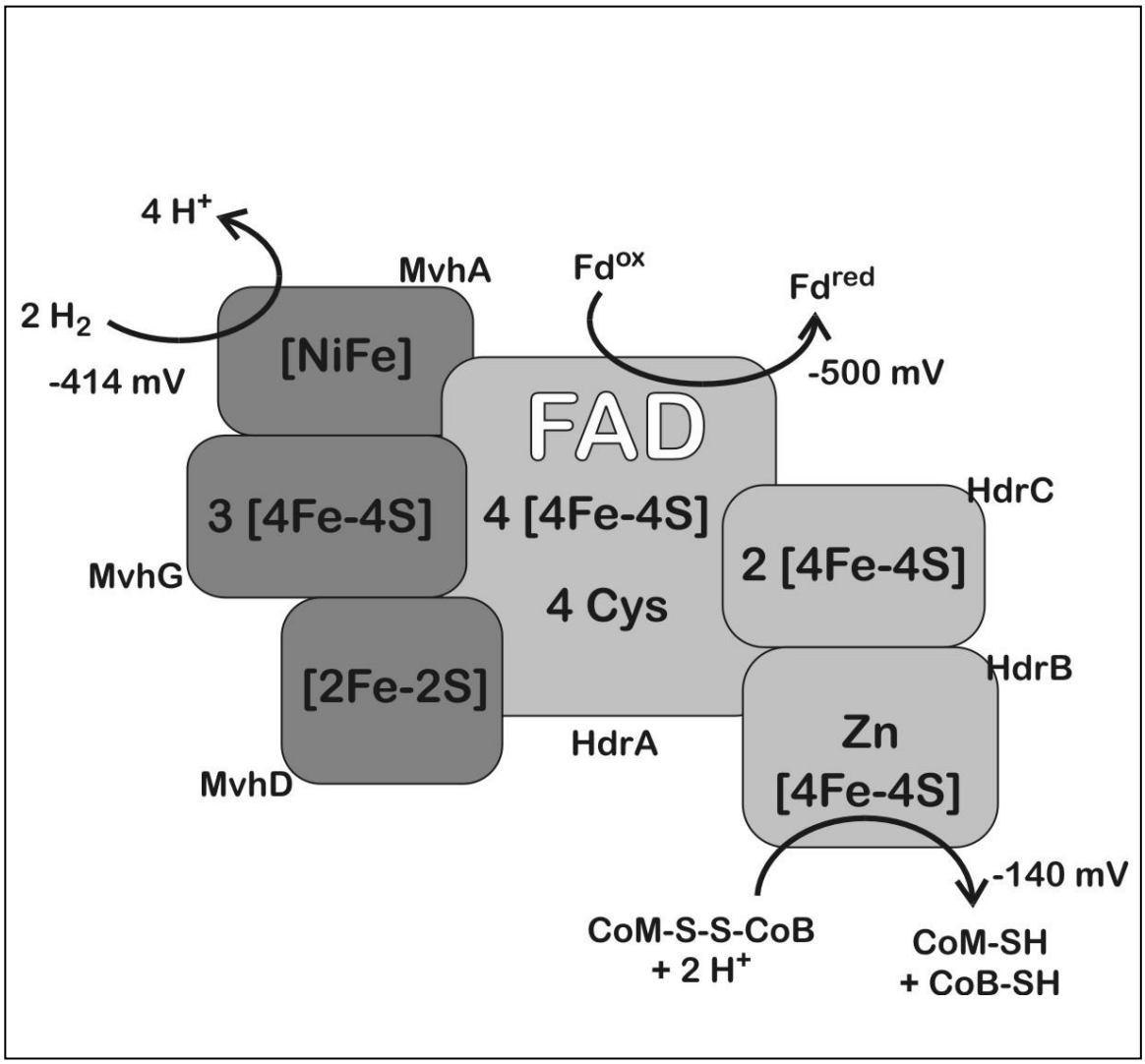


Figure 5.1: Model of the structure of the hydrogenase:heterodisulfide reductase complex from *Methanothermobacter marburgensis*.

5.1.3 Mechanism of electron bifurcation

For the electron bifurcation to take place, it seems to be essential that at least one flavin or quinone molecule is present. In the Hdr complexes, it has therefore been proposed that only the site of the FAD (flavin adenine dinucleotide) molecule is the site where the electron bifurcation takes place. FAD is very labile and easily lost from the protein complex. However, it is also easily reconstituted back into the enzyme; making it easy to prove the essentiality of the presence of FAD for this process. Thauer and coworkers proposed a model for the electron bifurcation based on the fact that flavoproteins (FP) can exhibit three different redox potentials, namely an E_o' for the FP/FPH_2 couple (n = 2), an E_o' for the $FP/FPH^\bullet$ couple (n = 1), and an E_o' for the $FPH^\bullet/FPH_2$ couple (n = 1). E_o' ($FP/FPH^\bullet$) is generally more positive and E_o' ($FPH^\bullet/FPH_2$) more negative than E_o' (FP/FPH_2). In the proposed model, the flavin is reduced by two electrons to the FPH_2 form with an intermediate potential. Subsequently it first forms $FPH^\bullet$, releasing the low-potential electron, followed by oxidation to FP, releasing the high-potential electron.

In *M. marburgensis*, the hydrogenase:heterodisulfide reductase complex reduces heterodisulfide at a very low rate by using electrons from the oxidation of hydrogen. This activity of heterodisulfide reductase increases many fold when ferredoxin is added to the kinetic assay. This increase in the enzyme activity lies in the fact of the tight coupling of the ferredoxin reduction and the heterodisulfide reduction during the events of the bifurcation process. The midpoint potential of the H_2/H^+ couple is about -400 mV under cell growth conditions whereas the midpoint potential of the heterodisulfide/(HS-CoM + HS-CoB) couple is about -140 mV. Under this condition, the expected flow of electrons should be automatically from the site of

hydrogen oxidation to that of heterodisulfide reduction. But, as reflected by the enzyme activity assay, it does not constitute the major process unless ferredoxin is simultaneously reduced despite the fact that ferredoxin^{RED}/ferredoxin^{OX} couple has a midpoint potential of -500 mV. Thus, high-potential electrons are generated that are used for heterodisulfide reduction, while low-potential electrons are generated that only reduce ferredoxin. Thermodynamically, these latter electrons could also reduce heterodisulfide, but apparently the enzyme prevents this from happening.

5.1.4 Electron bifurcation in other systems

It is remarkable that the electrons that reduce ferredoxin have a lower potential than those released by hydrogen oxidation. ATP hydrolysis coupled to an electron transfer step is the more classic way for an enzyme to change redox potentials. Typical examples are the Fe-protein in the nitrogenase systems and archerases (*7*). Electron bifurcation is widespread in nature, however, in particular, it is found in the electron transport chain. In complex I of the oxidative phosphorylation pathway, NADH delivers two electrons and a proton to an FMN molecule that is bound to the protein. The FMN donates each electron to a separate iron-sulfur cluster, but the two pathways combine into a single path that reaches to the quinone reduction site. Bifurcation of electrons happens at complex III during the oxidation of ubiquinol (QH_2). One electron ends up at cytochrome c whereas the other electron follows a path containing cytochrome b_H and cytochrome b_L and ends up reducing another quinone molecule. Through crystallographic study, the Rieske 2Fe cluster of complex III is proposed to play an important role in the bifurcation by

inducing a conformational change upon oxidation or reduction causing the cluster to change its position relative to the position of the cytochrome c.

5.1.5 Models for electron bifurcation

Based on the above scenario in the oxidative phosphorylation chain, three models can be postulated to describe the bifurcated flow of electrons in our enzyme system (**Figure 5.2**).

5.1.5.1 Model I

In model I, two electrons are transferred to the FAD. The resultant change in charge causes it to move closer (at least the flavin part) to the high-potential [4Fe-4S] cluster in the same subunit (HdrA). When it releases one electron to the cluster, the now „red-hot" FADH then moves toward the bound ferredoxin to give up its second electron and be ready to accept two new electrons.

5.1.5.2 Model II

In model II, when one electron is transferred to the high-potential [4Fe-4S] cluster, the cluster moves away from the flavin site forcing the next electron to be transferred to the bound ferredoxin.

5.1.5.3 Model III

In model III, once the iron-sulfur cluster is reduced, the subsequent electron transfers to the heterodisulfide reduction site is so slow that even the highly reactive “red hot” semiquinone $FADH^{\bullet}$ state is not able to transfer the second electron to the cluster since that would create a $[4Fe\text{-}4S]^{0}$ state, but instead it just has to transfer the electron to the bound ferredoxin. The „0“ state is generally not attainable for 4Fe clusters, since the midpoint potential for the 1+/0 couple is very low. In this model, no movement is essential.

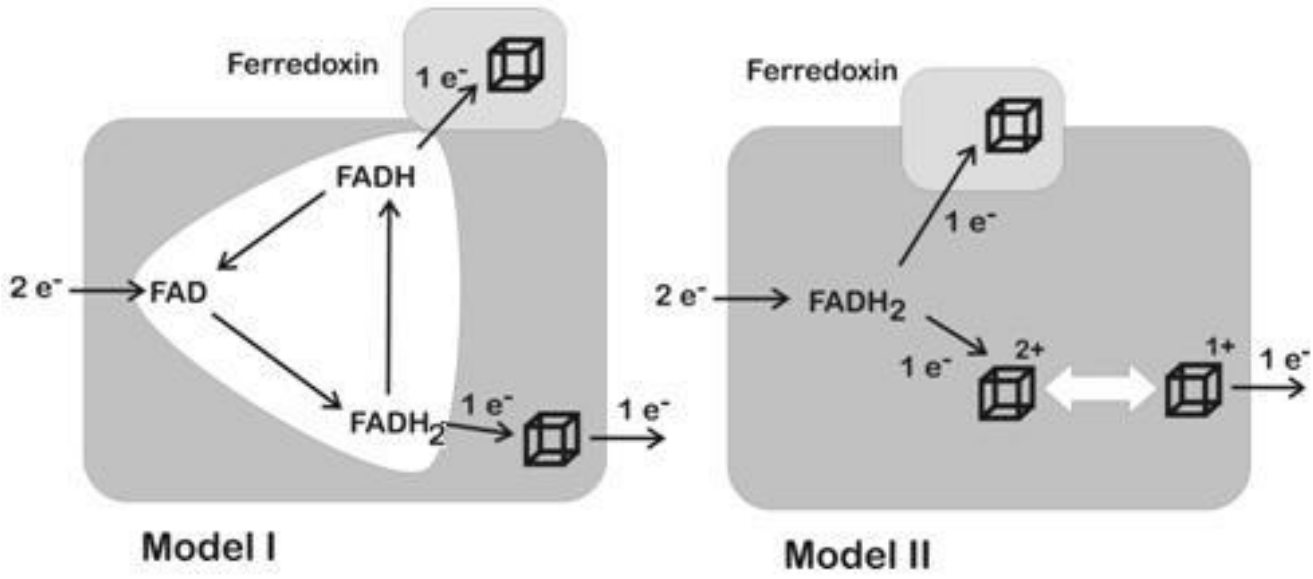


Figure 5.2: Models for electron bifurcation.

5.2 Objective of this study

To get a detailed understanding of the electron bifurcation process in the heterodisulfide reductase enzyme found in methanogenic Archaea we might have to study the whole hydrogenase:heterodisulfide reductase complex or just the heterodisulfide reductase. However to obtain basic information about this process it is important to obtain the smallest protein component or complex that still contains all the essential cofactors needed for electron bifurcation. The major reason behind this is more of a practical aspect rather than a conceptual understanding. The dominant method to be used to characterize the various redox processes involved during the bifurcation and to characterize the role of electron donors and acceptors is electron paramagnetic resonance (EPR) spectroscopy. Therefore there is an inherent requirement for simplicity of the system (less cofactors/iron-sulfur clusters) under investigation. Otherwise, in investigations with the larger enzyme complexes it would be harder (if not impossible) to interpret the EPR spectroscopic data. Also, establishment of the smallest yet functional subunit system should essentially establish the least but absolutely stringent requirements of the parts of the complex to be present for the electron bifurcation process to take place. As a result, such understanding should support or disprove our hypothesis behind this process.

So far, the spectroscopic data that is available, is all from either the hydrogenase:heterodisulfide reductase complex or the Hdr from *M. marburgensis* (*8-11*). Therefore our first focus was on the Hdr enzyme from this organism. The bifurcation has not been proven for the Hdr enzyme. The hydrogenase:heterodisulfide reductase complex can be purified using a three column-step purification method (*4*). The two enzymes can be separated

and the Hdr purified in another three column steps. To study the events that take place during the bifurcation process, it would be necessary to do mutational studies. Our collaborator John Leigh at the in University of Washington is well-equipped for work with the organism *M. maripaludis*. They already have a well-developed set of genetic tools available for this organism. Therefore it was also tested if the *M. marburgensis* Hdr can be overexpressed in *M. maripaludis*. This, however, was not possible, since the *M. maripaludis* strains expressing the Hdr subunits would not grow. However, the HdrA subunit by itself was successfully overexpressed. The purification and initial characterization of this subunit is described here.

If the HdrA only shows activity inside the completely folded enzyme we have to depend on the *M. Maripaludis* Hdr for site-directed mutagenesis. The Hdr in this organism is part of an even larger complex. When expressed with a His6-tag on the HdrB subunit, the full complex from *M. maripaludis* can be obtained in a single purification step using a Ni-NTA column. The full complex obtained this way contains heterodisulfide reductase (Hdr), hydrogenase (Vhu), formylmethanofuran dehydrogenase (Fwd) and formate dehydrogenase (Fdh). Even the polyferredoxin is part of this complex (FwdF subunit). There exist methods, however, to simplify the enzyme complex. When the cells are grown with H_2 as the electron source the Fdh is no longer part of the complex (HDR/Vhu/Fwd). We recently found that the Fwd enzyme is lost when an additional size-exclusion column step is performed. When formate is used as an electron source the hydrogenase is absent (HDR/Fwd/Fdh). It has not been tried yet to see if the HDR will separate from the Fwd/Fdh components.

The data presented here shows our first efforts in trying to see what enzyme complexes can be obtained and is very much a work in progress.

5.3 Materials and methods

The research work described here was carried out in the laboratory of Dr. Evert Duin. For all the experiments anaerobic conditions are required. To achieve this, all purification steps, sample handling and experiments were done in a glove box (Coy Laboratory Products, Inc., Grass Lake, USA) filled with a gas mixture consisting of 95% $N_2$and 5% H_2. Also, all buffers and solutions used in the procedures were degassed by boiling them under a nitrogen or argon atmosphere and subsequent cooling down under vacuum for 2 to 12 hours followed by overnight equilibration inside the glove box. In most cases, the buffers were filtered with 0.45 μm Millipore filter to remove particles that might affect the columns to be used for protein purification.

5.3.1 Purification of hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC) from *M. marburgensis*

5.3.1.1 Growth of *M. marburgensis* cells

M. marburgensis was grown at 65 °C in a 13 L glass fermenter (New Brunswick) containing 10 L of growth medium. The growth medium (*12*) contained 65 mM KH_2PO_4, 50 mM NH_4Cl, 30 mM Na_2CO_3, 0.5 mM nitrilotriacetic acid, 2 mM $MgCl_2$, 50 μM $FeCl_2$, 1 μM $CoCl_2$, 1 μM Na_2MoO_4, 5 μM $NiCl_2$, and 20 μM resazurin. It was made anaerobic by gassing with 80%

H_2/20% CO_2/0.1% H_2S at a rate of 1,200 ml/min. The resazurin was added as an indicator to the medium so that change in the color of the medium would indicate when sufficient anaerobic conditions were reached. After 1-2 hour of equilibration, when the optimum temperature and anaerobic condition was reached, the medium was inoculated with about 200 ml of fresh cell culture. The medium was agitated at 1000 rpm. After about 13 hour of incubation, at a ΔOD_{568} of ~4.5, the cells were harvested.

5.3.1.2 Harvest and sonication of *M. marburgensis* cells

The cells were harvested anaerobically by centrifugation at 15,000 rpm using a flow-through centrifuge (Hettich, contifuge 17 RS). The rotor was brought into the anaerobic tent. The cells were suspended in buffer A containing 50 mM Tris/HCl at pH 7.6, 2 mM DTT (Dithiothreitol), 2 mM CoM-SH (Coenzyme M), and 20 μM FAD. The suspended cells were then sonicated on ice 3 times for a total of 7 min (pulsing for 0.5 seconds). The cells were allowed to cool down in between the runs for a couple of minutes and at the end of the procedure. The sonicated cells were then centrifuged anaerobically at 35000 rpm for 20 minutes. The supernatant was carefully decanted into a beaker equilibrated inside the anaerobic tent.

5.3.1.3 Purification of hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC)

According to the protocol of Thauer *et al.* (*4*), the supernatant was applied to a DEAE-Sepharose column equilibrated with buffer A containing 50 mM Tris/HCl pH 7.6. According to the protocol, a NaCl step gradient was used in buffer A: 100 mL 0 M NaCl, 100 mL 0.2 M NaCl,

100 mL 0.3 M NaCl, and 100 mL 0.4 M NaCl. The last peak was collected and applied to a Q-Sepharose column equilibrated with buffer A. Again, a NaCl step gradient was used in buffer A: 100 mL 0 M NaCl, 100 mL 0.3 M NaCl, 100 mL 0.4 M NaCl, 100 mL 0.45 M NaCl, and 100 mL 0.54 M NaCl. The last peak was collected and concentrated by filtration using 10kDa filter to 2–3 mL, which was then applied to a Superdex 200 column equilibrated with buffer B (buffer A + 150 mM NaCl). The different fractions collected from this run were then analyzed using a 15% SDS PAGE gel.

5.3.2 Purification of heterodisulfide reductase (Hdr) from *M. maripaludis* cells

5.3.2.1 *M. maripaludis* cells

The Leigh group at the University of Washington has graciously supplied us wild type *M. maripaludis* cells.

5.3.2.2 Purification of heterodisulfide reductase

The *M. maripaludis* cells were sonicated on ice 3 times for a total of 7 min (pulsing for 0.5 seconds). The cells were allowed to cool down in between the runs for a couple of minutes and at the end of the procedure. The sonicated cells were then centrifuged anaerobically at 35000 rpm for about 20 minutes. The supernatant was carefully decanted into a beaker equilibrated inside the anaerobic tent. The supernatant was then applied to a nickel column equilibrated with buffer A containing 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 10 mM sodium dithionite, 100 mM NaCl and 10 mM imidazole. The protein was eluted by

washing the column with buffer B containing 25 mM HEPES, 10 mM sodium dithionite, 100 mM NaCl and 100 mM imidazole. The fractions were analyzed both with 17% SDS-PAGE and 8% native PAGE. The concentration of the protein content was determined using the method of Bradford with bovine serum albumin (Serva) as standard (*13, 14*). Also, just to check, EPR measurements were done on the complex at 77 K and 20 dB. To further purify the protein, the eluted protein was concentrated to about 2 mL by filtration using a 10 kDa filter paper and applied to a Superdex 200 column equilibrated with buffer containing 30 mM Tris-HCl pH 8.0 and 100 mM NaCl. The protein was eluted with the same buffer. The major peaks of the chromatography profile were analyzed with 17% SDS-PAGE and 8% native PAGE.

5.3.3 Purification of HdrA from *M. maripaludis* HdrA$_{marburgensis}$ cells

5.3.3.1 *M. maripaludis* HdrA$_{marburgensis}$ cells

The Leigh group at the University of Washington has graciously supplied us with the *M. maripaludis* strains with the *M. marburgensis* HdrA gene.

5.3.3.2 Purification of HdrA

The *M. maripaludis* cells were sonicated on ice 3 times for a total of 7 min (pulsing for 0.5 seconds). The cells were allowed to cool down in between the runs for a couple of minutes and at the end of the procedure. The sonicated cells were then centrifuged anaerobically at 35000 rpm for about 20 minutes. The supernatant was carefully decanted into a beaker equilibrated inside the anaerobic tent. The supernatant was then applied to a nickel column equilibrated with

buffer A containing 50 mM Tris-HCl pH 7.6 and 100 mM NaCl. The protein was eluted by washing the column with buffer B containing 50 mM Tris-HCl pH 7.6, 100 mM NaCl and 500 mM imidazole. The fractions were analyzed with 12% SDS-PAGE. The major peak was considered to be the peak of our interest. Thus, to check the concentration of the protein content of this peak the Bradford method was used with bovine serum albumin (Serva) as standard (*13, 14*). Also, to get the iron content of the protein fraction, a rapid colorimetric method was used (*15*) (see below). In addition, EPR measurements were done on the reduced protein sample. To further purify the protein, the eluted fraction was applied to a Q-Sepharose column equilibrated with buffer A containing 50 mM Tris-HCl pH 7.6. To elute the protein, a NaCl step gradient was used in the buffer A: 100 mL 0 M NaCl, 100 mL 0.3 M NaCl, 100 mL 0.4 M NaCl, 100 mL 0.45 M NaCl, and 100 mL 0.54 M NaCl. The peak of the interest was collected and concentrated by filtration using a 10 kDa filter to about 1 mL, which were then applied to a Superdex 200 column equilibrated with buffer B (buffer A + 150 mM NaCl). The different major fractions collected from each of the purification step were analyzed with 12% SDS PAGE.

5.3.4 Iron determination

The iron standard was prepared using 0.0523 M of ferrous ethylenediammonium sulfate in 0.01 M HCl for the calibration curve. 0.25 mL of freshly prepared iron releasing reagent which contained 0.6 M HCl and 0.142 M potassium permanganate ($KMnO_4$) were added to 0.5 mL of the protein sample and the standards. The digested mixture was incubated in a capped tube for 2 hours at 60°C. Following the digestion, 0.1 mL of reducing, iron chelating reagent which contained 6.5 mM ferrozine (disodium 3-(2-pyridyl)-5,6-bis(4-phenyl sulfonate)-1,2,4-triazine), 13.1 mM neocuprine (2,9-dimethyl-1,10-phenanthroline), 2 M ascorbic acid and 5 M

ammonium acetate were added to the digested mixture and mixed. The solution was left to stand at room temperature for at least 30 min. After this, the absorbance of standards and protein samples were measured at 562 nm. A standard curve was constructed by plotting the concentration of the standard versus their absorbance. From this curve, the concentration of the iron content was calculated.

5.3.5 UV-vis absorption analysis

The UV-vis absorption spectra of the protein samples were recorded under anaerobic conditions by using stoppered cuvettes in an Agilent 8453 UV-visible Spectrophotometer. To check the iron-sulfur cluster signal, absorbance in the 410-420 nm region was observed.

5.3.6 EPR measurements

CW EPR spectra were measured at X-band (9 GHz) frequency on a Bruker EMX spectrometer, fitted with the ER-4119-HS high sensitivity perpendicular-mode cavity. General EPR conditions were: microwave frequency, 9.385 GHz; microwave power incident to the cavity, 0.20 mW; field modulation frequency, 100 kHz; microwave amplitude, 0.6 mT. The Oxford Instrument ESR 900 flow cryostat in combination with the ITC4 temperature controller was used for measurements using a helium flow. Samples for EPR were prepared in quartz tubes that were sealed with a closed off rubber tube. The samples were frozen using liquid nitrogen.

5.4 Results

5.4.1 Purification of the hydrogenase:heterodisulfide reductase complex (MvhADG/HdrABC) from *M. marburgensis*

M. marburgensis cell supernatant was applied to DEAE-Sepharose column. The last peak (fraction no. 44-48) of the chromatography profile (**Figure 5.3**) was collected and applied to a Q-Sepharose column. The last peak (fraction no. 60-68) of its chromatography profile (**Figure 5.4**) was collected. The fractions of this peak were pooled together and applied to a Superdex 200 column. Fraction no. 18-25 was collected (**Figure 5.5**). Each of these fractions was then analyzed by 15% SDS-PAGE gel. From the gel (**Figure 5.6**), it appears that, fractions 19, 20 and 21 contain the MvhADG/HdrABC complex but that there are still other impurities present. From the chromatograph (**Figure 5.5**), it was also evident that the protein complex was present as a shoulder just in front of the most intense peak.

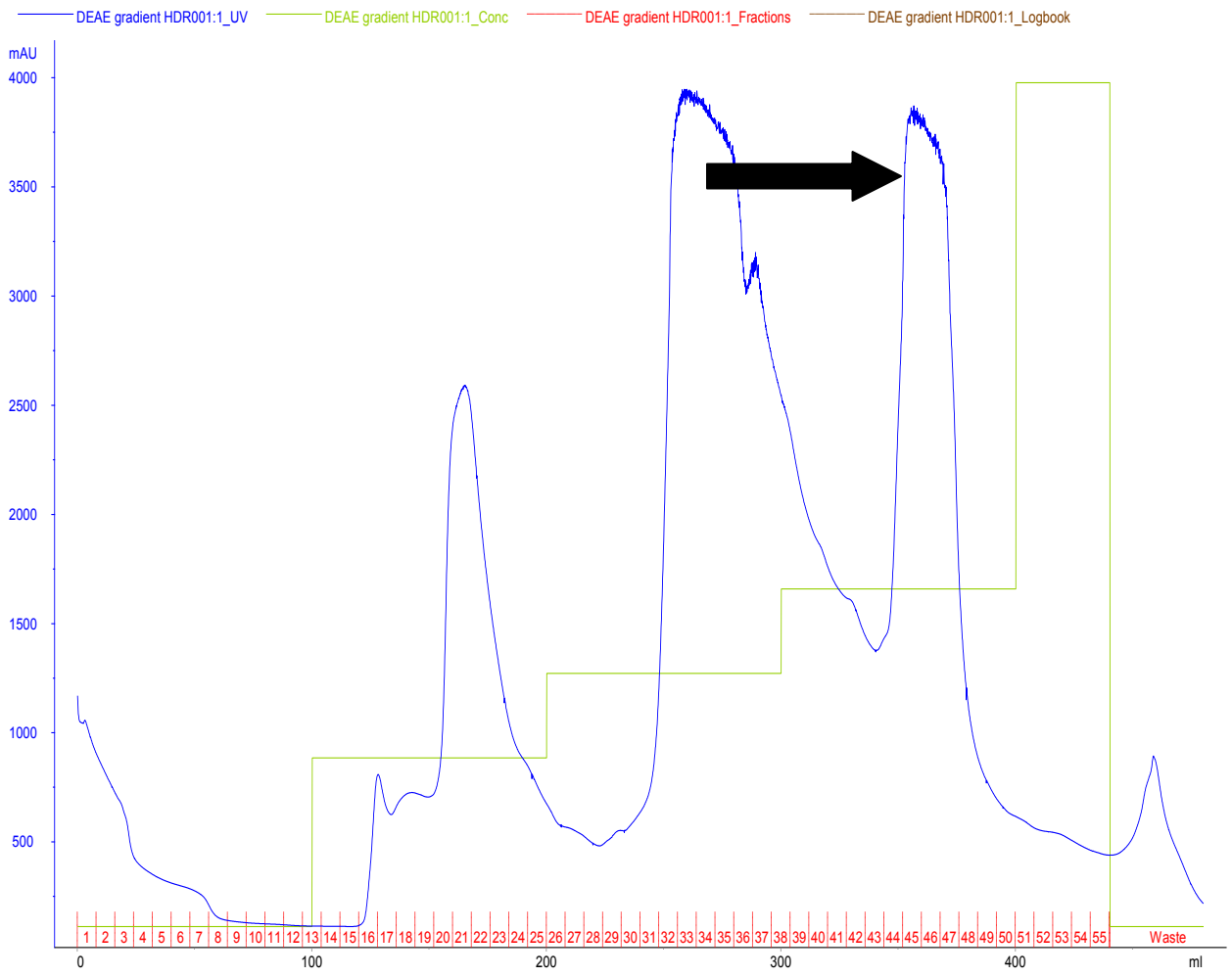


Figure 5.3: Chromatography profile of DEAE-Sepharose column for purification of MvhADG/HdrABC complex.

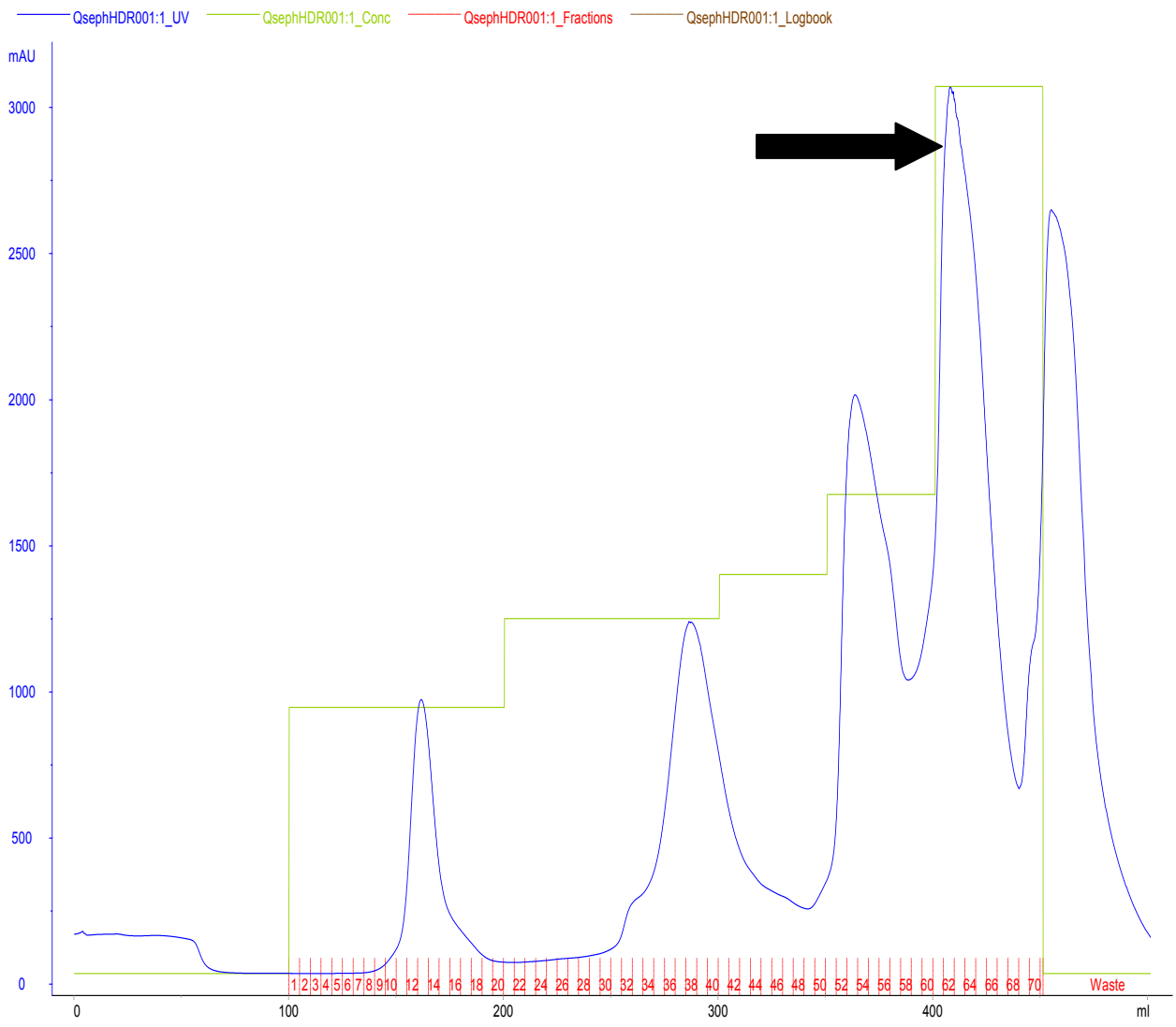


Figure 5.4: Chromatography profile of Q-Sepharose column for purification of MvhADG/HdrABC complex.

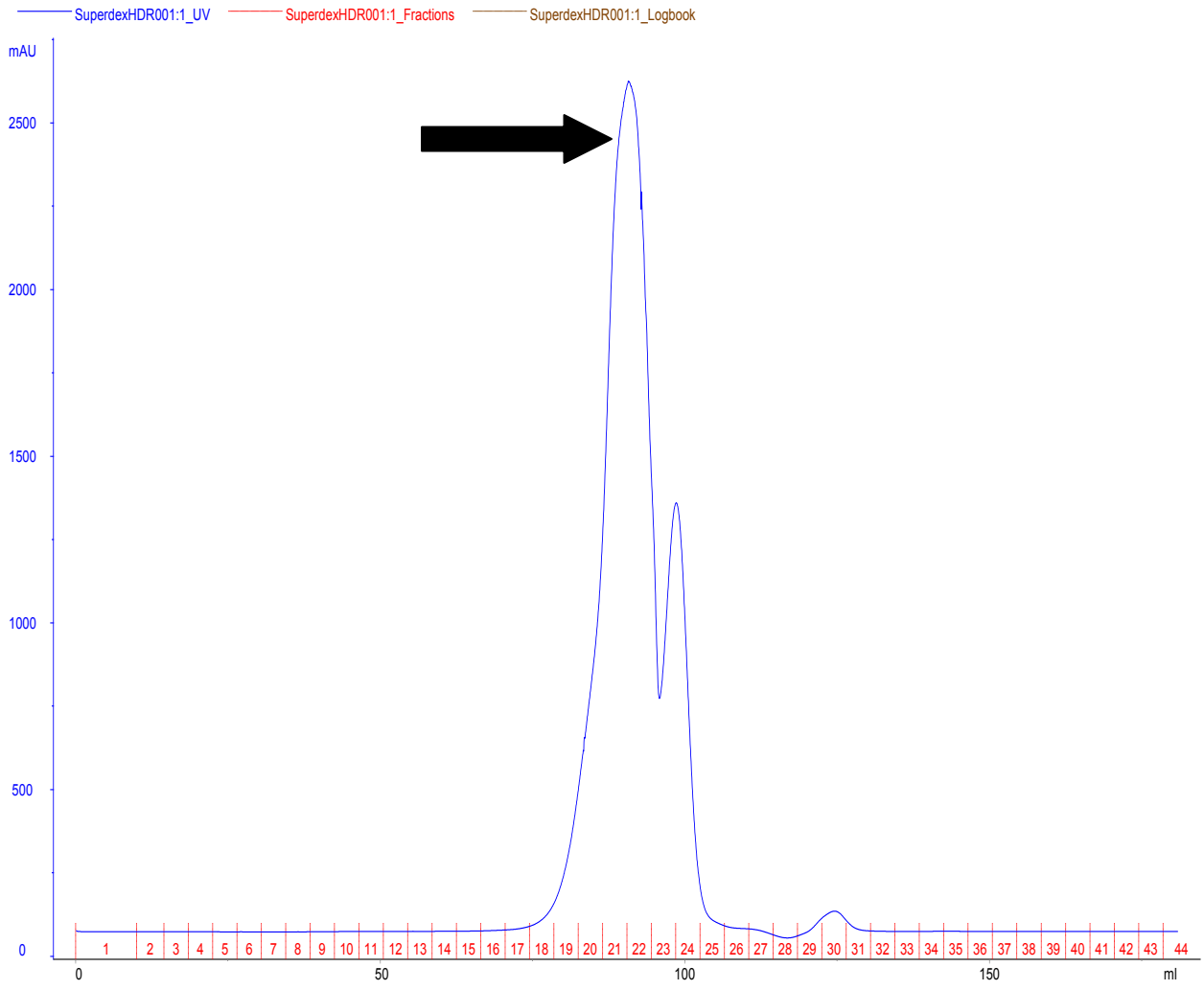


Figure 5.5: Chromatography profile of Superdex 200 column for purification of MvhADG/HdrABC complex.

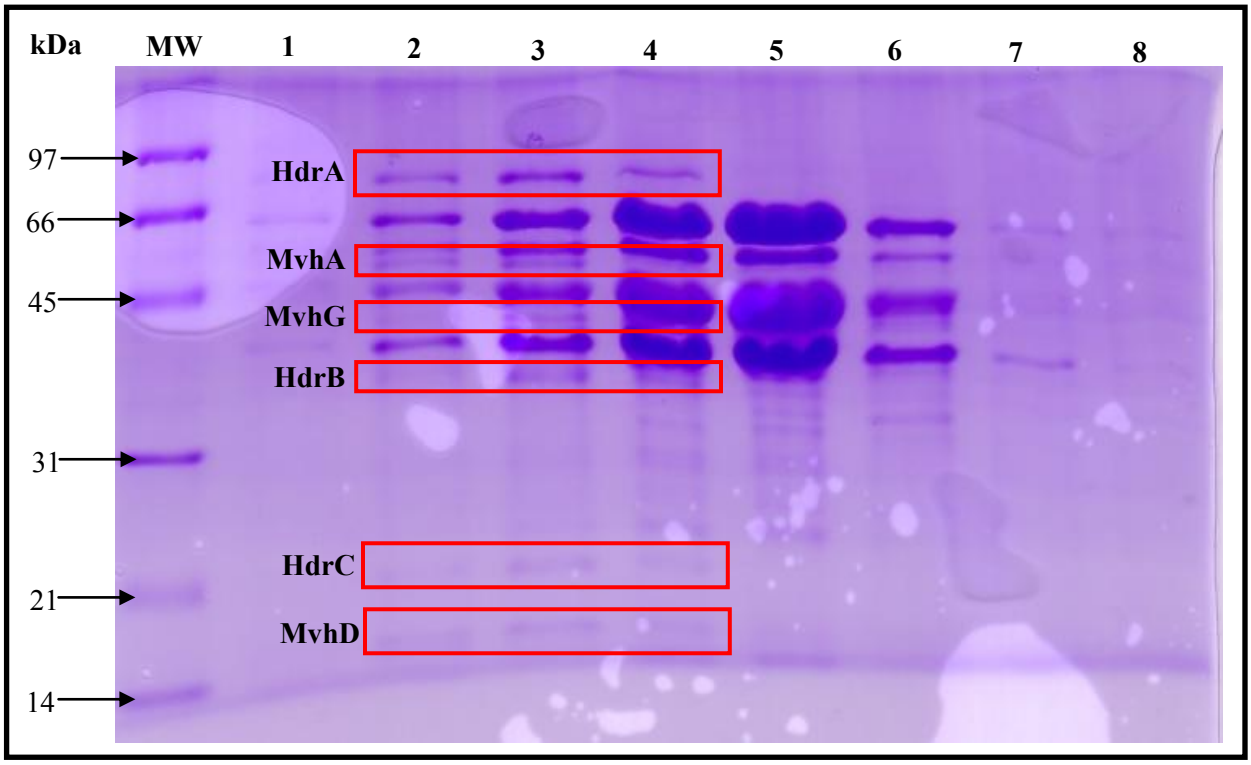


Figure 5.6: 15%SDS-PAGE analysis of the fractions from Superdex 200 column for MvhADG/HdrABC complex purification. Lane 1-8 corresponds to the fractions 18-25 in order. Lane MW is for the protein marker. The probable bands of fraction 19, 20 and 21 (lane 2, 3 and 4) for the subunits of the complex are enclosed with red rectangle and the corresponding names of the subunits are given in black bold letters to the left of the rectangles.

5.4.2 Purification of heterodisulfide reductase (Hdr) from *M. maripaludis* cells

The cell supernatant was applied to the nickel column and the protein was eluted using an imidazole gradient. The concentration of the protein content of this sample was determined using the Bradford method and was found to be 8.67 mg/mL. When EPR was used to check the iron-sulfur cluster signal of this protein complex, the signal (**Figure 5.8**) did not resemble any of the standard type signals (**Figure 5.7**). The obvious reason behind this is that the enzyme complex contains multiple clusters, a molybdenum/tungsten site and a nickel site since the Hdr is part of a multimeric complex (Hdr/Vhu/Fwd/Fdh). To isolate Hdr from this complex and as already tested by the Leigh group, size-exclusion chromatography was performed. In the chromatograph, the complex appeared to be separated into three major peaks (**Figure 5.9**). Fractions 22-23, fractions 24-25 and fractions 26-27 were collected separately. When these fractions were analyzed with SDS-PAGE (**Figure 5.10**) and native PAGE (**Figure 5.11**), it was not possible to identify the bands that belonged to the Hdr enzyme. Some of the more promising bands were also analyzed using mass spectrometry, but the obtained sequences did not correspond with those of the Hdr subunits.

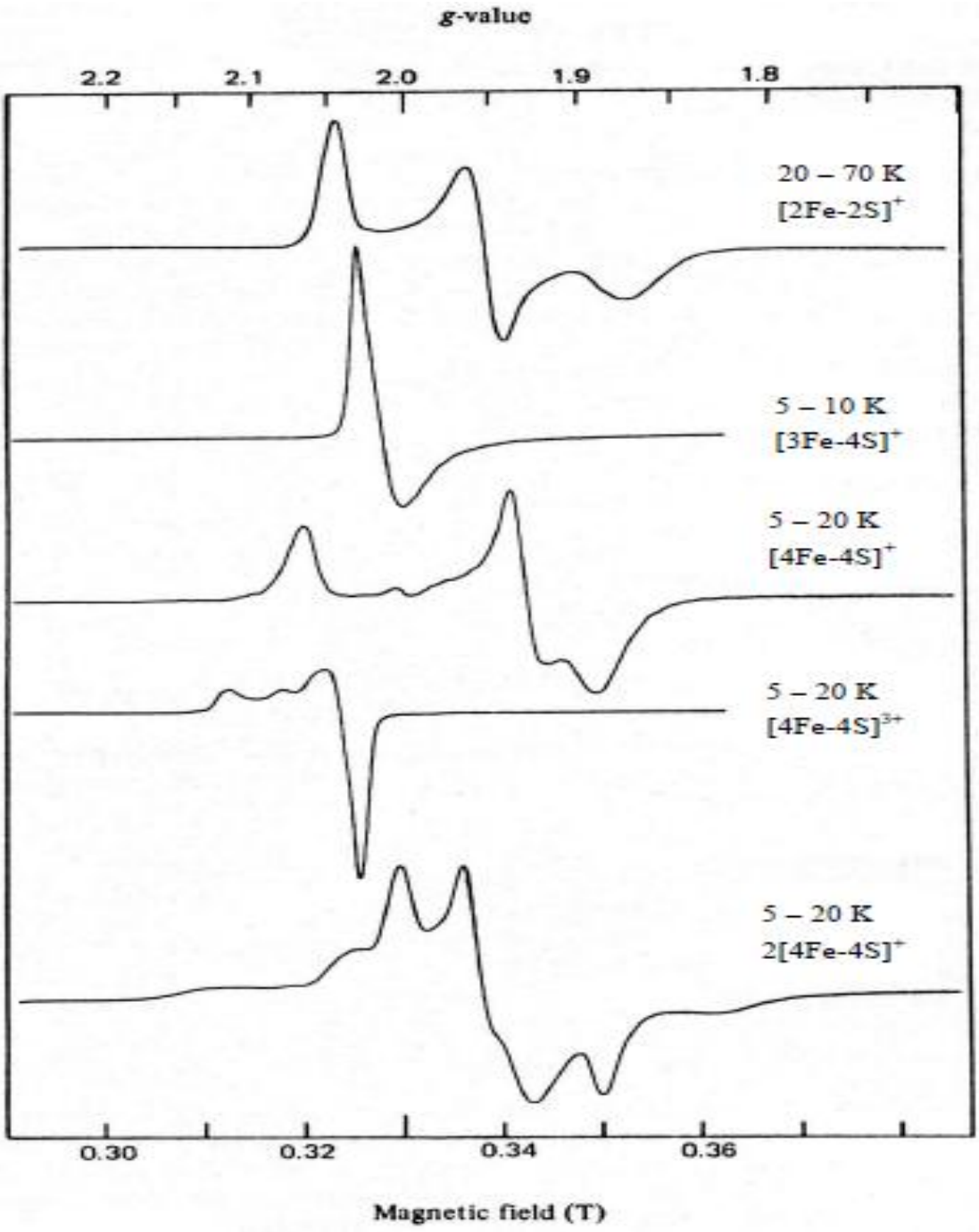


Figure 5.7: EPR spectra of iron-sulfur clusters. Adapted from reference (*16, 17*).

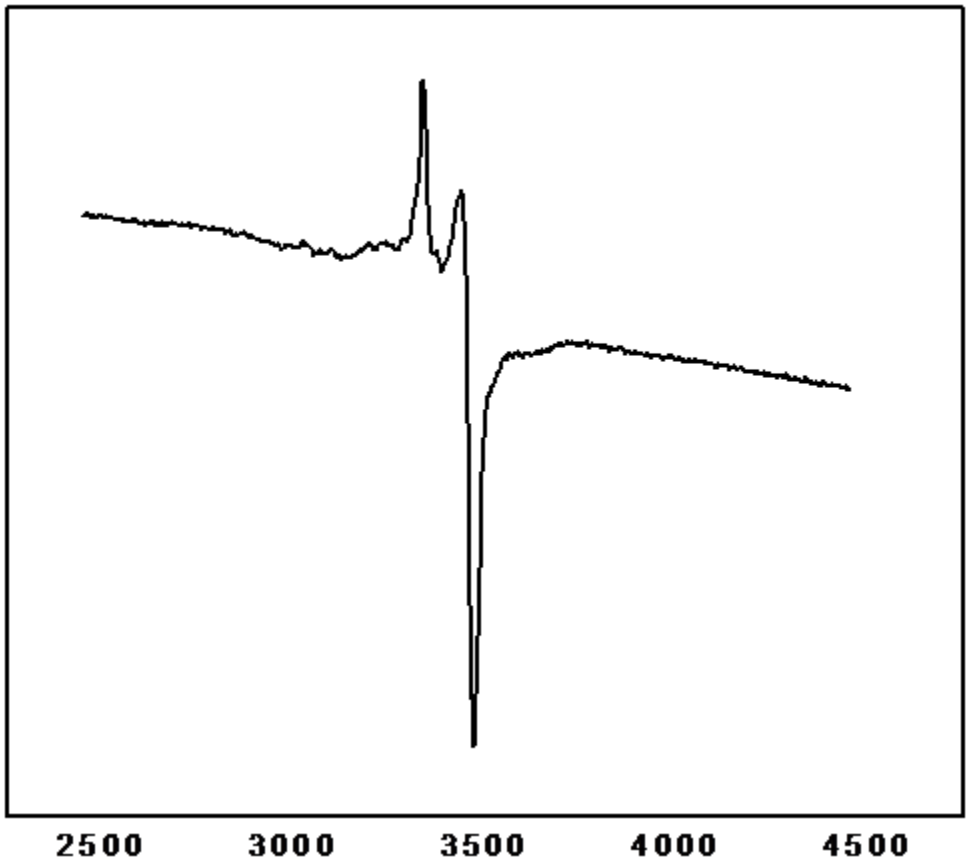


Figure 5.8: EPR spectrum of Hdr complex from *M. maripaludis*.

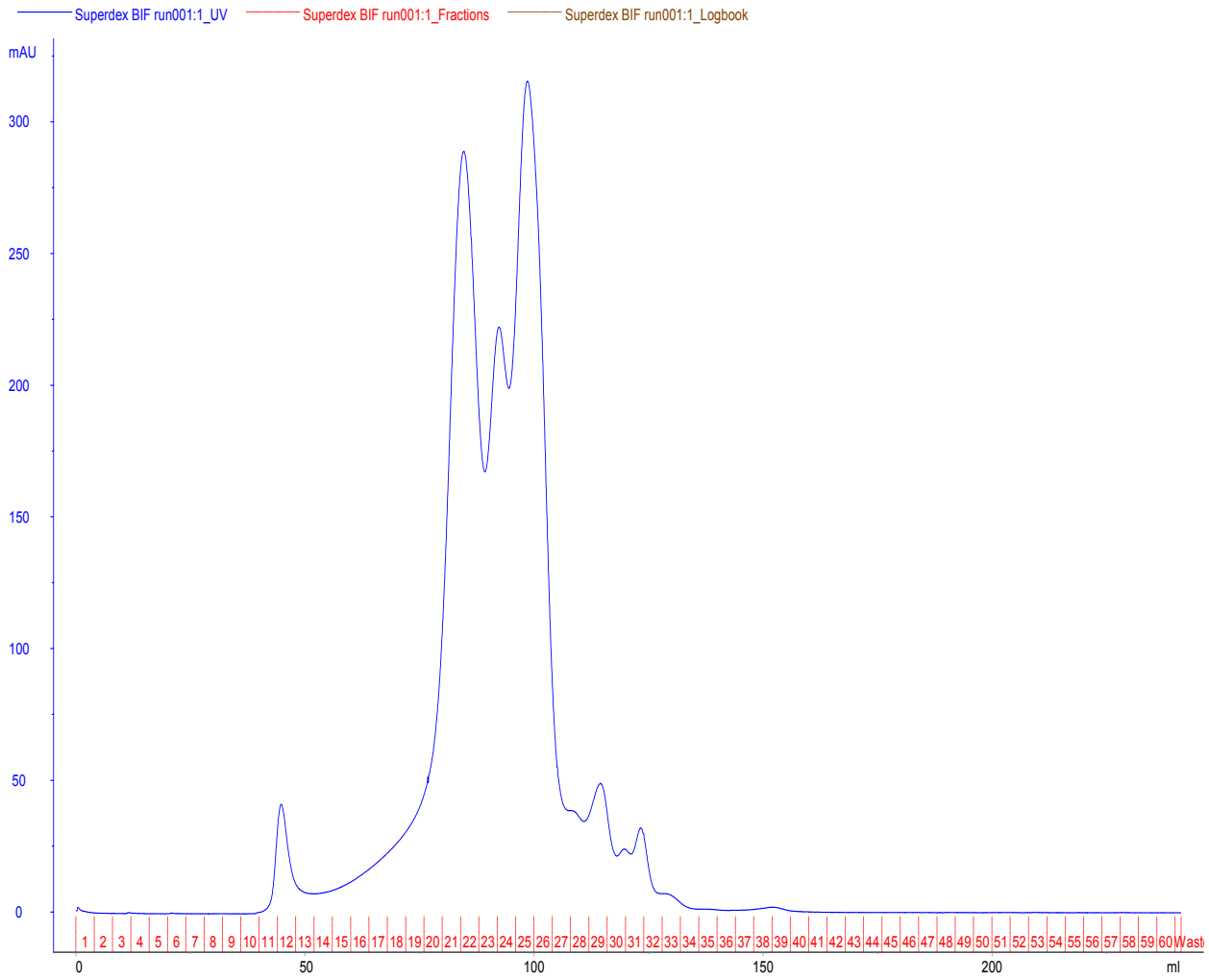


Figure 5.9: Chromatography profile of Superdex 200 column for purification of Hdr from *M. maripaludis*.

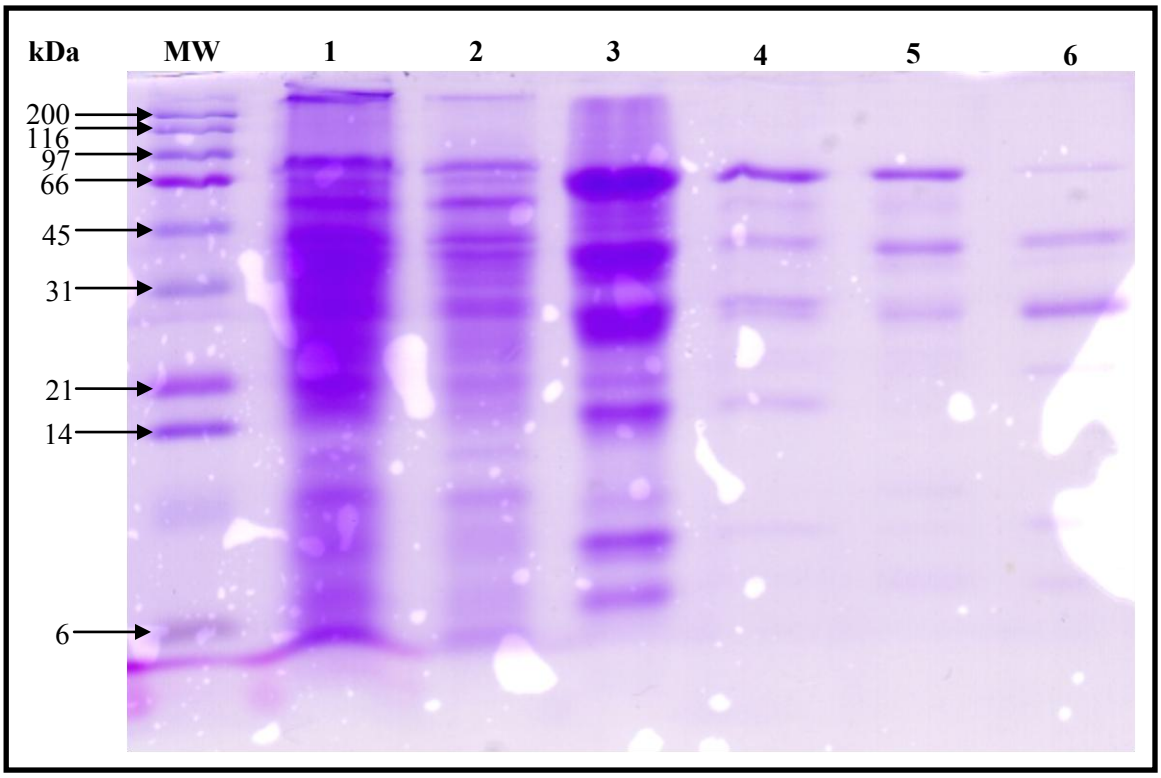


Figure 5.10: 17% SDS-PAGE analysis of the Hdr complex from *M. maripaludis* cells. Lane 1: Cell extract, Lane 2: Flow through from the nickel column, Lane 3: Protein sample eluted from the nickel column, Lane 4: Fraction no. 22-23 from Superdex 200 column, Lane 5: Fraction no. 24-25 from Superdex 200 column and Lane 6: Fraction no. 26-27 from Superdex 200 column. The lane MW is for protein marker.

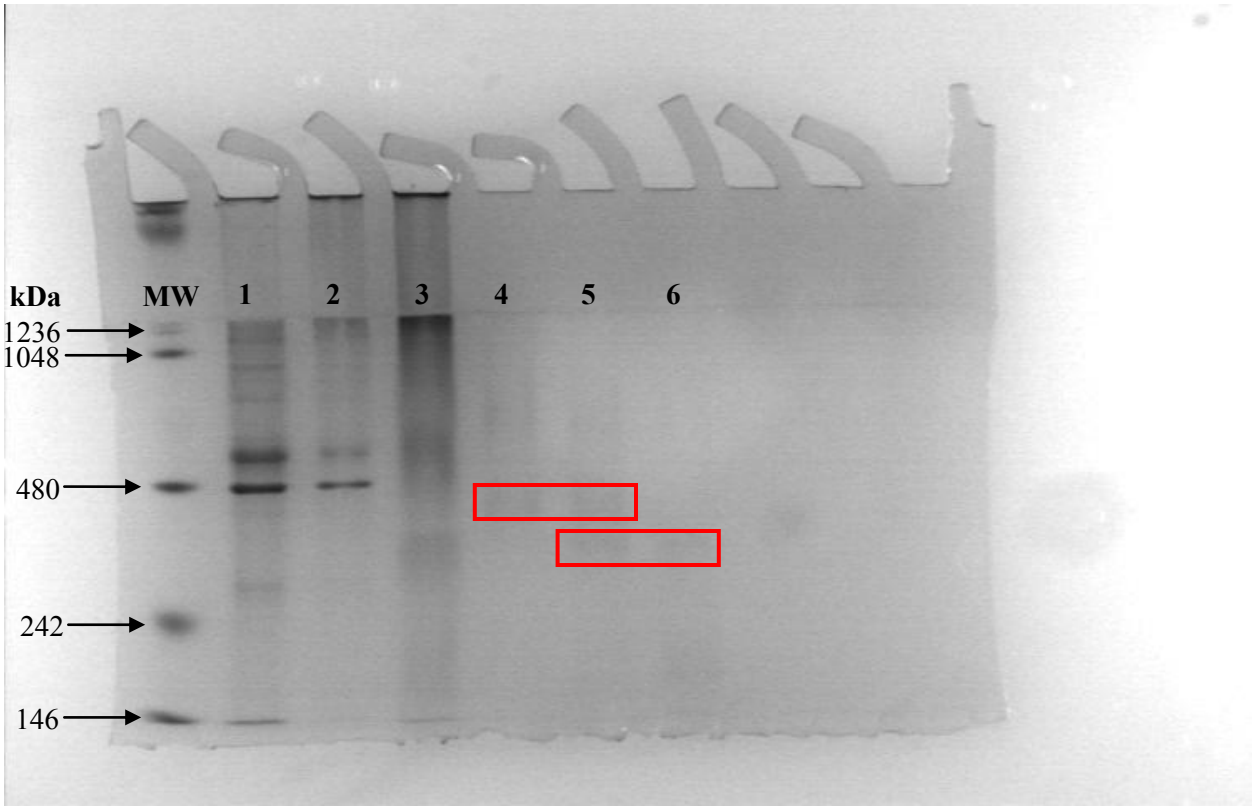


Figure 5.11: 8% native PAGE analysis of the Hdr complex from *M. maripaludis* cells. Lane 1: Cell extract, Lane 2: Flow through from the nickel column, Lane 3: Protein sample eluted from the nickel column, Lane 4: Fraction no. 22-23 from Superdex 200 column, Lane 5: Fraction no. 24-25 from Superdex 200 column and Lane 6: Fraction no. 26-27 from Superdex 200 column. The lane MW is for protein marker. The overlapping bands on the consecutive lanes are enclosed by red rectangles.

5.4.3 Purification of HdrA from *M. maripaludis* $HdrA_{marburgensis}$ cells

5.4.3.1 Purification of HdrA

The cell supernatant was applied to a nickel column and the protein was eluted using an imidazole gradient. The major peak (fraction no. 9-13) was collected (**Figure 5.12**). This fraction was analyzed with 12% SDS-PAGE (**Figure 5.13**). From the gel, it could be seen that, the protein sample is about 70% pure. The HdrA subunit has an estimated weight of 71 kDa. Further attempts (Q-Sepharose and Superdex 200 column) to purify this protein were not successful. The HdrA band could not be detected on SDS-PAGE after these purification steps.

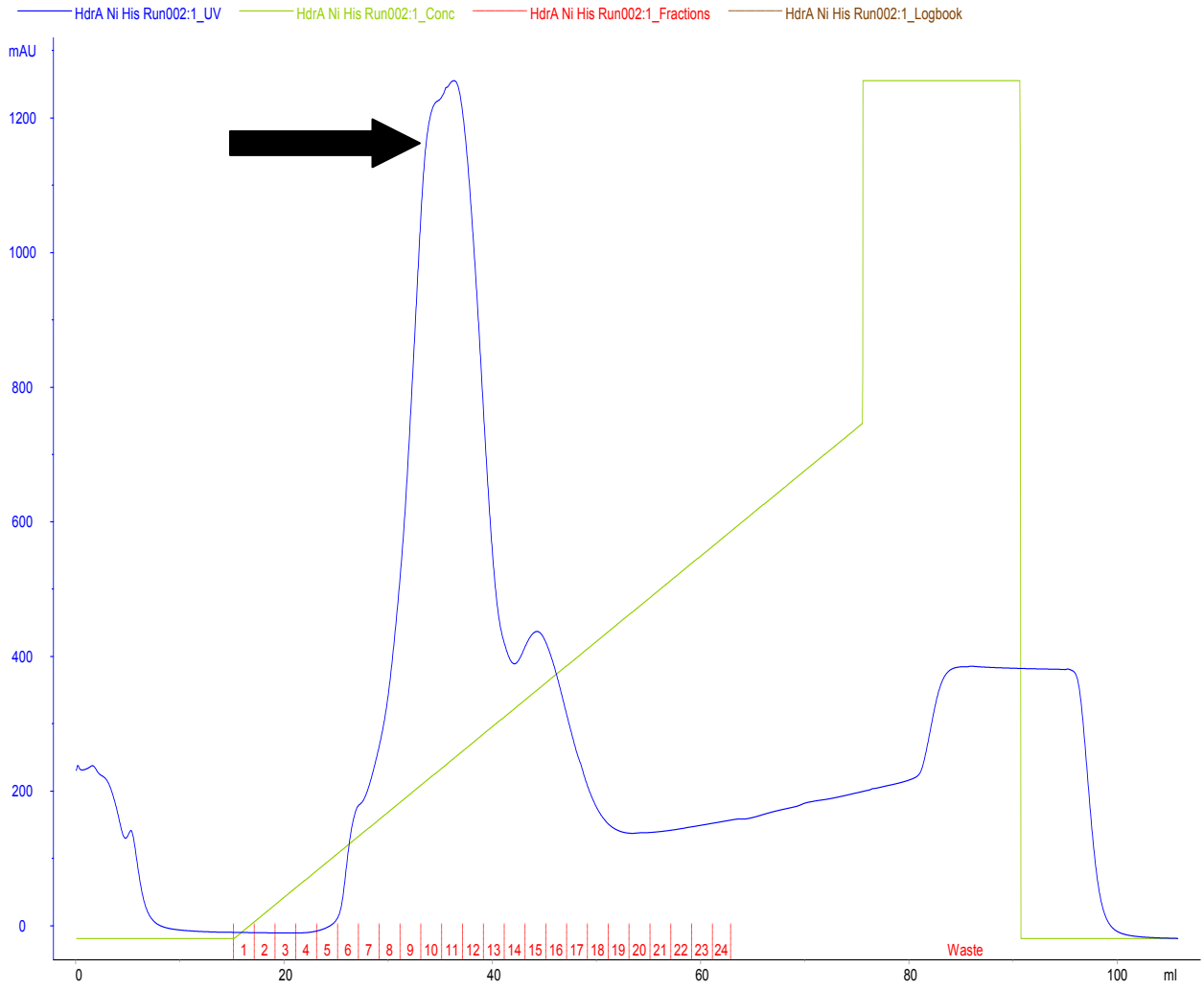


Figure 5.12: Chromatography profile of nickel column for purification of HdrA from *M. maripaludis* HdrA$_{marburgensis}$.

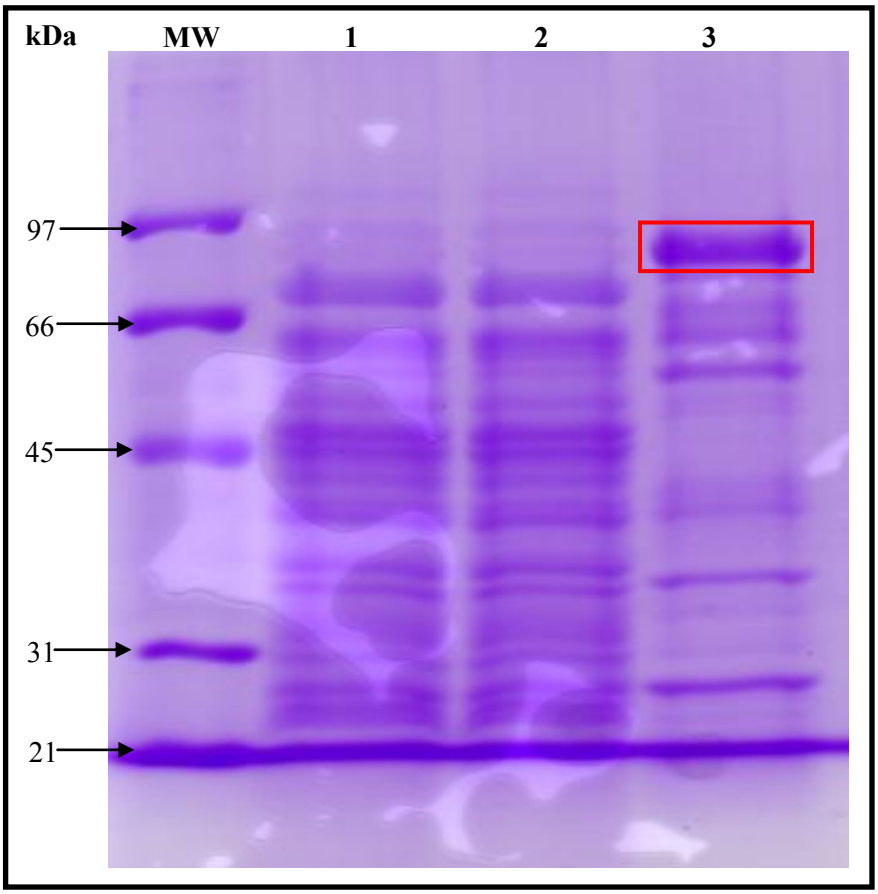


Figure 5.13: 12% SDS-PAGE analysis of the HdrA sample from *M. maripaludis* $HdrA_{marburgensis}$ cells. Lane 1: Cell extract, Lane 2: Flow through from the nickel column, Lane 3: Protein sample eluted from the nickel column. The lane MW is for protein marker. The band for HdrA protein is enclosed by a red rectangle.

5.4.3.2 Protein and Iron Determination

The concentration of the protein sample eluted from the nickel column was determined by using Bradford method and the concentration was determined to be 2.27 mg/mL or 31.79 μM (considering the molecular weight of HdrA as 71 kDa). Using the colorimetric method, the iron concentration of the sample was found to be about 258 μM. The EPR measurements (**Figure 5.15**) are in line with the presence of [4Fe-4S] cluster. Therefore the cluster content is about 2.

5.4.3.3 UV-vis absorption of the protein sample

Cubane iron-sulfur clusters display an absorption band at around 410-420 nm for the oxidized form. For HdrA an absorbance was detected in this region (**Figure 5.14**). When the protein sample was reduced by dithionite, this signal is lowered and after about three minutes, the signal is completely gone (**Figure 5.14**). The disappearance of the cluster signal upon reduction could either be due to the reduction itself or disintegration of the cluster. Since an EPR signal was obtained for the reduced protein the former appears to be the case.

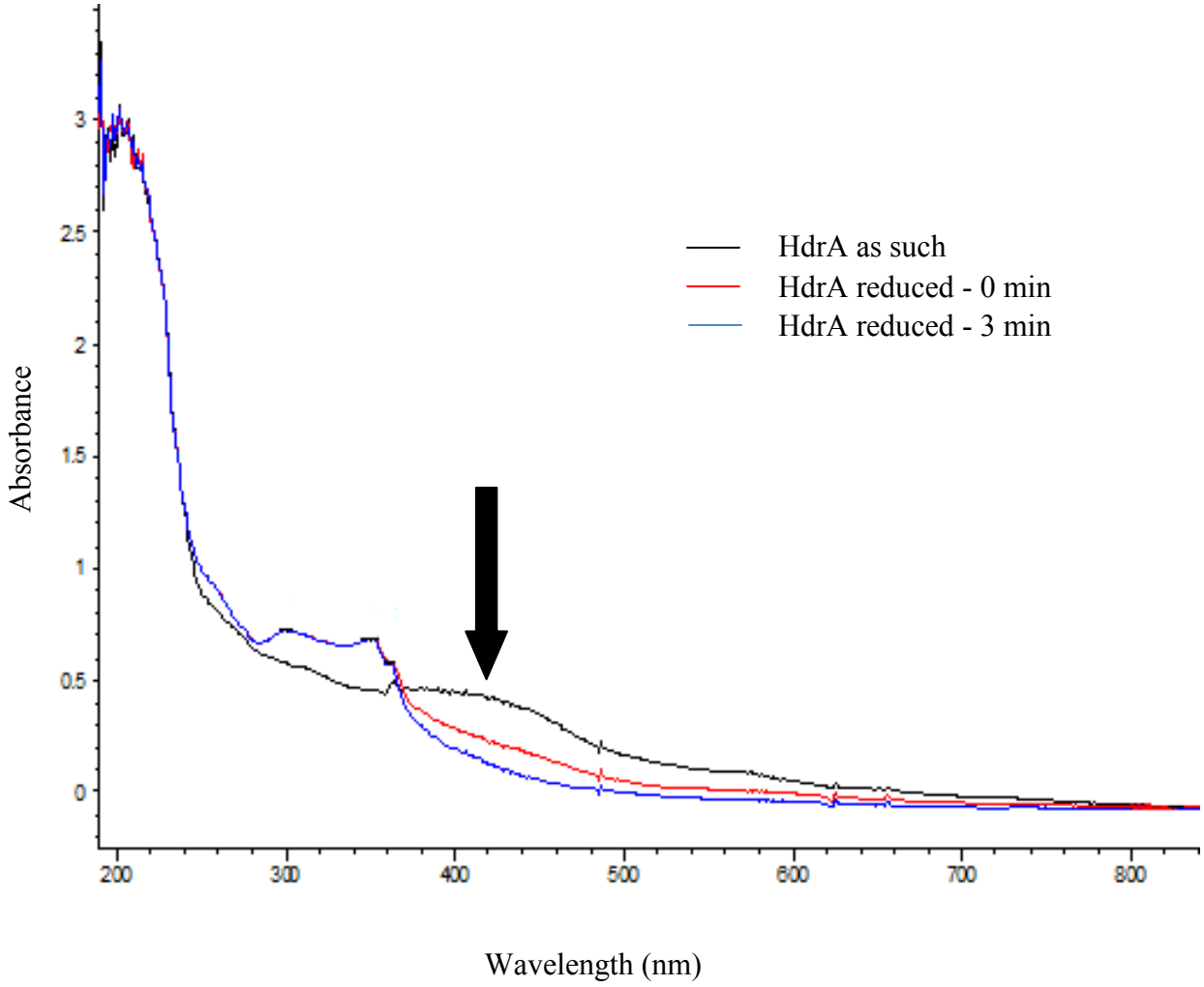


Figure 5.14: UV-vis absorption of the HdrA protein sample. Black arrow indicates the absorption of the iron-sulfur cluster at around 410 nm.

5.4.3.4 EPR measurement of the HdrA protein sample

EPR spectra of the HdrA protein sample in its reduced condition were recorded at different microwave powers: 20 dB, 30 dB, 40 dB and 50 dB (**Figure 5.15**). All these spectra were recorded at 8 K after the optimization of the temperature for the enhancement of the signal. From the figure, it was apparent that, the EPR signal for the HdrA protein sample is achieved optimally at 8 K and 30 dB. Under these optimal conditions, the EPR signal is comparable to that of the standard [4Fe-4S] cluster (**Figure 5.7**).

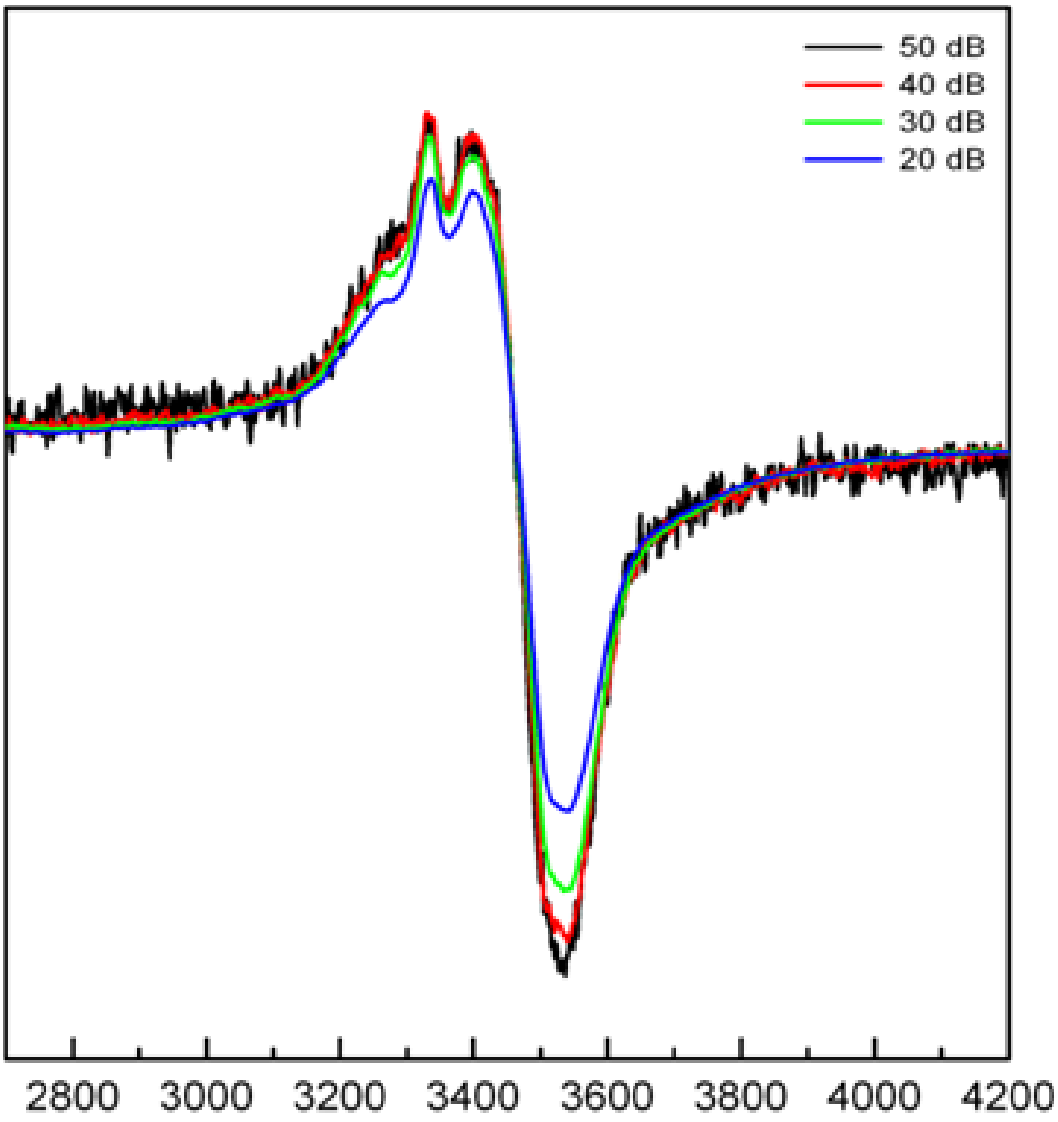


Figure 5.15: EPR spectra of HdrA protein sample at the temperature of 8 K and at different microwave frequencies.

5.5 Conclusions and future direction

The mechanism of electron bifurcation is still unknown. Understanding its mechanism and discovering its existence in the heterodisulfide reductase enzyme complex system in methanogenic archaea would provide the proof of the important role of electron bifurcation in the anaerobic energy transduction. Also, such understanding would also open the doors of harnessing the power of anaerobic energy conversion. Methane production and methane activation processes are performed by methanogenic and methylotrophic Archaea.

In this work, we were able to purify (70%) HdrA, a subunit of the Hdr enzyme from *M. marburgensis*. From preliminary data, it is observed that possibly at least two iron-sulfur clusters were present in the protein sample. Sequence data, however indicates that there could be 4 clusters present in this subunits. Further purification should bring this number up. Additional reconstitution procedures can also be performed. With the availability of the almost 100% pure protein in the future, it would be possible to find out the accurate type and species of iron-sulfur clusters. Also, such pure protein would allow us to investigate the mechanism of electron bifurcation through different approaches such as redox titration, structural characterization and freeze-quench study. This would also eventually prove whether such a subunit is sufficient to necessarily carry out the electron bifurcation or a larger unit of the complex is needed to successfully yield the desired output. For structural characterization, a pure enzyme in high enough amounts should be prepared to do the X-ray crystallization of the protein. From the structural information, it would be possible to derive the specific sites of the protein that are involved in this mechanism and how they are involved. However, it is also important to note

that, there is always a possibility of not having a crystal of the protein good enough to get the necessary structural information of the protein. Thus, mutational study can be done in the absence of structural information. Since, genetic tools for the mutational study are readily available for the organism *M. maripaludis*, it is worth continuing the efforts to purify the Hdr from its complex of the wild type *M. maripaludis*. Here, we were able to isolate three different parts of the Hdr complex but without being able to identify them. In the future, efforts should be made to purify Hdr from this complex by varying the conditions or increasing the steps of the chromatography and utilizing mass spectrophotometry to determine specific bands from the gels or specific fractions from the purifications.

Also, purification of the Hdr from *M. marburgensis* could prove worthwhile in the future; especially if it is found out that HdrA subunit is not sufficient to be functional and the whole protein is needed. Here, we were able to obtain the MvhADG/HdrABC complex with impurities. In the future, at each step of purification, H_2:CoM-S-S-CoB oxidoreductase activity should be measured for each fraction, so that collection of the fractions can be more accurate. A hydrograph for measuring H_2 concentrations has recently been purchased but not tested yet. Also, some of the chromatography conditions could be optimized to diminish the overlap of bands within the chromatography profile.

5.6 References

1. Li, F., Hinderberger, J., Seedorf, H., Zhang, J., Buckel, W., and Thauer, R. K. (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from Clostridium kluyveri, *Journal of bacteriology 190*, 843-850.
2. Herrmann, G., Jayamani, E., Mai, G., and Buckel, W. (2008) Energy conservation via electron-transferring flavoprotein in anaerobic bacteria, *Journal of bacteriology 190*, 784-791.
3. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., and Hedderich, R. (2008) Methanogenic archaea: ecologically relevant differences in energy conservation, *Nature reviews. Microbiology 6*, 579-591.
4. Kaster, A. K., Moll, J., Parey, K., and Thauer, R. K. (2011) Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea, *Proc Natl Acad Sci U S A 108*, 2981-2986.
5. Setzke, E., Hedderich, R., Heiden, S., and Thauer, R. K. (1994) H2: heterodisulfide oxidoreductase complex from Methanobacterium thermoautotrophicum. Composition and properties, *European journal of biochemistry / FEBS 220*, 139-148.
6. Costa, K. C., Wong, P. M., Wang, T., Lie, T. J., Dodsworth, J. A., Swanson, I., Burn, J. A., Hackett, M., and Leigh, J. A. (2010) Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase, *Proc Natl Acad Sci U S A 107*, 11050-11055.

7. Kim, J., Hetzel, M., Boiangiu, C. D., and Buckel, W. (2004) Dehydration of (R)-2-hydroxyacyl-CoA to enoyl-CoA in the fermentation of alpha-amino acids by anaerobic bacteria, *FEMS microbiology reviews 28*, 455-468.

8. Shokes, J. E., Duin, E. C., Bauer, C., Jaun, B., Hedderich, R., Koch, J., and Scott, R. A. (2005) Direct interaction of coenzyme M with the active-site Fe-S cluster of heterodisulfide reductase, *FEBS Lett 579*, 1741-1744.

9. Duin, E. C., Bauer, C., Jaun, B., and Hedderich, R. (2003) Coenzyme M binds to a [4Fe-4S] cluster in the active site of heterodisulfide reductase as deduced from EPR studies with the [33S]coenzyme M-treated enzyme, *FEBS Lett 538*, 81-84.

10. Madadi-Kahkesh, S., Duin, E. C., Heim, S., Albracht, S. P., Johnson, M. K., and Hedderich, R. (2001) A paramagnetic species with unique EPR characteristics in the active site of heterodisulfide reductase from methanogenic archaea, *European journal of biochemistry / FEBS 268*, 2566-2577.

11. Duin, E. C., Madadi-Kahkesh, S., Hedderich, R., Clay, M. D., and Johnson, M. K. (2002) Heterodisulfide reductase from Methanothermobacter marburgensis contains an active-site [4Fe-4S] cluster that is directly involved in mediating heterodisulfide reduction, *FEBS Lett 512*, 263-268.

12. Schonheit, P., Moll, J., and Thauer, R. K. (1980) Growth-Parameters (Ks, Mu-Max, Ys) of Methanobacterium-Thermoautotrophicum, *Arch Microbiol 127*, 59-65.

13. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical biochemistry 72*, 248-254.

14. Stoscheck, C. M. (1990) Quantitation of protein, *Methods in enzymology 182*, 50-68.

15. Fish, W. W. (1988) Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples, *Methods in enzymology 158*, 357-364.

16. Rekittke, I., Wiesner, J., Rohrich, R., Demmer, U., Warkentin, E., Xu, W., Troschke, K., Hintz, M., No, J. H., Duin, E. C., Oldfield, E., Jomaa, H., and Ermler, U. (2008) Structure of (E)-4-hydroxy-3-methyl-but-2-enyl diphosphate reductase, the terminal enzyme of the non-mevalonate pathway, *J Am Chem Soc 130*, 17206-17207.

17. Cammack, R., Patil, D. S., and Fernandez, V. M. (1985) Electron-spin-resonance/electron-paramagnetic-resonance spectroscopy of iron-sulphur enzymes, *Biochemical Society transactions 13*, 572-578.

Appendix Table A-1 Chemical shift assignments for the nuclei of free GIP (BMRB entry: 17254)

Residue no.	Amino acid	Nucleus	Chemical shift
1	MET	HA	4.054
1	MET	HB2	2.012
1	MET	HB3	1.862
1	MET	HG2	2.169
1	MET	HG3	2.169
1	MET	HE1	1.987
1	MET	HE2	1.987
1	MET	HE3	1.987
1	MET	C	177.041
1	MET	CA	51.058
1	MET	CG	34.178
1	MET	CE	24.553
2	SER	H	8.17
2	SER	HA	4.348
2	SER	HB2	3.712
2	SER	HB3	3.712
2	SER	C	174.058
2	SER	CA	58.303
2	SER	CB	63.868
2	SER	N	121.533
3	TYR	H	8.125
3	TYR	HA	4.561
3	TYR	HB2	2.956
3	TYR	HB3	2.846
3	TYR	HD1	7.006
3	TYR	HD2	7.006
3	TYR	HE1	6.738
3	TYR	HE2	6.738
3	TYR	C	174.731
3	TYR	CA	57.946
3	TYR	CB	38.884
3	TYR	CD1	133.117
3	TYR	CE1	118.111
3	TYR	N	122.382
4	ILE	H	8.015
4	ILE	HA	4.274

4	ILE	HB	1.626
4	ILE	HG12	1.363
4	ILE	HG13	0.993
4	ILE	HG21	0.776
4	ILE	HG22	0.776
4	ILE	HG23	0.776
4	ILE	HD11	0.737
4	ILE	HD12	0.737
4	ILE	HD13	0.737
4	ILE	C	173.509
4	ILE	CA	57.896
4	ILE	CB	39.192
4	ILE	CG1	26.835
4	ILE	CG2	16.956
4	ILE	CD1	12.588
4	ILE	N	127.482
5	PRO	HA	4.183
5	PRO	HB2	2.244
5	PRO	HB3	1.853
5	PRO	HG2	1.984
5	PRO	HG3	1.908
5	PRO	HD2	3.62
5	PRO	HD3	3.536
5	PRO	C	177.404
5	PRO	CA	63.439
5	PRO	CB	31.974
5	PRO	CG	27.338
5	PRO	CD	50.87
6	GLY	H	8.423
6	GLY	HA2	3.953
6	GLY	HA3	3.792
6	GLY	C	174.082
6	GLY	CA	45.146
7	GLN	H	7.986
7	GLN	HA	4.585
7	GLN	HB2	2.045
7	GLN	HB3	1.908
7	GLN	HG2	2.282
7	GLN	HG3	2.282

7	GLN	HE21	7.517
7	GLN	HE22	6.826
7	GLN	C	173.973
7	GLN	CA	53.522
7	GLN	CB	28.955
7	GLN	CG	33.526
7	GLN	CD	180.521
7	GLN	NE2	112.403
8	PRO	HA	4.423
8	PRO	HB2	2.225
8	PRO	HB3	1.836
8	PRO	HG2	1.961
8	PRO	HG3	1.923
8	PRO	HD2	3.75
8	PRO	HD3	3.589
8	PRO	C	176.778
8	PRO	CA	63.117
8	PRO	CB	32.034
8	PRO	CG	27.453
8	PRO	CD	50.58
9	VAL	H	8.296
9	VAL	HA	4.172
9	VAL	HB	2.018
9	VAL	HG11	0.893
9	VAL	HG12	0.893
9	VAL	HG13	0.893
9	VAL	HG21	0.893
9	VAL	HG22	0.893
9	VAL	HG23	0.893
9	VAL	C	176.386
9	VAL	CA	62.342
9	VAL	CB	32.781
9	VAL	CG1	20.775
9	VAL	N	120.566
10	THR	H	8.235
10	THR	HA	4.321
10	THR	HB	4.157
10	THR	HG21	1.124
10	THR	HG22	1.124

10	THR	HG23	1.124
10	THR	C	173.499
10	THR	CA	61.59
10	THR	CB	69.954
10	THR	CG2	21.537
10	THR	N	117.806
11	ALA	H	8.073
11	ALA	HA	4.54
11	ALA	HB1	1.262
11	ALA	HB2	1.262
11	ALA	HB3	1.262
11	ALA	C	174.666
11	ALA	CA	52.007
11	ALA	CB	19.875
11	ALA	N	126.558
12	VAL	H	8.484
12	VAL	HA	4.204
12	VAL	HB	1.984
12	VAL	HG11	0.852
12	VAL	HG12	0.852
12	VAL	HG13	0.852
12	VAL	HG21	0.852
12	VAL	HG22	0.852
12	VAL	HG23	0.852
12	VAL	C	174.99
12	VAL	CA	61.618
12	VAL	CB	33.104
12	VAL	CG1	20.922
12	VAL	N	120.498
13	VAL	H	8.118
13	VAL	HA	4.843
13	VAL	HB	1.817
13	VAL	HG11	0.769
13	VAL	HG12	0.769
13	VAL	HG13	0.769
13	VAL	HG21	0.817
13	VAL	HG22	0.817
13	VAL	HG23	0.817
13	VAL	C	176.141

13	VAL	CA	60.986
13	VAL	CB	33.144
13	VAL	CG1	21.551
13	VAL	CG2	20.761
13	VAL	N	124.007
14	GLN	H	9.054
14	GLN	HA	4.614
14	GLN	HB2	1.592
14	GLN	HB3	1.769
14	GLN	HG2	2.045
14	GLN	HG3	2.014
14	GLN	HE21	7.124
14	GLN	HE22	6.719
14	GLN	C	173.87
14	GLN	CA	53.934
14	GLN	CB	32.143
14	GLN	CG	33.046
14	GLN	CD	179.198
14	GLN	N	124.757
14	GLN	NE2	110.972
15	ARG	H	8.602
15	ARG	HA	5.036
15	ARG	HB2	1.732
15	ARG	HB3	1.618
15	ARG	HG2	1.493
15	ARG	HG3	1.39
15	ARG	HD2	3.089
15	ARG	HD3	3.089
15	ARG	HE	7.172
15	ARG	C	176.081
15	ARG	CA	55.278
15	ARG	CB	31.198
15	ARG	CG	27.71
15	ARG	CD	43.277
15	ARG	N	123.97
16	VAL	H	8.694
16	VAL	HA	4.304
16	VAL	HB	1.64
16	VAL	HG11	0.593

16	VAL	HG12	0.593
16	VAL	HG13	0.593
16	VAL	HG21	0.64
16	VAL	HG22	0.64
16	VAL	HG23	0.64
16	VAL	C	173.773
16	VAL	CA	61.22
16	VAL	CB	35.353
16	VAL	CG1	20.673
16	VAL	CG2	21.548
16	VAL	N	124.82
17	GLU	H	9.16
17	GLU	HA	4.939
17	GLU	HB2	1.907
17	GLU	HB3	1.907
17	GLU	HG2	1.92
17	GLU	HG3	1.764
17	GLU	C	173.81
17	GLU	CA	55.088
17	GLU	CB	31.524
17	GLU	CG	37.513
17	GLU	N	129.391
18	ILE	H	8.83
18	ILE	HA	4.134
18	ILE	HB	1.78
18	ILE	HG12	1.451
18	ILE	HG13	1.451
18	ILE	HG21	0.779
18	ILE	HG22	0.779
18	ILE	HG23	0.779
18	ILE	HD11	0.707
18	ILE	HD12	0.707
18	ILE	HD13	0.707
18	ILE	C	174.777
18	ILE	CA	60.643
18	ILE	CB	40.096
18	ILE	CG1	25.828
18	ILE	CG2	17.09
18	ILE	CD1	15.7

18	ILE	N	124.847
19	HIS	H	8.856
19	HIS	HA	4.745
19	HIS	HB2	3.157
19	HIS	HB3	3.104
19	HIS	HD2	7.17
19	HIS	HE1	8.299
19	HIS	C	173.997
19	HIS	CA	54.532
19	HIS	CB	28.489
19	HIS	CD2	119.487
19	HIS	CE1	136.51
19	HIS	N	128.321
20	LYS	H	8.545
20	LYS	HA	4.019
20	LYS	HB2	2.117
20	LYS	HB3	2.117
20	LYS	HG2	1.55
20	LYS	HG3	1.55
20	LYS	HE2	3.139
20	LYS	HE3	3.139
20	LYS	C	175.603
20	LYS	CA	57.826
20	LYS	CB	35.025
20	LYS	N	123.322
21	LEU	H	8.155
21	LEU	HA	4.535
21	LEU	HB2	2.006
21	LEU	HB3	2.006
21	LEU	HG	1.19
21	LEU	HD11	0.885
21	LEU	HD12	0.885
21	LEU	HD13	0.885
21	LEU	HD21	0.885
21	LEU	HD22	0.885
21	LEU	HD23	0.885
21	LEU	C	175.742
21	LEU	CA	53.469
21	LEU	CB	37.528

21	LEU	N	122.807
22	ARG	H	8.739
22	ARG	HA	4.728
22	ARG	HB2	1.764
22	ARG	HB3	1.611
22	ARG	HG2	1.458
22	ARG	HG3	1.458
22	ARG	HD2	3.201
22	ARG	HD3	3.085
22	ARG	HE	7.348
22	ARG	C	175.822
22	ARG	CA	56.008
22	ARG	CB	30.409
22	ARG	CG	27.531
22	ARG	CD	43.082
22	ARG	N	126.951
23	GLN	H	8.593
23	GLN	HA	4.477
23	GLN	HB2	1.935
23	GLN	HB3	1.727
23	GLN	HG2	2.153
23	GLN	HG3	2.153
23	GLN	HE21	7.468
23	GLN	HE22	6.799
23	GLN	C	175.431
23	GLN	CA	54.863
23	GLN	CB	30.434
23	GLN	CG	33.911
23	GLN	CD	180.314
23	GLN	N	128.95
23	GLN	NE2	111.693
24	GLY	H	9.009
24	GLY	HA2	3.597
24	GLY	HA3	3.961
24	GLY	C	174.958
24	GLY	CA	46.981
24	GLY	N	117.269
25	GLU	H	9.063
25	GLU	HA	4.228

25	GLU	HB2	2.153
25	GLU	HB3	1.795
25	GLU	HG2	2.223
25	GLU	HG3	2.178
25	GLU	C	175.972
25	GLU	CA	56.387
25	GLU	CB	29.895
25	GLU	CG	36.111
25	GLU	N	126.35
26	ASN	H	8
26	ASN	HA	4.79
26	ASN	HB2	2.824
26	ASN	HB3	2.648
26	ASN	HD21	7.631
26	ASN	HD22	7.023
26	ASN	C	173.634
26	ASN	CA	52.577
26	ASN	CB	41.154
26	ASN	CG	176.558
26	ASN	N	117.979
26	ASN	ND2	114.416
27	LEU	H	8.348
27	LEU	HA	4.898
27	LEU	HB2	1.585
27	LEU	HB3	1.585
27	LEU	HG	1.458
27	LEU	HD11	1.156
27	LEU	HD12	1.156
27	LEU	HD13	1.156
27	LEU	HD21	1.156
27	LEU	HD22	1.156
27	LEU	HD23	1.156
27	LEU	C	176.417
27	LEU	CA	52.631
27	LEU	CB	41.098
27	LEU	N	122.555
28	ILE	H	9.382
28	ILE	HA	4.206
28	ILE	HB	1.973

28	ILE	HG12	1.429
28	ILE	HG13	1.429
28	ILE	HG21	0.848
28	ILE	HG22	0.848
28	ILE	HG23	0.848
28	ILE	HD11	0.738
28	ILE	HD12	0.738
28	ILE	HD13	0.738
28	ILE	C	175.951
28	ILE	CA	60.342
28	ILE	CB	39.788
28	ILE	CG2	17.767
28	ILE	CD1	12.913
28	ILE	N	125.2
29	LEU	H	8.479
29	LEU	HA	4.191
29	LEU	HB2	2.207
29	LEU	HB3	2.207
29	LEU	HG	1.471
29	LEU	HD11	0.694
29	LEU	HD12	0.694
29	LEU	HD13	0.694
29	LEU	HD21	0.694
29	LEU	HD22	0.694
29	LEU	HD23	0.694
29	LEU	CA	60.29
29	LEU	CB	39.885
29	LEU	N	125.19
30	GLY	H	8.105
30	GLY	HA2	4.147
30	GLY	HA3	3.955
30	GLY	C	175.42
30	GLY	CA	45.44
30	GLY	N	101.065
31	PHE	H	7.368
31	PHE	HA	5.178
31	PHE	HB2	2.933
31	PHE	HB3	2.544
31	PHE	C	172.251

31	PHE	CA	56.738
31	PHE	CB	40.304
31	PHE	N	115.927
32	SER	H	8.478
32	SER	HA	4.762
32	SER	HB2	3.888
32	SER	HB3	3.508
32	SER	C	173.334
32	SER	CA	56.879
32	SER	CB	65.859
32	SER	N	114.993
33	ILE	H	8.41
33	ILE	HA	5.701
33	ILE	HB	1.831
33	ILE	HG12	1.363
33	ILE	HG13	1.363
33	ILE	HG21	0.768
33	ILE	HG22	0.768
33	ILE	HG23	0.768
33	ILE	HD11	0.336
33	ILE	HD12	0.336
33	ILE	HD13	0.336
33	ILE	C	175.618
33	ILE	CA	58.866
33	ILE	CB	43.721
33	ILE	CG1	25.547
33	ILE	CG2	20.198
33	ILE	CD1	14.791
33	ILE	N	113.297
34	GLY	H	9.26
34	GLY	HA2	4.635
34	GLY	HA3	3.627
34	GLY	C	172.247
34	GLY	CA	43.631
34	GLY	N	109.829
35	GLY	H	8.548
35	GLY	HA2	5.269
35	GLY	HA3	3.884
35	GLY	C	174.137

35	GLY	CA	44.268
35	GLY	N	106.944
36	GLY	H	6.664
36	GLY	HA2	4.693
36	GLY	HA3	3.978
36	GLY	C	177.68
36	GLY	CA	43.807
36	GLY	N	105.448
37	ILE	H	8.497
37	ILE	HA	4.124
37	ILE	HB	2.041
37	ILE	HG12	1.282
37	ILE	HG13	0.905
37	ILE	HG21	0.961
37	ILE	HG22	0.961
37	ILE	HG23	0.961
37	ILE	HD11	0.764
37	ILE	HD12	0.764
37	ILE	HD13	0.764
37	ILE	C	175.326
37	ILE	CA	64.61
37	ILE	CB	37.81
37	ILE	CG1	26.088
37	ILE	CG2	17.973
37	ILE	CD1	13.746
37	ILE	N	114.741
38	ASP	H	9.781
38	ASP	HA	4.631
38	ASP	HB2	2.93
38	ASP	HB3	2.575
38	ASP	C	175.329
38	ASP	CA	52.637
38	ASP	CB	39.794
38	ASP	N	117.982
39	GLN	H	7.383
39	GLN	HA	4.432
39	GLN	HB2	1.788
39	GLN	HB3	1.788
39	GLN	HG2	2.249

39	GLN	HG3	2.12
39	GLN	HE21	7.113
39	GLN	HE22	6.75
39	GLN	C	175.123
39	GLN	CA	53.809
39	GLN	CB	30.204
39	GLN	CG	33.853
39	GLN	CD	179.971
39	GLN	N	119.074
39	GLN	NE2	113.998
40	ASP	H	8.62
40	ASP	HA	4.838
40	ASP	HB2	2.621
40	ASP	HB3	2.793
40	ASP	C	176.434
40	ASP	CA	51.38
40	ASP	CB	41.544
40	ASP	N	123.108
41	PRO	HA	4.565
41	PRO	HB2	2.166
41	PRO	HB3	1.903
41	PRO	HG2	2.019
41	PRO	HG3	1.875
41	PRO	HD2	4.081
41	PRO	HD3	4.049
41	PRO	C	178.338
41	PRO	CA	64.346
41	PRO	CB	31.943
41	PRO	CG	26.957
41	PRO	CD	51.031
42	SER	H	8.411
42	SER	HA	4.15
42	SER	HB2	3.892
42	SER	HB3	3.892
42	SER	C	175.292
42	SER	CA	61.065
42	SER	CB	62.997
42	SER	N	115.249
43	GLN	H	7.592

43	GLN	HA	4.2
43	GLN	HB2	2.288
43	GLN	HB3	1.872
43	GLN	HG2	2.267
43	GLN	HG3	2.267
43	GLN	HE21	7.452
43	GLN	HE22	6.772
43	GLN	C	174.995
43	GLN	CA	54.837
43	GLN	CB	28.77
43	GLN	CG	33.649
43	GLN	N	118.424
43	GLN	NE2	113.049
44	ASN	H	7.229
44	ASN	HA	4.758
44	ASN	HB2	2.722
44	ASN	HB3	3.095
44	ASN	C	174.149
44	ASN	CA	49.94
44	ASN	CB	38.914
44	ASN	N	119.643
45	PRO	HA	4.146
45	PRO	HB2	1.871
45	PRO	HB3	1.033
45	PRO	HG2	1.59
45	PRO	HG3	0.892
45	PRO	HD2	3.92
45	PRO	HD3	3.451
45	PRO	C	176.973
45	PRO	CA	63.394
45	PRO	CB	31.815
45	PRO	CG	25.836
45	PRO	CD	50.421
46	PHE	H	7.591
46	PHE	HA	4.347
46	PHE	HB2	2.242
46	PHE	HB3	3.125
46	PHE	C	174.993
46	PHE	CA	57.466

46	PHE	CB	34.214
46	PHE	N	115.856
47	SER	H	6.803
47	SER	HA	4.194
47	SER	HB2	3.882
47	SER	HB3	3.558
47	SER	C	175.363
47	SER	CA	56.83
47	SER	CB	63.808
47	SER	N	112.15
48	GLU	H	8.996
48	GLU	HA	4.124
48	GLU	HB2	2.012
48	GLU	HB3	1.946
48	GLU	HG2	2.275
48	GLU	HG3	2.193
48	GLU	C	176.307
48	GLU	CA	57.98
48	GLU	CB	30.314
48	GLU	CG	36.446
48	GLU	N	124.463
49	ASP	H	7.961
49	ASP	HA	4.566
49	ASP	HB2	2.89
49	ASP	HB3	2.658
49	ASP	C	176.656
49	ASP	CA	52.844
49	ASP	CB	41.534
49	ASP	N	117.472
50	LYS	H	8.454
50	LYS	HA	4.191
50	LYS	HB2	1.892
50	LYS	HB3	1.892
50	LYS	HG2	1.704
50	LYS	HG3	1.704
50	LYS	HD2	1.426
50	LYS	HD3	1.426
50	LYS	C	176.387
50	LYS	CA	56.481

50	LYS	CB	30.466
50	LYS	CG	24.361
50	LYS	CD	29.706
50	LYS	N	120.286
51	THR	H	8.386
51	THR	HA	4.333
51	THR	HB	4.321
51	THR	HG21	1.115
51	THR	HG22	1.115
51	THR	HG23	1.115
51	THR	C	174.7
51	THR	CA	61.667
51	THR	CB	69.95
51	THR	CG2	21.564
51	THR	N	108.627
52	ASP	H	7.619
52	ASP	HA	4.494
52	ASP	HB2	3.138
52	ASP	HB3	2.776
52	ASP	C	176.844
52	ASP	CA	54.645
52	ASP	CB	41.523
52	ASP	N	122.291
53	LYS	H	8.771
53	LYS	HA	4.621
53	LYS	HB2	2.172
53	LYS	HB3	2.172
53	LYS	HG2	1.379
53	LYS	HG3	1.379
53	LYS	HD2	1.61
53	LYS	HD3	1.61
53	LYS	HE2	3.079
53	LYS	HE3	3.079
53	LYS	C	176.922
53	LYS	CA	56.139
53	LYS	CB	32.121
53	LYS	N	129.103
54	GLY	H	8.734
54	GLY	HA2	3.545

54	GLY	HA3	4.137
54	GLY	C	172.307
54	GLY	CA	45.208
54	GLY	N	106.524
55	ILE	H	9.016
55	ILE	HA	4.78
55	ILE	HB	2.086
55	ILE	HG12	1.272
55	ILE	HG13	1.131
55	ILE	HG21	0.729
55	ILE	HG22	0.729
55	ILE	HG23	0.729
55	ILE	HD11	0.455
55	ILE	HD12	0.455
55	ILE	HD13	0.455
55	ILE	C	174.992
55	ILE	CA	57.592
55	ILE	CB	35.815
55	ILE	CG2	18.546
55	ILE	CD1	8.457
55	ILE	N	120.479
56	TYR	H	8.788
56	TYR	HA	5.419
56	TYR	HB2	2.473
56	TYR	HB3	2.473
56	TYR	HD1	6.898
56	TYR	HD2	6.898
56	TYR	HE1	6.607
56	TYR	HE2	6.607
56	TYR	C	175.689
56	TYR	CA	55.444
56	TYR	CB	42.81
56	TYR	CD1	133.806
56	TYR	CE1	117.699
56	TYR	N	124.57
57	VAL	H	8.764
57	VAL	HA	4.362
57	VAL	HB	2.134
57	VAL	HG11	0.728

57	VAL	HG12	0.728
57	VAL	HG13	0.728
57	VAL	HG21	0.75
57	VAL	HG22	0.75
57	VAL	HG23	0.75
57	VAL	C	176.841
57	VAL	CA	62.641
57	VAL	CB	32.109
57	VAL	CG1	22.548
57	VAL	CG2	21.981
57	VAL	N	121.119
58	THR	H	8.654
58	THR	HA	4.399
58	THR	HB	4.478
58	THR	HG21	1.043
58	THR	HG22	1.043
58	THR	HG23	1.043
58	THR	C	175.144
58	THR	CA	62.131
58	THR	CB	69.162
58	THR	CG2	21.221
58	THR	N	118.563
59	ARG	H	7.166
59	ARG	HA	4.346
59	ARG	HB2	1.714
59	ARG	HB3	1.646
59	ARG	HG2	1.578
59	ARG	HG3	1.497
59	ARG	HD2	3.105
59	ARG	HD3	3.105
59	ARG	C	175.017
59	ARG	CA	56.26
59	ARG	CB	34.149
59	ARG	CG	27.197
59	ARG	CD	43.252
59	ARG	N	119.725
60	VAL	H	8.469
60	VAL	HA	4.316
60	VAL	HB	1.831

60	VAL	HG11	0.775
60	VAL	HG12	0.775
60	VAL	HG13	0.775
60	VAL	HG21	0.534
60	VAL	HG22	0.534
60	VAL	HG23	0.534
60	VAL	C	175.3
60	VAL	CA	62.063
60	VAL	CB	34.95
60	VAL	CG1	21.696
60	VAL	CG2	22.037
60	VAL	N	122.993
61	SER	H	8.112
61	SER	HA	4.278
61	SER	HB2	3.875
61	SER	HB3	3.653
61	SER	C	175.357
61	SER	CA	59.246
61	SER	CB	62.928
61	SER	N	122.681
62	GLU	H	9.447
62	GLU	HA	4.163
62	GLU	HB2	1.982
62	GLU	HB3	2.015
62	GLU	HG2	2.359
62	GLU	HG3	2.292
62	GLU	C	178.129
62	GLU	CA	58.275
62	GLU	CB	29.09
62	GLU	CG	36.193
62	GLU	N	131.457
63	GLY	H	9.619
63	GLY	HA2	4.044
63	GLY	HA3	3.679
63	GLY	C	174.146
63	GLY	CA	45.47
63	GLY	N	116.309
64	GLY	H	7.585
64	GLY	HA2	4.396

64	GLY	HA3	3.885
64	GLY	C	172.075
64	GLY	CA	45.323
64	GLY	N	105.751
65	PRO	HA	4.233
65	PRO	HB2	2.55
65	PRO	HB3	2.018
65	PRO	HG2	1.995
65	PRO	HG3	2.073
65	PRO	HD2	3.664
65	PRO	HD3	3.35
65	PRO	C	178.941
65	PRO	CA	64.883
65	PRO	CB	32.565
65	PRO	CG	27.896
65	PRO	CD	49.084
66	ALA	H	7.33
66	ALA	HA	4.088
66	ALA	HB1	1.51
66	ALA	HB2	1.51
66	ALA	HB3	1.51
66	ALA	C	177.609
66	ALA	CA	54.231
66	ALA	CB	19.389
66	ALA	N	119.976
67	GLU	H	8.221
67	GLU	HA	3.874
67	GLU	HB2	2.123
67	GLU	HB3	2.063
67	GLU	HG2	2.195
67	GLU	HG3	2.038
67	GLU	C	181.181
67	GLU	CA	59.638
67	GLU	CB	30.025
67	GLU	CG	37.148
67	GLU	N	121.916
68	ILE	H	8.032
68	ILE	HA	3.739
68	ILE	HB	1.796

68	ILE	HG12	1.526
68	ILE	HG13	1.169
68	ILE	HG21	0.897
68	ILE	HG22	0.897
68	ILE	HG23	0.897
68	ILE	HD11	0.761
68	ILE	HD12	0.761
68	ILE	HD13	0.761
68	ILE	C	177.365
68	ILE	CA	64.314
68	ILE	CB	38.095
68	ILE	CG1	28.864
68	ILE	CG2	16.975
68	ILE	CD1	13.227
68	ILE	N	119.497
69	ALA	H	7.125
69	ALA	HA	4.309
69	ALA	HB1	1.464
69	ALA	HB2	1.464
69	ALA	HB3	1.464
69	ALA	C	177.401
69	ALA	CA	52.807
69	ALA	CB	21.222
69	ALA	N	119.68
70	GLY	H	7.57
70	GLY	HA2	4.26
70	GLY	HA3	3.679
70	GLY	C	174.456
70	GLY	CA	44.907
70	GLY	N	103.907
71	LEU	H	7.747
71	LEU	HA	3.769
71	LEU	HB2	1.282
71	LEU	HB3	1.282
71	LEU	HG	0.673
71	LEU	HD11	-0.289
71	LEU	HD12	-0.289
71	LEU	HD13	-0.289
71	LEU	HD21	-0.289

71	LEU	HD22	-0.289
71	LEU	HD23	-0.289
71	LEU	C	174.486
71	LEU	CA	55.195
71	LEU	CB	41.364
71	LEU	CG	27.019
71	LEU	N	123.521
72	GLN	H	8.436
72	GLN	HA	4.565
72	GLN	HB2	1.857
72	GLN	HB3	1.857
72	GLN	HG2	2.177
72	GLN	HG3	2.177
72	GLN	HE21	7.28
72	GLN	HE22	6.687
72	GLN	C	175.23
72	GLN	CA	54.012
72	GLN	CB	31.964
72	GLN	CG	33.562
72	GLN	CD	180.248
72	GLN	N	123.821
72	GLN	NE2	110.86
73	ILE	H	8.402
73	ILE	HA	3.242
73	ILE	HB	1.562
73	ILE	HG12	0.841
73	ILE	HG13	1.404
73	ILE	HG21	0.792
73	ILE	HG22	0.792
73	ILE	HG23	0.792
73	ILE	HD11	0.96
73	ILE	HD12	0.96
73	ILE	HD13	0.96
73	ILE	C	177.196
73	ILE	CA	63.538
73	ILE	CB	38.014
73	ILE	CG1	28.351
73	ILE	CG2	17.977
73	ILE	CD1	13.988

73	ILE	N	119.852
74	GLY	H	8.92
74	GLY	HA2	3.846
74	GLY	HA3	2.565
74	GLY	C	173.551
74	GLY	CA	44.655
74	GLY	N	116.672
75	ASP	H	7.615
75	ASP	HA	4.445
75	ASP	HB2	2.439
75	ASP	HB3	2.439
75	ASP	C	174.754
75	ASP	CA	55.718
75	ASP	CB	40.626
75	ASP	N	122.336
76	LYS	H	8.444
76	LYS	HA	4.348
76	LYS	HB2	2.153
76	LYS	HB3	2.153
76	LYS	HG2	1.937
76	LYS	HG3	1.937
76	LYS	HD2	1.74
76	LYS	HD3	1.74
76	LYS	C	176.773
76	LYS	CA	54.659
76	LYS	CB	33.946
76	LYS	N	123.369
77	ILE	H	8.695
77	ILE	HA	3.853
77	ILE	HB	1.51
77	ILE	HG12	1.153
77	ILE	HG13	1.153
77	ILE	HG21	0.678
77	ILE	HG22	0.678
77	ILE	HG23	0.678
77	ILE	HD11	0.697
77	ILE	HD12	0.697
77	ILE	HD13	0.697
77	ILE	C	174.243

77	ILE	CA	62.219
77	ILE	CB	37.871
77	ILE	CG2	19.332
77	ILE	CD1	13.642
77	ILE	N	125.789
78	MET	H	9.101
78	MET	HA	4.449
78	MET	HB2	2.119
78	MET	HB3	1.648
78	MET	HG2	2.357
78	MET	HG3	2.357
78	MET	HE1	1.908
78	MET	HE2	1.908
78	MET	HE3	1.908
78	MET	C	177.971
78	MET	CA	55.524
78	MET	CB	32.469
78	MET	CG	31.375
78	MET	CE	16.22
78	MET	N	125.28
79	GLN	H	7.628
79	GLN	HA	5.237
79	GLN	HB2	1.762
79	GLN	HB3	1.828
79	GLN	HG2	2.08
79	GLN	HG3	2.008
79	GLN	HE21	7.612
79	GLN	HE22	6.701
79	GLN	C	176.886
79	GLN	CA	54.88
79	GLN	CB	35.022
79	GLN	CG	34.353
79	GLN	CD	179.161
79	GLN	N	116.831
79	GLN	NE2	111.553
80	VAL	H	8.487
80	VAL	HA	4.606
80	VAL	HB	1.978
80	VAL	HG11	0.866

80	VAL	HG12	0.866
80	VAL	HG13	0.866
80	VAL	HG21	0.866
80	VAL	HG22	0.866
80	VAL	HG23	0.866
80	VAL	C	174.671
80	VAL	CA	60.612
80	VAL	CB	34.705
80	VAL	CG2	22.254
80	VAL	N	120.584
81	ASN	H	9.703
81	ASN	HA	4.485
81	ASN	HB2	3.265
81	ASN	HB3	3.059
81	ASN	HD21	7.448
81	ASN	HD22	6.813
81	ASN	C	174.647
81	ASN	CA	54.365
81	ASN	CB	36.77
81	ASN	CG	177.786
81	ASN	N	127.611
81	ASN	ND2	111.245
82	GLY	H	8.443
82	GLY	HA2	3.988
82	GLY	HA3	3.411
82	GLY	C	173.416
82	GLY	CA	45.189
82	GLY	N	102.442
83	TRP	H	8.367
83	TRP	HA	4.566
83	TRP	HB2	3.265
83	TRP	HB3	3.218
83	TRP	HD1	7.347
83	TRP	HE1	10.133
83	TRP	HE3	7.556
83	TRP	HZ2	7.418
83	TRP	HZ3	7.051
83	TRP	HH2	7.088
83	TRP	C	175.926

83	TRP	CA	56.726
83	TRP	CB	29.469
83	TRP	CD1	128.059
83	TRP	CE3	120.917
83	TRP	CZ2	114.443
83	TRP	CZ3	121.082
83	TRP	CH2	124.323
83	TRP	N	123.318
83	TRP	NE1	129.705
84	ASP	H	8.593
84	ASP	HA	4.445
84	ASP	HB2	2.672
84	ASP	HB3	2.643
84	ASP	C	176.543
84	ASP	CA	55.801
84	ASP	CB	41.902
84	ASP	N	125.959
85	MET	H	8.046
85	MET	HA	4.72
85	MET	HB2	2.115
85	MET	HB3	1.639
85	MET	HG2	2.565
85	MET	HG3	2.317
85	MET	HE1	1.958
85	MET	HE2	1.958
85	MET	HE3	1.958
85	MET	C	176.746
85	MET	CA	54.155
85	MET	CB	33.284
85	MET	CE	18.93
85	MET	N	123.667
86	THR	H	8.51
86	THR	HA	4.195
86	THR	HB	4.165
86	THR	HG21	1.342
86	THR	HG22	1.342
86	THR	HG23	1.342
86	THR	C	175.349
86	THR	CA	64.954

86	THR	CB	69.597
86	THR	CG2	22.089
86	THR	N	115.987
87	MET	H	8.539
87	MET	HA	4.585
87	MET	HB2	2.104
87	MET	HB3	1.849
87	MET	HG2	2.525
87	MET	HG3	2.418
87	MET	HE1	2.126
87	MET	HE2	2.126
87	MET	HE3	2.126
87	MET	C	174.638
87	MET	CA	54.364
87	MET	CB	32.291
87	MET	CG	32.317
87	MET	CE	23.105
87	MET	N	123.127
88	VAL	H	7.965
88	VAL	HA	4.873
88	VAL	HB	2.366
88	VAL	HG11	0.89
88	VAL	HG12	0.89
88	VAL	HG13	0.89
88	VAL	HG21	0.913
88	VAL	HG22	0.913
88	VAL	HG23	0.913
88	VAL	C	176.856
88	VAL	CA	59.058
88	VAL	CB	35.025
88	VAL	CG1	18.861
88	VAL	CG2	22.54
88	VAL	N	113.502
89	THR	H	8.506
89	THR	HA	4.381
89	THR	HB	4.097
89	THR	HG21	1.224
89	THR	HG22	1.224
89	THR	HG23	1.224

89	THR	C	175.023
89	THR	CA	61.116
89	THR	CB	70.777
89	THR	CG2	22.138
89	THR	N	112.612
90	HIS	H	9.989
90	HIS	HA	3.864
90	HIS	HB2	3.546
90	HIS	HB3	3.286
90	HIS	HD2	6.868
90	HIS	HE1	7.73
90	HIS	C	177.411
90	HIS	CA	61.701
90	HIS	CB	29.11
90	HIS	CD2	124.364
90	HIS	CE1	137.237
90	HIS	N	122.908
91	ASP	H	9.185
91	ASP	HA	4.301
91	ASP	HB2	2.652
91	ASP	HB3	2.41
91	ASP	C	178.552
91	ASP	CA	57.525
91	ASP	CB	41.698
91	ASP	N	115.99
92	GLN	H	7.808
92	GLN	HA	3.791
92	GLN	HB2	2.361
92	GLN	HB3	1.868
92	GLN	HG2	2.38
92	GLN	HG3	2.38
92	GLN	HE21	7.456
92	GLN	HE22	6.883
92	GLN	C	179.313
92	GLN	CA	58.899
92	GLN	CB	28.859
92	GLN	CG	34.801
92	GLN	CD	180.317
92	GLN	N	118.454

92	GLN	NE2	111.173
93	ALA	H	7.967
93	ALA	HA	3.832
93	ALA	HB1	1.269
93	ALA	HB2	1.269
93	ALA	HB3	1.269
93	ALA	C	178.754
93	ALA	CA	55.28
93	ALA	CB	18.986
93	ALA	N	121.707
94	ARG	H	8.258
94	ARG	HA	3.567
94	ARG	HB2	1.833
94	ARG	HB3	1.676
94	ARG	HG2	1.563
94	ARG	HG3	1.339
94	ARG	HD2	3.182
94	ARG	HD3	3.182
94	ARG	C	180.137
94	ARG	CA	59.887
94	ARG	CB	29.895
94	ARG	CG	27.187
94	ARG	CD	43.276
94	ARG	N	117.066
95	LYS	H	8.41
95	LYS	HA	3.72
95	LYS	HB2	1.755
95	LYS	HB3	1.65
95	LYS	HG2	1.469
95	LYS	HG3	1.258
95	LYS	HD2	1.508
95	LYS	HD3	1.508
95	LYS	HE2	2.831
95	LYS	HE3	2.831
95	LYS	C	179.265
95	LYS	CA	59.559
95	LYS	CB	32.178
95	LYS	CG	25.88
95	LYS	CD	28.95

95	LYS	N	120.092
96	ARG	H	7.454
96	ARG	HA	3.859
96	ARG	HB2	1.707
96	ARG	HB3	1.603
96	ARG	HG2	1.475
96	ARG	HG3	1.355
96	ARG	HD2	2.437
96	ARG	HD3	2.437
96	ARG	C	178.665
96	ARG	CA	57.47
96	ARG	CB	29.387
96	ARG	CG	26.337
96	ARG	CD	42.469
96	ARG	N	116.31
97	LEU	H	7.602
97	LEU	HA	3.999
97	LEU	HB2	2.433
97	LEU	HB3	2.433
97	LEU	HG	1.351
97	LEU	HD11	1.025
97	LEU	HD12	1.025
97	LEU	HD13	1.025
97	LEU	HD21	1.025
97	LEU	HD22	1.025
97	LEU	HD23	1.025
97	LEU	C	176.33
97	LEU	CA	56.818
97	LEU	CB	34.197
97	LEU	N	117.299
98	THR	H	7.185
98	THR	HA	4.541
98	THR	HB	4.253
98	THR	HG21	1.115
98	THR	HG22	1.115
98	THR	HG23	1.115
98	THR	C	174.776
98	THR	CA	60.281
98	THR	CB	69.126

98	THR	CG2	22.632
98	THR	N	103.957
99	LYS	H	7.005
99	LYS	HA	4.136
99	LYS	HB2	1.766
99	LYS	HB3	1.621
99	LYS	HG2	1.485
99	LYS	HG3	1.293
99	LYS	HD2	1.615
99	LYS	HD3	1.615
99	LYS	C	178.407
99	LYS	CA	57.649
99	LYS	CB	31.969
99	LYS	CG	24.827
99	LYS	CD	28.933
99	LYS	N	124.162
100	ARG	H	8.693
100	ARG	HA	3.902
100	ARG	HB2	1.828
100	ARG	HB3	1.828
100	ARG	HG2	1.713
100	ARG	HG3	1.678
100	ARG	HD2	3.213
100	ARG	HD3	3.213
100	ARG	C	176.11
100	ARG	CA	58.488
100	ARG	CB	30.058
100	ARG	CG	27.514
100	ARG	CD	43.228
100	ARG	N	127.484
101	SER	H	7.597
101	SER	HA	4.233
101	SER	HB2	4.064
101	SER	HB3	3.716
101	SER	C	174.046
101	SER	CA	58.167
101	SER	CB	62.969
101	SER	N	109.086
102	GLU	H	7.165

102	GLU	HA	4.503
102	GLU	HB2	1.748
102	GLU	HB3	1.748
102	GLU	HG2	2.144
102	GLU	HG3	2.076
102	GLU	C	175.151
102	GLU	CA	54.722
102	GLU	CB	30.739
102	GLU	CG	36.083
102	GLU	N	121.519
103	GLU	H	8.733
103	GLU	HA	4.008
103	GLU	HB2	1.931
103	GLU	HB3	1.931
103	GLU	HG2	2.289
103	GLU	HG3	2.145
103	GLU	C	174.747
103	GLU	CA	57.945
103	GLU	CB	30.339
103	GLU	CG	37.076
103	GLU	N	123.304
104	VAL	H	7.589
104	VAL	HA	4.821
104	VAL	HB	1.704
104	VAL	HG11	0.481
104	VAL	HG12	0.481
104	VAL	HG13	0.481
104	VAL	HG21	0.508
104	VAL	HG22	0.508
104	VAL	HG23	0.508
104	VAL	C	175.446
104	VAL	CA	60.205
104	VAL	CB	34.629
104	VAL	CG1	21.194
104	VAL	CG2	20.779
104	VAL	N	117.592
105	VAL	H	8.764
105	VAL	HA	4.595
105	VAL	HB	1.866

105	VAL	HG11	0.947
105	VAL	HG12	0.947
105	VAL	HG13	0.947
105	VAL	HG21	0.905
105	VAL	HG22	0.905
105	VAL	HG23	0.905
105	VAL	C	173.303
105	VAL	CA	59.725
105	VAL	CB	34.292
105	VAL	CG1	21.267
105	VAL	CG2	22.733
105	VAL	N	121
106	ARG	H	8.842
106	ARG	HA	4.724
106	ARG	HB2	1.839
106	ARG	HB3	1.792
106	ARG	HG2	1.514
106	ARG	HG3	1.345
106	ARG	HD2	2.898
106	ARG	HD3	2.381
106	ARG	C	175.423
106	ARG	CA	55.515
106	ARG	CB	29.879
106	ARG	CG	28.355
106	ARG	CD	43.559
106	ARG	N	126.175
107	LEU	H	9.447
107	LEU	HA	5.247
107	LEU	HB2	1.686
107	LEU	HB3	1.686
107	LEU	HG	1.236
107	LEU	HD11	0.853
107	LEU	HD12	0.853
107	LEU	HD13	0.853
107	LEU	HD21	0.853
107	LEU	HD22	0.853
107	LEU	HD23	0.853
107	LEU	C	176.451
107	LEU	CA	53.338

107	LEU	CB	41.631
107	LEU	CG	25.912
107	LEU	N	125.682
108	LEU	H	7.906
108	LEU	HA	5.037
108	LEU	HB2	1.369
108	LEU	HB3	1.369
108	LEU	HG	1.097
108	LEU	HD11	0.709
108	LEU	HD12	0.709
108	LEU	HD13	0.709
108	LEU	HD21	0.709
108	LEU	HD22	0.709
108	LEU	HD23	0.709
108	LEU	C	176.295
108	LEU	CA	54.777
108	LEU	CB	44.285
108	LEU	CG	24.282
108	LEU	N	123.684
109	VAL	H	9.081
109	VAL	HA	5.611
109	VAL	HB	1.863
109	VAL	HG11	0.557
109	VAL	HG12	0.557
109	VAL	HG13	0.557
109	VAL	HG21	0.587
109	VAL	HG22	0.587
109	VAL	HG23	0.587
109	VAL	C	174.671
109	VAL	CA	58.007
109	VAL	CB	35.316
109	VAL	CG1	18.316
109	VAL	CG2	21.324
109	VAL	N	119.516
110	THR	H	8.926
110	THR	HA	5.106
110	THR	HB	3.941
110	THR	HG21	1.072
110	THR	HG22	1.072

110	THR	HG23	1.072
110	THR	C	173.001
110	THR	CA	60.459
110	THR	CB	70.979
110	THR	CG2	21.717
110	THR	N	113.689
111	ARG	H	8.732
111	ARG	HA	4.767
111	ARG	HB2	1.862
111	ARG	HB3	1.525
111	ARG	HG2	1.503
111	ARG	HG3	1.421
111	ARG	HD2	2.982
111	ARG	HD3	2.928
111	ARG	C	175.195
111	ARG	CA	54.677
111	ARG	CB	33.352
111	ARG	CG	25.826
111	ARG	CD	43.189
111	ARG	N	125.719
112	GLN	H	8.848
112	GLN	HA	4.437
112	GLN	HB2	2.039
112	GLN	HB3	1.98
112	GLN	HG2	2.412
112	GLN	HG3	2.362
112	GLN	HE21	7.605
112	GLN	HE22	6.863
112	GLN	C	176.206
112	GLN	CA	55.931
112	GLN	CB	29.528
112	GLN	CG	33.905
112	GLN	CD	180.088
112	GLN	N	123.698
112	GLN	NE2	111.936
113	SER	H	8.497
113	SER	HA	4.372
113	SER	HB2	3.782
113	SER	HB3	3.782

113	SER	C	174.299
113	SER	CA	58.362
113	SER	CB	63.861
113	SER	N	117.066
114	LEU	H	8.222
114	LEU	HA	4.333
114	LEU	HB2	1.572
114	LEU	HB3	1.572
114	LEU	HD11	0.812
114	LEU	HD12	0.812
114	LEU	HD13	0.812
114	LEU	HD21	0.866
114	LEU	HD22	0.866
114	LEU	HD23	0.866
114	LEU	C	175.995
114	LEU	CA	55.191
114	LEU	CB	42.341
114	LEU	CG	27.057
114	LEU	CD1	24.635
114	LEU	CD2	24.635
114	LEU	N	124.369
115	GLN	H	8.3
115	GLN	HA	4.223
115	GLN	HB2	2.023
115	GLN	HB3	1.918
115	GLN	HG2	2.298
115	GLN	HG3	2.298
115	GLN	HE21	7.511
115	GLN	C	175.897
115	GLN	CA	55.995
115	GLN	CB	29.328
115	GLN	CG	33.838
115	GLN	N	121.367
115	GLN	NE2	112.635
116	LYS	H	8.249
116	LYS	HA	4.206
116	LYS	HB2	1.756
116	LYS	HB3	1.688
116	LYS	HG2	1.374

116	LYS	HG3	1.374
116	LYS	HD2	1.616
116	LYS	HD3	1.616
116	LYS	HE2	2.926
116	LYS	HE3	2.926
116	LYS	C	176.281
116	LYS	CA	56.326
116	LYS	CB	33.063
116	LYS	CG	24.705
116	LYS	CD	29.179
116	LYS	CE	42.599
116	LYS	N	122.711
117	ALA	H	8.237
117	ALA	HA	4.263
117	ALA	HB1	1.313
117	ALA	HB2	1.313
117	ALA	HB3	1.313
117	ALA	C	177.84
117	ALA	CA	52.548
117	ALA	CB	19.121
117	ALA	N	125.141
118	VAL	H	8.052
118	VAL	HA	3.992
118	VAL	HB	1.998
118	VAL	HG11	0.864
118	VAL	HG12	0.864
118	VAL	HG13	0.864
118	VAL	HG21	0.864
118	VAL	HG22	0.864
118	VAL	HG23	0.864
118	VAL	C	175.808
118	VAL	CA	62.454
118	VAL	CB	32.845
118	VAL	CG1	20.981
118	VAL	N	119.587
119	GLN	H	8.358
119	GLN	HA	4.257
119	GLN	HB2	2.035
119	GLN	HB3	1.967

119	GLN	HG2	2.303
119	GLN	HG3	2.303
119	GLN	HE21	7.468
119	GLN	HE22	6.811
119	GLN	C	174.405
119	GLN	CA	56.019
119	GLN	CB	29.36
119	GLN	CG	33.792
119	GLN	CD	180.41
119	GLN	N	124.129
119	GLN	NE2	112.419
120	GLN	H	8.422
120	GLN	HA	4.262
120	GLN	HB2	2.032
120	GLN	HB3	1.977
120	GLN	HG2	2.329
120	GLN	HG3	2.329
120	GLN	HE21	7.458
120	GLN	HE22	6.778
120	GLN	C	175.928
120	GLN	CA	56.204
120	GLN	CB	29.475
120	GLN	CG	33.779
120	GLN	CD	180.187
120	GLN	N	122.016
120	GLN	NE2	112.351
121	SER	H	8.322
121	SER	HA	4.366
121	SER	HB2	3.82
121	SER	HB3	3.82
121	SER	C	177.038
121	SER	CA	58.585
121	SER	CB	63.75
121	SER	N	116.821
122	MET	H	8.314
122	MET	HA	4.459
122	MET	HB2	2.075
122	MET	HB3	1.977
122	MET	HG2	2.557

122	MET	HG3	2.499
122	MET	HE1	2.04
122	MET	HE2	2.04
122	MET	HE3	2.04
122	MET	C	175.929
122	MET	CA	55.661
122	MET	CB	32.817
122	MET	CE	16.856
122	MET	N	122.031
123	LEU	H	8.128
123	LEU	HA	4.357
123	LEU	HB2	1.586
123	LEU	HB3	1.586
123	LEU	HD11	0.838
123	LEU	HD12	0.838
123	LEU	HD13	0.838
123	LEU	HD21	0.838
123	LEU	HD22	0.838
123	LEU	HD23	0.838
123	LEU	C	176.122
123	LEU	CA	55.438
123	LEU	CB	42.249
123	LEU	N	123.409
124	SER	H	7.801
124	SER	HA	4.192
124	SER	HB2	3.774
124	SER	HB3	3.774
124	SER	C	178.548
124	SER	CA	59.9
124	SER	CB	64.869
124	SER	N	122.045

Appendix Table A-2 Chemical shift assignments for the nuclei of GIP-Glutaminase L peptide complex (BMRB entry: 17255)

For GIP in bound form:

Residue no.	Amino acid	Nucleus	Chemical shift
1	MET	HE1	1.969
1	MET	HE2	1.969
1	MET	HE3	1.969
1	MET	CE	24.605
2	SER	H	8.157
2	SER	HA	4.341
2	SER	HB2	3.697
2	SER	HB3	3.697
2	SER	CA	58.313
2	SER	CB	64.004
2	SER	N	121.525
3	TYR	H	8.112
3	TYR	HA	4.546
3	TYR	HB2	2.942
3	TYR	HB3	2.82
3	TYR	HD1	6.992
3	TYR	HD2	6.992
3	TYR	CA	57.993
3	TYR	CB	39.021
3	TYR	N	122.397
4	ILE	H	8.009
4	ILE	HA	4.258
4	ILE	HB	1.608
4	ILE	HG12	1.349
4	ILE	HG13	0.973
4	ILE	HG21	0.765
4	ILE	HG22	0.765
4	ILE	HG23	0.765
4	ILE	HD11	0.726
4	ILE	HD12	0.726
4	ILE	HD13	0.726
4	ILE	CA	57.781
4	ILE	CB	39.215
4	ILE	CG1	26.832
4	ILE	CG2	16.826

4	ILE	CD1	12.858
4	ILE	N	127.694
5	PRO	HA	4.164
5	PRO	HB2	2.235
5	PRO	HB3	1.842
5	PRO	HG2	1.956
5	PRO	HG3	1.888
5	PRO	HD2	3.588
5	PRO	HD3	3.558
5	PRO	CA	63.449
5	PRO	CB	31.983
5	PRO	CG	27.372
5	PRO	CD	50.786
6	GLY	H	8.416
6	GLY	HA2	3.773
6	GLY	HA3	3.938
6	GLY	CA	45.191
6	GLY	N	110.047
7	GLN	H	7.972
7	GLN	HA	4.564
7	GLN	HB2	2.016
7	GLN	HB3	1.884
7	GLN	HG2	2.26
7	GLN	HG3	2.26
7	GLN	HE21	7.503
7	GLN	HE22	6.815
7	GLN	CA	53.541
7	GLN	CB	29.191
7	GLN	CG	33.609
7	GLN	N	120.547
7	GLN	NE2	112.516
8	PRO	HA	4.406
8	PRO	HB2	2.21
8	PRO	HB3	1.822
8	PRO	HG2	1.931
8	PRO	HG3	1.959
8	PRO	HD2	3.728
8	PRO	HD3	3.57
8	PRO	CA	63.097
8	PRO	CB	32.022
8	PRO	CG	27.34

8	PRO	CD	50.506
9	VAL	H	8.292
9	VAL	HA	4.167
9	VAL	HB	1.995
9	VAL	HG11	0.874
9	VAL	HG12	0.874
9	VAL	HG13	0.874
9	VAL	HG21	0.846
9	VAL	HG22	0.846
9	VAL	HG23	0.846
9	VAL	CA	62.222
9	VAL	CB	32.996
9	VAL	CG1	20.97
9	VAL	CG2	18.866
9	VAL	N	120.591
10	THR	H	8.24
10	THR	HA	4.31
10	THR	HB	4.144
10	THR	HG21	1.108
10	THR	HG22	1.108
10	THR	HG23	1.108
10	THR	CA	61.631
10	THR	CB	70.009
10	THR	CG2	21.537
10	THR	N	117.745
11	ALA	H	8.052
11	ALA	HA	4.528
11	ALA	HB1	1.242
11	ALA	HB2	1.242
11	ALA	HB3	1.242
11	ALA	CA	51.899
11	ALA	CB	19.927
11	ALA	N	126.513
12	VAL	H	8.498
12	VAL	HA	4.203
12	VAL	HB	1.968
12	VAL	HG11	0.832
12	VAL	HG12	0.832
12	VAL	HG13	0.832
12	VAL	HG21	0.825
12	VAL	HG22	0.825

12	VAL	HG23	0.825
12	VAL	CA	61.544
12	VAL	CB	33.505
12	VAL	CG1	20.661
12	VAL	CG2	21.208
12	VAL	N	120.32
13	VAL	H	8.091
13	VAL	HA	4.846
13	VAL	HB	1.802
13	VAL	HG11	0.742
13	VAL	HG12	0.742
13	VAL	HG13	0.742
13	VAL	HG21	0.803
13	VAL	HG22	0.803
13	VAL	HG23	0.803
13	VAL	CA	60.976
13	VAL	CB	33.285
13	VAL	CG1	21.44
13	VAL	CG2	20.847
13	VAL	N	123.775
14	GLN	H	9.018
14	GLN	HA	4.587
14	GLN	HB2	1.556
14	GLN	HB3	1.764
14	GLN	HG2	2.043
14	GLN	HG3	1.99
14	GLN	HE21	7.109
14	GLN	HE22	6.723
14	GLN	CA	53.958
14	GLN	CB	31.895
14	GLN	CG	32.911
14	GLN	N	124.782
14	GLN	NE2	111.111
15	ARG	H	8.569
15	ARG	HA	5.057
15	ARG	HB2	1.715
15	ARG	HB3	1.584
15	ARG	HG2	1.474
15	ARG	HG3	1.39
15	ARG	HD2	3.062
15	ARG	HD3	3.062

15	ARG	CA	55.234
15	ARG	CB	31.427
15	ARG	CG	27.65
15	ARG	CD	43.281
15	ARG	N	123.905
16	VAL	H	8.669
16	VAL	HA	4.286
16	VAL	HB	1.632
16	VAL	HG11	0.578
16	VAL	HG12	0.578
16	VAL	HG13	0.578
16	VAL	HG21	0.624
16	VAL	HG22	0.624
16	VAL	HG23	0.624
16	VAL	CA	61.203
16	VAL	CB	35.308
16	VAL	CG1	20.648
16	VAL	CG2	21.48
16	VAL	N	124.788
17	GLU	H	9.037
17	GLU	HA	5.059
17	GLU	HB2	1.868
17	GLU	HB3	1.868
17	GLU	HG2	1.881
17	GLU	HG3	1.755
17	GLU	CA	55.121
17	GLU	CB	31.717
17	GLU	CG	37.584
17	GLU	N	129.823
18	ILE	H	8.919
18	ILE	HA	4.252
18	ILE	HB	1.539
18	ILE	HG12	1.447
18	ILE	HG13	1.447
18	ILE	HG21	0.751
18	ILE	HG22	0.751
18	ILE	HG23	0.751
18	ILE	HD11	0.705
18	ILE	HD12	0.705
18	ILE	HD13	0.705
18	ILE	CA	60.507

18	ILE	CB	40.565
18	ILE	CG1	27.611
18	ILE	CG2	19.662
18	ILE	CD1	15.163
18	ILE	N	123.907
19	HIS	H	8.807
19	HIS	HA	4.8
19	HIS	HB2	3.161
19	HIS	HB3	3.109
19	HIS	HD2	7.191
19	HIS	CA	54.686
19	HIS	CB	28.869
19	HIS	N	127.764
20	LYS	H	8.579
20	LYS	HA	4.086
20	LYS	HB2	1.556
20	LYS	HB3	1.556
20	LYS	HG2	1.533
20	LYS	HG3	1.533
20	LYS	HD2	1.005
20	LYS	HD3	1.005
20	LYS	HE2	3.135
20	LYS	HE3	3.135
20	LYS	CA	57.803
20	LYS	CB	35.415
20	LYS	CG	27.333
20	LYS	CD	27.562
20	LYS	N	122.931
21	LEU	H	8.855
21	LEU	HA	4.556
21	LEU	HB2	1.604
21	LEU	HB3	1.452
21	LEU	HG	1.378
21	LEU	HD11	0.907
21	LEU	HD12	0.907
21	LEU	HD13	0.907
21	LEU	HD21	0.832
21	LEU	HD22	0.832
21	LEU	HD23	0.832
21	LEU	CA	53.47
21	LEU	CB	45.533

21	LEU	CG	26.821
21	LEU	CD1	23.627
21	LEU	CD2	25.494
21	LEU	N	123.076
22	ARG	H	8.626
22	ARG	HA	4.671
22	ARG	HB2	1.753
22	ARG	HB3	1.592
22	ARG	HG2	1.449
22	ARG	HG3	1.449
22	ARG	HD2	3.145
22	ARG	HD3	3.064
22	ARG	HE	7.278
22	ARG	CA	56.081
22	ARG	CB	30.275
22	ARG	CG	27.38
22	ARG	CD	42.829
22	ARG	N	126.534
22	ARG	NE	84.017
23	GLN	H	8.628
23	GLN	HA	4.446
23	GLN	HB2	1.907
23	GLN	HB3	1.695
23	GLN	HG2	2.097
23	GLN	HG3	2.097
23	GLN	HE21	7.439
23	GLN	HE22	6.776
23	GLN	CA	54.76
23	GLN	CB	30.51
23	GLN	CG	33.866
23	GLN	N	129.278
23	GLN	NE2	111.736
24	GLY	H	8.969
24	GLY	HA2	3.962
24	GLY	HA3	3.573
24	GLY	CA	46.96
24	GLY	N	117.183
25	GLU	H	9.038
25	GLU	HA	4.214
25	GLU	HB2	1.772
25	GLU	HB3	1.772

25	GLU	HG2	2.214
25	GLU	HG3	2.141
25	GLU	CA	56.392
25	GLU	CB	30.001
25	GLU	CG	36.165
25	GLU	N	126.285
26	ASN	H	7.931
26	ASN	HA	4.823
26	ASN	HB2	2.762
26	ASN	HB3	2.644
26	ASN	HD21	7.628
26	ASN	HD22	6.983
26	ASN	CA	52.39
26	ASN	CB	41.24
26	ASN	N	117.868
26	ASN	ND2	114.311
27	LEU	H	8.284
27	LEU	HA	4.949
27	LEU	HB2	1.479
27	LEU	HB3	1.26
27	LEU	HG	1.38
27	LEU	HD11	1.288
27	LEU	HD12	1.288
27	LEU	HD13	1.288
27	LEU	HD21	0.739
27	LEU	HD22	0.739
27	LEU	HD23	0.739
27	LEU	CA	54.385
27	LEU	CB	44.516
27	LEU	CG	27.206
27	LEU	CD1	24.553
27	LEU	CD2	25.63
27	LEU	N	123.213
28	ILE	H	8.902
28	ILE	HA	4.628
28	ILE	HB	1.904
28	ILE	HG12	1.382
28	ILE	HG13	1.009
28	ILE	HG21	0.763
28	ILE	HG22	0.763
28	ILE	HG23	0.763

28	ILE	HD11	0.743
28	ILE	HD12	0.743
28	ILE	HD13	0.743
28	ILE	CA	59.684
28	ILE	CB	42.522
28	ILE	CG1	26.678
28	ILE	CG2	18.814
28	ILE	CD1	13.632
28	ILE	N	119.351
29	LEU	H	10.951
29	LEU	HA	4.332
29	LEU	HB2	1.607
29	LEU	HB3	1.607
29	LEU	HG	1.501
29	LEU	HD11	0.89
29	LEU	HD12	0.89
29	LEU	HD13	0.89
29	LEU	HD21	0.758
29	LEU	HD22	0.758
29	LEU	HD23	0.758
29	LEU	CA	55.691
29	LEU	CB	44.463
29	LEU	CG	27.32
29	LEU	CD1	27.183
29	LEU	CD2	27.305
29	LEU	N	124.679
30	GLY	H	9.334
30	GLY	HA2	4.006
30	GLY	HA3	4.154
30	GLY	CA	46.146
30	GLY	N	107.432
31	PHE	H	7.408
31	PHE	HA	5.002
31	PHE	HB2	2.885
31	PHE	HB3	3.647
31	PHE	CA	56.653
31	PHE	CB	39.742
31	PHE	N	117.124
32	SER	H	8.63
32	SER	HA	5.814
32	SER	HB2	3.558

32	SER	HB3	3.558
32	SER	CA	56.286
32	SER	CB	65.93
32	SER	N	112.923
33	ILE	H	8.694
33	ILE	HA	5.781
33	ILE	HB	1.692
33	ILE	HG12	1.387
33	ILE	HG13	1.387
33	ILE	HG21	0.749
33	ILE	HG22	0.749
33	ILE	HG23	0.749
33	ILE	HD11	0.191
33	ILE	HD12	0.191
33	ILE	HD13	0.191
33	ILE	CA	58.253
33	ILE	CB	43.397
33	ILE	CG1	26.682
33	ILE	CG2	20.174
33	ILE	CD1	13.604
33	ILE	N	113.209
34	GLY	H	9.05
34	GLY	HA2	3.776
34	GLY	HA3	4.817
34	GLY	CA	43.545
34	GLY	N	108.49
35	GLY	H	9.485
35	GLY	HA2	3.847
35	GLY	HA3	5.379
35	GLY	CA	43.984
35	GLY	N	107.605
36	GLY	H	6.608
36	GLY	HA2	3.947
36	GLY	HA3	4.696
36	GLY	CA	43.893
36	GLY	N	105.396
37	ILE	H	8.581
37	ILE	HA	4.151
37	ILE	HB	2.043
37	ILE	HG12	1.232
37	ILE	HG13	1.272

37	ILE	HG21	0.95
37	ILE	HG22	0.95
37	ILE	HG23	0.95
37	ILE	HD11	0.748
37	ILE	HD12	0.748
37	ILE	HD13	0.748
37	ILE	CA	64.826
37	ILE	CB	37.411
37	ILE	CG1	26.094
37	ILE	CG2	17.894
37	ILE	CD1	13.718
37	ILE	N	115.311
38	ASP	H	9.911
38	ASP	HA	4.615
38	ASP	HB2	2.902
38	ASP	HB3	2.571
38	ASP	CA	52.607
38	ASP	CB	40.145
38	ASP	N	118.124
39	GLN	H	7.504
39	GLN	HA	4.414
39	GLN	HB2	1.938
39	GLN	HB3	1.938
39	GLN	HG2	2.333
39	GLN	HG3	2.063
39	GLN	HE21	6.979
39	GLN	HE22	6.939
39	GLN	CA	53.483
39	GLN	CB	30.343
39	GLN	CG	34.236
39	GLN	N	119.958
39	GLN	NE2	115.609
40	ASP	H	8.562
40	ASP	HA	4.932
40	ASP	HB2	2.618
40	ASP	HB3	2.952
40	ASP	CA	51.198
40	ASP	CB	41.257
40	ASP	N	121.473
41	PRO	HA	4.535
41	PRO	HB2	2.163

41	PRO	HB3	1.837
41	PRO	HG2	2.03
41	PRO	HG3	1.875
41	PRO	HD2	4.121
41	PRO	HD3	4.059
41	PRO	CA	64.219
41	PRO	CB	31.64
41	PRO	CG	26.831
41	PRO	CD	50.882
42	SER	H	8.294
42	SER	HA	4.133
42	SER	HB2	3.9
42	SER	HB3	3.9
42	SER	CA	61.232
42	SER	CB	63.076
42	SER	N	115.571
43	GLN	H	7.289
43	GLN	HA	4.202
43	GLN	HB2	2.306
43	GLN	HB3	1.709
43	GLN	HG2	2.185
43	GLN	HG3	2.11
43	GLN	HE21	7.572
43	GLN	HE22	6.853
43	GLN	CA	54.826
43	GLN	CB	29.029
43	GLN	CG	33.862
43	GLN	N	118.268
43	GLN	NE2	113.059
44	ASN	H	7.07
44	ASN	HA	4.677
44	ASN	HB2	2.674
44	ASN	HB3	3.098
44	ASN	CA	50.103
44	ASN	CB	38.917
44	ASN	N	119.825
45	PRO	HA	4.114
45	PRO	HB2	1.846
45	PRO	HB3	0.988
45	PRO	HG2	1.542
45	PRO	HG3	0.804

45	PRO	HD2	3.933
45	PRO	HD3	3.416
45	PRO	CA	63.395
45	PRO	CB	31.896
45	PRO	CG	25.705
45	PRO	CD	50.501
46	PHE	H	7.635
46	PHE	HA	4.364
46	PHE	HB2	2.311
46	PHE	HB3	3.205
46	PHE	CA	58.119
46	PHE	CB	40.103
46	PHE	N	115.588
47	SER	H	6.711
47	SER	HA	4.205
47	SER	HB2	3.854
47	SER	HB3	3.43
47	SER	CA	56.518
47	SER	CB	63.853
47	SER	N	111.81
48	GLU	H	8.761
48	GLU	HA	4.117
48	GLU	HB2	1.977
48	GLU	HB3	1.938
48	GLU	HG2	2.262
48	GLU	HG3	2.209
48	GLU	CA	57.831
48	GLU	CB	30.398
48	GLU	CG	36.449
48	GLU	N	124.099
49	ASP	H	7.967
49	ASP	HA	4.571
49	ASP	HB2	2.872
49	ASP	HB3	2.6
49	ASP	CA	52.816
49	ASP	CB	41.614
49	ASP	N	118.167
50	LYS	H	8.452
50	LYS	HA	4.195
50	LYS	HB2	1.857
50	LYS	HB3	1.857

50	LYS	HG2	1.377
50	LYS	HG3	1.377
50	LYS	HD2	1.415
50	LYS	HD3	1.415
50	LYS	CA	56.273
50	LYS	CB	29.975
50	LYS	CG	24.015
50	LYS	CD	27.99
50	LYS	CE	42.441
50	LYS	N	120.392
51	THR	H	8.326
51	THR	HA	4.308
51	THR	HB	4.314
51	THR	HG21	1.093
51	THR	HG22	1.093
51	THR	HG23	1.093
51	THR	CA	61.561
51	THR	CB	69.861
51	THR	CG2	21.495
51	THR	N	108.193
52	ASP	H	7.642
52	ASP	HA	4.459
52	ASP	HB2	3.125
52	ASP	HB3	2.788
52	ASP	CA	54.751
52	ASP	CB	41.837
52	ASP	N	122.292
53	LYS	H	8.844
53	LYS	HA	4.65
53	LYS	HB2	2.183
53	LYS	HB3	1.746
53	LYS	HG2	1.33
53	LYS	HG3	1.33
53	LYS	HD2	1.615
53	LYS	HD3	1.615
53	LYS	HE2	3.103
53	LYS	HE3	3.103
53	LYS	CA	56.489
53	LYS	CB	32.362
53	LYS	CG	25.237
53	LYS	CD	29.771

53	LYS	CE	42.225
53	LYS	N	129.965
54	GLY	H	8.819
54	GLY	HA2	4.168
54	GLY	HA3	3.555
54	GLY	CA	45.275
54	GLY	N	106.511
55	ILE	H	9.223
55	ILE	HA	4.748
55	ILE	HB	2.08
55	ILE	HG12	1.043
55	ILE	HG13	1.043
55	ILE	HG21	0.693
55	ILE	HG22	0.693
55	ILE	HG23	0.693
55	ILE	HD11	0.399
55	ILE	HD12	0.399
55	ILE	HD13	0.399
55	ILE	CA	57.559
55	ILE	CB	36.017
55	ILE	CG2	18.011
55	ILE	CD1	8.017
55	ILE	N	120.313
56	TYR	H	8.952
56	TYR	HA	5.306
56	TYR	HB2	2.53
56	TYR	HB3	2.44
56	TYR	HD1	6.888
56	TYR	HD2	6.888
56	TYR	CA	55.628
56	TYR	CB	42.675
56	TYR	N	125.445
57	VAL	H	8.656
57	VAL	HA	4.502
57	VAL	HB	2.07
57	VAL	HG11	0.626
57	VAL	HG12	0.626
57	VAL	HG13	0.626
57	VAL	HG21	0.632
57	VAL	HG22	0.632
57	VAL	HG23	0.632

57	VAL	CA	62.146
57	VAL	CB	31.995
57	VAL	CG1	21.59
57	VAL	CG2	22.338
57	VAL	N	121.052
58	THR	H	8.719
58	THR	HA	4.307
58	THR	HB	4.247
58	THR	HG21	1.104
58	THR	HG22	1.104
58	THR	HG23	1.104
58	THR	CA	62.326
58	THR	CB	68.681
58	THR	CG2	22.48
58	THR	N	119.551
59	ARG	H	7.23
59	ARG	HA	4.302
59	ARG	HB2	1.66
59	ARG	HB3	1.64
59	ARG	HG2	1.546
59	ARG	HG3	1.593
59	ARG	HD2	3.114
59	ARG	HD3	2.951
59	ARG	HE	7.493
59	ARG	CA	55.987
59	ARG	CB	34.173
59	ARG	CG	27.264
59	ARG	CD	43.391
59	ARG	N	119.022
59	ARG	NE	85.721
60	VAL	H	8.428
60	VAL	HA	4.198
60	VAL	HB	1.805
60	VAL	HG11	0.746
60	VAL	HG12	0.746
60	VAL	HG13	0.746
60	VAL	HG21	0.493
60	VAL	HG22	0.493
60	VAL	HG23	0.493
60	VAL	CA	62.372
60	VAL	CB	34.958

60	VAL	CG1	21.47
60	VAL	CG2	21.665
60	VAL	N	121.675
61	SER	H	8.124
61	SER	HA	4.209
61	SER	HB2	3.797
61	SER	HB3	3.617
61	SER	CA	59.379
61	SER	CB	62.898
61	SER	N	122.728
62	GLU	H	9.285
62	GLU	HA	4.137
62	GLU	HB2	1.955
62	GLU	HB3	1.955
62	GLU	HG2	2.358
62	GLU	HG3	2.219
62	GLU	CA	58.156
62	GLU	CB	29.133
62	GLU	CG	36.177
62	GLU	N	130.89
63	GLY	H	9.626
63	GLY	HA2	3.651
63	GLY	HA3	4.03
63	GLY	CA	45.414
63	GLY	N	116.325
64	GLY	H	7.537
64	GLY	HA2	3.88
64	GLY	HA3	4.362
64	GLY	CA	45.295
64	GLY	N	106.344
65	PRO	HA	4.152
65	PRO	HB2	2.623
65	PRO	HB3	2.043
65	PRO	HG2	2.229
65	PRO	HG3	2.229
65	PRO	HD2	3.681
65	PRO	HD3	3.283
65	PRO	CA	65.246
65	PRO	CB	32.125
65	PRO	CG	28.746
65	PRO	CD	48.835

66	ALA	H	7.668
66	ALA	HA	3.864
66	ALA	HB1	1.405
66	ALA	HB2	1.405
66	ALA	HB3	1.405
66	ALA	CA	55.058
66	ALA	CB	19.158
66	ALA	N	120.001
67	GLU	H	8.483
67	GLU	HA	3.797
67	GLU	HB2	2.13
67	GLU	HB3	2.13
67	GLU	HG2	2.173
67	GLU	HG3	2.029
67	GLU	CA	59.972
67	GLU	CB	29.842
67	GLU	CG	37.015
67	GLU	N	121.308
68	ILE	H	8.068
68	ILE	HA	3.656
68	ILE	HB	1.814
68	ILE	HG12	1.566
68	ILE	HG13	1.124
68	ILE	HG21	0.878
68	ILE	HG22	0.878
68	ILE	HG23	0.878
68	ILE	HD11	0.725
68	ILE	HD12	0.725
68	ILE	HD13	0.725
68	ILE	CA	64.302
68	ILE	CB	38.187
68	ILE	CG1	29.075
68	ILE	CG2	17.005
68	ILE	CD1	13.073
68	ILE	N	120.686
69	ALA	H	7.584
69	ALA	HA	4.262
69	ALA	HB1	1.433
69	ALA	HB2	1.433
69	ALA	HB3	1.433
69	ALA	CA	52.936

69	ALA	CB	21.118
69	ALA	N	119.667
70	GLY	H	7.533
70	GLY	HA2	3.675
70	GLY	HA3	4.246
70	GLY	CA	44.925
70	GLY	N	103.682
71	LEU	H	7.806
71	LEU	HA	3.696
71	LEU	HB2	1.23
71	LEU	HB3	0.723
71	LEU	HG	0.578
71	LEU	HD11	-0.426
71	LEU	HD12	-0.426
71	LEU	HD13	-0.426
71	LEU	HD21	0.532
71	LEU	HD22	0.532
71	LEU	HD23	0.532
71	LEU	CA	55.139
71	LEU	CB	44.142
71	LEU	CG	26.68
71	LEU	CD1	26.063
71	LEU	CD2	23.966
71	LEU	N	123.765
72	GLN	H	8.356
72	GLN	HA	4.535
72	GLN	HB2	1.86
72	GLN	HB3	1.783
72	GLN	HG2	2.157
72	GLN	HG3	2.157
72	GLN	HE21	7.241
72	GLN	HE22	6.668
72	GLN	CA	54.032
72	GLN	CB	31.972
72	GLN	CG	33.47
72	GLN	N	123.599
72	GLN	NE2	110.957
73	ILE	H	8.313
73	ILE	HA	3.152
73	ILE	HB	1.496
73	ILE	HG12	1.444

73	ILE	HG13	1.444
73	ILE	HG21	0.724
73	ILE	HG22	0.724
73	ILE	HG23	0.724
73	ILE	HD11	0.956
73	ILE	HD12	0.956
73	ILE	HD13	0.956
73	ILE	CA	63.642
73	ILE	CB	38.011
73	ILE	CG1	28.791
73	ILE	CG2	17.827
73	ILE	CD1	14.033
73	ILE	N	119.802
74	GLY	H	9.063
74	GLY	HA2	2.602
74	GLY	HA3	3.913
74	GLY	CA	44.751
74	GLY	N	117.216
75	ASP	H	7.614
75	ASP	HA	4.407
75	ASP	HB2	2.396
75	ASP	HB3	1.943
75	ASP	CA	55.759
75	ASP	CB	40.474
75	ASP	N	122.366
76	LYS	H	8.322
76	LYS	HA	4.292
76	LYS	HB2	2.17
76	LYS	HB3	2.17
76	LYS	HG2	1.923
76	LYS	HG3	1.923
76	LYS	HD2	1.597
76	LYS	HD3	1.597
76	LYS	HE2	2.505
76	LYS	HE3	2.505
76	LYS	CA	54.396
76	LYS	CB	33.989
76	LYS	CG	25.24
76	LYS	CE	40.262
76	LYS	N	122.798
77	ILE	H	8.673

77	ILE	HA	3.799
77	ILE	HB	1.506
77	ILE	HG12	1.18
77	ILE	HG13	1.18
77	ILE	HG21	0.664
77	ILE	HG22	0.664
77	ILE	HG23	0.664
77	ILE	HD11	0.68
77	ILE	HD12	0.68
77	ILE	HD13	0.68
77	ILE	CA	62.126
77	ILE	CB	37.944
77	ILE	CG2	19.388
77	ILE	CD1	13.628
77	ILE	N	125.785
78	MET	H	9.09
78	MET	HA	4.425
78	MET	HB2	1.649
78	MET	HB3	1.649
78	MET	HG2	2.328
78	MET	HG3	2.328
78	MET	HE1	1.887
78	MET	HE2	1.887
78	MET	HE3	1.887
78	MET	CA	55.478
78	MET	CB	32.441
78	MET	CG	31.386
78	MET	CE	16.089
78	MET	N	125.255
79	GLN	H	7.607
79	GLN	HA	5.225
79	GLN	HB2	1.753
79	GLN	HB3	1.784
79	GLN	HG2	2.074
79	GLN	HG3	2.028
79	GLN	HE21	7.581
79	GLN	HE22	6.681
79	GLN	CA	54.835
79	GLN	CB	35.024
79	GLN	CG	34.401
79	GLN	N	116.394

79	GLN	NE2	111.454
80	VAL	H	8.497
80	VAL	HA	4.511
80	VAL	HB	1.956
80	VAL	HG11	0.826
80	VAL	HG12	0.826
80	VAL	HG13	0.826
80	VAL	HG21	0.839
80	VAL	HG22	0.839
80	VAL	HG23	0.839
80	VAL	CA	60.588
80	VAL	CB	34.915
80	VAL	CG1	22.066
80	VAL	CG2	21.302
80	VAL	N	120.39
81	ASN	H	9.572
81	ASN	HA	4.474
81	ASN	HB2	3.238
81	ASN	HB3	3
81	ASN	HD21	7.4
81	ASN	HD22	6.577
81	ASN	CA	54.337
81	ASN	CB	36.692
81	ASN	N	127.486
81	ASN	ND2	110.358
82	GLY	H	8.479
82	GLY	HA2	3.439
82	GLY	HA3	4.006
82	GLY	CA	45.208
82	GLY	N	102.662
83	TRP	H	8.355
83	TRP	HA	4.547
83	TRP	HB2	3.305
83	TRP	HB3	3.172
83	TRP	HD1	7.338
83	TRP	HE1	10.127
83	TRP	HE3	7.552
83	TRP	HZ2	7.46
83	TRP	HZ3	7.063
83	TRP	HH2	7.167
83	TRP	CA	57.081

83	TRP	CB	29.505
83	TRP	N	123.595
83	TRP	NE1	129.833
84	ASP	H	8.486
84	ASP	HA	4.397
84	ASP	HB2	2.642
84	ASP	HB3	2.573
84	ASP	CA	55.886
84	ASP	CB	41.823
84	ASP	N	126.297
85	MET	H	8.086
85	MET	HA	4.647
85	MET	HB2	2.142
85	MET	HB3	1.611
85	MET	HG2	2.614
85	MET	HG3	2.261
85	MET	HE1	1.931
85	MET	HE2	1.931
85	MET	HE3	1.931
85	MET	CA	54.125
85	MET	CB	33.489
85	MET	CG	33.162
85	MET	CE	18.415
85	MET	N	123.789
86	THR	H	8.577
86	THR	HA	4.202
86	THR	HB	4.162
86	THR	HG21	1.35
86	THR	HG22	1.35
86	THR	HG23	1.35
86	THR	CA	64.971
86	THR	CB	69.76
86	THR	CG2	22.199
86	THR	N	115.359
87	MET	H	8.456
87	MET	HA	4.605
87	MET	HB2	2.076
87	MET	HB3	1.8
87	MET	HG2	2.481
87	MET	HG3	2.397
87	MET	HE1	2.102

87	MET	HE2	2.102
87	MET	HE3	2.102
87	MET	CA	54.189
87	MET	CB	32.485
87	MET	CG	32.224
87	MET	CE	22.95
87	MET	N	123.097
88	VAL	H	7.846
88	VAL	HA	4.854
88	VAL	HB	2.38
88	VAL	HG11	0.857
88	VAL	HG12	0.857
88	VAL	HG13	0.857
88	VAL	HG21	0.872
88	VAL	HG22	0.872
88	VAL	HG23	0.872
88	VAL	CA	59.028
88	VAL	CB	34.933
88	VAL	CG1	18.871
88	VAL	CG2	22.427
88	VAL	N	113.068
89	THR	H	8.567
89	THR	HA	4.33
89	THR	HB	4.598
89	THR	HG21	1.207
89	THR	HG22	1.207
89	THR	HG23	1.207
89	THR	CA	61.359
89	THR	CB	70.8
89	THR	CG2	21.994
89	THR	N	112.758
90	HIS	H	10.017
90	HIS	HA	3.776
90	HIS	HB2	3.453
90	HIS	HB3	3.263
90	HIS	HD2	6.949
90	HIS	CA	61.727
90	HIS	CB	28.469
90	HIS	N	122.939
91	ASP	H	9.26
91	ASP	HA	4.286

91	ASP	HB2	2.65
91	ASP	HB3	2.346
91	ASP	CA	57.504
91	ASP	CB	41.965
91	ASP	N	115.47
92	GLN	H	7.792
92	GLN	HA	3.777
92	GLN	HB2	1.871
92	GLN	HB3	1.871
92	GLN	HG2	2.423
92	GLN	HG3	2.385
92	GLN	HE21	7.479
92	GLN	HE22	6.88
92	GLN	CA	59.038
92	GLN	CB	29.071
92	GLN	CG	34.912
92	GLN	N	117.704
92	GLN	NE2	111.261
93	ALA	H	7.952
93	ALA	HA	3.808
93	ALA	HB1	1.249
93	ALA	HB2	1.249
93	ALA	HB3	1.249
93	ALA	CA	55.281
93	ALA	CB	19.046
93	ALA	N	121.954
94	ARG	H	8.233
94	ARG	HA	3.396
94	ARG	HB2	1.876
94	ARG	HB3	1.876
94	ARG	HG2	1.524
94	ARG	HG3	1.351
94	ARG	HD2	3.329
94	ARG	HD3	3.133
94	ARG	HE	7.373
94	ARG	CA	60.056
94	ARG	CB	30.159
94	ARG	CG	26.791
94	ARG	CD	43.024
94	ARG	N	117.346
94	ARG	NE	82.115

95	LYS	H	8.53
95	LYS	HA	3.632
95	LYS	HB2	1.721
95	LYS	HB3	1.621
95	LYS	HG2	1.469
95	LYS	HG3	1.233
95	LYS	HD2	1.521
95	LYS	HD3	1.521
95	LYS	HE2	2.827
95	LYS	HE3	2.827
95	LYS	CA	59.739
95	LYS	CB	32.195
95	LYS	CG	26.033
95	LYS	CD	29.112
95	LYS	CE	41.89
95	LYS	N	119.979
96	ARG	H	7.587
96	ARG	HA	3.842
96	ARG	HB2	1.664
96	ARG	HB3	1.58
96	ARG	HG2	1.488
96	ARG	HG3	1.298
96	ARG	HD2	2.498
96	ARG	HD3	2.205
96	ARG	CA	57.556
96	ARG	CB	29.221
96	ARG	CG	26.247
96	ARG	CD	42.228
96	ARG	N	117.105
97	LEU	H	7.519
97	LEU	HA	3.957
97	LEU	HB2	1.739
97	LEU	HB3	1.669
97	LEU	HG	1.26
97	LEU	HD11	0.865
97	LEU	HD12	0.865
97	LEU	HD13	0.865
97	LEU	HD21	0.682
97	LEU	HD22	0.682
97	LEU	HD23	0.682
97	LEU	CA	56.692

97	LEU	CB	42.796
97	LEU	CG	27.379
97	LEU	CD1	24.261
97	LEU	CD2	26.302
97	LEU	N	117.133
98	THR	H	7.131
98	THR	HA	4.681
98	THR	HB	4.148
98	THR	HG21	1.174
98	THR	HG22	1.174
98	THR	HG23	1.174
98	THR	CA	60.287
98	THR	CB	69.202
98	THR	CG2	21.366
98	THR	N	105.101
99	LYS	H	6.889
99	LYS	HA	4.131
99	LYS	HB2	1.74
99	LYS	HB3	1.74
99	LYS	HG2	1.463
99	LYS	HG3	1.257
99	LYS	HD2	1.592
99	LYS	HD3	1.592
99	LYS	HE2	2.763
99	LYS	HE3	2.763
99	LYS	CA	57.633
99	LYS	CB	32.011
99	LYS	CG	24.723
99	LYS	CD	29.327
99	LYS	CE	41.653
99	LYS	N	123.891
100	ARG	H	8.682
100	ARG	HA	3.876
100	ARG	HB2	1.814
100	ARG	HB3	1.814
100	ARG	HG2	1.725
100	ARG	HG3	1.725
100	ARG	HD2	3.212
100	ARG	HD3	3.212
100	ARG	CA	58.589
100	ARG	CB	30.071

100	ARG	CG	27.562
100	ARG	CD	43.322
100	ARG	N	128.32
101	SER	H	7.55
101	SER	HA	4.178
101	SER	HB2	4.023
101	SER	HB3	3.685
101	SER	CA	58.202
101	SER	CB	62.929
101	SER	N	108.729
102	GLU	H	7.224
102	GLU	HA	4.474
102	GLU	HB2	1.715
102	GLU	HB3	1.715
102	GLU	HG2	2.059
102	GLU	HG3	2.059
102	GLU	CA	54.703
102	GLU	CB	31.086
102	GLU	CG	36.134
102	GLU	N	121.834
103	GLU	H	8.675
103	GLU	HA	3.966
103	GLU	HB2	1.893
103	GLU	HB3	1.893
103	GLU	HG2	2.229
103	GLU	HG3	2.103
103	GLU	CA	57.83
103	GLU	CB	30.546
103	GLU	CG	36.967
103	GLU	N	122.965
104	VAL	H	7.516
104	VAL	HA	4.663
104	VAL	HB	1.648
104	VAL	HG11	0.438
104	VAL	HG12	0.438
104	VAL	HG13	0.438
104	VAL	HG21	0.509
104	VAL	HG22	0.509
104	VAL	HG23	0.509
104	VAL	CA	60.264
104	VAL	CB	34.457

104	VAL	CG1	21.078
104	VAL	CG2	20.453
104	VAL	N	117.462
105	VAL	H	8.621
105	VAL	HA	4.636
105	VAL	HB	1.769
105	VAL	HG11	0.482
105	VAL	HG12	0.482
105	VAL	HG13	0.482
105	VAL	HG21	0.86
105	VAL	HG22	0.86
105	VAL	HG23	0.86
105	VAL	CA	59.716
105	VAL	CB	34.866
105	VAL	CG2	22.253
105	VAL	N	120.82
106	ARG	H	8.765
106	ARG	HA	4.748
106	ARG	HB2	1.728
106	ARG	HB3	1.777
106	ARG	HG2	1.461
106	ARG	HG3	1.407
106	ARG	HD2	3.076
106	ARG	HD3	3.076
106	ARG	CA	55.449
106	ARG	CB	30.06
106	ARG	CG	28.171
106	ARG	CD	43.238
106	ARG	N	125.489
107	LEU	H	9.46
107	LEU	HA	5.212
107	LEU	HB2	1.186
107	LEU	HB3	1.586
107	LEU	HD11	0.756
107	LEU	HD12	0.756
107	LEU	HD13	0.756
107	LEU	HD21	0.584
107	LEU	HD22	0.584
107	LEU	HD23	0.584
107	LEU	CA	53.216
107	LEU	CB	44.49

107	LEU	CG	26.967
107	LEU	CD1	25.593
107	LEU	CD2	26.876
107	LEU	N	125.4
108	LEU	H	7.88
108	LEU	HA	5.019
108	LEU	HB2	1.416
108	LEU	HB3	1.333
108	LEU	HG	1.203
108	LEU	HD11	0.693
108	LEU	HD12	0.693
108	LEU	HD13	0.693
108	LEU	HD21	0.67
108	LEU	HD22	0.67
108	LEU	HD23	0.67
108	LEU	CA	54.524
108	LEU	CB	44.27
108	LEU	CG	26.993
108	LEU	CD1	24.317
108	LEU	CD2	24.312
108	LEU	N	123.53
109	VAL	H	9.071
109	VAL	HA	5.548
109	VAL	HB	1.817
109	VAL	HG11	0.541
109	VAL	HG12	0.541
109	VAL	HG13	0.541
109	VAL	HG21	0.577
109	VAL	HG22	0.577
109	VAL	HG23	0.577
109	VAL	CA	57.915
109	VAL	CB	35.182
109	VAL	CG1	18.294
109	VAL	CG2	21.28
109	VAL	N	119.5
110	THR	H	8.905
110	THR	HA	5.085
110	THR	HB	3.9
110	THR	HG21	1.045
110	THR	HG22	1.045
110	THR	HG23	1.045

110	THR	CA	60.651
110	THR	CB	70.979
110	THR	CG2	21.734
110	THR	N	114.109
111	ARG	H	8.784
111	ARG	HA	4.751
111	ARG	HB2	1.842
111	ARG	HB3	1.842
111	ARG	HG2	1.467
111	ARG	HG3	1.409
111	ARG	HD2	2.97
111	ARG	HD3	2.903
111	ARG	HE	7.214
111	ARG	CA	54.668
111	ARG	CB	33.387
111	ARG	CG	25.81
111	ARG	CD	43.363
111	ARG	N	126.316
111	ARG	NE	84.728
112	GLN	H	8.861
112	GLN	HA	4.414
112	GLN	HB2	2.034
112	GLN	HB3	1.986
112	GLN	HG2	2.406
112	GLN	HG3	2.348
112	GLN	HE21	7.594
112	GLN	HE22	6.857
112	GLN	CA	55.824
112	GLN	CB	29.451
112	GLN	CG	33.858
112	GLN	N	123.79
112	GLN	NE2	111.755
113	SER	H	8.501
113	SER	HA	4.366
113	SER	HB2	3.751
113	SER	HB3	3.751
113	SER	CA	58.268
113	SER	CB	63.893
113	SER	N	117.088
114	LEU	H	8.218
114	LEU	HA	4.324

114	LEU	HB2	1.539
114	LEU	HB3	1.539
114	LEU	HG	1.451
114	LEU	HD11	0.802
114	LEU	HD12	0.802
114	LEU	HD13	0.802
114	LEU	HD21	0.845
114	LEU	HD22	0.845
114	LEU	HD23	0.845
114	LEU	CA	55.253
114	LEU	CB	42.257
114	LEU	CD1	23.667
114	LEU	CD2	24.728
114	LEU	N	124.389
115	GLN	H	8.291
115	GLN	HA	4.206
115	GLN	HB2	2.012
115	GLN	HB3	1.881
115	GLN	HG2	2.277
115	GLN	HG3	2.277
115	GLN	CA	56.065
115	GLN	CB	29.517
115	GLN	CG	33.738
115	GLN	N	121.34
116	LYS	H	8.239
116	LYS	HA	4.19
116	LYS	HB2	1.72
116	LYS	HB3	1.665
116	LYS	HG2	1.359
116	LYS	HG3	1.359
116	LYS	HD2	1.602
116	LYS	HD3	1.602
116	LYS	HE2	2.913
116	LYS	HE3	2.913
116	LYS	CA	56.336
116	LYS	CB	33.131
116	LYS	CG	24.747
116	LYS	CD	29.078
116	LYS	CE	42.129
116	LYS	N	122.701
117	ALA	H	8.221

117	ALA	HA	4.238
117	ALA	HB1	1.304
117	ALA	HB2	1.304
117	ALA	HB3	1.304
117	ALA	CA	52.538
117	ALA	CB	19.14
117	ALA	N	125.024
118	VAL	H	8.044
118	VAL	HA	3.976
118	VAL	HB	1.979
118	VAL	HG11	0.85
118	VAL	HG12	0.85
118	VAL	HG13	0.85
118	VAL	HG21	0.865
118	VAL	HG22	0.865
118	VAL	HG23	0.865
118	VAL	CA	62.506
118	VAL	CB	32.914
118	VAL	CG1	20.843
118	VAL	CG2	21.131
118	VAL	N	119.586
119	GLN	H	8.351
119	GLN	HA	4.242
119	GLN	HB2	2.018
119	GLN	HB3	1.938
119	GLN	HG2	2.282
119	GLN	HG3	2.282
119	GLN	HE21	7.471
119	GLN	HE22	6.81
119	GLN	CA	56.076
119	GLN	CB	29.543
119	GLN	CG	33.767
119	GLN	N	123.994
119	GLN	NE2	112.474
120	GLN	H	8.415
120	GLN	HA	4.226
120	GLN	HB2	2.033
120	GLN	HB3	1.95
120	GLN	HG2	2.295
120	GLN	HG3	2.295
120	GLN	HE21	7.446

120	GLN	HE22	6.813
120	GLN	CA	56.283
120	GLN	CB	29.575
120	GLN	CG	33.751
120	GLN	N	122.052
120	GLN	NE2	112.499
121	SER	H	8.308
121	SER	HA	4.35
121	SER	HB2	3.8
121	SER	HB3	3.8
121	SER	CA	58.662
121	SER	CB	63.771
121	SER	N	116.851
122	MET	H	8.305
122	MET	HA	4.45
122	MET	HB2	2.054
122	MET	HB3	1.943
122	MET	HG2	2.538
122	MET	HG3	2.458
122	MET	HE1	2.012
122	MET	HE2	2.012
122	MET	HE3	2.012
122	MET	CA	55.51
122	MET	CB	32.796
122	MET	CG	32.046
122	MET	CE	16.893
122	MET	N	122.049
123	LEU	H	8.121
123	LEU	HA	4.334
123	LEU	HB2	1.57
123	LEU	HB3	1.57
123	LEU	HD11	0.787
123	LEU	HD12	0.787
123	LEU	HD13	0.787
123	LEU	HD21	0.851
123	LEU	HD22	0.851
123	LEU	HD23	0.851
123	LEU	CA	55.269
123	LEU	CB	42.366
123	LEU	CG	26.891
123	LEU	CD1	23.163

123	LEU	CD2	25.049
123	LEU	N	123.282
124	SER	H	7.793
124	SER	HA	4.178
124	SER	HB2	3.758
124	SER	HB3	3.758
124	SER	CA	59.892
124	SER	CB	64.881
124	SER	N	121.994

For the Glutaminase L Peptide:

Residue no.	Amino acid	Nucleus	Chemical shift
1	LYS	HA	4.016
1	LYS	HB2	1.794
1	LYS	HB3	1.794
1	LYS	HG2	1.381
1	LYS	HG3	1.381
1	LYS	HD2	1.641
1	LYS	HD3	1.641
2	GLU	HA	4.266
2	GLU	HB2	1.968
2	GLU	HB3	1.843
2	GLU	HG2	2.192
2	GLU	HG3	2.192
3	ASN	H	8.62
3	ASN	HA	4.666
3	ASN	HB2	2.791
3	ASN	HB3	2.669
3	ASN	HD21	7.56
3	ASN	HD22	6.869
4	LEU	H	8.306
4	LEU	HA	4.262
4	LEU	HB2	1.565
4	LEU	HB3	1.565
4	LEU	HG	1.537
4	LEU	HD11	0.791
4	LEU	HD12	0.791
4	LEU	HD13	0.791

4	LEU	HD21	0.842
4	LEU	HD22	0.842
4	LEU	HD23	0.842
5	GLU	H	8.381
5	GLU	HA	4.196
5	GLU	HB2	1.96
5	GLU	HB3	1.838
5	GLU	HG2	2.189
5	GLU	HG3	2.189
6	SER	H	8.146
6	SER	HA	4.351
6	SER	HB2	3.77
6	SER	HB3	3.77
7	MET	H	8.318
7	MET	HA	4.476
7	MET	HB2	2.05
7	MET	HB3	1.952
7	MET	HG2	2.525
7	MET	HG3	2.442
7	MET	HE1	2.152
7	MET	HE2	2.152
7	MET	HE3	2.152
8	VAL	H	7.592
8	VAL	HA	3.979
8	VAL	HB	1.985
8	VAL	HG11	0.79
8	VAL	HG12	0.79
8	VAL	HG13	0.79
8	VAL	HG21	0.813
8	VAL	HG22	0.813
8	VAL	HG23	0.813