ABSTRACT

Title of Thesis: EFFECT OF LIGAND BINDING ON THE BACKBONE DYNAMICS OF LINEAR AND CIRCULAR CONSTRUCTS OF SH3 DOMAIN.

Ishwar S Chhikara, Master of Science, 2004

Thesis Directed By: Prof. David Fushman,
Department of Chemistry And Biochemistry

Peptide backbone cyclization is an important bio-engineering tool that has been widely used to control the stability, structure and biological role of small peptides. Here we have analyzed the effect of ligand binding on the backbone dynamics of a linear and circular construct of N-terminal SH3 domain of the Murine C-Crk adapter protein. We have also determined the residues on the protein surface involved in binding with the ligand by carrying out a series of HSQC titration experiments at different (ligand/protein) concentrations. This is the first time that the interface for the adapter protein Murine c-Crk N-SH3 domain in complex with C3G derived peptide (ligand) in solution has been determined.

EFFECT OF LIGAND BINDING ON THE BACKBONE DYNAMICS OF LINEAR AND CIRCULAR CONSTRUCTS OF SH3 DOMAIN

By

Ishwar S Chhikara

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2004

Advisory Committee:
Professor David Fushman, Chair
Professor Jin Hu
Professor Victor Munoz

Dedication

This work is dedicated to my caring parents, who raised me with love, prepared me to face challenges and provided me with unparalleled support in all my endeavors. It is to them I owe what I am and what I will be.

Acknowledgements

I thank my advisor, Prof. David Fushman for his guidance and encouragement throughout the duration of this work. He is a wonderful advisor who taught me the virtues of patience and persistence to succeed in research.

Thanks to Professor Julio Camarero for providing us with plasmid for $SH3_{lin\_wt}$ and the C3G peptide and $SH3_{circ\_EG}$ sample for this work. I thank him for his guidance during the protein purification. Thanks to Professor Lorimer for the use of the HPLC instrument and the preparative centrifuge.

I would also like to thank all past and present members of our lab, and fellow graduate students for making the past two years a memorable graduate school experience. Thanks to Dr. Frank Schumann, Dr. Michael Assfalg, Dr. Mariapina D' Onofrio, Jennifer B Hall and Dr. Paul Vasos for valuable discussions regarding the data analysis. Thanks to Dr. Ranjani Varadan for help with protein expression and purification.

Thanks to all the faculty members for wonderful learning experience.

Table of Contents

Dedication ii
Acknowledgements iii
List of Tables v
List of Figures vi
Chapter 1: Introduction and Specific Aims 1
Section 1.1 Linear versus cyclized protein backbone 1
1.2 Naturally occurring cyclic proteins and peptides 1
1.2.1 Circular Proteins found in microorganisms 2
1.2.2 Circular Proteins found in microorganisms 4
1.2.3 Circular peptides from mammals 5
1.3 Implications of cyclized protein backbone 5
1.4 Previous Research 6
1.4.1 Research done in our lab 8
1.5 Research Objectives 15
Chapter 2: Crystal Structure of SH3 domain in complex with C3G peptide 16
Section 2.1 Review of crystal structure of C3G peptide in complex with SH3 domain 16
Chapter 3 20
Section 3.1 Protein Expression And Purification 20
3.1.1 Design of linear and circular construct 20
3.1.2 Growth Media and conditions 20
3.1.3 Purification of $SH3_{lin\_wt}$ 21
3.1.4 $SH3_{circ\_EG}$ and C3G peptide 22
3.1.5 $SH3_{circ\_EG}$ and C3G peptide binding affinities 22
3.2 NMR experiments 23
3.2.1 Chemical shift perturbation mapping 23
3.2.2 NMR titration experiments 24
Chapter 4 25
4.1 Objective 25
4.1.1 Mapping the interface of C3G peptide bound to $SH3_{lin\_wt}$ 25
Chapter 5 31
5.1 R_1, R_2 and NOE 31
Section 5.2 The Backbone Dynamics 38
Chapter 6: Summary and Discussion 48
Chapter 7: Conclusions 50

List of Tables

Table 1 Amino Acid sequence of $SH3_{lin\_wt}$ and $SH3_{circ\_EG}$ 17

List of Figures

1.1 Ribbon diagram for $SH3_{lin\_wt}$ showing the Secondary structure elements 9

1.2 Ribbon diagram for $SH3_{circ\_EG}$ showing the secondary structure elements 10

1.3 Rex contribution to R_2 for various SH3 constructs 13

1.4 β sheets of the SH3 domain showing the circularization region 14

2.1 SH3 domain in complex with C3G peptide 17

2.2 Ribbon diagram showing the interface between the SH3 domain and the peptide 18

2.3 Schematic diagram showing the residues on the protein and ligand surface present at the interface 19

4.1 HSQC spectra for the $SH3_{lin\_wt}$ in complex with the ligand superimposed on the spectra for the free protein 26

4.2 HSQC spectra for the $SH3_{circ\_EG}$ in complex with the ligand superimposed on the spectra for the free protein 27

4.3 Chemical shift perturbation observed for the $SH3_{lin\_wt}$ residues upon ligand binding 28

4.4 Chemical shift perturbation observed for the $SH3_{lin\_wt}$ residues upon ligand binding 29

5.1 Comparison of the Experimental values of R_1 measured for the free $SH3_{lin\_wt}$ and $SH3_{lin\_wt}$ in complex with the peptide 32

5.2 Comparison of the Experimental values of R_2 measured for the free $SH3_{lin\_wt}$ and $SH3_{lin\_wt}$ in complex with the peptide 33

5.3 Comparison of the Experimental values of NOE measured for the free 34

$SH3_{lin\_wt}$ and $SH3_{lin\_wt}$ in complex with the peptide

5.4 Comparison of the Experimental values of R_1 measured for the free $SH3_{circ\_EG}$ and $SH3_{circ\_EG}$ in complex with the peptide 35

5.5 Comparison of the Experimental values of R_2 measured for the free $SH3_{circ\_EG}$ and $SH3_{circ\_EG}$ in complex with the peptide 36

5.6 Comparison of the Experimental values of NOE measured for the free $SH3_{circ\_EG}$ and $SH3_{circ\_EG}$ in complex with the peptide 37

5.7 Order parameters for the $SH3_{lin\_wt}$ (free) and $SH3_{lin\_wt}$ in complex with the peptide 39

5.8 Rex for the $SH3_{lin\_wt}$ (free) and $SH3_{lin\_wt}$ in complex with the peptide 41

5.9 Order parameters for the $SH3_{circ\_EG}$ (free) and $SH3_{circ\_EG}$ in complex with the peptide 42

5.10 Rex for the $SH3_{circ\_EG}$ (free) and $SH3_{circ\_EG}$ in complex with the peptide 46

Chapter 1: Introduction and Specific Aims

Generally we define proteins as linear chains of amino acids that fold in a solvent into a three-dimensional structure that defines their biological function. This chain has a free N terminal amino group and a C-terminal carboxyl group. Since all the peptide bonds are formed by the joining of amino group of one amino acid to the carboxylic group of the second amino acid it can be argued that the linear peptide chain is missing a link that would join the free C-terminal carboxylic group to the N-terminal amino group to give a circular peptide. In past few years many proteins with cyclized or circular backbone have been identified in many organisms [1].

Section 1.1 Linear versus cyclized protein backbone

The linear structure of proteins has many important biological ramifications since the free C and N terminus of linear peptides are often flexible and are the target points for the attack of proteolytic enzymes. The free carboxyl and amino groups are like any other peptide bond forming carboxyl and amine groups and are perfectly suited for the formation of peptide bond. In recent years many proteins have been reported which have this expected circular backbone [1].

1.2 Naturally occurring cyclic proteins and peptides

Small cyclic peptides such as cyclosporin have been known for quite some time but naturally occurring circular proteins have been identified and characterized very recently only. The history of the circular proteins in nature can be traced to native medicine observations made about thirty five years ago. In the early 1970's Gran identified Kalata B1 a twenty-nine amino acid peptide that has uterotonic

properties. [2-4] It is found in the extract from the African plant'Oldenlandia affinis'. It was after another twenty-five years that researchers realized that the peptide is backbone cyclized and it is a gene product. [5] Circular proteins found in nature are composed of 14 to 70 amino acids and almost all of them have well defined three-dimensional structures. The available data indicates that the primary function of most of these circular proteins is host defense [1].

1.2.1 Circular Proteins found in microorganisms

The cyclotides constitute the largest group of circular proteins known so far. It has more than 45 members. Members of this group contain 28–37 amino acids [6]. Kalata B1 was the first cyclotide whose structure was determined. It contains a `cystine knot' motif in which two disulfide bonds are arranged like a ring and peptide backbone segments connecting them is penetrated by a third disulfide bond [7]. Similar cystine knot motifs have been reported in other proteins, including small toxins and growth factors. These features – cystine knot and circular backbone – are part of the cyclic cystine knot (CCK) motif. It has been found that there is another small triple stranded beta sheet associated with this motif. The common feature of all these proteins initially reported were size (30 amino acids), six conserved cysteine residues forming three disulfide bonds, circular backbone and a plant origin. The knotted disulfide arrangement and circular backbone of the CCK motif are the factors that make the cyclotides extremely stable. They are resistant to enzymatic hydrolysis and thermal denaturation as shown, by the retention of biological activity after being boiled in native medicine applications [8]. Research done on acyclic permutants of kalata B1 has shown that although open chain analogues have the basic three-

dimensional structure, they are intrinsically less stable and have weaker internal hydrogen bonds than the circular parent proteins. Based on this finding it appears that cyclization plays an important role in stabilizing the three-dimensional structure rather than defining it [1].

MCoTI-I and MCoTI-II two trypsin inhibitors found in the seeds of *Momordica cochinchinensis,* a vine plant from the Cucurbitaceae family could also be classified as members of the family of circular, disulfide-rich proteins from plants [9]. The structure of MCoTI-II was found to contain the CCK motif, so this peptide can tentatively be regarded as a member of the cyclotide family [10]. Although it is topological similar and has the six conserved cysteine residues as known for previously reported cyclotides, it has substantial sequence differences with these previously known cyclotides. It was found that, the two MCoTI peptides are the only CCK family members that inhibit trypsin and they differ from the other cyclotides in the sense that a linear homologue co-exists with circular forms in the same tissue. The linear form, MCoTI-III, is present only in small amounts and it is homologous, and not identical, in sequence with the circular forms, so it is not necessary that it is a precursor of the circular forms. Research indicates that the driving force for cyclization of these peptides might be the resistance to proteolysis [10]. In 1999, one more circular trypsin inhibitor was discovered in plants, and it was named sunflower trypsin inhibitor 1 (SFTI-1) because it is found in sunflower seeds. SFTI-1 consists of just 14 amino acids, and it shows both sequence and conformational similarity with the Bowman–Birk inhibitors, which is a family of small serine proteinase inhibitors found in the seeds of legumes and several other plants. [11] Researchers compared

the solution structure and inhibitory activity of native SFTI-1 with that of an acyclic analogue which has the peptide backbone broken at the hairpin end. It was seen that the three-dimensional structures of the two molecules is almost identical to each other and to the crystal structure of SFTI-1 bound to trypsin, which indicates that the circular nature of SFTI-1 has potentially evolved to increase the in vivo lifetime and not to confer conformational stability onto the active loop region. In the case of Momordica peptides, enhanced proteolytic stability seems to be the major driving force for cyclization [11].

1.2.2 Circular Proteins found in microorganisms

Bacteriocin AS-48 (70-amino acids) is the largest naturally occurring circular protein currently known. This protein protects the producing strain against other bacteria by forming pores in the cytoplasmic membrane of sensitive cells. The head-to-tail ligation that gives rise to circular backbone seems to confer considerable stability on the cyclized molecule shown by its thermal denaturation temperature of 93°C [12].

Another circular protein found in bacteria is microcin J25 (MccJ25), which is a highly hydrophobic 21-residue peptide excreted by E. coli. Microcins constitute a miscellaneous group of low molecular mass antibiotic peptides which are produced by diverse strains of Enterobacteriaceae, but among all of them MccJ25 is the only one known to be circular. The antibiotic activity of MccJ25 is due to its ability to interfere with cell division. The circular backbone is required for activity, as it was found that linear form of the protein, produced by cleaving MccJ25 with the enzyme

thermolysin, was inactive against three tested *E. coli* strains and 40 times less active than circular MccJ25 against Salmonella Newport [12].

1.2.3 Circular peptides from mammals

Rhesus theta defensin-1 (RTD-1) is a new type of defensin i.e. a small, disulfide-rich peptide involved in host defense and is found in the leukocytes of rhesus macaques. It is made up of just 18 amino acids, which includes six cysteines and five arginines, and its backbone is cyclized via a peptide bond [13].

Native circular RTD-1 has an antibacterial activity threefold greater than a synthetic acyclic analogue. Unlike the acyclic form; it is observed that it maintains its activity in the presence of physiological NaCl [14]. Analysis of the three-dimensional structure of the acyclic form shows that cyclization does not change the overall structure of the peptide. Therefore, the difference in antimicrobial activity cannot be explained by structural changes brought about by cyclization, and might be due to stabilization of the molecule in vivo. Resistance of these peptides to exoproteases will provide them a big advantage in the protease-rich conditions in which they function [1].

1.3 Implications of cyclized protein backbone

From all these examples of naturally occurring cyclized proteins and peptides we can see that two main advantages of cyclization appear to be increased resistance to protease digestion and improved thermodynamic stability, both of which can greatly improve the biological activity of a protein in vivo. Other possible benefits of cyclization, like orienting the active residues in a preferred conformation for binding,

or diminishing the flexibility of peptide termini and thereby minimizing entropic losses on receptor binding, might be there in some cases, but these do not appear to be the primary role of backbone cyclization.

Taking a clue from nature biochemists have been trying to use the protein cyclization for various purposes like making better drugs and more stable proteins. Many interesting studies have been carried out to better understand the effects of cyclization.

1.4 Previous Research

The backbone cyclization of the polypeptide chain, i.e. the ligation of its N- and C- termini via peptide bond, is expected to reduce the available conformational space and therefore decrease its backbone entropy. As a consequence, we expect a more rigid backbone. This fact is well known for peptides and in the last decade backbone cyclization has become a widely used strategy to control structure and biological function of small peptides and to improve the *in vivo* stability of these bioactive molecules, [15-17] But the impact of backbone cyclization on the stability of proteins has not been fully explored. To cyclize the backbone of a folded protein, the N- and C-termini have to be in close proximity in order to create an amide bond between both termini. Interestingly, this prerequisite is a surprisingly common feature in protein folds, particularly in single protein domains [18]. The first successful semi synthesis of circular protein was carried out by Creighton and Goldenberg by exposing native BPTI to a chemical cross-linking agent [19]. Since then, several

chemical as well as recombinant techniques have been developed which provide easy access to circular proteins.

According to the general polymer theory ligating the C- and N-termini of a protein fold should lead to a better thermodynamic stability of the protein fold [20]. However, synthetic cyclic proteins sometimes are less stable than their linear counterparts, which might be attributable to the strain introduced by linking the C- and N-termini. Hence, both entropic and enthalpic contributions should be considered when engineering a protein fold through cyclization. One way to reduce the enthalpic effect due to the strain introduced by the cyclization is the insertion of a flexible glycine spacer of varying length at the ligation site [21, 22].

Backbone cyclization of a protein also provides the advantage of a stable linkage, thus overcoming the problem of chemical and conformational instabilities associated with the cyclization via formation of a disulfide bond. This approach has a potential to improve the in vivo stability of proteins for therapeutic applications and thermodynamic stability of enzymes for industrial processes and delivers a promising tool for rational drug design and protein engineering [23]. Also, many proteins in the cell have modular architecture, i.e. they are composed of various individually folded domains, and the function and interactions of these units are central for the regulation of various events, including signal transduction and transcriptional control. Isolation of the individual domains for structural and biochemical studies, currently a common ‘reductionist’ approach in structural biology, takes them out of the context of the whole protein, and could potentially increase the flexibility of the termini by removing the restricting influence of the neighboring domains. This could alter the

thermodynamic stability of the domain. Restricting the mobility of the termini by cyclization could then, to some extent, mimic the 'natural' situation in multidomain systems, and thus could be useful for the understanding of the functional features of the individual domains.

Therefore, understanding the effect of backbone circularization on protein structure, dynamics, and function could provide insights into the importance of C- and N-termini for protein structure and stability, and the potential applications of the backbone cyclization in protein engineering.

1.4.1 Research done in our lab

One of the research focuses of out lab is to understand the effect of backbone cyclization on protein structure and dynamics. In the previous work done in our lab we used the N-terminal SH3 domain of the adapter protein c-Crkfor these studies [24]. This domain contains 57 amino acids and has juxtaposed N- and C-termini which allows the circularization of the protein without perturbing the native fold. According to the crystal structure of the SH3 domain in the complex with a C3G-derived proline-rich peptide [25], the protein forms a compact structure comprising five β-strands and a short 3_{10}-helix.

The protein fold is stabilized by a network of electrostatic and hydrophobic interactions. The salt bridge located between the side chains of residues Glu135 and Lys164 has been shown to be essential for the stability of the SH3 fold [24]

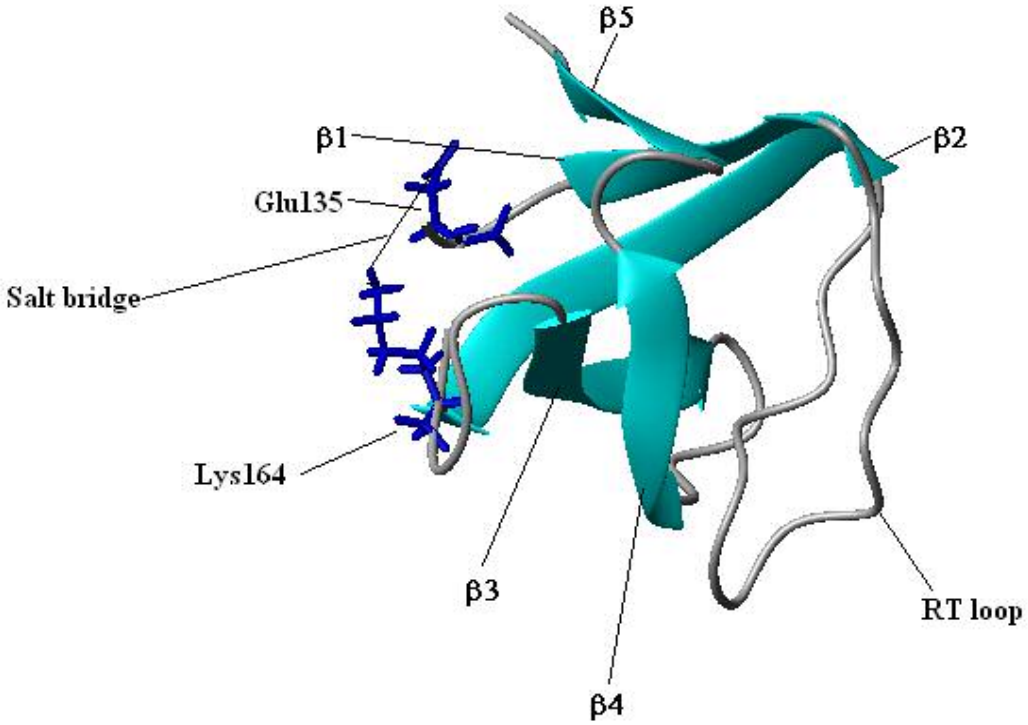

Figure 1.1 Ribbon Diagram for $SH3_{lin\_wt}$ showing the secondary structure elements and residues Glu135 and Lys164. [PDB id: 1CKA]

Deletion of Glu135 residue destabilizes the SH3 structure resulting in a partially folded protein. This also causes a ten-fold reduction in the ligand binding affinity. Other important contacts, inferred from the crystal structure, involve the side chains of residues Lys164, Trp170 and Val184 that are packed against each other in a stacked conformation, and a hydrogen bridge between the carbonyl group of Glu135 and the amide group of Ile161.

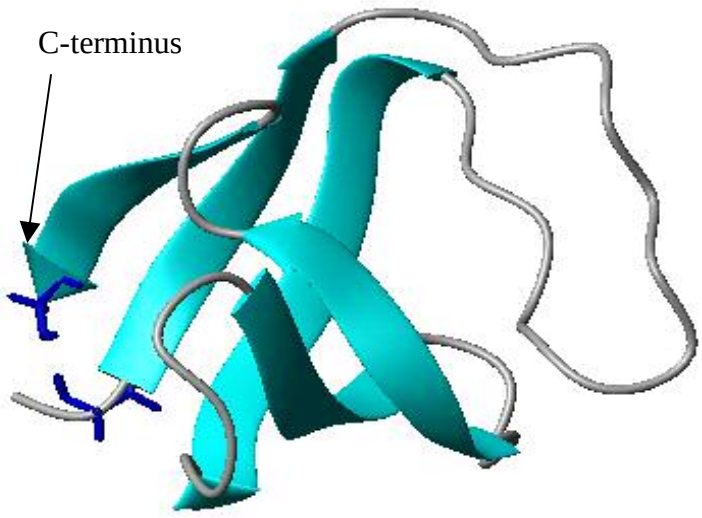

Figure 1.2 Ribbon diagram for $SH3_{circ\_EG}$ showing secondary structure elements. (PDB id: 1M3B)

The thermodynamics, folding kinetics and biological activities of different circular and linear c-Crk SH3 constructs have been studied by Camarero etal. [24]. It was found that backbone cyclization of a truncated SH3 domain lacking the key Glu135 residue stabilizes the fold and restores the affinity binding for the C3G-based poly-Pro ligand. In yet unpublished work our group studied in detail the structure and backbone dynamics of different circular SH3 constructs varying the cyclization site as well as the length of the loop used at the circularization site. The structure of all the circular constructs was found to be extremely similar to that of the full-length SH3 domain, and only minor structural differences were found in proximity to the cyclization site. In contrast, the backbone of all the circular constructs was significantly less mobile than that of the linear full-length SH3 domain.

In our group's previous work the most striking difference between the linear and circular constructs was observed in the slower, microsecond-time dynamics. The data analysis indicated significant conformational exchange contributions, R_{ex}, for Ile161, Asp163, Glu166, Glu167, Gln168, and Ala172 in the circular constructs (figure 1.3). These residues are located in the β2/β3 loop (also known as the N-Src loop) and in the adjacent parts of the β2 and β3 strands, most of them in close proximity to the circularization region (Fig. 2). The pattern of residues exhibiting conformational exchange motions was similar for all circular constructs and in $SH3_{circ\_EG}$ for both 500 and 600 MHz data. The R_{ex} contributions were most pronounced in the shorter circular constructs, $SH3_{circ\_\Delta}$ and $SH3_{circ\_EG}$, where the observed R_{ex} values were practically identical. The calculated R_{ex} values were approximately 3-fold smaller in $SH3_{circ\_wt}$ and even smaller in $SH3_{lin\_wt}$ construct.

Construct	Sequence
	134 190
$SH3_{lin\text{-}wt}$	AEYVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMIPVPYVEKYG
$SH3_{circ\text{-}\Delta}$	cyclo-[CYVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMIPVPYVEKYG]
$SH3_{circ\text{-}EG}$	cyclo-[CGYVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMIPVPYVEKYG]
$SH3_{circ\text{-}wt}$ *	cyclo-[CGAEYVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMIPVPYVEKYG]

Table1. The amino acid sequence and notations of different SH3 constructs

There are two plausible explanations for this phenomenon. It was likely that the observed increase in the exchange broadening reflected the presence of slow rearrangements which relieve a circularization-induced strain in the protein structure. Data analysis indicated that cyclization caused a displacement of the β1 strand which leads to a rearrangement in the proximal β2 strand; the magnitude of the effect

depends on the length of the circularization loop. Due to the way the strands are aligned and contact each other (Fig.2 and 3), a strain induced by this rearrangement in the C-terminal part of the β2 strand could then propagate to the proximal β3 strand, thus also affecting the β2/β3 loop.

Consistent with this explanation, an increase in the length of the circularization loop diminished number of residues showing the conformational exchange contributions. However, the observation that the effect did not disappear in the linear protein, and the R_{ex} terms were observed for the same residues, although significantly weaker than in the circular constructs, suggested that an additional mechanism could also be responsible for the observed conformational exchange. The location of the site involved in the conformational exchange suggested that this could be related to mutual reorientation of the side chains of Lys164 and Glu167, to satisfy the electrostatic attraction between their charges. Glu135 and Glu167 are positioned on the opposite sides from Lys164, and it seems feasible that in the absence of the Glu135-Lys164 salt bridge, the side chains of Glu167 and Lys164 would have some tendency to reorient. Since the aliphatic part of Lys164 side chain is involved in the hydrophobic core, the reorientation of Glu167 seemed more likely. It is worth mentioning also that Glu167 is located in the N-Src loop and, according to the crystal structure, is involved in the electrostatic interaction with the positive charge (Lys9) from the ligand.

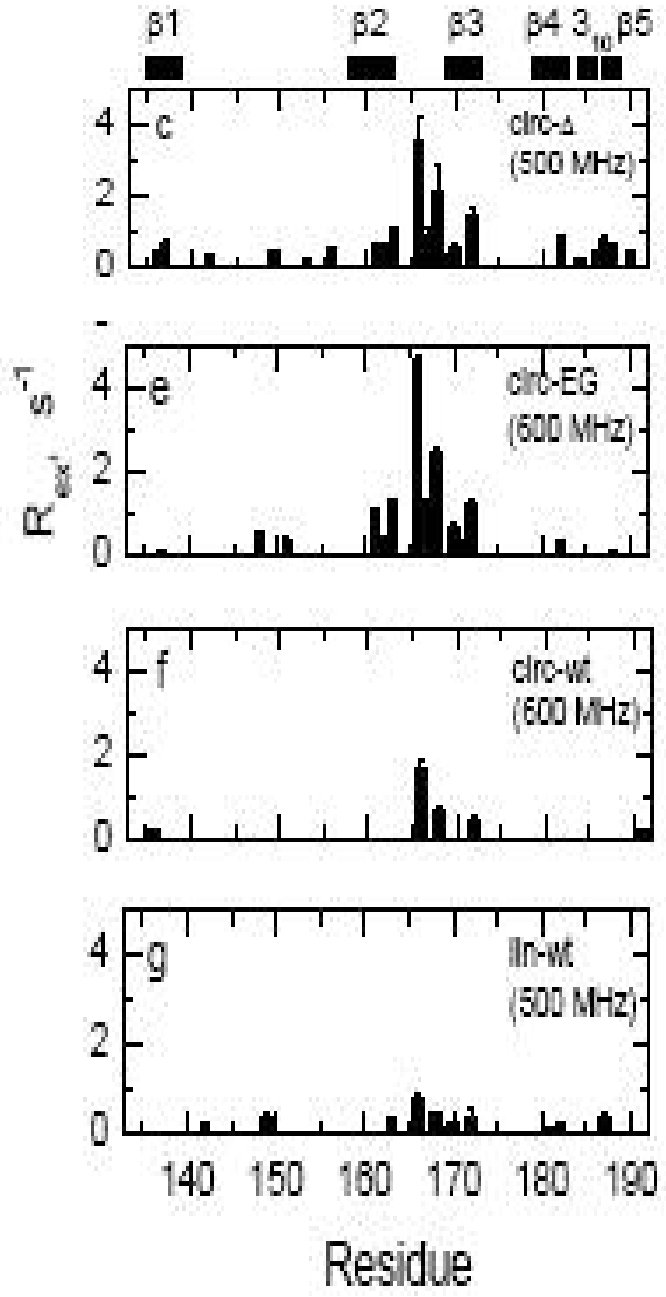

Fig 1.3 The conformational exchange contribution to the transverse relaxation rate for various SH3 constructs that have been previously studied in our lab. The sequence and notations have been shown in Table 1.

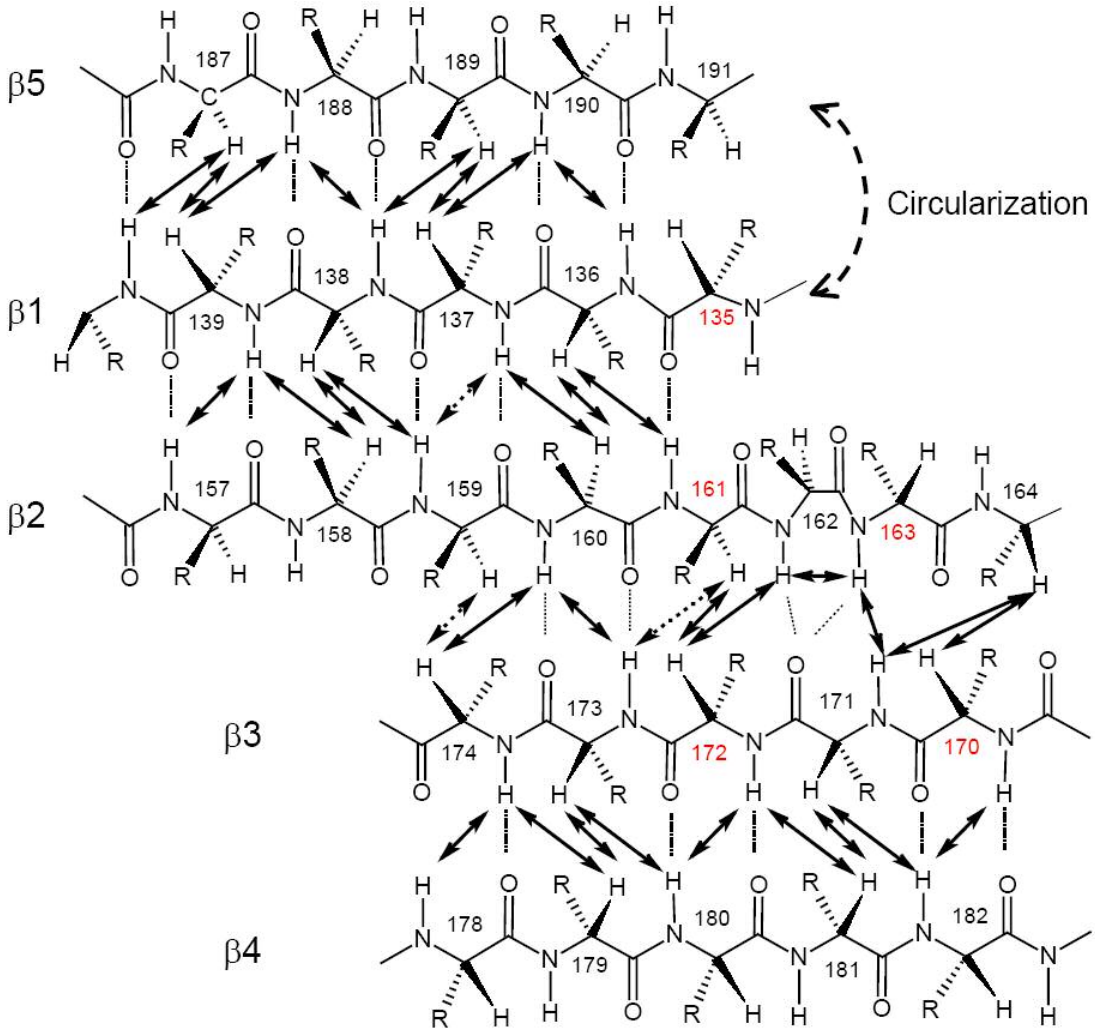

Fig 1.4 Schematic diagram showing the various β sheets in the SH3 domain along with the circularization site. Also shown is the observed NOE pattern.

1.5 Research Objectives

The rationale behind this work was to determine how the ligand binding will affect the conformational exchange motions in the two constructs. The R_{ex} represents μs to ms timescale motions and in the previous work R_{ex} contributions to the transverse relaxation rates were observed for some of the residues known to be involved in ligand binding. Our hypothesis was that the ligand binding to the residues Lys164 and Glu 167 which are in the region showing highest R_{ex} contribution will reduce any conformational exchange contribution arising due to the reorientations of the side chains of these residues by making that part of the molecule more rigid.

Another goal was to find out the residues on each SH3 domain surface involved in interactions with the ligand. We wanted to ascertain whether the solution structure of the complex and the crystal structure provide similar kinds of interactions between the protein and the ligand. We also wanted to find out if there are any differences between the linear and circular construct in terms of ligand binding interface.

Chapter 2: Crystal Structure of SH3 domain in complex with C3G peptide

SH3 domains consist of about 60 residues and they bind to small peptides containing proline residues arranged in characteristic PxxP motifs [26]. N-terminal SH3 domain of the adapter protein c-Crk domain contains 57 amino acids.

Section 2.1 Review of crystal structure of C3G peptide in complex with SH3 domain

Xiaodong and coworkers determined the crystal structure of Crk SH3-N in complex with a high affinity peptide derived from C3G (PPPALPPKKR) [27]. According to the crystal structure the protein forms a compact structure comprising five beta strands and a short 3_{10}-helix [28]. The protein fold is stabilized by a network of electrostatic and hydrophobic interactions.

The salt bridge located between the side chains of residues Glu135 and Lys164 is essential for the stability of the SH3 fold [24]. Deletion of Glu135 residue destabilizes the SH3 structure resulting in a partially folded protein. This also causes a ten-fold reduction in the ligand binding affinity. Other important contacts, inferred from the crystal structure, involve the side chains of residues Lys164, Trp170 and Val184 that are packed against each other in a stacked conformation, and a hydrogen bridge between the carbonyl group of Glu135 and the amide group of Ile161.

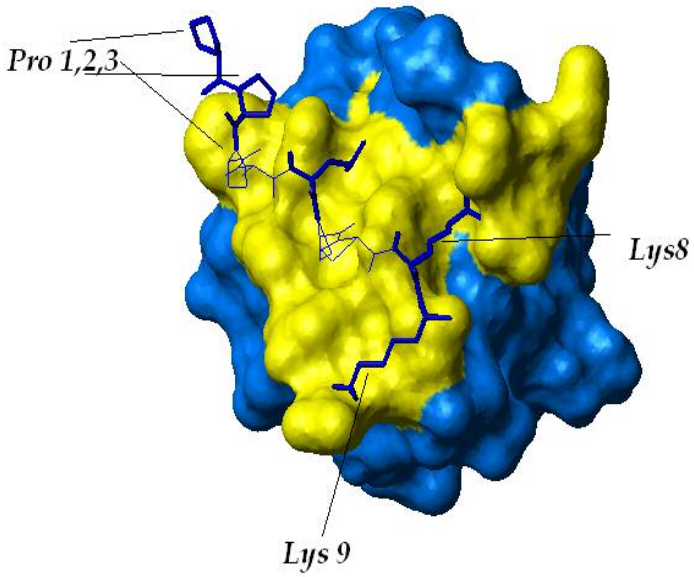

Figure 2.1 The N-terminal SH3 domain of the c-Crk protein in complex with its poly-Pro ligand showing the SH3 domain surface involved in interaction with the ligand(yellow). Ligand residues Lys8, Lys9 and Pro1, 2, 3 which are involved in binding with SH3 domain are also shown (neon). [PDB:1CKA]

As per the crystal structure three acidic residues Asp 147, Glu 149 and Asp 150 in the RT loop of the Murine c-Crk SH3-N domain are involved in the lysine specific interactions with the residue Lys8 on the ligand. These three residues approach each other closely, and present three oxygen atoms that form a nearly equilateral triangle with the lysine nitrogen at the geometric center and at a distance of 1.0 angstrom from the plane of the oxygen atoms. The oxygen-oxygen distances are in the range 4.3 to 4.6 while the lysine nitrogen- oxygen distances are in the range 2.7-2.8 Angstrom.

Glu 167, Glu 166, Trp 169, Phe143, Tyr 186, Pro 183, Pro 185 and Phe 141 are also involved in the interaction with the C3G peptide. The carboxyl group of Glu167 makes a direct salt bridge contact with the ε-amino group of Lys9 of the ligand, while

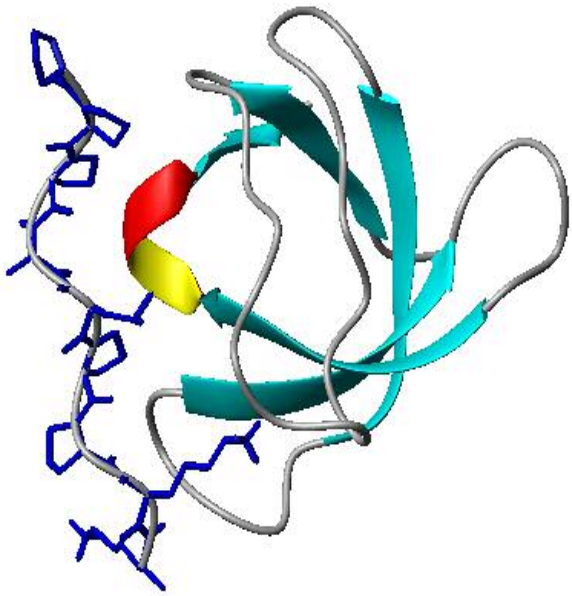

Figure 2.2: A schematic diagram showing the interface between the $SH3_{lin\_wt}$ and the C3G peptide (PDB: 1CKA)

the side chain carboxyl group of Glu166 is properly positioned for hydrogen bonding with the amide group of the same residue. Moreover, the indole group of Trp169 is well packed against the aliphatic component of the ligand residue Lys8, one of the critical residues for C3G peptide binding to Crk SH3 [28]. As a result, the ε-amino group of Lys8 is positioned such that it can reach and form a salt bridge with the carboxyl groups of Asp147, Glu149, and Asp150 located in the middle of the RT-loop of SH3.

All the residues of the SH3 domain involved in the interaction with the C3G derived peptide are shown in figure 2.3 [27].

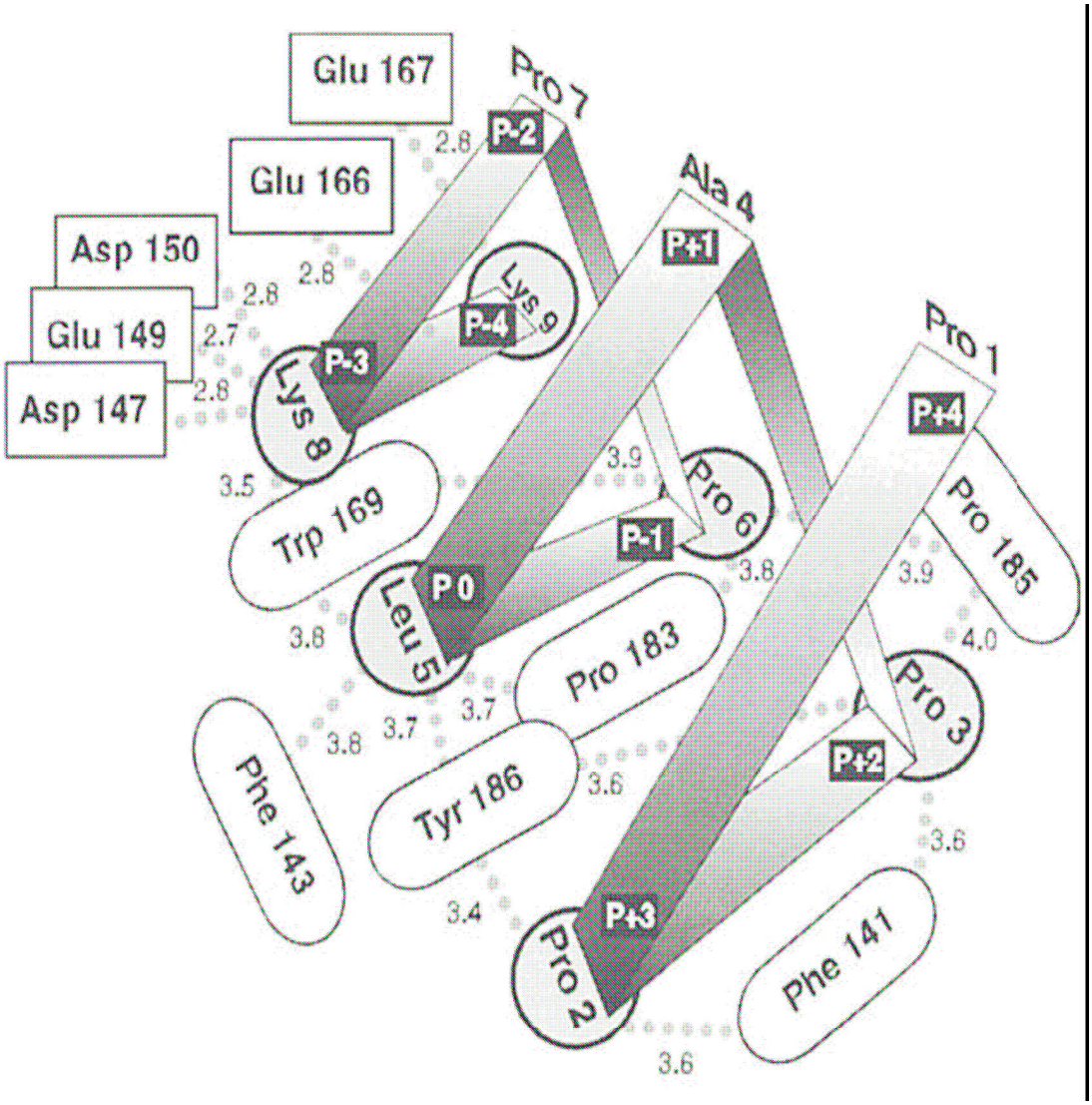

Figure 2.3 A schematic diagram showing the specific interactions between the residues on $SH3_{lin\_wt}$ domain surface with the residues on the ligand surface**. (**from reference [27]) The oval boxes show the hydrophobic residues and the rectangular boxes show the acidic residues on the SH3 domain surface. The ligand, C3G peptide is represented by left hand ribbon.

Chapter 3

Section 3.1 Protein Expression And Purification

3.1.1 Design of linear and circular construct

For this study we used linear full length N-terminal c-Crk SH3 and circular construct $SH3_{circ\_EG}$. The amino acid sequence of the linear construct, $SH3_{lin\text{-}wt}$ consisted of residues 134-190 from the wild type N-terminal SH3 domain of c-Crk, with an additional Gly added at the c-terminus. The circular $SH3_{circ\_EG}$ had the same sequence as the $SH3_{lin\text{-}wt}$ with the exception of two mutations. It contained Ala134Cys and Glu135Gly mutations and had the same number of amino acids as the linear construct $SH3_{lin\text{-}wt}$.

3.1.2 Growth Media and conditions

Plasmids DNA construct pTXB1- $SH3_{lin\_wt}$ (provided by Prof. Julio Camarero, Lawrence Livermore National Laboratory) used in this study was cloned into *E.coli* BL21 (DE3) pLysS cells. Starter cultures were grown for 6-8 hours at 37^oC to an OD_{600} ~0.6 using isolated colonies from a fresh plate. Cell cultures for unlabelled proteins were grown in the auto-inducing ZYP-5052 medium (composition from Dr. William Studier, Brookhaven National Laboratory) supplemented with ampicillin and chloramphenicol (at 100mg and 50 mg per liter culture respectively) at 37^oC in a shaker incubator with continuous agitation at 200rpm. For uniform ^{15}N isotope incorporation in proteins, cell cultures were grown in ZYP-5052 medium

replacing $(NH_4)_2SO_4$ with $^{15}NH_4Cl$ (1g/liter culture) and Na_2SO_4, such that $^{15}NH_4Cl$ provided the sole source of Nitrogen in the medium.

3.1.3 Purification of $SH3_{lin\_wt}$

All cell pellets were frozen at –80°C for about 30 minutes prior to cell lysis. Cells were resuspended in lysis buffer (50mM Tris pH 7.6, 0.02% Triton x100, 0.4 mg/ml lysozyme, protease inhibitors: 1mM PMSF, 50 μM TLCK, 5 μg/ml soyabean trypsin inhibitor, 2.5 μg/ml leupeptin). DNase1 (to 20 μg/ml) and $MgCl_2$ (to 10mM) were added to breakdown the DNA. The cell suspension was centrifuged at 25000rpm for 25 minutes in a preparative ultracentrifuge 45Ti rotor**.** The clarified supernatant (*ca*. 40 mL) was incubated with 5 mL of chitin-beads (New England Biolabs), previously equilibrated with column buffer (0.1 mM EDTA, 50 mM sodium phosphate, 250 mM NaCl, 0.1% Triton X-100 at pH 7.2), at 4°C for 30 min with gently shaking. The beads were extensively washed with column buffer (10 x 5 mL) and equilibrated with PBS (50 mM sodium phosphate, 100 mM NaCl at pH 7.2, 2 x 50 mL). The intein-fusion protein adsorbed on the beads was then cleaved by adding cysteamine or DTT (≈30 mM) for overnight at 25°C. The supernatant was separated by filtration and the beads were washed with additional PBS at pH 7.2 (4 x 5 mL). The supernatant and the washes were pooled and $SH3_{lin\_wt}$ was purified by preparative HPLC using a linear gradient of 20-45% B over 50 min. Analytical gradient HPLC was performed with 220 and 280 nm detection. Analytical HPLC was performed on a Jupiter C18 column (5 micron, 4.6 x 150 mm) at a flow rate of 1 mL/min. Semi-preparative HPLC was performed on tunable absorbance detector

using a Jupiter C18 column (15 micron, 15 x 250 mm), at a flow rate of 4 mL/min. All runs used linear gradients of 0.1% aqueous TFA (solvent A) in water *vs.* 90% acetonitrile plus 0.1% TFA in water (solvent B).The purified protein was characterized by NMR by comparing the observed HSQC spectra with the HSQC spectra of the earlier protein samples which had been verified by ESMS.

3.1.4 $SH3_{circ\_EG}$ and C3G peptide

$SH3_{circ\_EG}$ and C3G samples were provided to us by Prof. Julio Camarero (Chemical Biology and Nuclear Sciences Division, Lawrence Livermore National Laboratory, University of California, Livermore, CA 94551).The amino acid sequence for the peptide is: PPPALPPKKRY. The molecular weight of the peptide is 1377 Da.

3.1.5 $SH3_{circ\_EG}$ and C3G peptide binding affinities

The binding affinities of these two SH3 domains for the C3G-based poly-Pro ligand have been determined previously using a fluorescence-based binding assay [24]. The reported Kd values were 0.87± 0.06 µM ($SH3_{lin\text{-}wt}$), and 0.47 ± 0.05 µM ($SH3_{circ\text{-}EG}$), which were in total agreement with the values published in the literature. We could not determine the Kd values by NMR and used these previously reported Kd values while considering the binding affinities of the ligand.

3.2 NMR experiments

Protein and peptide concentrations were determined by UV absorption; SH3 linear and circular domains (wavelength = 280 nm), Extinction coefficient =15400 $M^{-1}cm^{-1}$; ligand (wavelength= 276 nm), Extinction coefficient =1345 $M^{-1}cm^{-1}$.

Samples for the NMR studies were prepared by dissolving the protein (concentration, 0.8 mM) in the buffer containing 20 mM sodium phosphate, 100 mM NaCl, 10% D_2O, 2mM DTT and 0.2% NaN_3. The pH was adjusted to 7.2. The NMR measurements were performed on Bruker DRX-600 (University of Maryland) spectrometer operating at ^1H resonance frequency of 600 MHz. The temperature was set to 304.2 degree Kelvin. NMR experiments performed for signal assignment included TOCSY, and heteronuclear ^1H ^{-15}N-HSQC.

3.2.1 Chemical shift perturbation mapping

The approach of chemical shift perturbation mapping was used to identify interfaces between SH3 domains and the C3G peptide. Backbone amide resonances in $SH3_{lin\_wt}$ and $SH3_{circ\_EG}$ were observed using ^1H-^{15}N HSQC. The combined amide chemical shift perturbation was calculated as $\Delta\delta = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{½}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift changes for ^1H and ^{15}N, respectively. Residues showing chemical shift perturbations were most likely to be involved in the formation of the interface between each SH3 domain and the C3G peptide.

3.2.2 NMR titration experiments

Binding of unlabelled C3G peptide to ^{15}N labeled SH3 domains was monitored by NMR titration experiments performed as a series of ^{1}H-^{15}N HSQCs. 0.7-0.8 mM ^{15}N labeled $SH3_{lin\_wt}$ and $SH3_{circ\_EG}$ samples were titrated with increasing amounts of an unlabelled C3G peptide solution. Binding was monitored through changes in the peak positions in the ^{1}H-^{15}N HSQC spectra and titrations were continued until no or very little chemical shift changes were observed (1.2:1 C3G peptide/SH3 domain molar ratio). Combined amide chemical shift differences were computed as described in section 3.2.1. 3.2.3 R_1, R_2 and NOE measurements

^{15}N transverse (R_2) and longitudinal (R_1) relaxation rates and the ^{15}N steady state NOE were measured using standard pulse sequences described in [29].The R_1 and R_2 rates were derived from a single exponential fit of the signal decay measured in a series of 2D HSQC planes. The 2D planes for these ^{15}N relaxation experiments were acquired with 7.2 kHz and 2 kHz in the ^{1}H and ^{15}N dimensions, respectively. The relaxation delays for the R_1 experiments were set to 4ms and 600 (x 4) ms. For the R_2 experiments the delay was set to 0 ms and 216ms (x4). For each 2D plane, 128 t_1 increments were collected, each consisting of 1024 complex points. Steady-state $\{^{1}H\}$-^{15}N NOEs were calculated from the ratio of signal intensities in the 2D NOE and NONOE experiments.

$$NOE = (Isat)/ (Ieq)$$

where Isat is the intensity of a peak with proton saturation (NOE) and Ieq is the intensity without proton saturation (NONOE).

Chapter 4

4.1 Objective

To determine the $SH3_{lin\_wt}$ residues involved in binding to the C3G peptide in solution with those predicted by the crystal structure

4.1.1 Mapping the interface of C3G peptide bound to $SH3_{lin\_wt}$

^{15}N labeled $SH3_{lin\_wt}$ was synthesized as described in Section 3.1. Chemical shift perturbation mapping was used to identify the residues involved in interaction with the C3G peptide. Since chemical shifts are extremely sensitive indicators of the electronic environments of the nuclei, changes in chemical shifts between free $SH3_{lin\_wt}$ and $SH3_{lin\_wt}$ in complex with the C3G peptide and $SH3_{circ\_EG}$ and it's complex with the peptide observed under identical experimental conditions would indicate a change in the microenvironment of the nuclei as a result of the interactions between the ligand and the protein.

The characteristic NMR-fingerprint region (^{1}H-^{15}N HSQC) for the two SH3 constructs studied here shows a common single set of well spread cross peaks corresponding to one folded conformation.

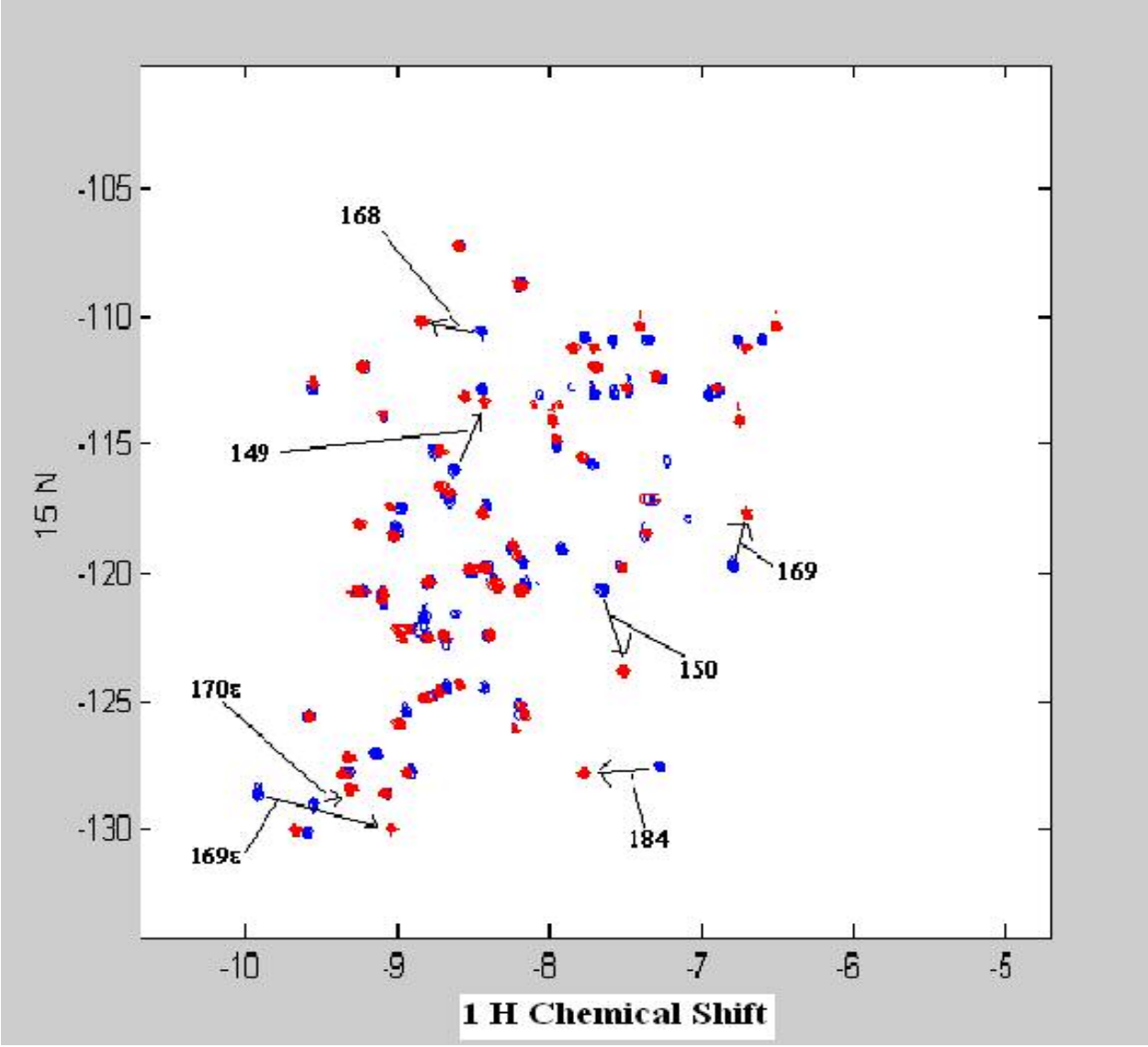

Fig 4.1 HSQC spectra for the $SH3_{lin\_wt}$ in complex with C3G peptide (red) superimposed on the HSQC spectra for the free $SH3_{lin\_wt}$ (blue), the residues showing significant shifts have been highlighted. 169ε and 170ε are the indole NH's of Trp169 and Trp170 respectively.

Previous assignments for the free $SH3_{lin\_wt}$ and $SH3_{circ\_EG}$ (unpublished data) were used as starting points for the assignments of peaks in the HSQC spectrum. The final

assignment of the peaks when all the protein is in complex with the SH3 peptide was confirmed by analyzing the TOCSY spectrum for the final protein sample.

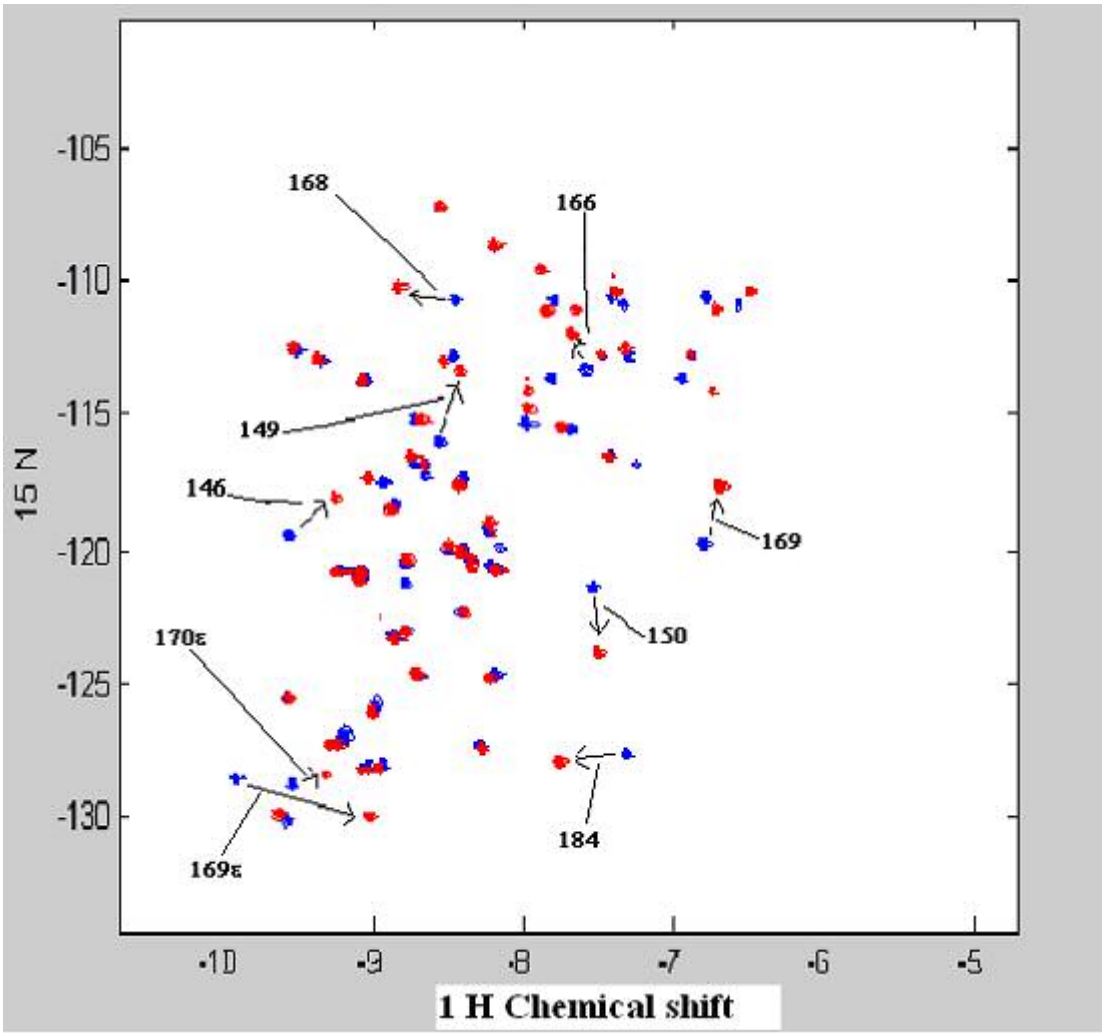

Fig. 4.2 HSQC spectra for the $SH3_{circ\_EG}$ in complex with C3G peptide (red) superimposed on the HSQC spectra for the free $SH3_{circ\_EG}$ (blue), the residues showing significant shifts have been highlighted. 169ε and 170ε are the indole NH's of Trp169 and Trp170 respectively.

Changes in chemical shifts in ^{15}N-^{1}H HSQC spectra of $SH3_{lin\_w}t$ and $SH3_{circ\_EG}$ in comparison with the corresponding C3G peptide bound constructs are plotted in figure 4.1-4.4

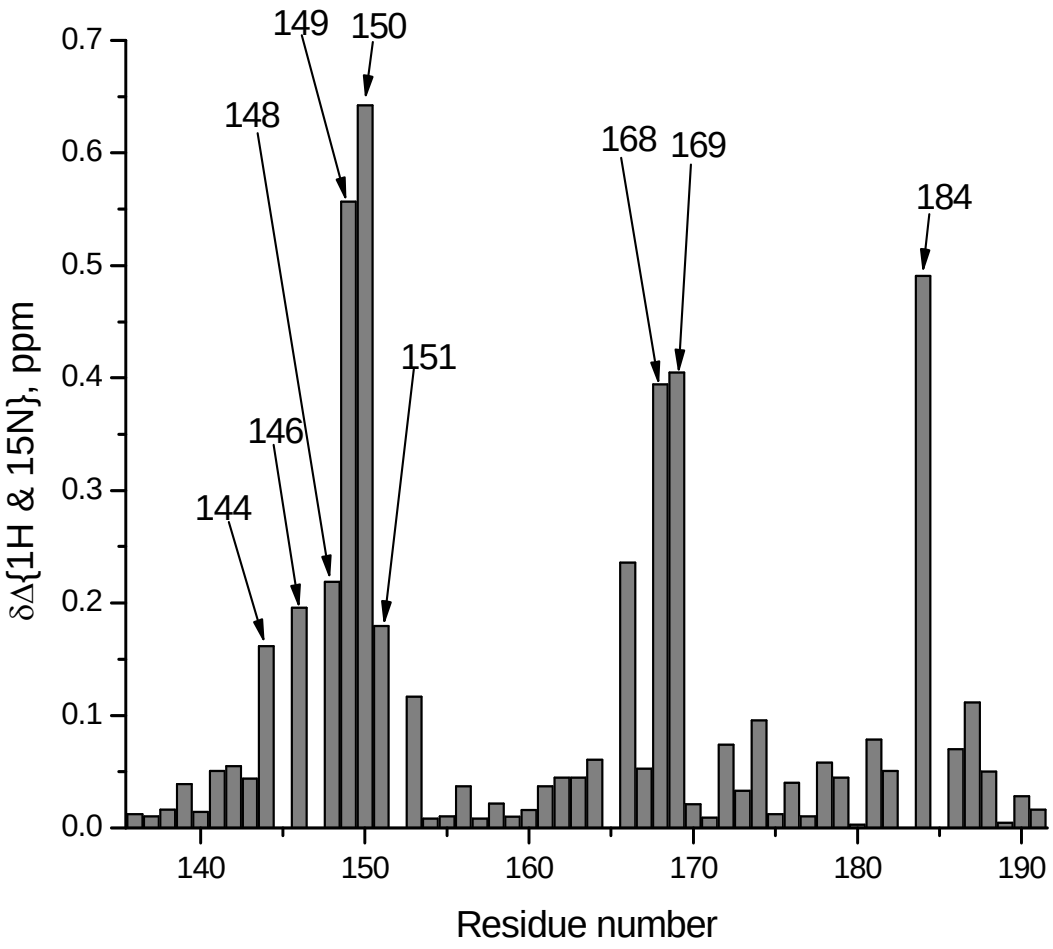

Fig 4.3 Chemical shift perturbation observed for the $SH3_{lin\_wt}$ residues upon ligand binding

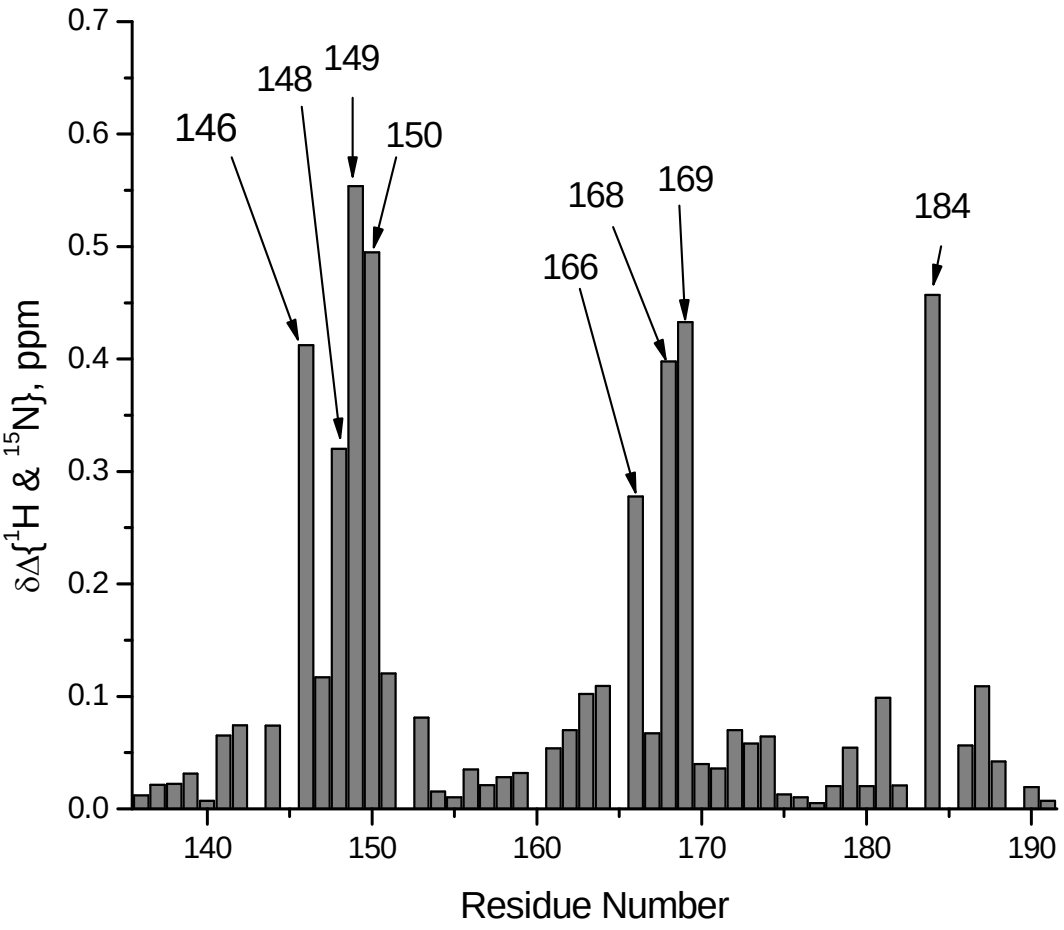

Fig 4.4 Chemical shift perturbation observed for the $SH3_{circ\_EG}$ residues upon ligand binding

The crystal structure of C3G derived peptide bound to the $SH3_{lin\_wt}$ domain has been described in Section 2. Experiments were performed at neutral buffer conditions (PBS, pH 7.2). Since prolines are not observed in the HSQC spectra we don't see residues Pro 185, Pro 183.

We can see significant shifts for Val 184, Glu 149, Asp 150, Glu 166, Gln 168, and Trp 169. These residues are either directly involved in ligand binding or are neighbors of the residues on the SH3 domain surface which are involved in ligand binding. Residues 147 and 145 could not be assigned to any observed peak in the final HSQC spectrum. Similar trend of chemical shift perturbations is observed for the $SH3_{lin\_wt}$ as well as $SH3_{circ\_EG}$ construct confirming similar protein-ligand interface. So we can see that even in solution significant perturbations are observed for protein residues which were indicated to be directly involved in binding and their neighbors based on the crystal structure.

Another observation worth noticing is the fact that the circular $SH3_{circ\_EG}$ and the linear construct $SH3_{lin\_wt}$ show perturbations for the same residues indicating that both of them form similar interface with the ligand.

Chapter 5

Backbone dynamics of the free and bound states of the two SH3 domains

5.1 R_1, R_2 and NOE

In order to characterize the effect of circularization on the backbone dynamics of the free $SH3_{lin\_wt}$ and $SH3_{circ\_EG}$ domains and their complexes with C3G peptide , we measured ^{15}N relaxation rates, R_1 and R_2, and the steady-state $\{^1H\}$-^{15}N NOE, as described in Methods. The experimental data are presented in Fig. 5.1-5.6. 48 to 50 well-resolved cross peaks from backbone amides were observed in the ^1H-^{15}N correlation maps.

The spectra were processed in XWINNMR. The peak assignment was done using XEASY. I made use of the in lab suite of AUTOPICK and RELAXFIT to calculate the R_1, R_2 and NOE. The program returns R_1, R_2 and NOE as output.

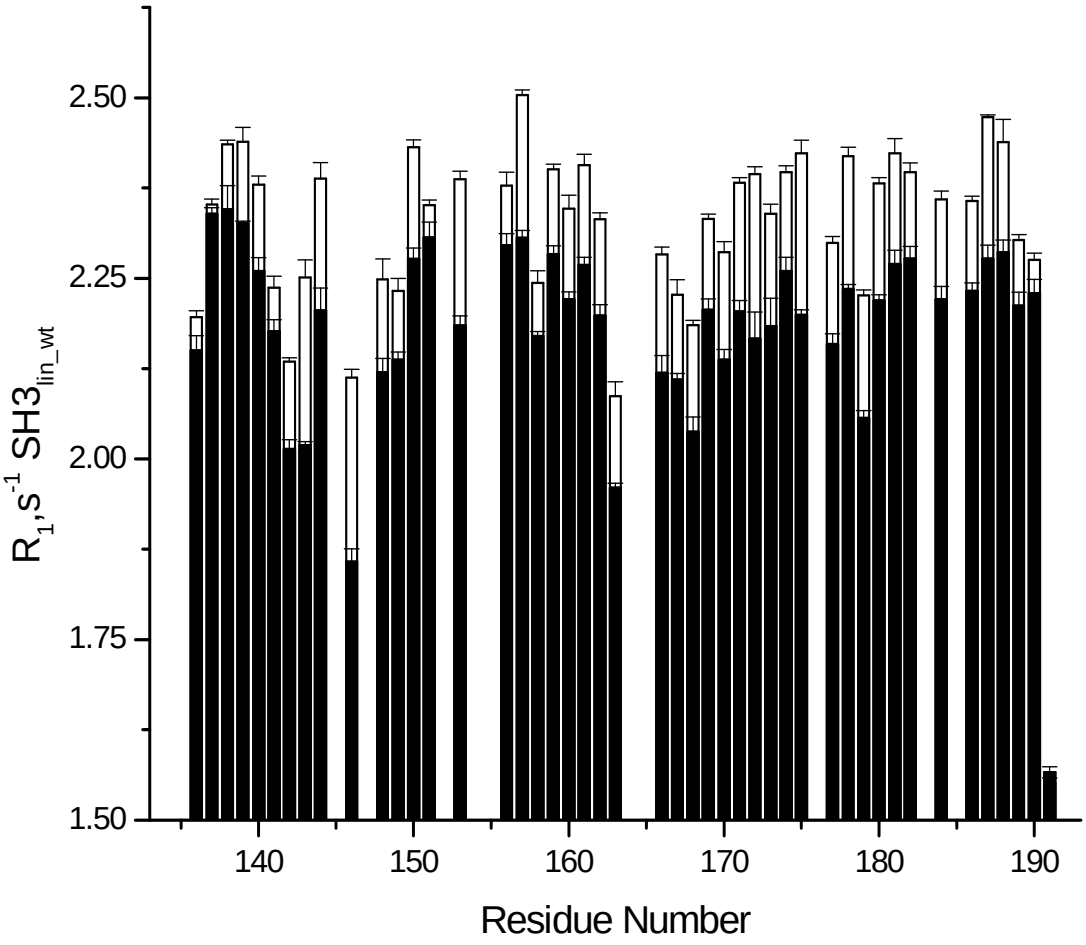

Fig 5.1 ^{15}N Relaxation data for $SH3_{lin\_wt}$: Comparison of Experimental values of R_1 measured for the free $SH3_{lin\_wt}$ (open bars) and $SH3_{lin\_wt}$ in complex with C3G peptide (solid bars)

The error bars here and in Figures 5.2-5.6 indicate the experimental errors in the data processing due to uncertainty in signal to noise estimation.

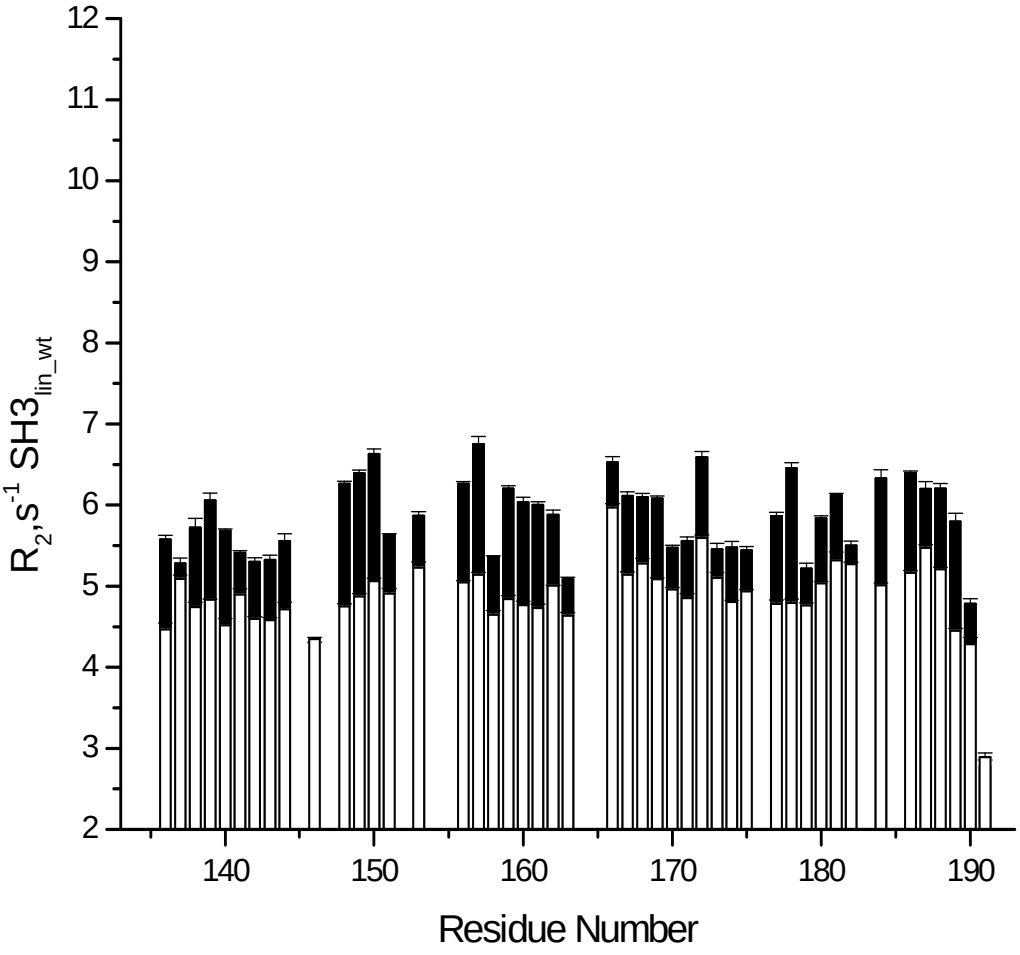

Fig 5.2 Comparison of the R_2 values measured for the free $SH3_{lin\_wt}$ (open bars) and $SH3_{lin\_wt}$ in complex with C3G peptide (solid bars)

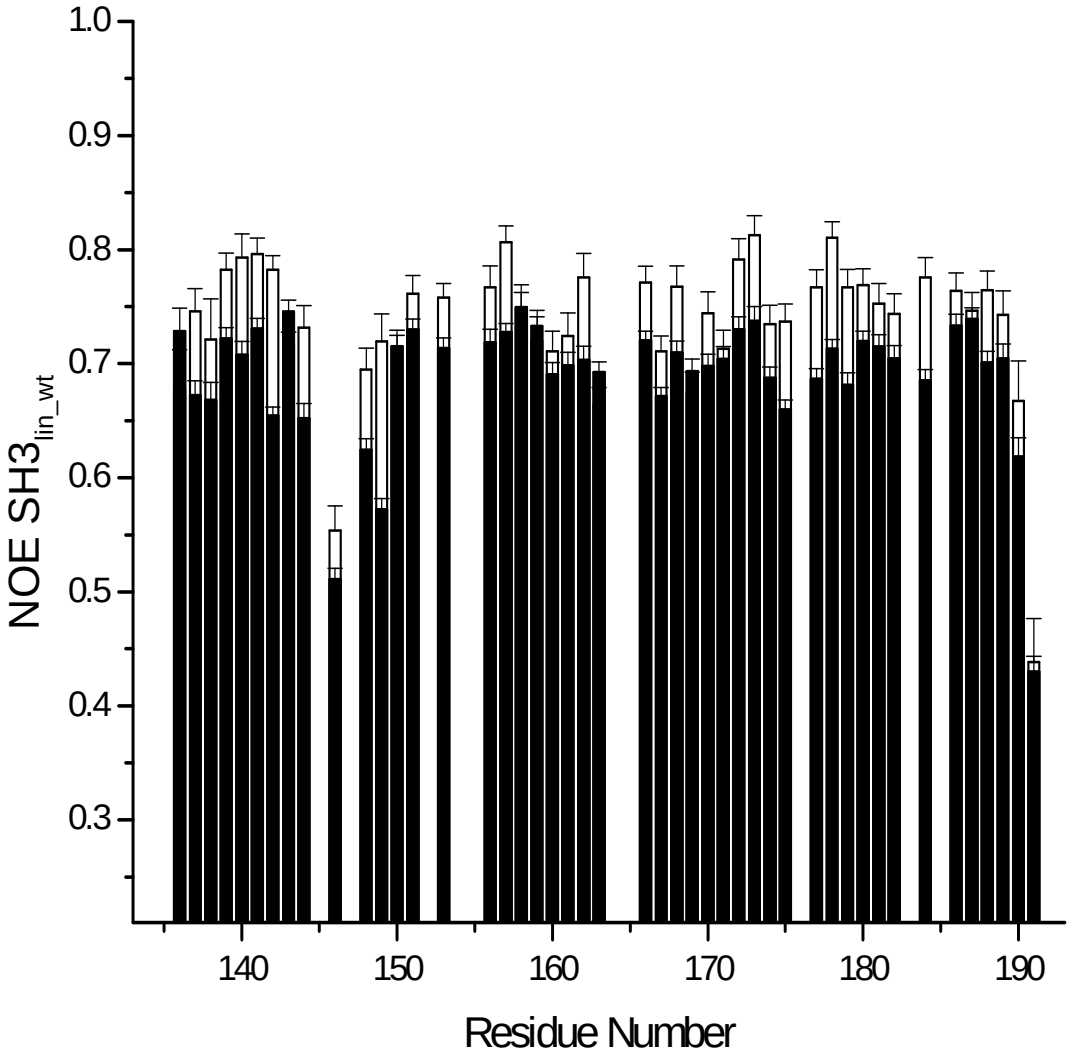

Fig 5.3 Comparison of the NOE values measured for the free $SH3_{lin\_wt}$ (solid bars) and $SH3_{lin\_wt}$ in complex with C3G peptide (open bars)

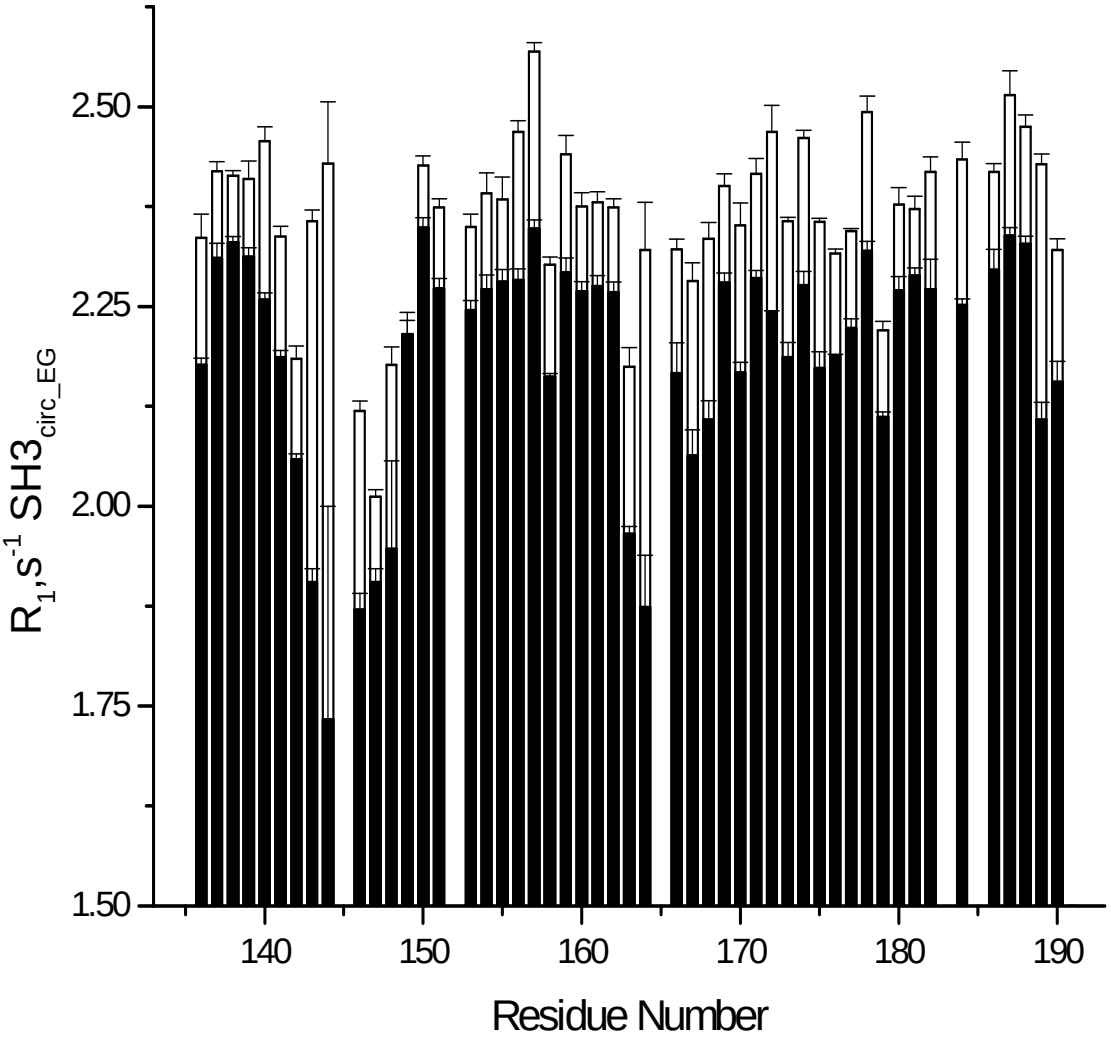

Fig 5.4 Comparison of Experimental values of R_1 measured for the free $SH3_{circ\_EG}$ (open bars) and $SH3_{circ\_EG}$ in complex with C3G peptide (solid bars)

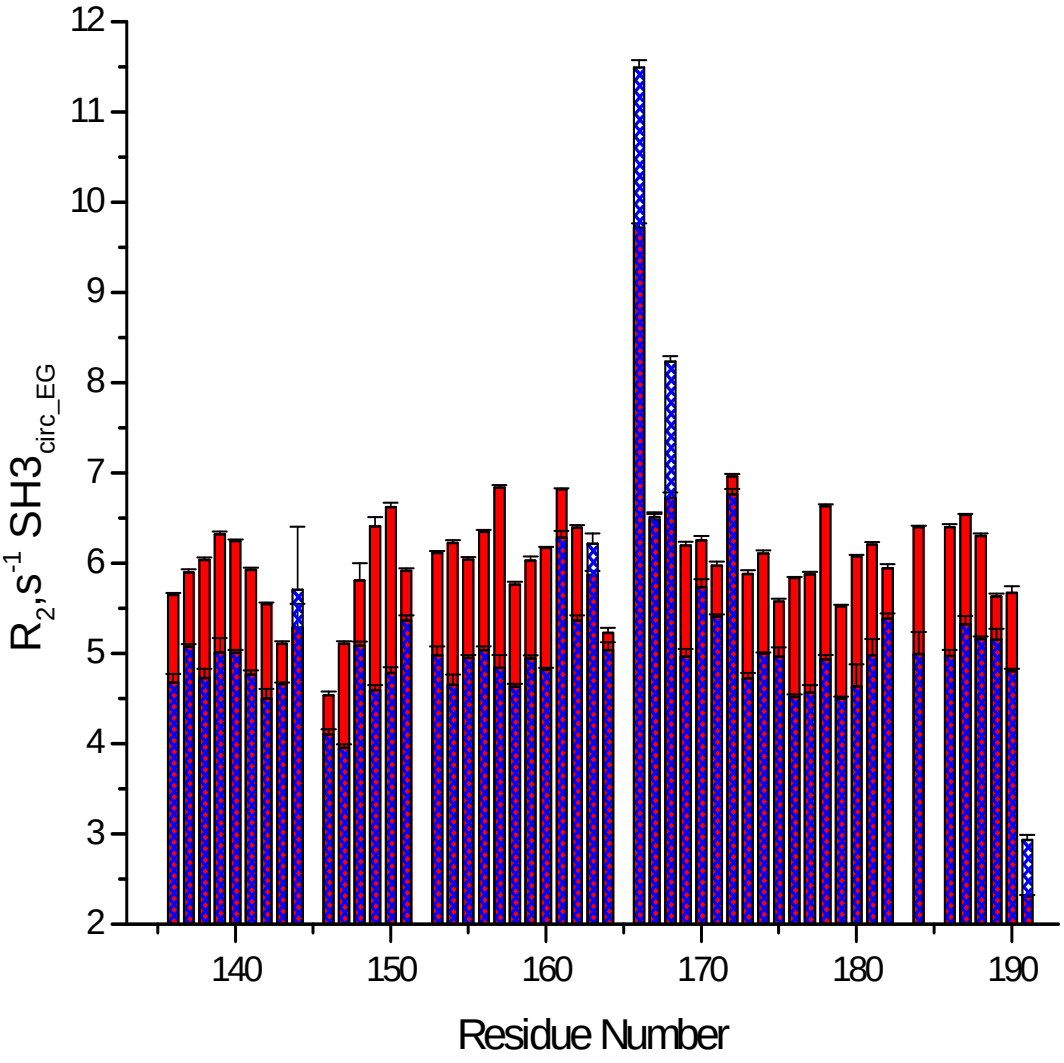

Fig 5.5 Comparison of the R_2 values measured for the free $SH3_{circ\_EG}$ (blue) and $SH3_{circ\_EG}$ in complex with C3G peptide (red)

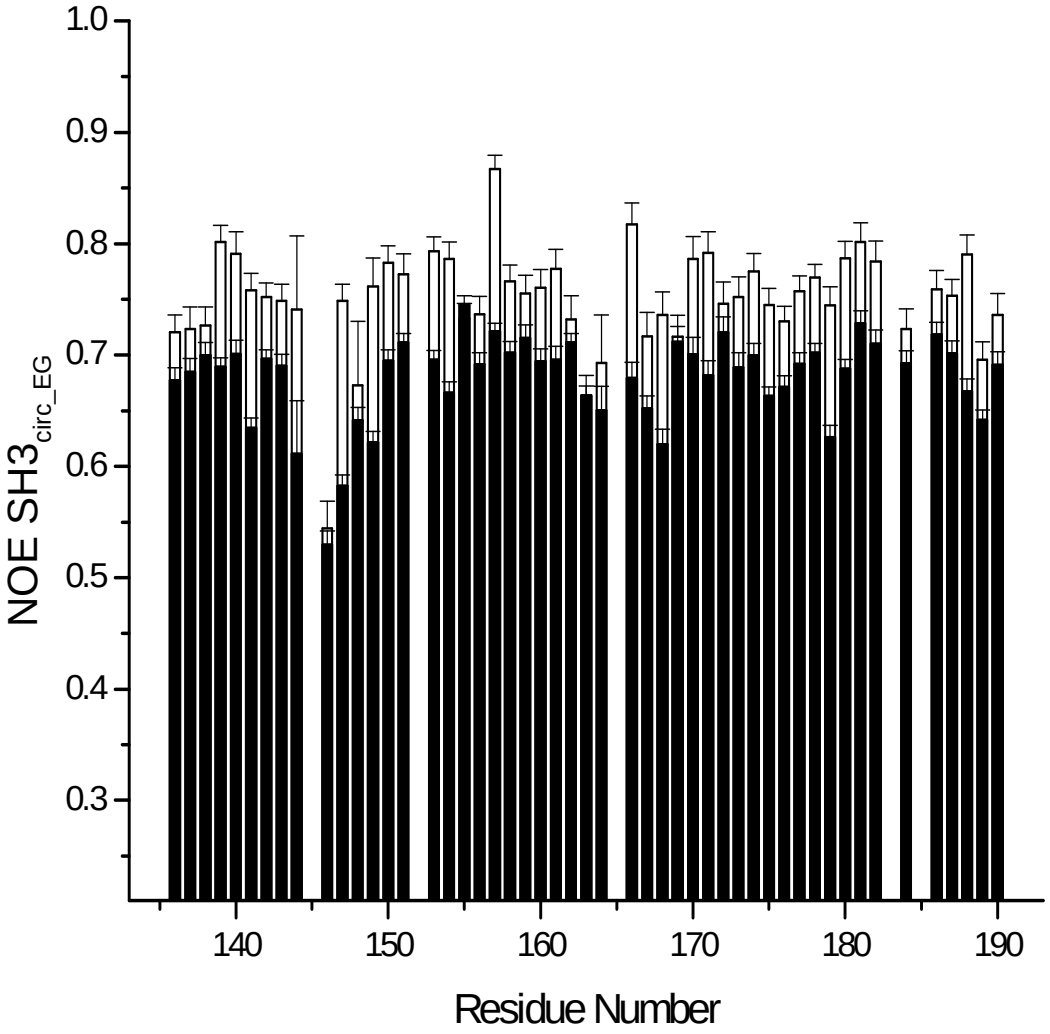

Fig 5.6 Comparison of the NOE values measured for the free $SH3_{circ\_EG}$ (solid bar) and $SH3_{circ\_EG}$ in complex with C3G peptide (open bar)

As can be seen the R_1 values show a decrease for all the residues on ligand binding as we would expect based on the increased mass of the protein ligand complex compared to the free SH3 domains. The R_2 values show an increase for all the residues upon ligand binding as is theoretically expected. The NOE also shows an

increase upon ligand binding. This behavior of R_1, R_2 and NOE is consistent with the increased tumbling time of the complex compared to the free protein.

The R_2 values of the residues 166-168 for the free $SH3_{circ\_EG}$ are higher than the values for the bound state. On the basis of our previous (yet unpublished) data we had concluded that the elevated R_2 values for these residues were most probably due to the conformational exchange. This decrease in the R_2 values on ligand binding for these residues can be attributed to reduced contribution from the conformational exchange.

Section 5.2 The Backbone Dynamics

The backbone dynamics were characterized using the "model-free" approach [30, 31]. The data analysis was done considering the overall tumbling of the protein to be isotropic, as suggested by Schumann et al. (Unpublished data) based on the ratio (1.00:1.07:1.14) of the inertia tensor components. The model-free analysis of ^{15}N relaxation data was performed using our program DYNAMICS [29] that yielded the following values of the overall tumbling time τ_c: 3.91 ns ($SH3_{lin\_wt}$ + C3G peptide), 3.97 ns ($SH3_{circ\_EG}$ + C3G peptide) for 600 MHz data.

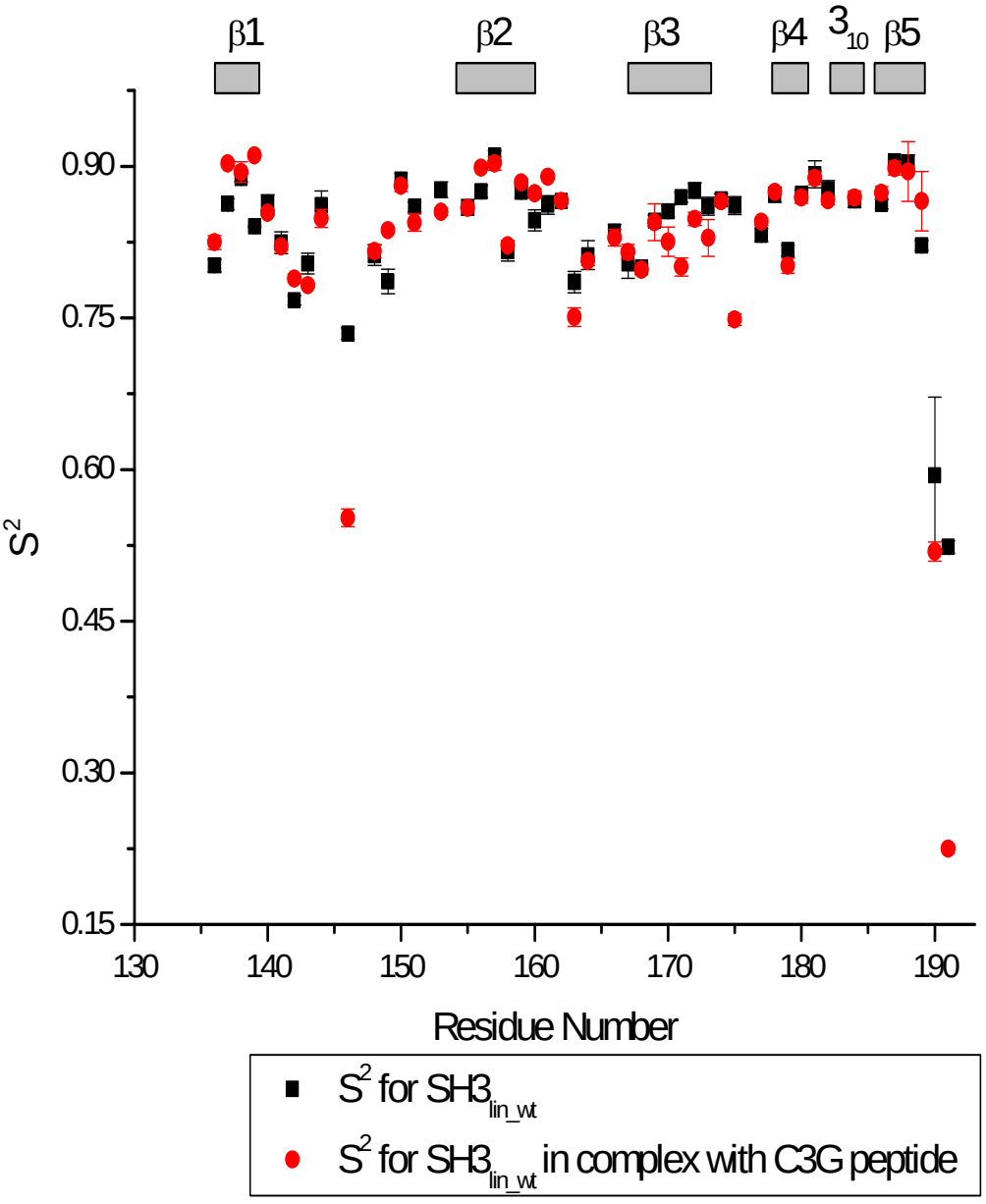

Fig 5.7 Order parameter calculated for the $SH3_{lin\_wt}$ (black square) and $SH3_{lin\_wt}$ in complex with C3G peptide (red circle). At the top of this figure and figures 5.8-5.10 are shown the secondary structure elements of each SH3 domain.

The error bars in figures 5.7-5.10 the experimental error propagated to the dynamics calculations.

The values reported (unpublished work) for the free SH3 domains were 3.18 ns ($SH3_{lin\text{-}wt}$), 3.06 ns ($SH3_{circ\text{-}EG}$) for 600 MHz. These numbers show the increase in the τc values as we would expect based on the mass increase of the protein ligand complex compared to the free protein. The mass of the complex is 1377 Da greater than the mass of the free protein. The model-free parameters derived from the ^{15}N relaxation data are shown in Figure 5.7-5.10

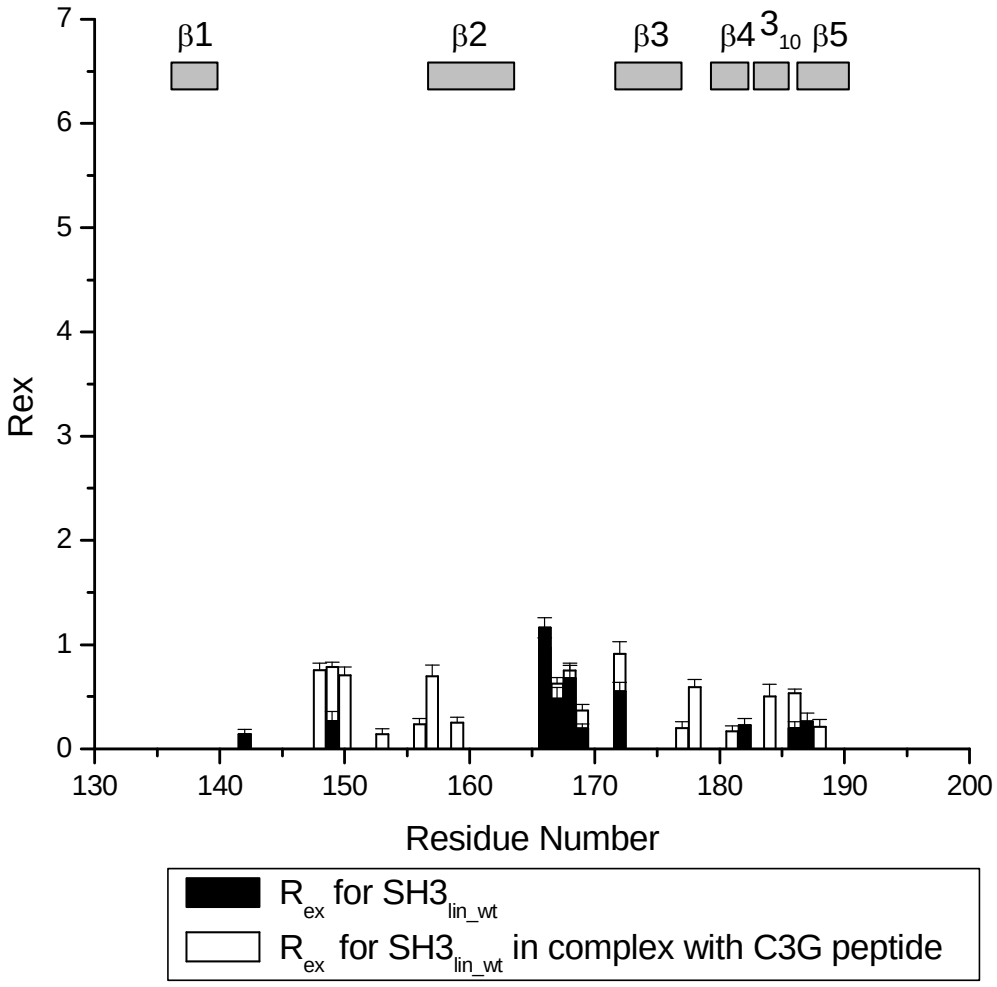

Fig 5.8 Rex calculated for the $SH3_{lin\_wt}$ (solid bars) and $SH3_{lin\_wt}$ in complex with C3G peptide (open bars)

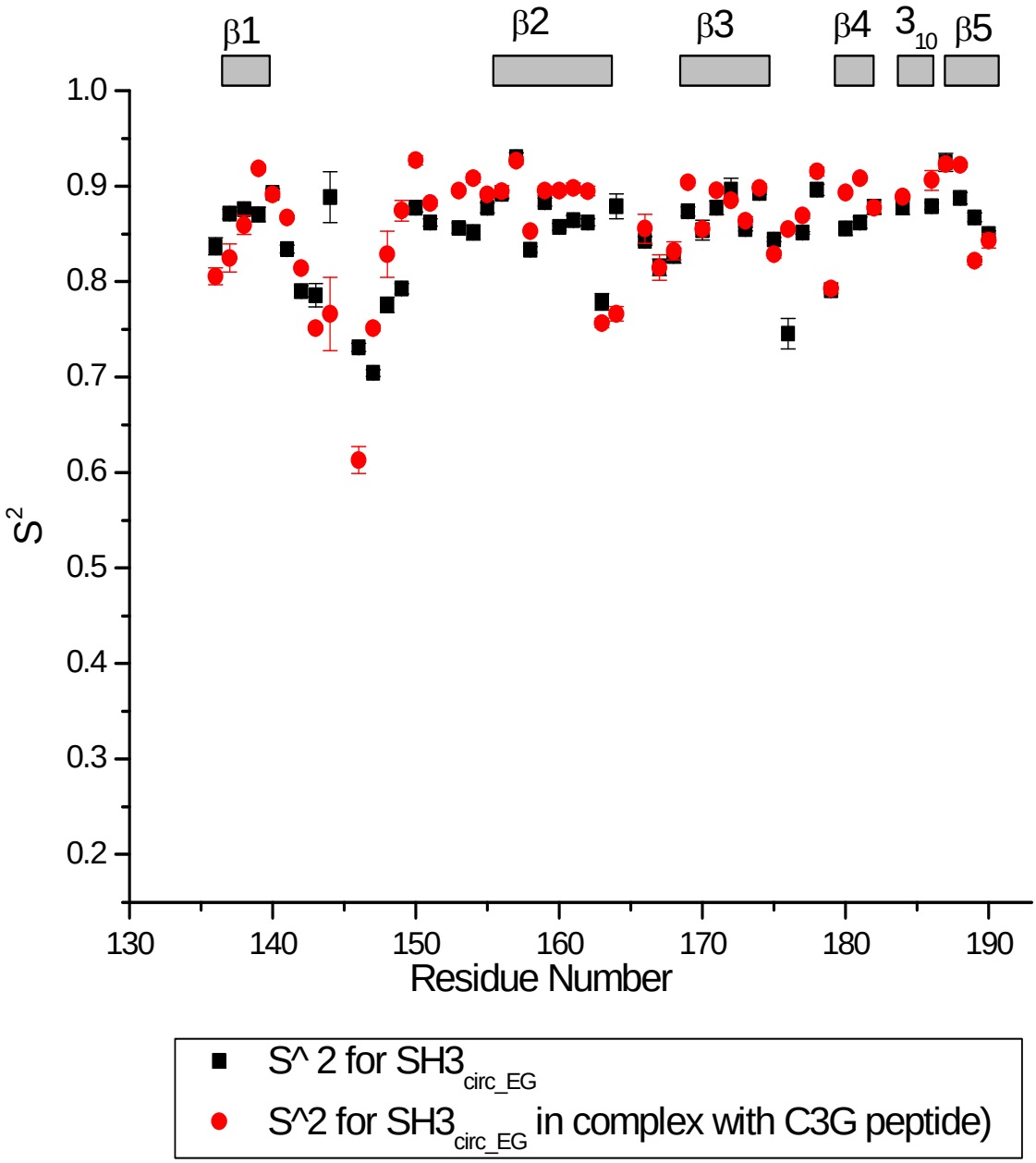

Fig 5.9 Order parameter calculated for the $SH3_{circ\_EG}$ (black square) and $SH3_{circ\_EG}$ in complex with C3G peptide (red circle)

Our in house suit of MATLAB program DYNAMICS returns τc, Rex and the order parameter S^2 as output along with the error bars. The order parameter is a measure of the rigidity of the protein backbone. The Rex term is introduced to account for the contributions to the R_2 values which are presumed to arise from conformational averaging on the micro to millisecond time scale but are not eliminated by the CPMG pulse train in their measurement.

$$R_1=d^2/4[\ J(\omega_H-\omega_N) + 3J(\omega_N) + 6J\ (\omega_H+\omega_N)] + c^2\ J(\omega_N) \quad (1)$$

$$R_2=1/T_2= (d^2/8)\ [4J(0) + J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H+\omega_N)] + (C^2/6)$$

$$[3J\ (\omega_N) + 4J(0)] + Rex \quad (2)$$

$$NOE = (d^2/4R_1)\ (\gamma_H/\gamma_N)[\ 6J(\omega_H+\omega_N)- J(\omega_H+\omega_N) \quad (3)$$

$$\text{where } d= (\mu_0/4\pi)\ \gamma_H\gamma_N\ (h/2\pi)\ (r_{NH})^{-3} \quad (4)$$

$$c=\omega_N\ (\sigma_{//}-\sigma_{\perp})/\ (3)\text{^}1/2 \quad (5)$$

μ_0=permeability of the free space, γ_H and γ_N are the gyro magnetic ratios of ^1H and ^{15}N (2.6752×10^8 and -2.712×10^7 rad s^{-1} T^{-1} respectively), ω_H and ω_N are the Larmor frequencies of ^1H and ^{15}N, respectively. r_{NH} is the N-H bond length ($\approx 1.02\ A^0$) and $J(\omega_i)$ are the spectral densities at the angular frequencies ω_i. $\sigma_{//}-\sigma_{\perp}$ is the anisotropy of the ^{15}N chemical shift tensor.

The amplitudes and effective correlation times of the internal motions of protein are calculated from the R_1, R_2 and NOE data using the model-free formalism. In this analysis $J(\omega)$, the spectral density function is modeled depending on whether the rotational diffusion tensor is isotropic or anisotropic. For our analysis as already mentioned we consider the molecule isotropic. When the internal motions of the NH bond occur on two fast but significantly different timescales they are characterized by

two effective correlation times' τ_f and τ_s corresponding to fast and slow motions respectively where $\tau_f << \tau_s << \tau_c$.

$$J(\omega)=(2/5)\{[(S^2\ \tau_c/(1+(\omega\tau_c)^2)+[(1-S_f^2)\ \tau_f'\ /(1+(\omega\tau_f')^2)]+$$
$$[(S_f^2- S^2)\ \tau_s'/(1+(\omega\tau_s')^2)]\} \quad (6)$$

$$\text{where } (1/\tau_f') = (1/\tau_f)+(1/\tau_c) \quad (7)$$

$$(1/\tau_s') = (1/\tau_s)+(1/\tau_c) \quad (8)$$

$S^2 = S_f^2 . S_s^2$ is the square of the generalized order parameter which characterizes the amplitude of internal motions of each NH bond, and S_f^2 and S_s^2 are the squares of the order parameters for the internal motions on the fast and slower time scales. τ_c is the overall tumbling or overall correlation time. τ_c is a measure of the time it takes for the molecule to loose its sense of initial direction. Motions represented by the generalized order parameter are in the ps to ns time scale thus it gives information about the ps to ns time scale protein backbone dynamics. The order parameter is a measure of the spatial restriction of the NH bond. $S^2 = 1$ means completely restricted motion, and $S^2 = 0$ means completely free motion.

The squared order parameter values, S^2, in the range of 0.8 - 0.95, observed are characteristic for restricted backbone dynamics of a well-folded protein. The overall pattern of higher and lower values of the order parameter is similar for the two constructs, in the free and bound states of the two SH3 domains reflecting the location of the residues in the protein core or in the flexible/unstructured loops.

Lower values of S^2 in the β1/β2, β2/β3, and β3/β4 loops indicate greater flexibility in these regions of the sequence. The termini are highly flexible in the linear construct as can be seen from the S^2 values.

As expected, the flexibility of the termini is significantly reduced upon circularization in the $SH3_{circ\_EG}$ construct, as indicated by the high value of S^2 for Tyr190.

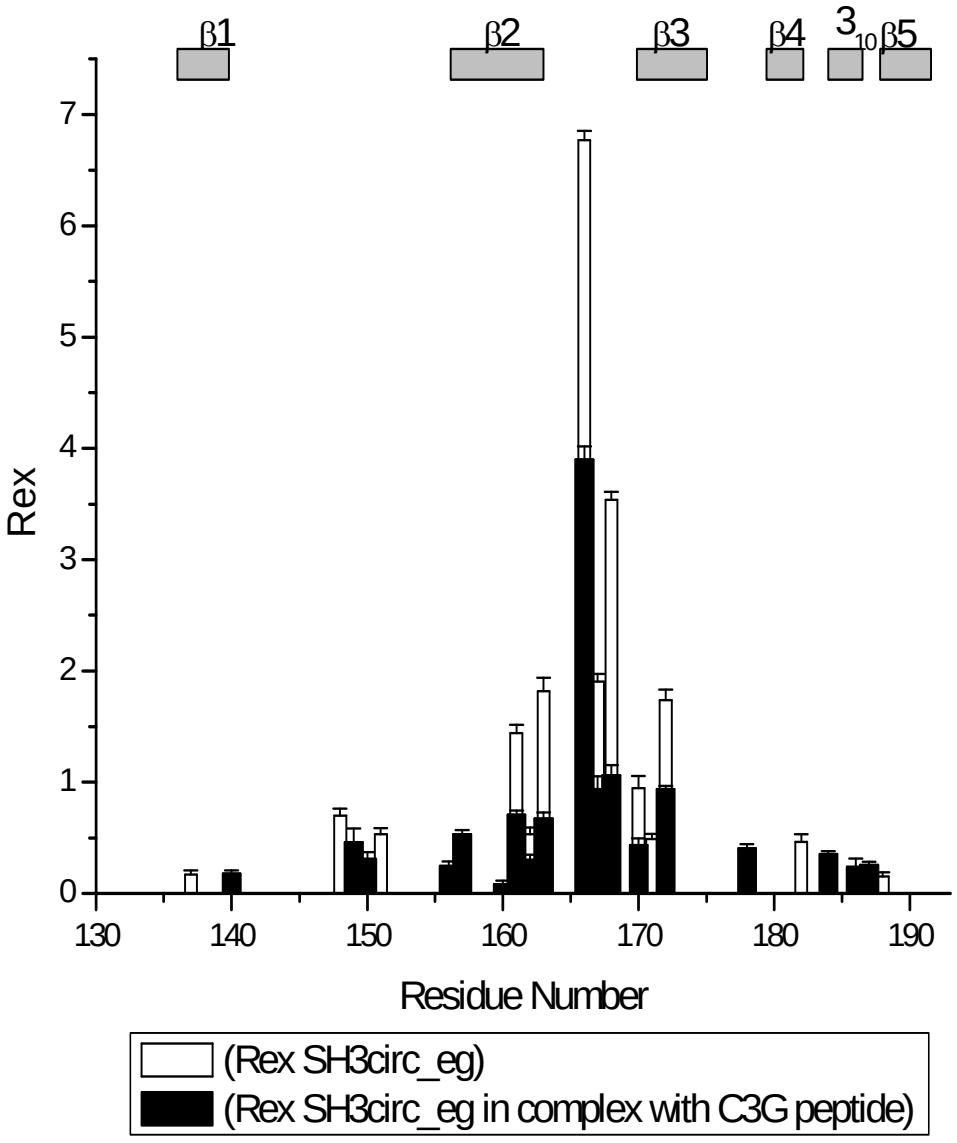

Fig 5.10 Rex calculated for the $SH3_{circ\_EG}$ (open bar) and $SH3_{circ\_EG}$ in complex with C3G peptide (solid bar)

As can be seen from Fig 5.10 the Rex values for the bound state of the $SH3_{circ\_EG}$ are lower compared to the free $SH3_{circ\_EG}$. Hence the ligand binding leads

to a reduction in the conformational exchange contribution to the transverse relaxation.

Chapter 6: Summary and Discussion

Using a combination of NMR methods, we characterized the effect of the ligand binding on the backbone dynamics of the SH3 domain. We also characterized interface of the Murine c-Crk N-SH3 domain in complex with the ligand.

This is the first in solution experimental evidence of the residues on the SH3 domain surface involved in the interface between the SH3 domain and the C3G peptide. The results agree with the crystal structure of the complex. Another important result is the observation that cyclic form $SH3_{circ\_EG}$ and the linear wild type form $SH3_{lin\_wt}$ have same residues involved in the ligand binding.

From the data analysis we see that the Rex contribution to R_1 is reduced by almost 50 percent but still significant Rex is observed for the $SH3_{circ\_EG}$ construct. The carboxyl group of Glu167 makes a direct salt bridge contact with the ε-amino group of Lys9 of the ligand, while the side chain carboxyl group of Glu166 is properly positioned for hydrogen bonding with the amide group of the same residue. Moreover, the indole group of Trp169 is well packed against the aliphatic component of the ligand residue Lys8, one of the critical residues for C3G peptide binding to Crk SH3 [29]. As a result, the ε-amino group of Lys8 is positioned such that it can reach and form a salt bridge with the carboxyl groups of Asp147, Glu149, and Asp150 located in the middle of the RT-loop of SH3.

The Rex contribution is not completely wiped out; the observed Rex i.e. the exchange broadening reflects the presence of slow rearrangements which relieve a circularization-induced strain in the protein structure. As previous data analysis had indicated, the cyclization causes a displacement of the β1 strand which leads to a rearrangement in the proximal β2 strand; the magnitude of the effect depends on the length of the circularization loop. Due to the way the strands are aligned and contact each other (Fig.2), a strain induced by this rearrangement in the C-terminal part of the β2 strand then propagates to the proximal β3 strand, thus also affecting the β2/β3 loop.

Chapter 7: Conclusions

Observed chemical shift perturbations for the $SH3_{lin\_wt}$ upon ligand binding are consistent with the crystal structure indicating that interface of the protein-ligand complex in solution is similar to that in crystal. Also, cyclic form SH3circ_EG and the linear wild type form SH3lin_wt appear to have similar interface in the protein ligand complex.

The analysis of backbone dynamics parameters indicates very small difference in the order parameters between the bound and free constructs of the two domains. But the conformational exchange contribution to transverse relaxation rate for the circular construct is significantly reduced upon ligand binding, indicating that it reduces the available conformational space for the SH3circ_EG. It supports our hypothesis that ligand binding suppresses the possible reorientations of the side chains of residues Lys164 and Glu167. But the fact that we observe residual Rex indicates a possible presence of slow rearrangements which might be involved in relieving a circularization-induced strain in the protein structure as has been previously suggested.

References

1. Trabi, M. and D.J. Craik, *Circular proteins - no end in sight.* Trends in Biochemical Sciences, 2002. **27**(3): p. 132-138.
2. Gran, L., *An oxytocic principle found in Oldenlandia affinis DC. An indigeneous congolese drug "kalata-kalata" used to accelerate discovery.* Meddelelser fra Norsk Farmaceutisk Selskap, 1970. **32**: p. 173-180.
3. Gran, L., *Effect of a Polypeptide Isolated from Kalata-Kalata (Oldenlandia-Affinis Dc) on Estrogen Dominated Uterus.* Acta Pharmacologica Et Toxicologica, 1973. **33**(5-6): p. 400-408.
4. Sletten, K., and Gran, L., *Some molecular properties of kalatapeptide B-1. A uterotonic polypeptide isolated from Oldenlandia affinis DC.* Meddelelser fra Norsk Farmaceutisk Selskap, 1973. **7**(8): p. 69-82.
5. Saether, O., et al., *Elucidation of the Primary and 3-Dimensional Structure of the Uterotonic Polypeptide Kalata B1.* Biochemistry, 1995. **34**(13): p. 4147-4158.
6. Craik, D.J., et al., *Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif.* Journal of Molecular Biology, 1999. **294**(5): p. 1327-1336.
7. Daly, N.L., et al., *Chemical synthesis and folding pathways of large cyclic polypeptide: Studies of the cystine knot polypeptide kalata B1.* Biochemistry, 1999. **38**(32): p. 10606-10614.
8. Gran, L., F. Sandberg, and K. Sletten, *Oldenlandia affinis (R&S) DC. A plant containing uteroactive peptides used in African traditional medicine.*J Ethnopharmacol, 2000. **70**(3): p. 197-203.
9. Hernandez, J.F., et al., *Squash trypsin inhibitors from Momordica cochinchinensis exhibit an atypical macrocyclic structure.* Biochemistry, 2000. **39**(19): p. 5722-30.
10. Felizmenio Quimio, M.E., N.L. Daly, and D.J. Craik, *Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from Momordica cochinchinensis.* J Biol Chem, 2001. **276**(25): p. 22875-82.
11. Luckett, S., et al., *High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds.* J Mol Biol, 1999. **290**(2): p. 525-33.
12. Blond, A., et al., *The cyclic structure of microcin J25, a 21-residue peptide antibiotic from Escherichia coli.* Eur J Biochem, 1999. **259**(3): p. 747-55.
13. Trabi, M., H.J. Schirra, and D.J. Craik, *Three-dimensional structure of RTD-1, a cyclic antimicrobial defensin from rhesus macaque leukocytes.* Biochemistry, 2001. **40**(14): p. 4211-4221.
14. Tang, Y.Q., et al., *A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins.* Science, 1999. **286**(5439): p. 498-502.
15. Hruby, V.J., *Conformational Restrictions of Biologically-Active Peptides Via Amino-Acid Side-Chain Groups.* Life Sciences, 1982. **31**(3): p. 189-199.
16. Hruby, V.J., F. Alobeidi, and W. Kazmierski, *Emerging Approaches in the Molecular Design of Receptor-Selective Peptide Ligands - Conformational,*

Topographical and Dynamic Considerations. Biochemical Journal, 1990. **268**(2): p. 249-262.
17. Ovchinnikov, Y.A. and V.T. Ivanov, *Conformational States and Biological-Activity of Cyclic Peptides.* Tetrahedron, 1975. **31**(18): p. 2177-2209.
18. Thornton, J.M. and B.L. Sibanda, *Amino and Carboxy-Terminal Regions in Globular-Proteins.* Journal of Molecular Biology, 1983. **167**(2): p. 443-460.
19. Goldenberg, D.P. and T.E. Creighton, *Folding Pathway of a Circular Form of Bovine Pancreatic Trypsin-Inhibitor.* Journal of Molecular Biology, 1984. **179**(3): p. 527-545.
20. Jacobson, H. and W.H. Stockmayer, *Intramolecular Reaction in Polycondensations .1. The Theory of Linear Systems.* Journal of Chemical Physics, 1950. **18**(12): p. 1600-1606.
21. Deechongkit, S. and J.W. Kelly, *The effect of backbone cyclization on the thermodynamics of beta-sheet unfolding: Stability optimization of the PINWW domain.* Journal of the American Chemical Society, 2002. **124**(18): p. 4980-4986.
22. Martinez, J.C., et al., *Thermodynamic analysis of alpha-spectrin SH3 and two of its circular permutants with different loop lengths: Discerning the reasons for rapid folding in proteins.* Biochemistry, 1999. **38**(2): p. 549-559.
23. Satoh, T., et al., *Synthetic peptides derived from the fourth domain of CD4 antagonize CD4 function and inhibit T cell activation.* Biochemical and Biophysical Research Communications, 1996. **224**(2): p. 438-443.
24. Camarero, J.A., et al., *Rescuing a destabilized protein fold through backbone cyclization.* Journal of Molecular Biology, 2001. **308**(5): p. 1045-1062.
25. Knudsen, B.S., S.M. Feller, and H. Hanafusa, *4 Proline-Rich Sequences of the Guanine-Nucleotide Exchange Factor C3g Bind with Unique Specificity to the First Src Homology-3 Domain of Crk.* Journal of Biological Chemistry, 1994. **269**(52): p. 32781-32787.
26. Kuriyan, J. and D. Cowburn, *Modular peptide recognition domains in eukaryotic signaling.* Annual Review of Biophysics and Biomolecular Structure, 1997. **26**: p. 259-288.
27. Wu, X.D., et al., *Structural Basis for the Specific Interaction of Lysine-Containing Proline-Rich Peptides with the N-Terminal Sh3 Domain of C-Crk.* Structure, 1995. **3**(2): p. 215-226.
28. Wu, X.D., et al., *High- Resolution Crystal-Structure of a Crk Sh3 Domain with Its Specific Ligand.* Protein Engineering, 1995. **8**: p. 38-38.
29. Fushman, D., S. Cahill, and D. Cowburn, *The main chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: Analysis of ^{15}N relaxation with monomer/dimer equilibration.* J. Mol. Biol., 1997. **266**: p. 173-194.
30. Lipari, G. and A. Szabo, *Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity.*Jour nal of the American Chemical Society, 1982. **104**(17): p. 4546-4559.
31. Lipari, G. and A. Szabo, *Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of*

Experimental Results. Journal of the American Chemical Society, 1982. **104**(17): p. 4559-4570.