Local and global conformational changes in signaling proteins studied by solution NMR spectroscopy

Permalink
https://escholarship.org/uc/item/4s2478p3

Author
Kruse, Nicole K.

Publication Date
2006

Peer reviewed|Thesis/dissertation
Local and Global Conformational Changes in Signaling Proteins Studied by Solution NMR Spectroscopy

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Chemistry by Nicole K. Kruse

Committee in charge:
Professor Patricia A. Jennings, Chair
Professor Robert E. Continetti
Professor Alexandra C. Newton
Professor Joseph P. Noel
Professor Stanley J. Opella

2006
The dissertation of Nicole K. Kruse is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2006
# TABLE OF CONTENTS

Signature Page ........................................................................................................ iii

Table of Contents ...................................................................................................... iv

List of Figures and Tables ......................................................................................... v

List of Abbreviations ................................................................................................. vii

Acknowledgements ..................................................................................................... ix

Vita .............................................................................................................................. xi

Abstract ..................................................................................................................... xii

Chapter 1. Introduction ................................................................................................. 1

Chapter 2. The Carboxyl Tail of PKC is Dynamically Unstructured .......................... 17

Chapter 3. PKC Undergoes a Pin1 Catalyzed Isomerization:
A Complex Twist to the Tail ..................................................................................... 46

Chapter 4. The Unfolding of IL-1β Monitored by PFG-Diffusion NMR:
Looking for a Molehill Next to the Mountain ...................................................... 82

Chapter 5. Conclusion and Future Directions .......................................................... 111

Chapter 6. Materials and Methods ........................................................................... 129

Appendix I. Measuring Diffusion by NMR ................................................................. 141

Appendix II. Acquiring High Quality $^{19}$F NMR Spectra .......................................... 170

References .................................................................................................................. 177
LIST OF FIGURES AND TABLES

Figure 1.1   Cartoon illustrating cellular signaling ....................................................... 4
Figure 1.2   Cartoon illustrating the structure of G-protein coupled receptors ............ 5
Figure 1.3   Crystal Structure of PKA ........................................................................ 7
Figure 1.4   Source of the observable chemical shift in NMR..................................... 9
Figure 1.5   Previous model of PKC regulation .......................................................... 12
Figure 2.1   Domain architecture and family classification of PKC isoforms .............. 20
Figure 2.2   Sequence alignment of PKC isoforms ..................................................... 24
Figure 2.3   Helical wheel of a portion of PKC-CT ..................................................... 26
Figure 2.3   \(^1^H,^{15}N\) HSQC spectra of PKC-CT constructs ........................................ 29
Figure 2.5   Portion of the strip plot from WT PKC-CT ............................................. 32
Table  2.1 Chemical shift assignments of the WT-PKC-CT construct .................... 33
Table  2.2 Chemical shift assignments of the T641E- PKC-CT construct ................. 34
Table  2.3 Chemical shift assignments of the S660E- PKC-CT construct ............... 35
Figure 2.6   CSI analysis of PKC-CT constructs ......................................................... 38
Figure 2.7   \(^{13}C\)\(^\alpha\) chemical shift difference between WT, T641E and S660E .......... 40
Figure 3.1   Time dependent changes observed for all PKC-CT constructs .......... 49
Figure 3.2   Structure of cis and trans prolyl-peptide bonds .................................. 551
Figure 3.3   Time-dependent trypsin cleavage assay ................................................. 53
Figure 3.4   Pin1 catalyzes changes in PKC-CT ......................................................... 55
Figure 3.5   The interaction between PKC-CT and Pin1 is specific ................................ 58
Figure 3.6   NMR evidence of a minor cis population in PKC-CT\(^{13}\) ....................... 61
Figure 3.7   TOCSY spectra of WT-PKC-CT\(^{13}\) pre and post Pin1 catalysis .............. 62
Figure 3.8   TOCSY and ROESY spectra of WT PKC-CT\(^{13}\) post Pin1 .................... 63
Figure 3.9   TOCSY spectra of pT-PKC-CT\(^{13}\) pre and post Pin1 catalysis ............ 64
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>one dimensional</td>
</tr>
<tr>
<td>2D</td>
<td>two dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>BIS-TRIS</td>
<td>2,2-Bis(hydroxymethyl)-2,2',2''-nitrilotriethanol</td>
</tr>
<tr>
<td>BPTI</td>
<td>Bovine Pancreatic Trypsin Inhibitor</td>
</tr>
<tr>
<td>COSY</td>
<td>correlation spectroscopy</td>
</tr>
<tr>
<td>CSI</td>
<td>chemical shift index</td>
</tr>
<tr>
<td>CSK</td>
<td>Carboxyl tail Src Kinase</td>
</tr>
<tr>
<td>DSS</td>
<td>2,2-dimethyl-2-silapentane-5-sulfonate</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia Coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>GdnHCl</td>
<td>guanidine hydrochloride</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GPCRs</td>
<td>G protein-coupled receptors</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-s-transferase</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-Hydroxyethyl)-1Piperazinethanesulfonic Acid</td>
</tr>
<tr>
<td>HSP70</td>
<td>70 kD Heat Shock Protein</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear Single Quantum Coherence</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>IL-1β</td>
<td>Interleukin-1β</td>
</tr>
<tr>
<td>ICE</td>
<td>Interleukin Converting Enzyme</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-Dthiogalactosidase</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>NaAc</td>
<td>sodium acetate</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaPO₄</td>
<td>sodium phosphate</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>NH$_4$OAc</td>
<td>ammonium acetate</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>nuclear overhauser effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>nuclear overhauser effect spectroscopy</td>
</tr>
<tr>
<td>PDK-1</td>
<td>Phosphoinositide dependent kinase</td>
</tr>
<tr>
<td>PFG</td>
<td>pulsed field gradient</td>
</tr>
<tr>
<td>Pin1</td>
<td>peptidyl-prolyl $cis$-$trans$ isomerase</td>
</tr>
<tr>
<td>PKA</td>
<td>Protein Kinase A</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein Kinase C</td>
</tr>
<tr>
<td>PKC-CT</td>
<td>Carboxyl Tail (residues 628-673) of PKC-$\beta$ III</td>
</tr>
<tr>
<td>PKC-CT$_{13}$</td>
<td>13 residue peptide (residues 633-646) based on PKC-CT</td>
</tr>
<tr>
<td>PONDR</td>
<td>predictors of natural disordered regions algorithm</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>ROESY</td>
<td>rotating frame overhauser effect spectroscopy</td>
</tr>
<tr>
<td>TRIS</td>
<td>Tris(Hydroxymethyl) Aminomethane</td>
</tr>
<tr>
<td>TOCSY</td>
<td>total correlation spectroscopy</td>
</tr>
<tr>
<td>TPPI</td>
<td>time proportional phase incrementation</td>
</tr>
<tr>
<td>Tween 20</td>
<td>Polyoxylethylene 20-Sorbitan Monolaurate</td>
</tr>
<tr>
<td>WATERGATE</td>
<td>WATER suppression by GrAdient-Tailored Excitation</td>
</tr>
</tbody>
</table>

**Other NMR Terms**

- BP-LED
- DIPSI
- Echo-AntiEcho
- (H)CC(CO)NH-TOCSY
- HNCA
- HNCACB
- MLEV-17
- States-TPPI
- WALTZ-16
ACKNOWLEDGEMENTS

Like so many milestones in life, the completion of a Ph.D. is not possible without the continuous help and support of numerous individuals.

I would first like to thank my advisor Patricia Jennings. Especially in the middle of my graduate career, when projects seemed to be going nowhere, she was the voice of unrelenting optimism, and never lost faith in my scientific ability. thankfully she would not let me give up. I would also like to thank my committee members, Professors Stanley Opella, Joseph Noel, Alexandra Newton and Robert Continetti, for their valuable help and input throughout the years.

There have also been a number of Jennings lab members, past and present who deserve acknowledgement, including; Christine Mullen, Dave Heidary, John Finke, Jamie Covault, Rosa Fayos, Mette Andersen, Yuki Horiuchi, Leigh Plesniak, Lilly Wong, Dominique Capraro, Ryan Bertsch, Melinda Roy, Jamie Mills, Ben Andrews, Andrea Shoenfish. I would particularly like to thank those people I have shared an office with, Juan-Carlos Rodriguez, Kendra Hailey and especially William Anderson. These three individuals bore the brunt of my daily struggles, have been there for numerous helpful brain-storming conversations, and helped to celebrate the successes. I also want to thank Giuseppe Melacini and Xuemei Huang and Richard Shoemaker for their invaluable NMR assistance. Additionally, I thank Alexandra Newton and Hilde Abrahamsen for their collaborative work with the PKC/Pin1 project.

Last but certainly not least, thank you to my wonderful family for their unending love and support over the years. Especially my parents, who instilled in me from an early age, that I could do anything I wanted to. Despite their initial reaction at the stop light in Valley one Sunday evening when driving me back to Midland my freshman year, of “Why on earth would you want to go to grad school?”, they have since continually supported me on this roller-coaster of a ride.
The research in Chapters 2, 3, and will be submitted for publication as “The PKC-Pin1 Interaction; A Complex Twist to the Tail” N.K Kruse, H. Abramsen, A.C. Newton, P.A. Jennings. Chapter 4 will be submitted for publication as “The Unfolding of IL-1β Monitored by PFG-Diffusion NMR: Looking for a Molehill Next to the Mountain” N. K Kruse, R. K Shoemaker, P.A. Jennings. I was the primary researcher and author for these publications, and the co-authors listed, Hilde Abrahamsen, Alexandra C. Newton, Richard K. Shoemaker, and Patricia A. Jennings directed and supervised the research.
EDUCATION

2000   B.S. Chemistry, and Biology – Magna cum Laude
       Midland Lutheran College, Fremont, NE

2003   M.S., Chemistry
       University of California, San Diego

2006   Ph.D., Chemistry
       University of California, San Diego

PUBLICATIONS


2.  N. K. Kruse, H. Abrahamsen, A. C. Newton, P. A. Jennings “The PKC-Pin1 Interaction; A Complex Twist to the Tail”, manuscript in preparation


HONORS AND AWARDS

1999   Award for Excellence in Chemistry from the Nebraska Section of The American Chemical Society

2000-2003  Graduate Assistance in Areas of national Need (GAANN) fellowship

2003-2005  NIH Growth Regulation and Oncogenesis Training Grant (NIH/NCI #T32 CA09523)

2004   First Prize for Outstanding Poster at the Third International Conference on Structural Biology and Functional Genomics
ABSTRACT OF THE DISSERTATION

Local and Global Conformational Changes in Signaling Proteins
Studied by Solution NMR Spectroscopy

by

Nicole K. Kruse
Doctor of Philosophy in Chemistry
University of California, San Diego, 2006
Professor Patricia A. Jennings, Chair

The proteins involved in cell signaling must be tightly regulated in order for the highly orchestrated process to continue in a harmonious manner. One common method of regulating activity is through conformational changes of the protein. Because regions involved in these conformational changes are often intrinsically unstructured, solution NMR spectroscopy provides a powerful tool-box from which to study these processes, on both a residue specific and global level.

The first aim of this research was to explore the conformation of the carboxyl tail of Protein Kinase C. This region of PKC is key to the maturation and regulation of this ubiquitous kinase. It was discovered that regardless of phosphorylation state the last 48 residues of PKC (PKC-CT) demonstrate a helical propensity, but overall this region is intrinsically unstructured, and the functional importance of this property is addressed.

Additionally, by investigating time dependent changes which were repeatedly observed in the spectra, an intrinsic proline isomerization in PKC-CT was revealed. More importantly, it was discovered that this isomerization can be catalyzed by the peptidyl-prolyl isomerase Pin1.
A new model for the maturation and regulation of PKC, addressing some of the previously unknown aspects of the process is proposed.

The second aim of this research was to implement additional NMR methods for studying conformational changes in proteins. PFG-Diffusion experiments are an elegant means for studying the global structure of a protein under a variety of solution conditions. However, the precision to which one can determine the hydrodynamic radius via this method is not well established. There are many conformations involved in the biology of the pleiotropic cytokine Interleukin 1-β and thus provides a well behaved yet biologically relevant system on which to implement these experiments. Because all known sources of error were addressed and the measurements were acquired in such replicate, it was possible to determine the lower limit to conformational changes which can be detected. It was also shown that an external sample can be used for a viscosity control. Finally, methods for simplifying spectra through specific labeling with $^{19}$F were explored, and experimental protocols for acquiring artifact free spectra were implemented.
Chapter 1

Introduction
The Structure-Function Paradigm

The accepted paradigm in biochemistry and structural biology in the 20th century was that a protein’s particular three-dimensional structure was a prerequisite for its unique function. In the early 1900’s the predominate view was the “lock and key” mechanism where the protein and binding partner, either small molecule or protein, are rigid entities which fit together with unique surface complementarity (Fischer 1894 as translated in (Lemieux and Spohr 1994)). In the 1950’s it was recognized this model did not fit all systems; some proteins undergo configurational adaptability (Karush 1950). Thus the concept of “induced fit” was introduced (Koshland 1958), where in addition to its unique structure, conformational changes were important for function. Since then, it has been readily accepted that proteins are not rigid molecules, but inherently flexible and that while the three dimensional structure is undoubtedly important for function, so are these flexible conformational changes (Petsko and Ringe 1984).

The 21st century has called for an even further re-examination of this structure-function paradigm. Evidence is accumulating to support the notion of biologically functional intrinsically unstructured proteins. This class of systems spans folded proteins which contain unstructured regions necessary for function to proteins which are entirely unstructured under native conditions. It is estimated that 30-51% of the proteins in the eukaryotic genome contain large disordered segments (Dunker et al. 2000; Ward et al. 2004). The prominence of these intrinsically unstructured, yet functional systems is reflected in the growing number of reviews published on this topic (Wright and Dyson 1999; Uversky et al. 2000; Dunker et al. 2001; Dunker et al. 2002; Dunker et al. 2005; Fink 2005). These intrinsically unstructured proteins are recognized for their wide range of functions including, but not limited to cytoskeletal, ribosomal, transport, biosynthesis and degradation. Interestingly, some of the most predominant functions among these disordered regions are those of protein-protein interactions (Dunker et al. 2002) and cell signaling and regulation (Iakoucheva et al. 2002).
Cell signaling: Regulation through conformational changes

As illustrated in the cartoon in Figure 1.1, cell signaling is the process by which an external stimulus is transmitted from outside the cell to inside the cell, and is a perfect example of the functional importance of conformational changes within the protein players. Extracellular ligands, (hormones, neurotransmitters, cytokines, and other small molecules) bind to a cell surface receptor, triggering conformational changes in the receptor. These conformational changes activate a cascade of events known as signal transduction, which result in a variety of cellular responses. One of the best characterized receptor families of proteins are the G protein-coupled receptors (GPCRs). While representing a diverse family of proteins, all GPCRs consist of seven transmembrane helices connected by alternating extracellular and intracellular loops, (Figure 1.2). The exact conformational changes involved in ligand induced activation are not well understood. However, through a variety of biophysical and biochemical characterizations of multiple GPCRs including Rhodopsin (Hubbell et al. 2003), the Muscarinic Acetylcholine Receptor (Han et al. 2005), and the β2-adrenergic receptor (Gether et al. 1997) it is generally accepted (Jensen et al. 2001) that upon ligand binding, there are changes in the relative packing of the transmembrane helices, particularly TM III and TM VI; which in turn cause rearrangements of the intracellular loops, opening a cleft which promotes binding and activation of a specific G protein heterotrimer.

After the membrane receptors, some of the main proteins involved in the signal transduction of the external stimulus are kinases and phosphatases. The interplay between these two types of proteins is one of the most ubiquitous means of regulating cellular activity. The activity of a substrate protein can either be up or down-regulated through the simple addition of a phosphate group by a kinase. The removal of this covalent switch by the phosphatase has the reverse effect on activity. Thus the highly controlled activity of these kinases and phosphatases is crucial to correct cellular function. Just like the traditional party game telephone, if not well controlled, the initial signal gets confused, and the end result is not what was originally intended. Hence, these signaling proteins employ a multitude of regulatory mechanisms. As eluded to previously, one common mechanism of regulation is through
Figure 1.1

Cartoon illustrating the principles behind cell signaling. An extracellular ligand binds to a receptor on the surface of the cell. This binding event causes conformation changes in the receptor which is transmitted into the cytosol, initiating a cascade of events ultimately resulting in the desired changes in gene expression or cellular response.

Like the party game "telephone" this entire process must be controlled and highly regulated, or the initial signal gets confused, and the end result is not what was intended. As described throughout the text, one such mechanism of regulation is through conformational changes of the various proteins.
Figure 1.2

Cartoon representation of the general structure of the G-protein coupled receptor (GPCR) family of proteins. This conserved structure consists of seven transmembrane helices (colored rods) connected by alternating extracellular and intracellular loops. (grey lines)

Binding of a signaling molecule to the extracellular loops changes the relative packing of the transmembrane helices, particularly helix 3 (in orange) and helix 6 (in purple), which in turn cause rearrangements of the flexible intracellular loops, opening a cleft promoting binding and activation of a specific G protein heterotrimer.
conformational changes of the kinase itself. Because all eukaryotic kinases catalyze the transfer of the γ-phosphate from ATP to a hydroxyl group on a substrate protein, the active conformations of these enzymes are remarkably similar. Figure 1.3 illustrates the crystal structure of the active kinase domain from the prototypical kinase PKA (Zheng et al. 1993).

This same overall fold is conserved among all kinases, and as shown consists of two sub-domains or lobes; the smaller N-terminal lobe and a larger C-terminal lobe. The active site lies in a deep cleft between the two lobes. As shown for Protein Kinase A (PKA), when phosphorylated, the activation loop is stabilized in an open extended conformation allowing for substrate binding (Taylor and Radzio-Andzelm 1994). Compared to the similar active structures however, the inactive conformations are markedly different across the different classes of kinases (Huse and Kuriyan 2002; Nolen et al. 2004). One important difference which has been the subject of several discussions (Goldsmith and Cobb 1994; Morgan and De Bondt 1994; Taylor and Radzio-Andzelm 1994; Johnson et al. 1996; Huse and Kuriyan 2002; Adams 2003; Nolen et al. 2004) is that of the activation loop (highlighted in yellow in Figure 1.3). Consistent with a major function of the intrinsically unstructured proteins or regions is signal transduction, these key regulatory regions are often devoid of electron density in crystal structures (Goldberg et al. 1996; Sicheri et al. 1997; Lei et al. 2000; Wybenga-Groot et al. 2001; Biondi et al. 2002; Meng et al. 2002; Ogawa et al. 2002). In addition to the highly conserved activation loop, there are numerous of other less understood conformational changes which lie outside the core kinase domain and are specific to each kinase.

Experimental techniques for studying conformational changes

There are numerous biophysical methodologies available to provide insight into these regulatory conformational changes and unstructured regions (Dunker et al. 2001). Techniques such as Raman optical activity (Zhu et al. 2005), infrared absorption and circular dichroism (CD) (Keiderling and Xu 2002), fluorescence (Cantor and Schimmel 1980) protease digestion (Markus 1965), H/D exchange (Englander et al. 1972; Zhang and Smith 1993) small angle
Crystal structure of the catalytic domain from the prototypical kinase PKA (Zheng et al. 1993). This same global fold is observed for all kinases. The smaller N-terminal lobe is shown in blue, and the larger C-terminal lobe is in red. The activation loop, a key regulatory element for most kinases is highlighted in yellow, with the phosphorylation site shown in stick representation. The active site catalytic triad conserved among all kinases is highlighted in green stick representation.

For reference later (Chapters 2 and 3) the region corresponding to the carboxyl tail of Protein Kinase C (PKC-CT) is shown in light blue, and the turn motif phosphorylation site highlighted in purple.
scattering (Millett et al. 2002), ultracentrifugation (Kirschner and Schachman 1971); are valuable, but are limited in their resolution. Additionally, flexible and disordered regions occasionally will freeze into a static structure, and can still be identified through X-ray crystallography as flexible regions by a high B factor (Petsko and Ringe 1984). However, this provides only a single snap-shot of one particular member of the conformational ensemble. Usually though, such disorder usually results in missing electron density in the crystal structure, and little information other than the fact that this region is disordered is available.

The NMR Tool Box

Nuclear Magnetic Resonance spectroscopy (NMR) provides a powerful experimental method for examining functionally dynamic and disordered regions with high resolution and accuracy. Unlike X-ray crystallography, the structures calculated from NMR utilize data collected from samples in solution, where functionally relevant motions are not restricted. Part of the appeal of NMR as an experimental biophysical method is the sizeable range of information which can be obtained (Wüthrich 1986; Derome 1987; Cavanagh 1996; Levitt 2001).

As illustrated in Figure 1.4, the principal observable in NMR spectroscopy is that of “chemical shift” which is a variable nuclear precession frequency in the main magnetic field dependent upon magnetic shielding that is ultimately determined by the chemical environment of the nucleus under examination. Because the chemical shift is sensitive to the molecular environment it can be used to gain structural information (Wishart and Sykes 1994b), and propensities to form structure (Radhakrishnan et al. 1998). In addition to the chemical shift interaction, nuclei that are connected via chemical bonds are coupled through scalar, or “J” coupling, which is a perturbation of the chemical-shift precession frequency caused by the interaction between the nuclear spins and the spins of the bonding electrons. Nuclei that are proximal through space, though not connected through chemical bonds, may also be associated using methods that exploit the fact that magnetic dipoles interact through space.
Figure 1.4

Cartoon illustrating the principal behind the NMR observable property of “chemical shift”. When in a magnetic field \((B)\), the magnetic moment \((\mu - \text{black arrow})\) of a spinning nucleus (shown in yellow) will begin to precess \((\omega - \text{blue arrow})\) about this magnetic field at a rate proportional to the strength of the magnetic field according to the equation \(\omega = -\gamma B\), where \(\gamma\) is the gyromagnetic ratio of the nucleus.

The relative rate of this precession is the NMR observable “chemical shift”. The chemical environment of a given nucleus will alter the effective external magnetic field \((B)\), leading to slightly different precession frequencies or chemical shifts for any given NMR active nucleus in a molecule. Therefore, the chemical shift observed is an excellent probe for molecular environment and structural properties.
As experimental NMR spectroscopy has evolved in the past 6 decades, a variety of multiple-pulse and multi-dimensional NMR experiments have been developed to allow the spectroscopist to manipulate the spins nearly at will to garner an enormous amount of detailed structural information from the molecules under study. For example in COSY or TOCSY type experiments, the scalar coupling between two NMR active nuclei can be used to “choreograph” the transfer of magnetization from one nucleus to another, through bonds, to learn which spins are connected covalently. NOESY or ROESY experiments can be utilized to exploit the through-space dipolar cross-relaxation to yield through space information, including detailed internuclear distance information between the spins which is useful when attempting to characterize structure and conformational changes. Unique to NMR is the ability to study dynamics on a variety of timescales by measuring the rate at which the NMR signals relax after excitation (Ishima and Torchia 2000). The combination of chemical-shift, spin-spin coupling, and relaxation provides the researcher with a powerful tool box to explore structure, function, and dynamics.

NMR methods have also been developed to implement precisely calibrated pulsed magnetic-field gradients (PFGs) to spatially encode the nuclear precession frequencies based upon their physical location in the sample. In addition to forming the basis for the Magnetic Resonance Imaging (MRI) tools, which have revolutionized clinical medicine, the ability to spatially encode and decode the nuclear spins using PFG methods allows the NMR spectroscopist to accurately measure the macroscopic molecular motion of the molecules within the samples (Stejskal and Tanner 1965). This motion, commonly referred to as molecular diffusion, depends on the size of the molecule and the nature of the sample, and can yield important information about the structure.

**Research aim 1: Conformation of the carboxyl tail of Protein Kinase C**

The work described within this dissertation covers two specific aims. The first aim was to employ the power of solution NMR to study the structure and conformational changes in the carboxyl tail of Protein Kinase C (PKC). PKC is implicated in numerous essential cellular
events including differentiation, cellular activation, modulation of secretion, apoptosis, cytoskeletal rearrangements, and the immune response, reviewed in (Blobe et al. 1996; Toker 1998), and thus when PKC is improperly regulated, it plays a role in numerous diseases. The initial discovery in the early 1980’s that PKC serves as the major receptor for the potent carcinogenic phorbol esters (Castagna et al. 1982), prompted extensive research into the factors governing the activity and regulation of PKC. These intense efforts have led to a growing understanding of how secondary messengers such as Ca^{2+}, diacylglycerol and phospholipids, as well as phosphorylation events regulate the activity of PKC, and has been extensively reviewed (Nishizuka 1986; Newton 1995; Mellor and Parker 1998; Parekh et al. 2000; Newton 2001; 2003).

Using the βII isozyme as a paradigm, Newton and co-workers assembled the following model (Figure 1.5) illustrating the lifetime and various conformations of PKC (Newton 2003). Newly synthesized PKC associates with a membrane compartment of the cell (Sonnenburg et al. 2001). In this newly synthesized, membrane associated state, PKC is in an open conformation in which the autoinhibitory sequence is removed from the substrate binding cavity (Dutil and Newton 2000). The carboxyl tail of PKC is also exposed in this state and provides a docking site for the up-stream kinase Phosphoinositide Dependent Kinase 1 (PDK-1). PDK-1 preferentially binds the un-phosphorylated tail, and subsequently phosphorylates the activation loop (Thr500 in PKC-βII) (Dutil et al. 1998; Gao et al. 2001). Once PDK-1 is released from the carboxyl tail, PKC proceeds to rapidly autophosphorylate two conserved sites in the tail (Thr641, and Ser660 in βII) (Dutil et al. 1994; Behn-Krappa and Newton 1999). These phosphorylations result in an enzyme which is in a more thermally stable, protease and phosphatase resistant conformation (Edwards and Newton 1997). This “mature” PKC is mostly localized in the cytosol, (Mochly-Rosen and Gordon 1998) then translocates to the membrane in response to cellular signals which elevate Ca^{2+} and diacylglycerol levels (Sakai et al. 1997). This membrane tethered, active PKC is again in an open conformation which is highly sensitive to dephosphorylation and down regulation (Hansra et al. 1996). The dephosphorylated species can accumulate in the detergent insoluble fraction of the cell.
Figure 1.5

Model of PKC regulation as proposed in (Newton 2003). Newly synthesized PKC associates with the membrane, upon which the auto-inhibitory sequence (green rectangle) is removed from the active site, and PDK-1 docks onto the carboxyl terminal tail (shown in red), subsequently phosphorylating the activation loop. Once PDK-1 is released from the tail, PKC undergoes two autophosphorylation events on the C-tail (T641, and S660 in βII). In this cartoon, phosphorylations are indicated by yellow circles. This mature yet in-active kinase is localized in the cytosol in a stable closed conformation. Upon generation of secondary messengers, (diacylglycerol and Ca^{2+} or phorbol esters), PKC translocates back to the membrane where it adopts an open, conformation (highlighted with the yellow star) and phosphorylates its substrates, initiating cascades of events responsible for a plethora of cellular responses. After prolonged activation, PKC is dephosphorylated, and has one of two possible fates; either proteolytic degradation through the ubiquitin-proteosome pathway or HSP70 can bind to the carboxyl tail and rescue PKC from degradation.
(Edwards et al. 1999) and is eventually proteolyzed through the ubiquitin-mediated pathway (Lu et al. 1998). Alternatively, upon de-phosphorylation of the turn motif (Thr641), the carboxyl tail can once again become a docking site for another PKC interacting protein. The 70 kD Heat Shock Protein (HSP70) has been shown to bind to the carboxyl tail of PKC and, through an unknown mechanism rescue PKC from proteolytic degradation and return it to the functional pool of PKC to be re-phosphorylated (Gao and Newton 2002).

Work in the Newton lab has shown that the PKC-CT construct (residues 628-673 of PKCβII) is not only necessary, but sufficient for binding to PDK-1 and HSP70 both in vivo and in vitro, and that glutamate mutations (T641E, and S660E) mimic the phosphorylated residues (Gao et al. 2001; Gao and Newton 2002). Owing to its importance in protein/protein interactions and the evolution of activity, we became interested in characterizing the solution conformations of PKC-CT. Chapter 2 of this dissertation presents the work aimed utilizing solution NMR methods to address the structure of this key regulatory portion of PKC and the effects of phosphorylation/mutation. It was determined that regardless of phosphorylation state, there was a propensity for PKC-CT to sample the region of Φ/Ψ space consistent with helical structure which could be stabilized in the presence of additional tertiary contacts with the catalytic domain. However, in the absence of such contacts, the carboxyl tail of PKC is intrinsically disordered which has the functional importance of facilitating binding to multiple protein partners that recognize overlapping but not identical determinants.

Chapter 3 of this dissertation continued the structural characterization of PKC-CT and describes our discovery that the peptidyl-prolyl isomerase Pin1 catalyzes the proline isomerization of the carboxyl tail. Due to the residue specific nature of the NMR techniques employed here, an unexpected activity of Pin1 was revealed, which is supported by, and helps explain some of the results from the in vivo investigations. This discovery provides an explanation for the discrimination of PDK-1 and HSP70 between the chemically identical nascently un-phosphorylated and de-phosphorylated carboxyl tail of PKC. A new model for the regulation of PKC which includes the Pin1 catalyzed isomerization events is presented.
Research aim 2: Implementing new NMR methodologies

The second aim of the research in this dissertation was to implement additional NMR methods to study conformational fluctuations in proteins. As mentioned previously, NMR is well suited to providing high resolution information on unstructured and disordered conformations of proteins. However, it can also be used to gain insight into the global size and shape of molecules. As discussed above, proteins are flexible molecules with dynamic motions. These motions vary in both extent of motion and time scale from the fast vibrations of individual atoms, to the slow conformational rearrangements of structural elements. However, often overlooked are the translational Brownian motions of the molecule through space. Often simply referred to as diffusion, these motions are dependent on the size and shape of the molecule. As introduced earlier, Pulsed Field Gradient (PFG) NMR offers a versatile method in which to determine the hydrodynamic radius of a molecule through measurement of the diffusion coefficient. This experimental technique has been used to examine both native and non-native conformations of a variety of proteins (Jones et al. 1997; Pan et al. 1997; Wilkins et al. 1999; Weljie et al. 2003) However, the extent to which the precision of these measurements can be made is not well documented.

With the eventual goal of being able to use the PFG-diffusion methods to study conformational changes in proteins, chapter 4 of this dissertation was aimed at implementing the PFG-diffusion methods on a well behaved yet biologically relevant system, and in the process, optimize the precision with which diffusion coefficients could be measured for proteins, ultimately to determine the extent to which one could estimate the hydrodynamic radius.

Interleukin1-β (IL-1β) is a pleiotropic cytokine with multiple biological functions. The size, shape and solubility of IL-1β are constantly changing. IL-1β is initially synthesized as a precursor protein (Pro-IL-1β) (Black et al. 1988), which is cleaved by Interleukin Converting Enzyme (ICE), also known as Caspase1 (Dinarello 1998). The mature IL-1β is then secreted by a non-classical mechanism in which it is hypothesized to be partially unfolded (Rubartelli et al. 1990). Interestingly, IL-1β has been found in the fibrils of atherosclerosis (Dinarello and
Wolff 1993), as well as the amaloid plaques of Alzheimer’s Disease (Griffin et al. 1995; Griffin et al. 2000). So in addition to serving as an excellent theoretical model system, the various conformations of IL-1β are biologically relevant.

In chapter 4, the GdnHCl induced unfolding of IL-1β was monitored by PFG-diffusion NMR, and the hydrodynamic radius of IL-1β was determined for the range of conformations from fully folded to unfolded. For this work, it was determined that the change in hydrodynamic radius must be greater than 1.4% in order to be detected by PFG-diffusion methods. Because such care was taken to eliminate and control for every known source of error, it is proposed that this represents the limit of precision which can be expected for other PFG-diffusion investigations. Additionally, these studies suggest that the previous method of utilizing a small molecule as an internal standard for viscosity (Chen et al. 1995; Jones et al. 1997) may be problematic, and that use of an external standard which more closely mimics the protein sample under investigation is not only possible, but in some cases preferred. In order to facilitate the use of these PFG-diffusion experiments for use in studying conformational changes in the future, Appendix I contains a step by step protocol for setting up these diffusion studies as well as information regarding data analysis.

One of the hurdles that must be overcome when attempting to study disordered proteins, or conformational changes in larger proteins by NMR spectroscopy, is resonance overlap caused by the characteristic lack of dispersion among the individual resonances or sheer number of resonances respectively. A technique which can be applied to overcome this hurdle is specific labeling (Mcintosh and Dahlquist 1990; Kelly et al. 1999). Traditional specific labeling for biomolecular NMR involves incorporating $^{15}$N or $^{13}$C only at selected residues. While these nuclei are important to the biomolecular NMR community, another nucleus, $^{19}$F is commonly overlooked. However, there are several useful properties of the $^{19}$F nucleus (Gerig 1989; Danielson and Falke 1996) make it highly attractive when studying subtle conformational changes. $^{19}$F NMR has successfully been used to study conformational changes in proteins, especially those too large to study by traditional NMR methods (Luck and Falke 1991b; a; Luck et al. 2000; Salopek-Sondi and Luck 2002).
Included in the future directions section of Chapter 5 is work which was done to enable $^{19}$F detection at the UCSD biomolecular NMR facility. It was determined that in spite of not having a probe with a dedicated $^{19}$F channel, by using a spin-echo sequence it is possible to collect high quality $^{19}$F spectra on a variety of different proteins. Appendix II is included as a guideline for re-configuring the DMX500 spectrometer to acquire such $^{19}$F spectra.

Chapter 7 outlines all the specific experimental procedures utilized throughout the body of this work. This includes all protein purification protocols, as well as the specific parameters used for the NMR experiments.

The work presented in this dissertation illustrates the application of a diverse set of solution NMR spectroscopy tools, applied in the investigation of structure and conformational changes in biological systems. A variety of experimental methods were applied with the unifying goal of gaining insight into a wide range of conformational changes important to the function of signaling proteins.
Chapter 2

The Carboxyl Tail of PKC is Intrinsically Disordered
Introduction

There are numerous methods of post-translational modification employed by eukaryotic cells for regulatory purposes, the most ubiquitous being that of phosphorylation. Extra-cellular signals activate a multitude of kinases which covalently add a phosphate group to Ser, Thr, and Tyr residues of numerous different proteins to either up or down regulate their activity. Indicative of the high prevalence of regulation by phosphorylation, 2% of the human genome encodes for kinases, or 575 different proteins. The structurally conserved kinase domain (Figure 1.3) is the third most populous domain in the genome (Lander et al. 2001). The activity of most kinases is actually in part regulated by phosphorylation of the kinase itself. Probably the most abundant and best understood of these kinase phosphorylations is that of the kinase activation loop. Phosphorylation at this site, in some cases, results in the correct positioning of the activation loop to allow substrate binding (Taylor and Radzio-Andzelm 1994). In other cases, this phosphorylation of the activation loop is required for efficient phosphoryl transfer (Adams 2003). However, as examined in this chapter, the activation loop is not the only regulatory phosphorylation site on kinases.

Since the discovery in the early 1980’s that Protein Kinase C (PKC) serves as the major receptor for the potent carcinogenic phorbol esters (Castagna et al. 1982), the factors governing the activity and regulation of PKC have been actively researched. These past two decades have led to some understanding on how phosphorylation events, in addition to secondary messengers such as Ca^{2+}, diacylglycerol and phospholipids allosterically activate PKC, and has been extensively reviewed (Nishizuka 1986; Newton 1995; Mellor and Parker 1998; Parekh et al. 2000; Newton 2001; 2003). Despite all the efforts, the complete story on regulation, including maturation and degradation has remained unclear, partially due to the complication of different isoforms, with unique cellular functions and regulation. The PKC family consists of 10 different isoforms differing mainly in the composition of their N-terminal regulatory region. A cartoon illustrating the domain architecture and classification of the various PKC isoform families is given in Figure 2.1. The differences in composition of the N-terminal regulatory region dictate the secondary messenger response, and thus serves as the
Figure 2.1

Cartoon illustrating the domain architecture and sub-classification of the PKC isoform families. All isoforms have a kinase domain (blue) with three highly conserved phosphorylation sites (yellow circles) The ten different isoforms of PKC are divided into sub-families based on the domain architecture of the regulatory region. All isoforms have a pseudosubstrate region (green) which serves as an autoinhibitory sequence. In the conventional isoforms, directly following this is a C1 domain (purple) which is sensitive to diacylglycerol, and a C2 domain (orange) which is sensitive to Ca$^{2+}$, thus these isoforms are responsive to both. Novel isoforms have a functional C2, domain, but the C1 domain is altered in such a manner in that it doesn’t bind diacylglycerol. Therefore, these isozymes respond to Ca$^{2+}$, but are insensitive to diacylglycerol. Atypical PKC’s lack a C2 domain, and contain a non-ligand binding C1 domain, and thus are not regulated by either Ca$^{2+}$ or diacylglycerol. Of particular importance, and the focus of the work here is the Carboxyl tail V5 region.
basis upon which the different isoforms are further divided into three different families. The conventional isoforms ($\alpha$, $\beta$I, $\beta$II, $\gamma$) represent the best characterized of the various PKC isoforms. At the extreme N-terminal end, there is a pseudosubstrate region which serves as an autoinhibitory sequence. Directly following this is a C1 domain which is sensitive to diacylglycerol, and a C2 domain which is sensitive to Ca$^{2+}$, thus these conventional isozymes are responsive to both. Novel isoforms ($\delta$, $\epsilon$, $\theta$, $\eta$) have a functional C2 domain, but the C1 domain is altered in such a manner in that it doesn’t bind diacylglycerol. Therefore, these isozymes respond to Ca$^{2+}$, but are insensitive to diacylglycerol. Atypical ($\zeta$, $\iota$, $\lambda$) PKC’s lack a C2 domain, and contain a non-ligand binding C1 domain, and thus are not regulated by either Ca$^{2+}$ or diacylglycerol.

Another key difference between the different isoforms is the carboxyl-tail V5 domain. Despite being small (~50 residues) these regions are of incredible functional importance. It was first recognized as a functional domain due to the sequence variance between the $\alpha$, $\beta$, and $\gamma$ isoforms in this region (Coussens et al. 1986). The later determination that a single gene encodes for the two alternatively spliced $\beta$I and $\beta$II isozymes which differ only in this V5 region lends additional support to the functional importance of this domain (Coussens et al. 1987). There have been multiple studies in which chimeric PKC’s have been created varying only in this V5 region, and these studies demonstrate that this carboxyl tail V5 region is responsible for a variety of the isozyme specific responses, such as subcellular localization and translocation (Gokmen-Polar and Fields 1998; Stebbins and Mochly-Rosen 2001; Babwah et al. 2003; Wang et al. 2004). Adding to the functional importance of this domain, as highlighted in Figure 2.1 and 2.2, this V5 region contains two phosphorylation sites highly conserved among all PKC isoforms.

As described in the introduction chapter, using the $\beta$II isozyme as a paradigm, Newton and co-workers assembled a model (Figure 1.5) for the lifetime and regulation of PKC (Newton 2003). In the newly-synthesized membrane associated state, the carboxyl tail of PKC is exposed, providing a docking site for the up-stream kinase PDK-1. PDK-1 preferentially binds this un-phosphorylated tail, and subsequently phosphorylates the activation loop (Thr500 in
PKC-βII) (Dutil et al. 1998; Gao et al. 2001). Release of PDK-1 is the rate limiting step in PKC maturation (Gao et al. 2001). Once PDK-1 is released from the carboxyl tail, PKC proceeds to rapidly autophosphorylate two conserved sites in the tail (Thr641, and Ser660 in βII) (Dutil et al. 1994; Behn-Krappa and Newton 1999). This “mature” PKC is mainly localized in the cytosol, (Mochly-Rosen and Gordon 1998) but then translocates to the membrane in response to cellular signals which elevate Ca\(^{2+}\) and diacylglycerol levels (Sakai et al. 1997). This membrane tethered, active PKC is again in an open conformation which is highly sensitive to dephosphorylation and down regulation (Hansra et al. 1996). This dephosphorylated species can accumulate in the detergent insoluble fraction of the cell (Edwards et al. 1999) and is eventually proteolyzed through the ubiquitin-mediated pathway (Lu et al. 1998). Alternatively, upon de-phosphorylation of the turn motif (Thr641), the carboxyl tail can once again become a docking site for another PKC interacting protein. The 70 kD heat shock protein (HSP70) has been shown to bind to the carboxyl tail of PKC and, through an unknown mechanism rescue PKC from proteolytic degradation and return it to the functional pool of PKC to be re-phosphorylated (Gao and Newton 2002).

Despite this abundance of biological information on PKC, there is comparatively little structural information. There have been crystal and NMR structures of a handful of the regulatory domains (Xu et al. 1997; Sutton and Sprang 1998; Verdaguer et al. 1999; Ochoa et al. 2001). However, at the time this project was initiated, there was no structure of the catalytic domain from any of the 10 PKC isoforms. Most of the available structural information regarding the catalytic domain was obtained through homology modeling of PKC onto the prototypical kinase PKA to which there is a 40% sequence identity (Orl and Newton 1994). However, as shown in Figure 2.2, the carboxyl tail of PKA ends prematurely to that of PKC. As explained above; this region has been shown to be important for isozyme specific activity, and also serves as the docking site for the up-stream kinase PDK-1, an interaction crucial for activating PKC, as well as the docking site for HSP70, an interaction which prolongs the lifetime of PKC by somehow rescuing it from proteolytic degradation. Work in the Newton lab
Figure 2.2

Sequence alignment of the 10 different PKC isoforms in addition to PKA Highlighted in yellow are the conserved phosphorylation sites including the activation loop as well as the sites on the carboxyl tail. Notice that the sequence of PKA, the prototypical kinase to which PKC has been mapped to for structural information ends prematurely to the rest. (Figure adapted from Newton 2001)
has shown that the PKC-CT construct (residues 628-673 of PKCβII) is not only necessary but sufficient for binding to PDK-1 and HSP70 both in vivo and in vitro, and that glutamate mutations (T641E, and S660E) mimic the phosphorylated residues (Gao et al. 2001; Gao and Newton 2002). Owing to importance in protein/protein interactions and the regulation of activity, we became interested in PKC-CT. Specifically, we wanted to investigate the structural basis for this phosphorylation dependent interaction. There have been previous accounts where phosphorylated residues stabilize α-helices (Pullen et al. 1995) as well as the contrary, where the phosphorylations de-stabilize helical structures (Szilak et al. 1997). If plotted on a helical wheel, the carboxyl sequence of PKC was hypothesized to adopt a helical structure (Figure 2.3). When phosphorylated at Ser660, the additional negative charge could destabilize this putative helix both by disrupting the hydrophobic face as well as through un-favorable electrostatic interactions with nearby Glu657.

The work described in this chapter was aimed at addressing the structure of this key regulatory portion of PKC and the effects of phosphorylation/mutation. By employing solution NMR methods such as chemical shift index (CSI) analysis to study the WT and the two phospho-mimetic constructs (T641E and S660E), we found that regardless of phosphorylation state, there was a propensity for PKC-CT to sample the region of Φ/Ψ space consistent with helical structure which could be stabilized in the presence of additional tertiary contacts with the catalytic domain. However, in the absence of such contacts, the carboxyl tail of PKC is intrinsically disordered which has the functional importance of facilitating binding to multiple protein partners that recognize overlapping but not identical determinants.
A portion of the PKC-CT sequence can be mapped to a helical wheel. As indicated, introduction of a negative charge through phosphorylation at Ser660 could de-stabilize the helix both by disrupting the hydrophobic face, as well as potential un-favorable electrostatic interactions with nearby Glu657.
Results

Optimizing protein expression levels

High levels of protein expression are necessary in order for heteronuclear NMR experiments to be feasible. Initial attempts at expression of the GST-PKC-CT construct resulted in final PKC-CT yields of only ~0.5 mg/L. Careful examination of the plasmid sequence which was initially sub-cloned from the rat genome revealed the presence of multiple rare Escherichia Coli (E. coli) codons. When this same plasmid was transformed into cells such as the BL21-Codon Plus E. coli cells from Stratagene which contain extra transfer RNA’s for these rare codons, the expression of all PKC-CT constructs was significantly enhanced. Additionally, it was found that IPTG was toxic to this expression system, and by reducing the concentration of IPTG used for induction, the final yields of PKC-CT increased about 8 fold.

Initial comparison of PKC-CT and phospho-mimetic constructs

The $^1$H-$^{15}$N HSQC spectra of the WT and two phospho-mimetic constructs (T641E and S660E) of PKC-CT at 600 MHz are given in Figure 2.4. All expected resonances are present. Two further things are apparent upon first inspection. First and foremost, in all spectra the backbone amide cross-peaks are all clustered within a narrow range of just over 0.7 ppm in $^1$H. This lack of dispersion among the resonances is consistent with a predominately unstructured sample. Additionally the overall pattern remains the same in all three constructs, indicating that the phospho-mimetic mutations have not drastically altered the conformation of PKC-CT.

Sequence specific resonance assignments

The backbone ($^1$H$^N$, $^{15}$N, and $^{13}$C$^\alpha$) as well as complete $^{13}$C side-chain resonances were assigned by analysis of HNCA, HNCACB, and (H)CC(CO)NH-TOCSY spectra (Sattler et al. 1999). The (H)CC(CO)NH-TOCSY spectra was used primarily for residue type information, as under the conditions employed here, it contains the $^{13}$C chemical shift for the entire side chain spin system. Given the number of similar residues in PKC-CT, the HNCACB
Figure 2.4

$^1$H-$^{15}$N HSQC spectra of WT and two phospho-mimetic PKC-CT constructs. From first inspection it is clear that there is conservation of the overall chemical shift pattern, consistent with no major structural changes upon phospho-mimetic mutation. In addition, the lack of $^1$H chemical shift dispersion in all spectra is indicative of an unstructured protein.
spectra proved to be especially useful for residue specific assignments, due to this experiment containing the inter- and intra-residue cross peaks for both the $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ resonances. Illustrating the useful combination of the HNCACB and (H)CC(CO)NH-TOCSY experiments, strip plots highlighting the resonances and connectivities for residues Asp644-Gln653 for the WT-PKC-CT construct are given in Figure 2.5. Sequence specific assignments for all expected resonances of all three constructs were obtained, the backbone amide crosspeaks are labeled in the HSQC spectra in Figure 2.4. A few minor crosspeaks in the HSQC’s were not assigned, and are probably the result of conformational heterogeneity. The complete list of obtained resonance assignments for the WT, T641E, and S660E, are included in Tables 2.1, 2.2 and 2.3, respectively.

**CSI analysis**

The $^{13}\text{C}^{\alpha}$ chemical shifts from the HNCA spectra were used to assess secondary structure propensities. The Chemical Shift Index (CSI) method of analysis provides guidelines to which the experimentally determined $^{13}\text{C}^{\alpha}$ chemical shifts can be compared allowing for classification of each residue into one of three secondary structure elements, $\alpha$-helical, $\beta$-sheet, or random coil (Wishart and Sykes 1994a). Values known as secondary shifts ($\Delta\delta$) are the difference between the experimentally determined $^{13}\text{C}^{\alpha}$ chemical shifts and published random coil values. In general, for stable secondary structural elements, $^{13}\text{C}^{\alpha}$ chemical shifts experience approximately 2.5 ppm downfield shift in $\alpha$-helices, and about a 2.0 ppm upfield shift in $\beta$-sheets as compared to a random coil chemical shift (Wishart and Case 2001). Analysis of secondary shifts has also been particularly useful in identifying conformational preferences towards secondary structure in intrinsically unstructured proteins (Radhakrishnan et al. 1998; Bussell and Eliezer 2001; Yao et al. 2001; Lin et al. 2002). Given that the HSQC spectra of all PKC-CT constructs indicate disorder, the CSI analysis should be particularly insightful in this regard. Because the secondary shift values were predicted to be small, and $^{13}\text{C}^{\alpha}$ chemical shifts can also be affected by neighboring residues, the sequence dependent corrections of (Schwarzinger et al. 2001) were used in the analysis.
Figure 2.5

Individual $^{15}$N slices for specific residues (Asp644 – Gln653) obtained from the HNCACB and (H)CC(CO)NH-TOCSY experiments for the WT-PKC-CT peptide. The HNCACB spectrum is in red (positive peaks) and blue (negative peaks) while the (H)CC(CO)NH-TOCSY is in black. The residue and corresponding $^{15}$N frequency are labeled at the top of each pair of strips. As described in the text, the HNCACB provides inter-residue connectivities between the alpha (red) and beta (blue) carbons, and are drawn in for clarity. For each cross peak in the HSQC, the (H)CC(CO)NH provides the $^{13}$C chemical shifts for the entire side chain of the preceding residue, which are marked with an X in the figure.
Table 2.1 Chemical shift assignments of the WT-PKC-CT construct *

<table>
<thead>
<tr>
<th>Residue</th>
<th>$^1$HN</th>
<th>$^{15}$N</th>
<th>$^{13}$C$_{\alpha}$</th>
<th>$^{13}$C$_{\beta}$</th>
<th>$^{13}$C$_{\gamma}$</th>
<th>$^{13}$C$_{\delta}$</th>
<th>$^{13}$C$_{\varepsilon}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn 628</td>
<td>8.41</td>
<td>120.0</td>
<td>53.1</td>
<td>38.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 629</td>
<td>8.30</td>
<td>121.2</td>
<td>58.0</td>
<td>39.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 630</td>
<td>8.34</td>
<td>121.8</td>
<td>54.4</td>
<td>40.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 631</td>
<td>8.16</td>
<td>120.9</td>
<td>56.5</td>
<td>29.9</td>
<td>26.6</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>Phe 632</td>
<td>8.20</td>
<td>119.7</td>
<td>57.9</td>
<td>38.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 633</td>
<td>8.03</td>
<td>120.8</td>
<td>58.0</td>
<td>39.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr 634</td>
<td>8.05</td>
<td>116.5</td>
<td>61.9</td>
<td>69.3</td>
<td>21.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 635</td>
<td>8.27</td>
<td>123.5</td>
<td>55.9</td>
<td>30.5</td>
<td>27.2</td>
<td>43.1</td>
<td></td>
</tr>
<tr>
<td>His 636</td>
<td>8.41</td>
<td>121.3</td>
<td>53.8</td>
<td>28.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 637</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 638</td>
<td></td>
<td></td>
<td>63.1</td>
<td>32.1</td>
<td>27.4</td>
<td>50.4</td>
<td></td>
</tr>
<tr>
<td>Val 639</td>
<td>8.36</td>
<td>121.1</td>
<td>62.3</td>
<td>32.6</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 640</td>
<td>8.52</td>
<td>127.0</td>
<td>54.8</td>
<td>42.0</td>
<td>27.0</td>
<td>24.83/23.60</td>
<td></td>
</tr>
<tr>
<td>Thr 641</td>
<td>8.28</td>
<td>118.7</td>
<td>59.8</td>
<td>69.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 642</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 643</td>
<td></td>
<td></td>
<td>63.4</td>
<td>31.7</td>
<td>27.4</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Asp 644</td>
<td>8.45</td>
<td>119.4</td>
<td>54.5</td>
<td>40.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 645</td>
<td>8.25</td>
<td>119.9</td>
<td>55.8</td>
<td>29.3</td>
<td>33.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 646</td>
<td>8.44</td>
<td>122.4</td>
<td>56.9</td>
<td>30.0</td>
<td>36.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 647</td>
<td>8.23</td>
<td>122.4</td>
<td>62.7</td>
<td>32.4</td>
<td>21.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 648</td>
<td>8.34</td>
<td>125.9</td>
<td>61.1</td>
<td>37.8</td>
<td>27.54/17.37</td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td>Arg 649</td>
<td>8.49</td>
<td>126.1</td>
<td>55.8</td>
<td>30.8</td>
<td>27.2</td>
<td>43.2</td>
<td></td>
</tr>
<tr>
<td>Asn 650</td>
<td>8.58</td>
<td>121.1</td>
<td>53.1</td>
<td>38.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 651</td>
<td>8.26</td>
<td>121.5</td>
<td>61.6</td>
<td>38.4</td>
<td>27.41/17.93</td>
<td>13.6</td>
<td></td>
</tr>
<tr>
<td>Asp 652</td>
<td>8.48</td>
<td>124.1</td>
<td>54.4</td>
<td>40.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 653</td>
<td>8.42</td>
<td>121.6</td>
<td>56.1</td>
<td>28.9</td>
<td>33.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 654</td>
<td>8.49</td>
<td>117.2</td>
<td>59.3</td>
<td>63.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 655</td>
<td>8.41</td>
<td>122.5</td>
<td>56.7</td>
<td>29.5</td>
<td>35.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 656</td>
<td>8.09</td>
<td>120.1</td>
<td>57.5</td>
<td>39.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 657</td>
<td>8.30</td>
<td>122.7</td>
<td>56.8</td>
<td>29.8</td>
<td>36.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 658</td>
<td>8.04</td>
<td>109.4</td>
<td>45.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 659</td>
<td>8.10</td>
<td>120.1</td>
<td>57.7</td>
<td>39.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 660</td>
<td>8.21</td>
<td>117.4</td>
<td>58.1</td>
<td>63.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 661</td>
<td>8.26</td>
<td>122.6</td>
<td>57.8</td>
<td>39.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 662</td>
<td>8.07</td>
<td>122.0</td>
<td>62.4</td>
<td>32.5</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 663</td>
<td>8.50</td>
<td>122.7</td>
<td>53.2</td>
<td>38.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 664</td>
<td>8.38</td>
<td>117.0</td>
<td>59.0</td>
<td>63.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 665</td>
<td>8.47</td>
<td>122.5</td>
<td>29.7</td>
<td>35.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 666</td>
<td>8.13</td>
<td>120.1</td>
<td>57.7</td>
<td>39.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 667</td>
<td>8.00</td>
<td>123.7</td>
<td>54.6</td>
<td>42.0</td>
<td>24.8</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>Lys 668</td>
<td>8.30</td>
<td>124.2</td>
<td>54.1</td>
<td>32.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 669</td>
<td></td>
<td></td>
<td>63.2</td>
<td>31.8</td>
<td>27.4</td>
<td>50.6</td>
<td></td>
</tr>
<tr>
<td>Glu 670</td>
<td>8.62</td>
<td>122.1</td>
<td>56.4</td>
<td>30.2</td>
<td>36.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 671</td>
<td>8.43</td>
<td>123.6</td>
<td>62.2</td>
<td>32.4</td>
<td>21.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys 672</td>
<td>8.59</td>
<td>127.3</td>
<td>56.1</td>
<td>32.7</td>
<td>24.6</td>
<td>28.8</td>
<td>42.0</td>
</tr>
<tr>
<td>Ser 673</td>
<td>8.15</td>
<td>124.0</td>
<td>60.1</td>
<td>64.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Estimated errors are 0.01 ppm $^1$H, 0.1 ppm $^{15}$N, 0.2 ppm $^{13}$C$_{\alpha}$, and 0.5 $^{13}$C$_{\text{others}}$. 
Table 2.2 Chemical shift assignments of the T641E-PKC-CT construct *

<table>
<thead>
<tr>
<th>Residue</th>
<th>$^1$HN</th>
<th>$^{15}$N</th>
<th>$^{13}$C$_{\alpha}$</th>
<th>$^{13}$C$_{\beta}$</th>
<th>$^{13}$C$_{\gamma}$</th>
<th>$^{13}$C$_{\delta}$</th>
<th>$^{13}$C$_{\varepsilon}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn 628</td>
<td>8.38</td>
<td>120.0</td>
<td>53.2</td>
<td>38.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 629</td>
<td>8.27</td>
<td>121.1</td>
<td>57.9</td>
<td>39.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 630</td>
<td>8.32</td>
<td>121.7</td>
<td>54.5</td>
<td>40.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 631</td>
<td>8.12</td>
<td>120.8</td>
<td>56.6</td>
<td>30.1</td>
<td>26.9</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>Phe 632</td>
<td>8.18</td>
<td>119.7</td>
<td>58.0</td>
<td>39.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 633</td>
<td>8.02</td>
<td>120.7</td>
<td>58.0</td>
<td>39.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr 634</td>
<td>8.03</td>
<td>116.3</td>
<td>61.9</td>
<td>36.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 635</td>
<td>8.24</td>
<td>123.4</td>
<td>56.0</td>
<td>30.6</td>
<td>27.3</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>His 636</td>
<td>3.38</td>
<td>121.3</td>
<td>53.6</td>
<td>29.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 637</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 638</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 639</td>
<td>8.29</td>
<td>120.9</td>
<td>62.3</td>
<td>33.1</td>
<td>21.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 640</td>
<td>8.40</td>
<td>126.7</td>
<td>54.8</td>
<td>42.3</td>
<td>27.0</td>
<td>24.97/23.74</td>
<td></td>
</tr>
<tr>
<td>Glu 641</td>
<td>8.40</td>
<td>123.6</td>
<td>54.1</td>
<td>29.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 642</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 643</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 644</td>
<td>8.41</td>
<td>119.8</td>
<td>54.5</td>
<td>40.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 645</td>
<td>8.30</td>
<td>120.1</td>
<td>56.0</td>
<td>29.5</td>
<td>33.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 646</td>
<td>8.48</td>
<td>122.2</td>
<td>56.8</td>
<td>29.9</td>
<td>36.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 647</td>
<td>8.16</td>
<td>122.2</td>
<td>62.7</td>
<td>32.5</td>
<td>21.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 648</td>
<td>8.28</td>
<td>125.5</td>
<td>61.1</td>
<td>38.0</td>
<td>27.54/17.86</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td>Arg 649</td>
<td>8.45</td>
<td>125.9</td>
<td>55.8</td>
<td>30.9</td>
<td>27.4</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>Asn 650</td>
<td>8.53</td>
<td>120.9</td>
<td>53.2</td>
<td>38.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 651</td>
<td>8.22</td>
<td>121.4</td>
<td>61.5</td>
<td>38.7</td>
<td>27.3/17.89</td>
<td>13.7</td>
<td></td>
</tr>
<tr>
<td>Asp 652</td>
<td>8.45</td>
<td>124.1</td>
<td>54.4</td>
<td>41.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 653</td>
<td>8.40</td>
<td>121.6</td>
<td>56.1</td>
<td>29.0</td>
<td>33.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 654</td>
<td>8.47</td>
<td>117.1</td>
<td>59.4</td>
<td>63.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 655</td>
<td>8.39</td>
<td>122.4</td>
<td>57.0</td>
<td>29.8</td>
<td>36.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 656</td>
<td>8.06</td>
<td>120.0</td>
<td>58.1</td>
<td>30.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 657</td>
<td>8.28</td>
<td>122.6</td>
<td>56.9</td>
<td>30.0</td>
<td>36.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 658</td>
<td>8.01</td>
<td>109.3</td>
<td>45.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 659</td>
<td>8.08</td>
<td>120.1</td>
<td>57.9</td>
<td>39.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 660</td>
<td>8.18</td>
<td>117.3</td>
<td>57.9</td>
<td>63.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 661</td>
<td>8.22</td>
<td>122.5</td>
<td>62.5</td>
<td>39.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 662</td>
<td>8.04</td>
<td>121.8</td>
<td>53.2</td>
<td>32.7</td>
<td>21.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 663</td>
<td>8.47</td>
<td>122.6</td>
<td>59.0</td>
<td>38.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 664</td>
<td>8.34</td>
<td>116.9</td>
<td>57.0</td>
<td>63.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 665</td>
<td>8.45</td>
<td>122.5</td>
<td>57.0</td>
<td>29.8</td>
<td>36.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 666</td>
<td>8.01</td>
<td>120.0</td>
<td>57.8</td>
<td>39.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 667</td>
<td>7.97</td>
<td>123.6</td>
<td>54.7</td>
<td>42.5</td>
<td>26.9</td>
<td>23.74/25.1</td>
<td></td>
</tr>
<tr>
<td>Lys 668</td>
<td>6.28</td>
<td>124.1</td>
<td>54.2</td>
<td>32.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 669</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 670</td>
<td>8.58</td>
<td>122.0</td>
<td>56.4</td>
<td>30.4</td>
<td>36.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 671</td>
<td>8.38</td>
<td>123.4</td>
<td>62.3</td>
<td>32.6</td>
<td>62.3</td>
<td>21.2</td>
<td></td>
</tr>
<tr>
<td>Lys 672</td>
<td>8.55</td>
<td>127.2</td>
<td>56.1</td>
<td>32.9</td>
<td>24.8</td>
<td>29.12/33.57/42.2</td>
<td></td>
</tr>
<tr>
<td>Ser 673</td>
<td>8.11</td>
<td>123.9</td>
<td>60.0</td>
<td>64.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Estimated errors are 0.01 ppm $^1$H, 0.1 ppm $^{15}$N, 0.2 ppm $^{13}$C$_{\alpha}$, and 0.5 $^{13}$C$_{\text{others}}$
Table 2.3 Chemical shift assignments of the S660E-PKC-CT construct *

<table>
<thead>
<tr>
<th>Residue</th>
<th>$^1$HN</th>
<th>$^15$N</th>
<th>$^{13}$C$_{\alpha}$</th>
<th>$^{13}$C$_{\beta}$</th>
<th>$^{13}$C$_{\gamma}$</th>
<th>$^{13}$C$_{\delta}$</th>
<th>$^{13}$C$_{\varepsilon}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn 628</td>
<td>8.38</td>
<td>120.2</td>
<td>53.1</td>
<td>38.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 629</td>
<td>8.26</td>
<td>121.4</td>
<td>58.0</td>
<td>39.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 630</td>
<td>8.32</td>
<td>122.0</td>
<td>54.5</td>
<td>40.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 631</td>
<td>8.11</td>
<td>121.0</td>
<td>56.7</td>
<td>30.3</td>
<td>30.2/26.72</td>
<td>43.2</td>
<td></td>
</tr>
<tr>
<td>Phe 632</td>
<td>8.17</td>
<td>119.9</td>
<td>58.0</td>
<td>39.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 633</td>
<td>8.01</td>
<td>120.9</td>
<td>58.0</td>
<td>39.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr 634</td>
<td>8.02</td>
<td>116.5</td>
<td>61.9</td>
<td>69.6</td>
<td>21.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 635</td>
<td>8.23</td>
<td>123.6</td>
<td>56.0</td>
<td>30.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His 636</td>
<td>8.36</td>
<td>121.5</td>
<td>53.7</td>
<td>29.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 637</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 638</td>
<td></td>
<td></td>
<td>62.9</td>
<td>31.9</td>
<td>27.2</td>
<td>32.1</td>
<td></td>
</tr>
<tr>
<td>Val 639</td>
<td>8.30</td>
<td>121.1</td>
<td>62.2</td>
<td>32.7</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 640</td>
<td>8.47</td>
<td>127.0</td>
<td>54.9</td>
<td>42.2</td>
<td>27.1</td>
<td>24.93/23.52</td>
<td></td>
</tr>
<tr>
<td>Thr 641</td>
<td>8.23</td>
<td>118.8</td>
<td>59.8</td>
<td>69.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 642</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63.3</td>
<td>32.1</td>
<td>27.5</td>
</tr>
<tr>
<td>Pro 643</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63.3</td>
<td>32.1</td>
<td>27.5</td>
</tr>
<tr>
<td>Asp 644</td>
<td>8.41</td>
<td>119.4</td>
<td>54.4</td>
<td>40.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 645</td>
<td>8.20</td>
<td>120.0</td>
<td>55.8</td>
<td>29.5</td>
<td>33.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 646</td>
<td>8.41</td>
<td>122.5</td>
<td>56.9</td>
<td>30.3</td>
<td>36.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 647</td>
<td>8.18</td>
<td>122.4</td>
<td>62.5</td>
<td>32.7</td>
<td>21.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 648</td>
<td>8.30</td>
<td>126.0</td>
<td>61.0</td>
<td>38.2</td>
<td>27.24/17.65</td>
<td>12.9</td>
<td></td>
</tr>
<tr>
<td>Arg 649</td>
<td>8.45</td>
<td>126.2</td>
<td>55.8</td>
<td>31.2</td>
<td>26.9</td>
<td>43.0</td>
<td></td>
</tr>
<tr>
<td>Asn 650</td>
<td>8.55</td>
<td>121.3</td>
<td>53.2</td>
<td>38.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 651</td>
<td>8.22</td>
<td>121.6</td>
<td>61.5</td>
<td>38.8</td>
<td>26.94/17.72</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td>Asp 652</td>
<td>8.44</td>
<td>124.2</td>
<td>54.4</td>
<td>41.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln 653</td>
<td>8.37</td>
<td>121.7</td>
<td>56.2</td>
<td>29.4</td>
<td>33.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 654</td>
<td>8.46</td>
<td>117.4</td>
<td>59.3</td>
<td>63.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 655</td>
<td>8.39</td>
<td>122.7</td>
<td>57.1</td>
<td>30.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 656</td>
<td>8.04</td>
<td>120.1</td>
<td>57.2</td>
<td>39.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 657</td>
<td>8.27</td>
<td>122.9</td>
<td>56.9</td>
<td>30.2</td>
<td>36.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 658</td>
<td>8.10</td>
<td>110.0</td>
<td>45.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 659</td>
<td>8.06</td>
<td>120.3</td>
<td>58.1</td>
<td>39.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glu 660</strong></td>
<td>8.49</td>
<td>122.4</td>
<td>56.7</td>
<td>30.1</td>
<td>35.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 661</td>
<td>8.19</td>
<td>121.4</td>
<td>57.6</td>
<td>39.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 662</td>
<td>8.02</td>
<td>122.6</td>
<td>62.3</td>
<td>32.8</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 663</td>
<td>8.48</td>
<td>123.0</td>
<td>53.4</td>
<td>38.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 664</td>
<td>8.35</td>
<td>117.3</td>
<td>59.0</td>
<td>63.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 665</td>
<td>8.45</td>
<td>122.7</td>
<td>57.1</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 666</td>
<td>8.07</td>
<td>120.1</td>
<td>57.5</td>
<td>39.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 667</td>
<td>7.96</td>
<td>123.7</td>
<td>54.6</td>
<td>42.2</td>
<td>26.9</td>
<td>24.86/23.67</td>
<td></td>
</tr>
<tr>
<td>Lys 668</td>
<td>8.25</td>
<td>124.3</td>
<td>54.1</td>
<td>32.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 669</td>
<td></td>
<td></td>
<td>63.1</td>
<td>31.9</td>
<td>27.2</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Glu 670</td>
<td>8.58</td>
<td>122.2</td>
<td>56.4</td>
<td>30.3</td>
<td>36.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 671</td>
<td>8.38</td>
<td>123.7</td>
<td>62.3</td>
<td>32.7</td>
<td>21.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys 672</td>
<td>8.55</td>
<td>127.4</td>
<td>56.1</td>
<td>32.9</td>
<td>24.4</td>
<td>29.0</td>
<td>42.0</td>
</tr>
<tr>
<td>Ser 673</td>
<td>8.10</td>
<td>124.1</td>
<td>59.9</td>
<td>64.7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Estimated errors are 0.01 ppm $^1$H, 0.1 ppm $^{15}$N, 0.2 ppm $^{13}$C$_{\alpha}$, and 0.5 $^{13}$C$_{\text{others}}$. 
The results of the CSI analysis are best viewed as a bar graph as shown in Figure 2.6 (A). Stable secondary structure elements (α-helices or β-sheets) are definable if four or more consecutive residues have similar secondary shifts outside the random coil range (± 0.7 ppm). All PKC-CT constructs have only scattered residues which fall outside the random coil range. However some regions may be interpreted as having a propensity to form secondary structure. In particular the region from Asn650-Phe666 consistently shows positive secondary shifts, indicating a conformational preference towards α-helical character. Additionally, the region surrounding the turn motif phosphorylation site (Thr641) there are alternating secondary shifts consistent with the sampling of turns in the conformational ensemble. This conserved phosphorylation site in PKC was initially called the “turn motif” due to the corresponding phosphorylation site in PKA which resides at the apex of a tight turn, as shown in Figure 1.3 (Knighton et al. 1991; Zheng et al. 1993). This CSI data would be consistent with that observation. Alternatively, in the presence of contacts with the catalytic domain this region could be stabilized into helices (Millhauser 1995).

These data indicate that while there are regions in PKC-CT which are sampling turns and α-helices, the overall the conformational ensemble is highly dynamic and the macroscopic average is intrinsically unstructured. However, where there are conformational preferences in the isolated 48 residue carboxyl tail, it is possible that when in context of the full length 673 residue PKC, tertiary contacts, especially when phosphorylated may stabilize these structural propensities.

**Phospho-mimetic constructs retain intrinsic disorder**

Changes in $^{13}\text{C}^\alpha$ shifts upon perturbation via mutation, denaturant etc. can give residue specific information regarding the effect of the perturbant on $\Phi/\Psi$ torsion angles in terms of relative stability of the secondary structural units (Roy et al. 2005), or propensity to form such structures (Radhakrishnan et al. 1998). The $^{13}\text{C}^\alpha$ chemical shifts of the phospho-mimetic and WT constructs can be compared to determine if phosphorylation will have any effect on the propensity for PKC-CT to form stable secondary structure. The $^{13}\text{C}^\alpha$ chemical shifts of T641E
Figure 2.6

(A) Graph of the CSI analysis for all three PKC-CT constructs studied. As shown, only a handful residues fall outside the random coil region indicated by the red-shaded box, and this is not enough to classify stable secondary structure elements. However, there are regions which indicate a propensity to form secondary structural elements. The region surrounding the turn motif (Thr641) has secondary shifts ($\Delta\delta$) alternating in sign, consistent with turn formation, which could be further stabilized into $\alpha$-helices in the context of the full length protein. Also the region from 650-665 has secondary shifts which are consistently positive, indicating $\alpha$-helical propensity.

(B) The crystal structure of PKC-\(\iota\) (Messerschmidt et al. 2005) highlighting the carboxyl tail (blue) The phosphorylated residues are shown in red, and the residues they form hydrogen bonds with on the upper lobe are shown in purple. The residues making hydrophobic contacts are also shown highlighted. The helical regions from this structure are indicated on the top portion of (A) in orange cylinders.
Δδ = $^{13}$C$^\alpha$ chemical shift (experimental) - sequence corrected random coil (published)

A

B
and S660E relative to those of WT PKC-CT are graphed in Figure 2.7. This graph further demonstrates that the phospho-mimetic mutations do not alter the structure of PKC-CT. The site of mutation for each has been excluded from the graph due to the difference in random coil values between a Glu and Ser/Thr. The only residues with differences outside the experimental error of 0.2 ppm are adjacent to the site of mutation and are consistent with the sequence dependent perturbations to $^{13}\text{C}^{\alpha}$ chemical shifts of (Schwarzinger et al. 2001).

**Discussion**

PKC is implicated in numerous essential cellular events including tumorgenesis and differentiation, cellular activation, modulation of secretion, apoptosis, cytoskeletal rearrangements, and the immune response, reviewed in (Blobel et al. 1996; Toker 1998). While there is a large body of biological work aimed at elucidating the regulatory mechanisms of this significant kinase, there is relatively little structural information. After the discoveries that PKC-CT is both necessary and sufficient for binding to PKD-1 and HSP70 both *in vivo* and *in vitro*, indicating that it behaves like a somewhat separate entity in the context of the protein (Gao et al. 2001; Gao and Newton 2002), we set out to learn more about the structural basis for PKC regulation through the carboxyl tail. By using this “divide and conquer” approach, we now have a portion of PKC that is not only biologically relevant but highly amenable to solution NMR studies.

**The functionally relevant PKC-CT is intrinsically unstructured**

The lack of chemical shift dispersion in the $^{1}\text{H}-^{15}\text{N}$ HSQC spectra (Figure 2.4) was an initial indication that these constructs were unstructured. The chemical shift of a given resonance is dependent on its chemical environment. An HSQC fingerprint with such a narrow range of chemical shifts such as seen with the PKC-CT constructs indicates that the chemical environment is influenced primarily by the chemical nature of its own side chain and immediate sequential residues rather than through long range effects due to secondary or tertiary structure. When in a protein with well defined structure, each nuclei is in a much more
Figure 2.7

Bar graph illustrating the differences in $^{13}$C$\alpha$ chemical shift of the two phospho-mimetic mutants from the WT sample. If there was a change in propensity to form secondary structure upon phosphorylation, it would be reflected in this graph. However, the only differences in $^{13}$C$\alpha$ chemical shift outside of the experimental error of 0.2 ppm are the sites of mutation (not included in the graph) and immediately surrounding residues. The phospho-mimetic mutants retain the intrinsic disorder of WT-PKC-CT structure.
distinctive and diverse chemical environment than compared to nuclei in a random coil. For example the RIIα subunit of PKA, is highly helical (Newlon et al. 2001). Despite the similar number of residues of the two constructs, RIIα (45 residues) and PKC-CT (48 residues), the HSQC spectrum of RIIα has 1H N chemical shifts spanning from 6.9 – 9.2 ppm, a range of over 2 ppm. Conversely, as shown in Figure 2.4 the 1H N chemical shifts for all of the PKC-CT constructs span from 7.6 – 8.4 ppm, a range of just over 0.7 ppm.

Consistent with the lack of chemical shift dispersion in the HSQC spectra, the CSI analysis indicates that PKC-CT is primarily random coil. However, in addition to providing guidelines from which to determine stable secondary structural elements, CSI analysis has been shown to be useful in highlighting conformational preferences for secondary structure (Radhakrishnan et al. 1998; Yao et al. 2001; Roy et al. 2005). The CSI analysis of the PKC-CT constructs (Figure 2.6 (A)) shows a region of consistently positive secondary shifts in the region surrounding the hydrophobic motif phosphorylation site (Ser660); consistent with the sampling of an α-helix in the conformational ensemble. The region surrounding the turn motif phosphorylation site (Thr641) exhibits secondary shifts alternating in sign, which would be consistent with the initiation of turn formation, which could be further stabilized into helices (Millhauser 1995).

If the conformational ensemble had a high population of either of these structures, there would be a αN(i,i+2) NOE correlation for turns, and αN(i,i+3), αN(i,i+4) NOE correlations for the α-helices (Wüthrich 1986). However, only sequential NOE correlations between i and i+1 residues are found in NOESY spectra of PKC-CT (data not shown). This absence of medium and long range NOE correlations, does not rule out the possibility of stable secondary structure, but just implies that if present, they are highly dynamic in nature, as conformational exchange minimizes the buildup of NOE signal through cross relaxation.

Importantly, the T641E and S660E mutations which have been show previously (Gao et al. 2001; Gao and Newton 2002) to mimic the phosphorylated residues in vivo and in vitro, have no effect on even the propensity towards structural formation in the conformational
ensemble of PKC-CT, as was initially hypothesized (Figure 2.3). As highlighted in Figure 2.7, the changes in $^{13}\text{C}_{\alpha}$ chemical shift upon mutation are all within the experimental error.

**Comparison with recently solved structures**

Since the initiation of this study, structures of catalytic domains from two different PKC isoforms have been published, that of $\theta$ and $\iota$ (Xu et al. 2004; Messerschmidt et al. 2005) respectively. Despite these two isoforms belonging to a different family classification, novel and atypical, it is interesting to compare their findings to those presented here for the conventional PKC-$\beta$II.

The first PKC isoform to be successfully crystallized is that of PKC-$\theta$ (Xu et al. 2004). In this structure, the majority of the carboxyl tail is not visible due to missing electron density. The only portion which can be seen is that of the hydrophobic motif phosphorylation site Ser695, (corresponding to Ser660 in PKC-CT). As predicted though modeling to the PKA structure, (Figure 1.3), this site is seen tucked in a hydrophobic grove adjacent to the N-lobe on the back side of the active site. The regions surrounding the hydrophobic motif phosphorylation site, including the turn motif phosphorylation site are disordered. This is consistent with the results from this study in which PKC-CT was determined to be dynamically unstructured.

The more recently solved PKC-$\iota$ structure (Messerschmidt et al. 2005) reveals a little more information about the carboxyl tail. In this structure, there is density for more of the carboxyl tail. As shown in Figure 2.6 (B) both phosphorylation sites are visible, as well as some of the connecting residues. The region directly preceding the turn motif Thr655 (corresponding to Thr641 in PKC-CT) is disordered, but interestingly, the phosphorylation site is structured. This site was initially called the turn motif due to the corresponding phosphorylation site in PKA sitting at the apex of a tight turn (Figure 1.3) which served to help anchor the carboxyl tail to the upper lobe of the catalytic domain. In the PKC-$\iota$ structure however, this turn is not present. Rather than the phosphate group making contacts with the surrounding residues to stabilize the turn as in PKA, in PKC-$\iota$ the phosphate forms ionic contacts with residues in strands on the back of the active site anchoring it to the top of the
catalytic domain (Figure 2.6 (B)). Interestingly, as described above, the CSI analysis presented here indicates the potential to form either turns or nascent \( \alpha \)-helices at this location. Any secondary structural elements at this location are different from what was observed in the structure of these other two isozymes, \( \theta \) and \( \iota \), providing a possible structural basis for the difference in isozyme specificity from the V5 domain.

Unlike the PKC-\( \theta \) structure, the region in-between the two conserved phosphorylation sites is well defined in the PKC-\( \iota \) structure. The region directly following the turn motif phosphorylation site adopts a short helix, followed by a short connecting region and then another short \( 3_{10} \) helix just prior to the hydrophobic motif, which is again docked against the upper lobe. Rather than a phosphorylated Ser at this position, PKC-\( \iota \) has a Glu at this position (Figure 2.2). Glu574 corresponding to Ser660 in PKC-\( \beta \) is hydrogen bonded to the outer surface of helix C, and the surrounding hydrophobic residues interact with the hydrophobic groove of the upper lobe (Figure 2.6 (B)). These helical structures of PKC-\( \iota \) are indicated on the upper portion of the PKC-\( \beta \) CSI data graph (Figure 2.6 (A)), and correlate well with the regions of helical propensity describe earlier. It is entirely possible that the hydrogen bonds between the phosphorylated residues and the upper lobe of the kinase domain stabilize these helical structures.

Given that the phosho-mimetic PKC-CT constructs retained the intrinsic disorder of the WT-PKC-CT, it is highly probable that the role of these essential phosphorylations in PKC activity and regulation may be to anchor the carboxyl tail to the catalytic domain, stabilizing an active conformation of the kinase. However, when these contacts are not present, either from not being phosphorylated in the full length protein, or lack of the kinase domain as presented here, the carboxyl tail of PKC is intrinsically unstructured. The CSI data presented here is the only structural information available for the carboxyl tail when these contacts are not present, and indicates that under these conditions, the carboxyl tail of PKC is intrinsically unstructured. Given the importance of the non-phosphorylated carboxyl tail as the binding sight for PDK-1 and HSP70, this finding is particularly relevant in understanding the structural basis for these key regulatory protein/protein interactions.
Functional implications for dynamic disorder

With the growing number of protein structures being solved during the 20th century, the field of structural biology developed the fundamental theory that a defined folded structure is necessary for protein function. While there is no doubt that protein function is intimately linked to its function, this central dogma is being challenged. For years, there have been crystal structures of proteins reported in which functional regions which are missing from the electron density. However, partially due to the developments in NMR methodologies for studying unfolded protein structures (reviewed in (Dyson and Wright 2004)), the significance of these disordered regions is becoming more apparent. Coincidentally, the turn of the century has called for a re-assessment of the structure-function paradigm (Wright and Dyson 1999; Dunker et al. 2001). Using a bioinformatics, data-mining approach of a PONDR (predictors of natural disordered regions) algorithm, the existence of thousands of natively disordered proteins has been proposed (Romero et al. 2001). Dunker and co-workers have applied this algorithm to proteomes from more than 30 different organisms and classified the resulting disordered proteins according to function. The function most represented among intrinsically disordered proteins is that of molecular recognition, particularly protein-protein interaction (Dunker et al. 2002). The functional importance of this disorder as it relates to PKC is two fold.

First, intrinsic disorder provides the flexibility to bind to multiple partners. The lack of a pre-defined structure facilitates promiscuity, by accommodating multiple binding conformations (Kriwacki et al. 1996; Dyson and Wright 2002), particularly in signaling proteins (Iakoucheva et al. 2002). As stated previously, one of the motivations for studying PKC-CT was because it serves as the docking site for both PDK-1 and HSP70, as well as the yet unidentified phosphatases, and as will be shown in chapter 3, Pin1. Additionally, the flexibility of region no doubt facilitates the auto-phosphorylation of the two sites. The intrinsic disorder of this region provides the flexibility necessary to accommodate the different binding sites of these highly different proteins.

A second somewhat contradictory implication is that intrinsic disorder allows for highly specific interactions coupled with low affinity. As proposed for highly specific protein-DNA
interactions (Spolar and Record 1994) for a favorable free energy ($\Delta G$) of binding, the entropic ($\Delta S$) cost of the interaction must be overcome by favorable enthalpic ($\Delta H$) contributions. This aspect of intrinsic disorder is particularly important for the PKC/PDK-1 interaction. The release of PDK-1 from the carboxyl tail of PKC is the rate-limiting step in PKC maturation (Gao et al. 2001), and thus must be a relatively low affinity interaction. Yet, PDK-1 is proposed to be constitutively active (Toker and Newton 2000), and thus its substrate recognition must be highly specific. The disordered nature of the carboxyl tail of PKC provides the necessary specificity at the favorable expense of affinity.

Conclusion

The CSI analysis of WT-PKC-CT and phospho-mimetic mutations to the highly conserved phosphorylation sites establish that this region of PKC have a slight $\alpha$-helical propensity, yet the ensemble is predominately populated with random-coil conformations. Furthermore, the intrinsic disorder is retained upon phosphorylation. Comparison with recently solved crystal structures demonstrates that the role of these phosphorylations is probably to anchor the carboxyl tail to the upper lobe of the catalytic domain, upon which secondary structural elements are stabilized. The data presented here serves as the only structural information to date for the biologically relevant species of PKC in which the stabilizing contacts between the phosphorylated carboxyl tail and the catalytic domain are not present. Given the importance of the non-phosphorylated species as a binding partner for the regulatory proteins PDK1 and HSP70, this state is particularly relevant. The functional implications of this structural disorder, mainly allowing the carboxyl tail to serve as a docking site for multiple proteins, were discussed.

The text of this chapter in part will be submitted for publication as “The PKC-Pin1 Interaction; A Complex Twist to the Tail”, manuscript in preparation. I was the primary researcher and author for this publication, and Hilde Abrahamesn, Alexandra C Newton and Patricia A. Jennings the co-authors directed and supervised the research.
Chapter 3

PKC Undergoes a Pin1 Catalyzed Proline Isomerization:

A Complex Twist to the Tail
Introduction

The Protein Kinase C enzymes constitute a ubiquitous family of kinases which are responsible for numerous cellular functions including cell growth and differentiation, synaptic transmission, neural development, axonal regeneration, endocrine and exocrine secretion, and smooth muscle contraction and relaxation, as reviewed in (Liu 1996). Much research effort in the past two decades have led to some understanding on the activity and regulation of PKC, and has been extensively reviewed (Nishizuka 1986; Newton 1995; Mellor and Parker 1998; Parekh et al. 2000; Newton 2001; 2003).

As illustrated in Figure 1.5, PKC undergoes multiple conformational changes throughout its lifetime. The newly synthesized, membrane associated state of PKC is in an open conformation in which the autoinhibitory sequence is removed from the substrate binding cavity (Dutil and Newton 2000), and the carboxyl tail is exposed for PDK-1 binding (Gao et al. 2001). Following autophosphorylation at two conserved sites in the carboxyl tail, PKC is released into the cytosol in thermally stable, protease and phosphatase resistant conformation (Edwards and Newton 1997). In response to cellular signals which elevate Ca\(^{2+}\) and diacylglycerol levels PKC translocates back to the membrane (Sakai et al. 1997). This membrane tethered, active PKC is again in an open conformation which is highly sensitive to phosphatase action (Hansra et al. 1996). Interestingly, the susceptibility of the carboxyl tail to proteolytic cleavage is dependent on Ca\(^{2+}\) concentration. Ca\(^{2+}\) concentrations sufficient to recruit PKC to the membrane yet insufficient to fully activate PKC render the carboxyl tail subject to trypsin cleavage. However, Ca\(^{2+}\) concentrations sufficient for full activation result in a carboxyl tail which is resistant to trypsin cleavage (Keranen and Newton 1997). Further underscoring the importance of the conformation of the carboxyl tail in the PKC lifecycle, the upstream kinase, PDK-1 and chaperone HSP70 discriminate between the chemically identical un-phosphorylated, and de-phosphorylated carboxyl tails, with PDK-1 preferring the former (Gao et al. 2001), and HSP70 the later (Gao and Newton 2002).

Chapter 2 of this dissertation explored the conformational dynamics of PKC-CT. Through those investigations it was discovered that the carboxyl tail of PKC has \(\alpha\)-helical
propensity, but in general is unstructured, regardless of phosphorylation state. Highlighting the
dynamic nature of the carboxyl tail, is the fact that time dependent spectral changes as those
presented in Figure 3.1 were continually observed in all PKC-CT constructs. The same subset
of peaks repeatedly would disappear, move, or appear. Once the assignments were
completed, and it became apparent that the resonances affected all localized to the N-terminal
region of PKC-CT, we became interested in this time-dependent phenomenon as a possibly
biologically relevant event rather than just simple degradation reaction in vitro.

In addition to the two conserved phosphorylation sites in PKC-CT, multiple proline
residues are conserved amongst the 10 PKC isozymes (Figure 2.3). The unique cyclic
structure of proline makes it the only naturally occurring amino acid for which both the cis and
trans peptide bond conformations are thermodynamically accessible (Figure 3.2). The
activation energy between the cis and trans conformations is around 70-80 kJ/mol (Cheng
and Bovey 1977; Grathwohl and Wuthrich 1981), and the interconversion between the two
forms is an intrinsically slow process. However, this unique property of the proline residue
lends the ability to populate discrete conformational populations with only modest
interconversion energy. Proline isomerization has long been recognized as potentially the rate
limiting step in protein folding (Brandts et al. 1975; Schmid and Baldwin 1978), but only
recently is its role in functional regulation being realized (Yaron and Naider 1993). The
observed time dependent changes would be consistent with the activation energy of proline
isomerization.

Particularly important to the signal transduction field is the rising predominance of
phosphorylation–directed proline isomerization (Zhou et al. 1999; Lu et al. 2002). The peptidyl
prolyl isomerase Pin1 is a member of the parvulin family of peptidyl-prolyl isomerases, but is
the only known isomerase to have specificity for phospho-Ser/Thr-Pro bonds. Pin1 has been
shown to display approximately 1000x selectivity for such sites (Yaffe et al. 1997). Notably, the
Pin1 catalyzed isomerization of prolyl bonds directly following phosphorylated Ser or Thr
residues has been shown to be essential for cell cycle progression (Lu et al. 1996).
Figure 3.1

$^1$H-$^{15}$N HSQC spectra of WT-PKC-CT illustrating the time dependent spectral changes which were observed for all PKC-CT constructs. The black spectrum was acquired on day 1, while the red spectrum was acquired 15 days later. Resonances, which experience a change in chemical environment resulting in either a change in chemical shift or extreme line broadening, are labeled. For clarity, these residues are also highlighted in red in the sequence below the spectra. As seen, these affected residues are localized to the N-terminal region of the protein.
Careful examination of the PKC sequence (Figure 2.3) revealed the presence of a possible Pin1 recognition site at the highly conserved turn motif (Thr641) phosphorylation site. Most of the PKC isozymes contain a proline after this phosphorylation site. What is also striking is the high conservation of an additional proline N-terminal to this site corresponding to either Pro637 or Pro638 in PKC βII.

The research in this chapter was directed towards investigating the possibility of a Pin1 catalyzed cis/trans isomerization within the carboxyl tail of PKC. Solution NMR is uniquely suited for studying proline isomerization on a residue specific basis. With respect to the chemical shift time scale, proline isomerization is usually a slow process leading to individual resonance frequencies observable for both the cis and trans states. The inter-residue proton distances are strongly affected by isomerization and characteristic NOE correlation patterns readily distinguish the cis vs. the trans states allowing for conformational assignment (Wüthrich 1986). (Figure 3.2) The observable inter-residue NOE correlation for the trans X-proline peptide bond will be between, the δH_{(Pro)} and the αH_{(x)} whereas in the cis conformation, the observable NOE correlation will be between, the αH_{(Pro)} and the αH_{(x)}. An additional asset to studying proline isomerization via NMR methodology is that simple integration of peaks has the potential to yield relative populations of the two isomers.

Using these methods, it was not only discovered that there is an intrinsic proline isomerization in the carboxyl tail of PKC, but more importantly, that this isomerization can be catalyzed by Pin1. Furthermore due to the residue specific nature of NMR, it was demonstrated that Pin1 is catalytically active towards a non pThr-Pro site. Despite its growing prominence in regulation of signal transduction, Pin1 had never been proposed to interact with PKC prior to this work. Additionally, a new model for the PKC lifecycle which includes regulation not only by phosphorylation, but also isomerization. This model includes our discovery of the Pin1 mediated conformational regulation of PKC and provides the first explanation for the difference in molecular recognition of the chemically equivalent, un-phosphorylated and de-phosphorylated carboxyl tail of PKC by regulatory/docking partners.
Figure 3.2
Cartoon illustrating the different possible isomers of the peptide bond preceding a proline residue. As highlighted by the red circles, the inter-proton distances and thus observable NOE correlations are isomer specific. When in the *trans* conformation, the observable inter-residue NOE correlation is between, the $\delta H_{(Pro)}$ and the $\alpha H_{(Pro-1)}$; whereas in the *cis* conformation, the observable NOE correlation is between, the $\alpha H_{(Pro)}$ and the $\alpha H_{(Pro-1)}$. 
Results

Time-dependent Trypsin cleavage indicates cis/trans heterogeneity

Trypsin is an isomer specific protease, and will cleave after a Lys or Arg, only if the peptide bond located two residues away is in the trans conformation, and thus is often used as a probe for proline isomerization in protein folding studies (Brandts and Lin 1986). One of the five prolines in the PKC-CT sequence, Pro637, lends itself to this type of analysis. Trypsin will only cleave after Arg635 if the His636-Pro637 peptide bond is in the trans conformation. Figure 3.3 is an SDS-Page gel of time points taken from a trypsin cleavage reaction of WT PKC-CT. As shown, the majority of the sample is cleaved by the trypsin within ten minutes. The un-cleaved population remains such even after the last time point of three hours. However, if left overnight, this population is eventually cleaved, indicating that the trypsin was active, but the PKC-CT was in an un-cleavable cis conformation. When given enough time to re-establish the cis/trans equilibrium, the sample is ultimately cleaved.

Pin1 catalyzes an isomerization in PKC-CT

The trypsin cleavage assay is an elegant probe for the isomeriazation state of one of the proline bonds in PKC-CT, however there are four additional prolines which are inaccessible by this method. Of particular interest is the Thr641-Pro642 bond, given the specificity for Pin1 to recognize and isomerize pThr-Pro bonds. ¹H NMR is uniquely suited for this type of investigation by allowing for simultaneous yet site specific monitoring of all proline residues in PKC-CT. The fingerprint region of TOCSY spectra of the phospho-mimetic T641E PKC-CT construct pre and post Pin1 catalysis is shown in Figure 3.4 (A). The spectrum in black was acquired on a sample without Pin1, while the spectrum in red was acquired on a sample to which catalytic amounts of Pin1 had been added. The fingerprint region contains the cross-peaks between the Hn and the Hα protons, therefore, like in the HSQC spectrum, each peak represents a single residue. Interestingly, the peaks corresponding to the same N-terminal residues which shifted or disappeared over time in Figure 3.1 experience similar changes upon
Figure 3.3

SDS-Page gel of a time-course trypsin cleavage assay revealing the presence of a minor cis population of the His636-Pro637 bond. As explained in the text and indicated in the sequence above the gel, trypsin will only cleave after an Arg if the peptide bond two residues away is in the trans conformation. The red oval on the gel is highlighting the fact that while the majority of the PKC-CT population is cleaved within 10 minutes, a substantial population remains uncleavable. However if left overnight, it will eventually be cleaved (highlighted by the blue oval), presumably because there was enough time to isomerize from the un-cleavable cis conformation to the cleavable trans population.
Figure 3.4

$^1$H TOCSY Spectra of T641E PKC-CT pre (black) and post (red) Pin1 catalysis. (A) The fingerprint region of the spectra, the crosspeaks corresponding to the same N-terminal residues which moved or disappeared over time in Figure 3.1, experience similar changes upon Pin1 catalysis and are labeled. (B) The region of the TOCSY spectra containing the Pro $\delta$-$\alpha$ crosspeaks. Interestingly, the Pro affected by Pin1 is not the expected Pro642 which follows the phosphorylation site, but rather Pro637, and minor changes are observed for Pro638.
Pin1 catalysis. The easiest change to identify and compare is the new set of peaks appearing at 7.6 ppm in both spectra. Curiously though, as shown in Figure 3.4 (B), the peptide bond affected by Pin1 is not the expected phospho-threonine mimetic Glu641-Pro642 bond, but rather the His636-Pro637 bond. The affected cross-peaks in these spectra correspond to residues N-terminal to Pro637 and Pro638. Of additionally curiosity is that the NOE correlations for this post Pin1 catalyzed sample (data not shown) are between the $\text{H}_\delta$ of Pro, and the and the $\text{H}_\alpha$ of the His, consistent with still being in the Trans conformation (Wüthrich 1986). Similar changes to these shown for the phospho-mimetic T641E-PKC-CT construct were seen with the WT-PKC-CT construct.

**PKC specifically interacts with Pin1**

To determine whether this unexpected change in PKC upon Pin1 catalysis was due to a specific interaction between the two proteins or just an acceleration of the intrinsic isomerization of this His-Pro moiety, studies were directed toward investigating the effects of PKC on Pin1. Figure 3.5 (A) is an overlay of $^1\text{H}$-$^{15}\text{N}$ HSQC’s of labeled Pin1 isolated (black) and with unlabeled T641E-PKC-CT (red). The cross-peaks which shifted outside the experimental error in either the $^1\text{H}$ or $^{15}\text{N}$ dimension upon PKC addition are labeled using the published assignments from (Jacobs et al. 2002). As shown in Figure 3.5 (B), when mapped back onto the crystal structure of Pin1 (Ranganathan et al. 1997), these residues localize to the two possible substrate binding sites, the WW domain recognition site (shown in red), and the active site of the PPIase domain (shown in blue).

**NMR evidence supports the Trypsin cleavage experiment**

Yaffe et al. have shown that glutamate mutations do not fully mimic phosphorylated residues in terms of Pin1 catalytic activity (Yaffe et al. 1997), prompting interest in investigating whether the unexpected catalytic actions of Pin1 with the bacterially expressed PKC-CT were a result of this incomplete phospho-mimic utilized in these studies. While the carboxyl tail of PKC is auto phosphorylated *in vivo*, it is not a quantitative reaction *in vitro*, making it necessary
Figure 3.5

(A) $^{1}\text{H}-^{15}\text{N}$ HSQC spectra of Pin1 without (black) and with (red) unlabeled T641E PKC-CT. Residues which shift upon the addition of Pin1 are labeled according to the published assignments of (Jacobs et al. 2002). When mapped back onto the crystal structure (Ranganathan et al. 1997) in (B), these residues clearly localize to the two possible substrate binding sites for Pin1; the WW domain (red) and the active site of the PPlase domain (blue). This is consistent with the interaction between Pin1 and PKC_CT being specific.
to switch to synthetic peptides in which a phosphorylated Thr could be included. A 13 residue peptide based on the sequence of PKC-CT (PKC-CT\textsubscript{13}) was designed to encompass both proline sites as well as a few flanking residues which might be necessary for binding and recognition by Pin1. A benefit to using this synthetic peptide was that sample concentrations were not limited by low expression levels, and it was possible to obtain 2 mM NMR samples.

As shown in Figure 3.6, with these more concentrated samples, NMR evidence is obtained for a minor \textit{cis} population, as indicated by the trypsin cleavage experiment. Figure 3.6 contains overlaid TOCSY (in red) and NOESY (in black) spectra of the WT, un-phosphorylated PKC-CT\textsubscript{13}. Three sets of proline residues are identifiable and labeled. The first, labeled in black, correspond to the His636-Pro637-Pro638 triad. Lines are included in the figure to connect the TOCSY peaks to the NOESY peaks from the H\textsubscript{δ} of the Pro, to the H\textsubscript{α} of the preceding residue (either His636 or Pro637). These NOE correlations are indicative of a \textit{trans} conformation of the peptide bonds (Wüthrich 1986). The second set of prolines, labeled in green, correspond to the Thr641-Pro642-Pro643 triad. Again, the NOE correlations indicate the \textit{trans} conformation and are highlighted. Interestingly, as highlighted in blue, there is an additional set of proline peaks, which through the presence of an NOE correlation from the H\textsubscript{α} of the Pro to the H\textsubscript{α} of the His, can be assigned as a \textit{cis} conformation to the His636-Pro637 bond (Wüthrich 1986). The Pro637-Pro638 bond for this population is still in the \textit{trans} conformation.

Integration of the Pro637 $\delta$-$\alpha$ crosspeaks in the TOCSY spectra for each population reveal that about 10% of the total PKC-CT\textsubscript{13} of this sample is in the \textit{cis} conformation. The same heterogeneity was seen in peptides corresponding to the T641E phospho-mimetic, as well as the pT641 peptide.

**Pin1 has multiple catalytic effects on PKC-CT\textsubscript{13}**

Shown in Figure 3.7 is the same TOCSY spectrum from Figure 3.6 (in red), overlaid with a TOCSY spectrum acquired on a WT-PKC-CT\textsubscript{13} sample to which catalytic amounts of Pin1 were added (in blue). Clearly, Pin1 is catalyzing similar changes in this construct as those previously observed for the 48 residue PKC-CT. The residues N-terminal to Pro637
were affected, while Thr641 and Pro642 remain un-altered. The crosspeaks corresponding to the \textit{trans} conformation of His636-Pro637 have shifted, and the crosspeaks corresponding to the \textit{cis} population have completely disappeared. Figure 3.8 is the TOCSY spectrum with Pin1 (in blue) overlaid with a ROESY spectrum (green) from the same sample. A ROESY spectrum was required because the catalytic actions of Pin1 altered the correlation time of PKC-CT\(_{13}\) to the effect that it was at the cross-over point of the NOE relaxation pathway as evidenced by, positive, negative and missing NOE peaks (Claridge 1999; Levitt 2001). This figure illustrates the complexity of the Pin1 activity on PKC. Clearly assignable and thus labeled (in black and red), are two new populations of the His-Pro bond. \(\delta\)-\(\alpha\) ROE correlations are present for both, consistent with both conformations being \textit{trans}. Additionally there are a handful of extra peaks which are consistent with multiple conformations of the residues N-terminal to the proline (labeled in yellow), but due to lack of ROE correlations, can not be specifically assigned. Further complicating the situation is the fact that similar to the 48 residue PKC-CT, there are insoluble, fibrilar species in the NMR tube. This dynamic heterogeneity makes it impossible to define exactly this new conformation of PKC-CT\(_{13}\) created by the catalytic actions of Pin1. Importantly though, it can be said that Pin1 isomerized the His636-Pro637 bond, predominantly the complete removal of the \textit{cis} conformation. However, the \textit{trans} Thr641-Pro642 bond remains unchanged. Similar effects were observed for the T641E phospho-mimetic peptide.

Unexpectedly however, Pin1 had no visible activity towards the phosphorylated peptide. Shown in Figure 3.9 are TOCSY spectra of the pT641 PKC-CT13 peptide pre (red) and post (blue) Pin1 catalysis. In addition to observing no changes in the NMR spectra, there were no fibrilar precipitates formed after adding the Pin1. It is noteworthy to mention at this point, that as seen in these spectra, this phosphorylated peptide still retains the intrinsic minor \textit{cis} population to the His636-Pro637 bond as was seen in the un-phosphorylated peptide (Figure 3.7).
Figure 3.6

Overlaid $^1$H TOCSY (red) and NOESY (black) spectra of WT-PKC-CT$_{13}$. The peaks corresponding to the Thr641-Pro642 residues are labeled in green. The NOE correlations present are between the Pro$\delta$, and Thr$\alpha$, consistent with this peptide bond being in the trans conformation.

Clearly visible and labeled are multiple conformations to the His636-Pro637 bond. The population corresponding to the trans conformation is labeled in black. The minor population which can be assigned as being cis from the presence of the Pro$\alpha$ - His$\alpha$ NOE correlations is labeled in blue. From peak integration, it was determined that approximately 10% of the sample contains this cis peptide bond.
Figure 3.7

Overlaid $^1\text{H}$ TOCSY spectra of WT-PKC\_CT$_{13}$ pre (red) and post (blue) Pin1 catalysis. Resonances which were affected by Pin1 are labeled. As seen, the effects of Pin1 on PKC-CT$_{13}$ are similar to those observed with PKC-CT, with the changes being to the residues N-terminal to Pro637, while Pro642 remains unaffected.
Overlaid $^1$H TOCSY spectra post Pin1 catalysis (blue), and the corresponding ROESY spectra (green). Importantly, the cis isomer has been completely removed from the population, and as indicated, there are two new conformations to the His636-Pro637 bond, which according to the presence of the Pro-$\delta$-His$\alpha$ ROE correlations are both trans. These presumably differ in the pucker state of the indole ring, and possibly represent intermediate states between either the cis-trans isomerization or between the soluble-insoluble species. There are also multiple peaks (labeled in yellow) which are consistent with the N-terminal residues, but due to conformational heterogeneity and lack of ROE correlations can not be specifically assigned. Resonances which were affected by Pin1 are labeled. As seen, the effects of Pin1 on PKC-CT$_{13}$ are similar to those observed with PKC-CT, with the changes being to the residues N-terminal to Pro637, while Pro642 remains unaffected.
Figure 3.9

Overlaid $^1$H TOCSY spectra of pT-PKC-CT$_{13}$ pre (red) and post (blue) Pin1 catalysis. Similar to the un-phosphorylated peptide, Pin1 has no visible effect on the pThr641-Pro642 bond. However, unlike the changes seen in the un-phosphorylated peptide, Pin1 has only a minor (if any) effect on the cis His636-Pro637 species.
Calcium stabilizes the *cis* conformation of PKC-CT$_{13}$

Previously, Keranen and Newton, have shown that *in vitro*, the carboxyl tail of PKC undergoes a conformational change upon activation by Ca$^{2+}$ which renders it un-cleavable by trypsin (Keranen and Newton 1997). Knowing that trypsin is isomer specific (Brandts and Lin 1986), (Figure 3.3) as well as published reports of instances in which Ca$^{2+}$ stabilizes the *cis* conformation of peptidyl-prolyl bonds (Ng and Weis 1998; Perera et al. 1998), it seemed important to ask the question of whether or not the presence of this classic activator of PKC stabilizes the *cis* conformation of the carboxyl tail. Shown in Figure 3.10 are TOCSY spectra of both the un-phosphorylated (A) and phosphorylated (B) PKC-CT$_{13}$, in the absence (black) and presence (red) of added Ca$^{2+}$. Strikingly, the disappearance or extreme broadening of the crosspeaks corresponding to all proline residues is observed, while the remaining peaks stay at similar intensities. The presence of the Ca$^{2+}$ has drastically altered the *cis/trans* interconversion rate for both sets of prolines (Pro637-Pro638, and Pro642-Pro643) to the point where it is in the intermediate exchange regime. Taken together with the previously reported difference in trypsin cleavage susceptibility upon activation this work suggests that the active conformation of the carboxyl tail of PKC is one in which at least one if not both sets of prolines are in the *cis* conformation.

Evidence that HSP70 reverses the actions of Pin1

HSP70 prolongs the lifetime of PKC by rescuing it from degradation; however the exact mechanism by which this is accomplished remains unknown. What is known, is that HSP70 discriminates between nascently un-phosphorylated PKC, and PKC which has been de-phosphorylated, preferring to bind to the later (Gao and Newton 2002). Figure 3.11 is preliminary evidence that this rescue is accomplished by HSP70 reversing the catalytic actions of Pin1 in the carboxyl tail. The black $^{1}$H-$^{15}$N spectrum in Figure 3.11 is S660E PKC-CT which has experienced the time dependent changes described in Figure 3.1, which were found to be catalyzed by Pin1 in as shown in Figure 3.4. Not only are peaks missing, but more striking on first glance is the characteristic set of peaks at 7.6 ppm $^{1}$H and 126 ppm $^{15}$N.
Figure 3.10

TOCSY spectra of both the non-phosphorylated (A) and phosphorylated (B) PKC-CT13, in the absence (black) and presence (red) of 1.5 M Ca\(^{2+}\). The intensity of the crosspeaks corresponding to all proline residues (highlighted by the blue circles) are dramatically reduced. The remainder of the resonances retain the same intensity and just have minor chemical shift perturbations which would be expected, due to the increased ionic strength. This would be consistent with the Ca\(^{2+}\) stabilizing the Cis isomers of the proline residues to the point where the interconversion rates are now on the order of the intermediate exchange of the NMR time scale.
$^1$H-$^{15}$N HSQC of PKC-CT isolated (black) and with approximately equimolar unlabeled HSP70. The black spectrum is displaying the signatures of the isomerized PKC-CT, peaks are missing, as well as the additional peaks at 7.6 ppm $^1$H. Interestingly, as highlighted by the colored arrow, addition of HSP70 partially reverses these changes. The only other major difference in the spectra is highlighted by the black arrow and is a decrease in intensity to the crosspeak corresponding to Leu640, a residue which previously has been determined important in the HSP70-PKC interaction (Gao and Newton 2002).
Interestingly, when unlabeled HSP70 was added to the sample (red spectrum), only a few changes are noted, particularly the disappearance of this peak at 7.6 ppm as highlighted by the arrow. One of the few other peaks altered by the presence of HSP70 is that of Leu640, which through mutational studies has been shown to be important for the PKC-HSP70 interaction (Gao and Newton 2002). From this, one can propose that the role of HPS70 is to reverse the actions of Pin1. Accordingly, HSP70 has previously been shown to have isomerase activity (Schiene-Fischer et al. 2002; Swain and Gierasch 2002).

**Discussion**

The activity of PKC is controlled, in part, by three highly conserved phosphorylation events on the activation loop (T500), the turn motif (T641) and the hydrophobic motif (S660) (Dutil et al. 1994; Keranen et al. 1995). However the exact mechanism of this regulation has remained elusive. Chapter 2 of this dissertation explored the possibility of the carboxyl tail phosphorylations (the turn motif, and the hydrophobic motif) altering the structure of the tail, thus providing the mechanism for PDK-1 and HSP70 binding specificity. However all NMR evidence collected, indicates that the carboxyl tail is dynamically unstructured regardless of phosphorylation state. Interestingly, what was noticed were continual, reproducible time dependent changes in all spectra. The work presented in this chapter further explored these changes and their potential biological significance.

The conformation of the carboxyl tail is not controlled by phosphorylation alone, but also by proline *cis/trans* isomerization. Importantly, this isomerization can be catalyzed by the peptidyl-prolyl isomerase Pin1, an interaction which was previously unknown. The discovery of the isomerization in the carboxyl tail of PKC, not only is further insight into the complex regulation of PKC, but importantly provides a novel explanation for why the nascent un-phosphorylated and de-phosphorylated carboxyl-terminus have different molecular recognition surfaces. Additionally due to the residue specific nature of the NMR experiments utilized here, it was discovered that this Pin1 catalyzed isomerization is more complex than initially predicted, with Pin1 isomerizing an unexpected site on PKC-CT.
An intrinsic cis/trans isomerization in the carboxyl tail of PKC

The NMR experiments in chapter 2 of this dissertation indicated that while PKC-CT has regions of α-helical propensity, the conformational ensemble is dynamic in nature. One possible explanation for this could be the fact that the isolated carboxyl tail needs contacts with the catalytic domain in order to stabilize these helices. While this may still be true, an additional explanation explored here was that of proline cis/trans isomerization.

The unique structural arrangement of the proline side chain makes the population of the cis isomer of the peptide bond energetically possible. The conversion between the trans and cis isomers of the N-terminal peptide bond formed with a proline residue (shown in Figure 3.2) is appropriately termed proline cis/trans isomerization. This isomerization has a moderately high energy of activation (70-80 kJ/mol) and is an intrinsically slow process (Cheng and Bovey 1977; Grathwohl and Wuthrich 1981). Therefore when time dependent changes such as those seen in Figures 3.1 were observed on the time scale of days to weeks, and particularly when it was realized that the residues affected were localized to the N-terminal region of PKC-CT, the possibility of proline isomerization being the source of this conformational change was investigated.

As stated in the text above, trypsin will only cleave the peptide bond following an Arg or Lys, if the peptide bond two residues following the cleavage site is in the trans conformation. Thus, a time dependent trypsin cleavage reaction is an elegant assay for a possible isomerization of the His636-Pro637 bond. PKC-CT will only be cleaved after Arg635 if this bond is trans. As shown in Figure 3.3, the majority of the WT-PKC-CT sample is cleaved within the first ten minutes of the reaction. However a substantial population remains uncleaved even three hours later, but if left overnight, will eventually be cleaved. One unlikely explanation for this un-cleaved portion could be due to diminished activity of the trypsin after the initial 10 minutes. However, had this been the case, one would expect the intensity of the band on the gel to diminish over the initial three hours of the experiment. The other, more probable explanation is that this minor population of PKC-CT has a cis peptide bond between His636-Pro637, and thus is un-cleavable by trypsin. Once the cis/trans equilibrium was re-
established over night, resorting a trans population to PKC-CT, it was once again a suitable substrate for trypsin and thus cleaved.

This interpretation of the trypsin cleavage experiment is further supported by NMR spectroscopic evidence with the PKC-CT13 peptides. Figure 3.7 clearly reveals a cis isomer of the His636-Pro637 bond, as identified by the presence of the inter-residue α-α NOE correlation. Through integration of the TOCSY cross-peaks, it was determined that approximately 10% of the PKC-CT13 population contains this cis peptide bond. This is consistent with previous reports of between 9.5%-16.5% cis content of His-Pro bonds depending on solution pH (Reimer et al. 1997; Reimer et al. 1998). It is interesting to point out that according to their work with 5’mer peptides Reimer et al. all report a similar 9.4% cis content for a Thr-Pro bond (Reimer et al. 1998). The possibility of a cis conformation at the Thr641-Pro642 site can not be ruled out, but if present, it only represents a small fraction of the population, and this small percentage does not change with the phosphorylated peptide as would be predicted (Schutkowski et al. 1998).

**Pin1 catalyzes changes in PKC-CT**

Careful examination of the PKC sequence reveals the presence of a proline following the conserved turn motif (Thr641) phosphorylation site, not only in the βII isozyme, but a number of the other isozymes as well (Figure 2.3). With the rising predominance of signaling proteins being regulated by the Pin1 controlled proline isomerization of a phospho-Thr-Pro motif (Wulf et al. 2005), the possibility of Pin1 catalyzing an isomerization in the carboxyl tail of PKC was investigated.

The initial studies to investigate the PKC-Pin1 interaction were conducted on the same 48 residue PKC-CT construct as the prior heteronuclear work, using the T641E mutant to mimic the phosphorylated residue preceding the proline. These experiments revealed that Pin1 did have an effect on PKC-CT. As shown in Figure 3.4, Pin1 catalyzed the same changes in PKC that were seen over time in the heteronuclear work. In accordance with Pin1 being a catalyst, these changes were observed within a matter of hours rather than days, as
the prior un-catalyzed time dependent changes. The effects of Pin1, are strikingly similar to those which occurred over time. They localize to the same N-terminal residues, with the residues displaying the same shifting or disappearing behavior. Additionally the same characteristic peak at 7.6 ppm appears in the homonuclear spectra, and with the use of the NOESY spectra, can now be assigned to Phe633.

However, it was surprising that these changes caused by Pin1 were not the result of an isomerization of the predicted phospho-threonine mimetic Glu641-Pro642 bond. As shown in Figure 3.4 (B), the cross peaks corresponding to Pro642 remain identical pre and post Pin1 catalysis. What is observed is a shifting of the cross peaks corresponding to Pro637, and to a less extent Pro638. This is consistent with the fact that all residues which experience changes are N-terminal to Pro638; whereas the residues C-terminal to Pro642, are no different in the two spectra. What is even more unexpected is that the NOE correlations, as shown in Figure 3.5 (B) indicate that this His-Pro bond is still in the trans conformation. A possible explanation for this would be that this proline now has a different pucker to the pyrrolidine ring (Cai et al. 1995). There have been multiple proposals for the intermediate transition state between trans to cis, many of them differing in the pucker state of the ring (Fischer 2000; Hur and Bruice 2002). Lending support to this explanation is the fact that after Pin1 catalysis, a substantial portion of the PKC-CT sample is found in fibrilar aggregates. Because the solution NMR methods employed here are not able to report on these insoluble species, there is no data to say conclusively, but it is highly probable that the His-Pro bond in this portion of the sample has been isomerized, and what is left in solution is in an intermediate conformation.

The Interaction of PKC with Pin1 is specific

Due to the unexpected nature of the changes seen in PKC upon the addition of Pin1, it seemed important to investigate whether this was a specific interaction between the two, or if this was just an acceleration of the intrinsic isomerization of the His-Pro moiety due to local pH changes altering the protonation of the histidine (Reimer et al. 1997), or a simple desolvation effect (Dugave and Demange 2003) which have both been shown to accelerate isomerization.
rates. To accomplish this, the changes in Pin1 upon adding T641E PKC-CT were monitored. As shown in Figure 3.5, the addition of unlabeled T641E PKC-CT to $^{15}$N labeled Pin1 causes changes in a handful of the crosspeaks in the $^1$H-$^{15}$N HSQC spectra of Pin1. Using the published assignments (Jacobs et al. 2002), the affected residues can be identified, and mapped back to the crystal structure (Ranganathan et al. 1997). These residues localize to the two possible binding sites of Pin1, the WW domain, and the catalytic site of the PPLase domain. If the changes in PKC were due just to a simple chemical acceleration of the intrinsic isomerization as mentioned above, it would be unlikely that such localized changes in functionally relevant regions would be seen in Pin1. This is strong support that the PKC/Pin1 interaction is specific, and that Pin1 is catalytically active towards the carboxyl tail of PKC, despite the unexpected specificity.

**The catalytic actions of Pin1 towards PKC are complex**

Based on the finding that the interaction between T641E-PKC-CT and Pin1 was specific and not just random acceleration of the intrinsic isomerization, studies were initiated to investigate possible reasons behind the unexpected isomerization of the His636-Pro637 bond rather than the phospho-mimetic Glu641-Pro642 bond, namely whether or not the glutamate mutation was a true phospho-mimetic in terms of the Pin1 recognition and activity. Yaffe et al. have shown that Pin1 does not have the same activity towards peptides with glutamate in place of the phospho-Thr (Yaffe et al. 1997). Additionally, Schutkowski et al. demonstrated that just the extra charge of the glutamate side chain did not fully mimic that of a phospho-Ser/Thr in terms of intrinsic cis/trans isomerization propensity (Schutkowski et al. 1998). Therefore the constructs were changed to synthetic 13 residue peptides in which an actual phospho-Thr could be incorporated in the T641 position, while continuing the investigations with the un-phosphorylated WT peptide and T641E peptides. Initially working with the un-phosphorylated peptide, the results were similar to those obtained with the 48 residue PKC-CT, yet there are some notable differences. As shown in Figure 3.7, the residues affected were N-terminal to Pro637, while Pro642 remained un-altered. However, of particular interest
is that Pin1 affected both the cis and trans populations of PKC-CT13. Notably, the spectra are completely devoid of any peaks corresponding to a cis population. Conversely, as shown in Figure 3.8, there are now at least two new conformations corresponding to a trans His636-Pro637 bond. Additionally, like with PKC-CT there is still the presence of the fibrilar aggregates after the addition of Pin1 to the sample, for which it is impossible to determine the isomerization state. Similar results were seen with the T641E peptide.

Surprisingly, there were no changes observed for the phospho-Thr peptide (Figure 3.9). It is possible, that there is a minor change to the cis His636-Pro637 conformation, but the remainder of the peaks overlay remarkably well. Therefore it can be concluded that unlike the situations of PKC interacting with HSP70 and PDK-1, a glutamate mutation to Thr641 does not mimic the phosphorylation in terms of the Pin1/PKC interaction.

While the specificity of Pin1 has been shown to be dependent on a phosphorylated Ser/Thr preceding the proline, the binding surface of the WW domain has been shown to be an extended surface recognizing more than just the phosphorylated residue (Verdecia et al. 2000). In fact, Pin1 screening against peptide libraries, revealed that Pin1 preferentially binds to peptides with arginine or aromatic residues N-terminal to the proline, and leucine or isoleucine residues C-terminal to the proline. With these selection criteria in mind, the His636-Pro637 site would be the preferred binding site of Pin1 over that of the Thr641-Pro642. As shown in the PKC sequence below the spectra in Figure 3.1, the region directly N-terminal to the histidine contains multiple phenyalanines as well as an arginine, C-terminal to Pro637 is a leucine. The residues surrounding Pro642 do not meet any of these criteria.

**Unexpected activity is supported by in vivo investigations**

While these NMR experiments were being conducted, in vivo experiments were also carried out by our collaborators to determine the biological significance of the PKC/Pin1 interaction (Abrahamsen et al 2006). Using GST pull-down experiments from COS7 cells transfected with full length PKC βII and GST tagged Pin1 constructs, it was determined that PKC does indeed interact with Pin1 in vivo. PKC was found to pull down with not only full
length Pin1, but also the isolated WW or PPlase domains. It is noteworthy to point out that unlike previous studies which demonstrated that the WW domain has higher affinity for substrates than the PPlase domain (Verdecia et al. 2000), PKC was found to have higher affinity binding to the PPlase domain. Moreover, the phosphorylation state of PKC can be judged by gel mobility shift (Dutil et al. 1994; Keranen et al. 1995), and in these studies the PPlase domain preferentially bound the un-phosphorylated PKC. Furthermore, despite the requirement of Pin1 for efficient dephosphorylation of PKC, this dephosphorylation actually results in a 5X increase in the affinity of Pin1 for PKC. These results are consistent with the NMR studies presented here in which Pin1 was catalytically active towards PKC which was not phosphorylated at the turn motif. Interestingly this increase in affinity is specific for the de-phosphorylated species, and not present for the nascent un-phosphorylated species.

**Further insight into the role of isomerization in the C-terminal tail of PKC**

While trying to find conditions which might stabilize the PKC-CT constructs so further investigations into the actions of Pin1 would be possible, it was found that Ca\textsuperscript{2+} and other divalent metals can stabilize the \textit{cis} isomer (Ng and Weis 1998; Perera et al. 1998). It seemed relevant to determine whether Ca\textsuperscript{2+} stabilizes the \textit{cis} isomer of PKC-CT\textsubscript{13}, providing a new explanation for why activating concentrations of Ca\textsuperscript{2+} render the carboxyl tail of PKC resistant to trypsin proteolysis (Keranen and Newton 1997). Using a concentration of Ca\textsuperscript{2+} as close as possible to the stoichiometry of the activating conditions used in (Keranen and Newton 1997), dramatic changes were seen in the TOCSY spectra of PKC-CT\textsubscript{13}. Figure 3.10 are the overlaid TOCSY spectra for the WT (a) and pT (b) peptides, in 0 M Ca\textsuperscript{2+} (black) and 1.5 M Ca\textsuperscript{2+} (red). As expected, due to the increased ionic strength, all resonances have slightly different chemical shifts when referenced to DSS. Striking, however, is that while the majority of the resonances have similar intensities both with and without the Ca\textsuperscript{2+}, all peaks corresponding to the prolines are dramatically reduced. This would be consistent with the prolines now being in an intermediate exchange regime with respect to the chemical shift time scale. This leads to
the proposal that Ca\(^{2+}\) stabilizes the *cis* isomer of both Pro637 and Pro642 and this is the active conformation of the carboxyl tail of PKC.

**Proposed model of how isomerization of the carboxyl tail of PKC regulates activity**

By combining the results of the *in vivo* work conducted by our collaborators concurrently with these NMR investigations, a new model for PKC regulation can be proposed.

The *in vivo* studies, elucidated that the biological role of Pin1 is to promote the down-regulation of PKC (Abrahamsen et al. 2006). If the active conformation of the carboxyl tail of PKC is the *cis* isomer of the prolines as proposed (Figure 3.10) it would seem logical that down regulation would be accomplished by the removal of these isomers. However, *in vivo* work mentioned earlier is somewhat contradictory, in that not only Pin1 is required for efficient dephosphorylation of PKC, but it also has a higher affinity for the de-phosphorylated species. Therefore, it is proposed here that the actions of Pin1 towards PKC-βII are two-fold. The first action is to covert the active *cis* isomer of the pThr641-Pro642 bond into the *trans* species which is now more susceptible to dephosphorylation (Weiwad et al. 2000; Zhou et al. 2000). While there is no direct NMR evidence to support this claim, there is no evidence to dispute it either. The initial NMR studies of with PKC-CT did not provide evidence for the *cis* isomer of the His636-Pro637 bond, yet the trypsin cleavage experiment (Figure 3.3) indicated it was present. It wasn’t until sample concentrations were high enough that there was NMR evidence to support this. If Pin1 isomerizes this *cis* pT641-Pro642 conformation to the *trans* conformation, it would not be noticed in the current spectra. The data presented here are thus consistent with Pin1 isomerizing the pT641-Pro642 site from the active *cis* conformation to the *trans* conformation prompting the efficient de-phosphorylation of Thr641. Given the higher affinity of Pin1 for phosphorylated substrates (Yaffe et al. 1997), while still phosphorylated, the His636-Pro637 site is untouched by Pin1 and remains *cis*. Consistent with this proposal is the lack of isomerization of the His636-Pro637 bond in the phosphorylated peptide.

Though, once de-phosphorylated, the second action of Pin1 would be the isomerization of the *cis* His636-Pro637 bond. Figure 3.8 was direct evidence that this is
possible. The biological work explained earlier demonstrated that Pin1 has a higher affinity for the dephosphorylated species, but that this high affinity interaction relies on PKC being active and then de-phosphorylated. The non-activated PKC would still be in the _trans_ conformation, and thus not an efficient substrate for Pin1. In _vivo_, once PKC is de-phosphorylated, it accumulates in the detergent-insoluble fraction of the cell (Edwards et al. 1999), from which it is ultimately degraded. Pin1 was shown to be necessary for this ubiquitin mediated degradation (Abrahamsen et al. 2006). The catalytic actions of Pin1 on the PKC-CT peptides prompted the formation of insoluble fibrilar aggregates in the NMR tube. This suggests the possibility that these fibrils are more than a mere artifact as initially thought, but actually biologically relevant. Proline isomerization has previously been shown to promote fibril formation in β-2-microglobulin (Eakin et al. 2006; Jahn et al. 2006), and perhaps more relevant is the fact that Pin1 de-regulation is implicated in the amyloid plaques of Alzheimer’s disease (Lu et al. 1999; Lu et al. 2003; Pastorino et al. 2006). The isomerized carboxyl tail could be the nucleation point for the remainder of the protein. Furthermore, one can propose from these results that the changes seen the peaks corresponding to the _trans_ population are not from the catalytic actions of Pin1 directly, but rather the transient equilibrium between insoluble and soluble species. Supporting this proposal is that the biological role of HSP70 is to rescue PKC from degradation. Figure 3.11 demonstrated that this is partially accomplished by removing the peaks corresponding to this intermediate species. While HSP70 was not able to completely reverse the insolubility _in vitro_, it is possible that its presence _in vivo_ could keep PKC soluble, allowing it to be re-phosphorylated rather than degraded.

These conclusions are summarized in a new model of PKC regulation presented in Figure 3.12. As previously described, newly synthesized PKC is recruited to the membrane. In order to phosphorylate the activation loop (Thr500) PDK-1 docks onto the carboxyl tail of PKC, which is in the _trans_ conformation. Once PKD-1 is released, PKC precedes to autophosphorylate the two sites on the carboxyl tail, the turn motif (Thr641) and the hydrophobic motif (Ser660). This fully phosphorylated species accumulates in a compact protease resistant, thermally stable conformation in the cytosol. Upon the generation of the
Figure 3.12

New proposed model for the regulation of PKC including phosphorylation and isomerization of the carboxyl tail. (A) Newly synthesized PKC associates with the membrane, where PDK-1 docks onto the carboxyl tail, in which all prolines are in the trans conformation, and phosphorylates the activation loop (T500). (B) Upon the release of PDK-1, PKC subsequently autophosphorylates two sites on the carboxyl tail, the turn motif (T641) and the hydrophobic motif (S660). (C) PKC returns to the cytosol in its primed yet inactive conformation. (D) Generation of secondary messengers, Diacylglycerol and Ca$^{2+}$ promote the translocation of PKC back to the membrane, where it adopts an open conformation, and the Ca$^{2+}$ stabilizes the active cis conformation of the carboxyl tail. (E) After prolonged activation of PKC, Pin1 will eventually catalyze the cis to trans isomerization of the pThr641-Pro642 bond, rendering this site sensitive to de-phosphorylation. (F) After dephosphorylation, PKC is once again a substrate for Pin1, only the recognition site is the cis His636-Pro637 bond. Once this site is isomerized, this conformation of PKC is highly susceptible to aggregation and eventually proteolyzed via the ubiquitin proteosome pathway (H), or HSP70, can bind to the carboxyl tail preventing aggregation, thus rescuing PKC from ultimate degradation.
secondary messengers, Diacylglycerol and Ca^{2+}, PKC translocates back to the membrane, where the Ca^{2+} stabilizes the cis isomers of both sets of proline bonds His636-Pro637 and pThr641-Pro642. The down regulation of PKC is initiated by the Pin1 catalyzed isomerization of the pThr641-Pro642 bond from the cis to the trans conformation. This trans conformation is now more susceptible to phosphatase dependent dephosphorylation (Zhou et al. 2000). Once dephosphorylated, PKC is again a substrate for Pin1, except this time the site of isomerization is the His636-Pro637 bond. As with the pThr641-Pro642 bond, Pin1 isomerizes the cis to trans conversion of the His636-Pro637 site. This isomerization creates a conformation of the carboxyl tail which is susceptible to aggregation, and PKC localizes to the detergent insoluble fraction of the cell, where it is eventually proteolized via the ubiquitin proteosome. However, if present, HSP70 can rescue PKC from this degradation by reversing the actions of Pin1 at this site.

**Conclusion**

The research in this chapter, not only demonstrated an intrinsic proline isomerization in the carboxyl tail of PKC, but importantly that this biologically relevant conformational change can be catalyzed by the proline isomerase Pin1. This work initiated subsequent *in vivo* investigations in which the biological significance of this new PKC interacting partner was determined. Additionally, due to the residue specific power of NMR spectroscopy, it was determined that the catalytic actions of Pin1 towards PKC were not as straightforward as would be predicted. However, through the collaborative efforts of the structural work presented here, and the *in vivo* biological work (Abrahamsen et al 2006), a new model is proposed for the regulation of PKC (Figure 3.12). By including the isomerization state of the carboxyl tail of PKC, the question of how PDK1 and HSP70 discriminate between chemically identical carboxyl tails is addressed, and importantly, more light is shed into the mechanism by which HSP70 rescues PKC from degradation.
The text of this chapter in part will be submitted for publication as “The PKC-Pin1 Interaction: A Complex Twist to the Tail”, manuscript in preparation. I was the primary researcher and author for this publication, and Hilde Abrahamesn, Alexandra C Newton and Patricia A. Jennings the co-authors directed and supervised the research.
Chapter 4

The Unfolding of Interleukin 1-β Monitored by PFG-Diffusion NMR:
Looking for A Molehill Next to A Mountain
**Introduction**

Molecules in solution are dynamic in nature, and experience a variety of motional degrees of freedom including vibrational, rotational, and translational motion. Translational Brownian molecular motion is often simply called diffusion. Diffusion is dependent on the physical parameters of both the molecule and its solution environment. The diffusion coefficient, $D$, is described by the Stokes Einstein equation 4-1:

\[
D = \frac{kT}{f}
\]

where, $k$ is the Boltzmann constant, $T$ is the temperature, and $f$ is the frictional coefficient, which for the simple case of a spherical particle can be described by equation 4-2:

\[
f = 6 \pi \eta r_h
\]

where $\eta$ is the viscosity of the solvent, and $r_h$ is the hydrodynamic radius of the particle. Therefore, the diffusion coefficient can be used as a probe of size and shape of individual molecules as well as aggregates. This is particularly relevant and interesting in protein systems which undergo conformational changes in which the secondary structure remains largely unchanged, but the overall packing or tertiary structure is altered. For example, when calmodulin binds Ca$^{2+}$ there is almost no change in the amount of secondary structure, but rather a major rearrangement of the $\alpha$-helices surrounding the Ca$^{2+}$ binding site. (Babu et al. 1988; Kuboniwa et al. 1995; Zhang et al. 1995)

Pulsed-field gradient (PFG) NMR is one method used to measure translational diffusion by utilizing gradient pulses to spatially encode the position of each nuclear spin. In a magnetic field, NMR active spins precess at their Larmor frequency according to the equation 4-3, where $\omega_0$ is the Larmor frequency, $\gamma$ is the gyromagnetic ratio, and $B_0$ is the strength of the magnetic field.

\[
\omega_0 = - \gamma B_0
\]

When an external gradient is applied, the resulting magnetic field at any given coordinate can be described as the sum of the main field and the field due to the applied gradient (equation 4-4) where $r$ describes the position of a specific spin.
By combining equations 4-4 and 4-5, one sees that the application of a gradient field alters the Larmor precession of the NMR active nuclei in a spatially dependent and controlled manner.

The simplest PFG - echo experiment (Stejskal and Tanner 1965) is given in Figure 4.1 for the purpose of explaining the principles behind PFG-diffusion NMR. The initial 90° pulse (a) creates the transverse magnetization, which precesses about the magnetic field at a rate proportional to the magnitude of the magnetic field. When the first “encoding” gradient pulse is applied (b) the coherence created by the 90° pulse is lost due to the individual spins experiencing a slightly different magnetic field from the gradient pulse. After a delay period \( \Delta \), a second “decoding” gradient is applied (d). Due to the inversion of the spins from the 180° pulse (c), this gradient pulse has the opposite effect as the first gradient pulse. If the molecule has not experienced significant translational motion during the diffusion delay, the second gradient pulse will exactly reverse the loss of coherence. However, if the molecule diffused appreciably during the delay, the effect of the applied gradient will no longer be exactly opposite of the previous and full coherence will not be retained, resulting in attenuation to the signal.

Since the systems of interest in this investigation are proteins, modifications to this simple pulse-program were necessary, and this sequence is given in Figure 4.2. First, due to long correlation times, transverse or \( T_2 \) relaxation is more detrimental to the signal than longitudinal or \( T_1 \) relaxation, so the gradient echo sequence above was modified to the stimulated echo sequence (Tanner 1970). In this sequence, a pair of 90° pulses replace the usual 180° pulse in-between the gradient pulses so that during the diffusion delay \( (\Delta) \) magnetization is stored longitudinally. Second, when gradients are applied, this changing magnetic field, in addition to affecting the Larmor frequency of the spins of interest, creates extra currents in the surrounding coils of the probe called “eddy currents” which distort the spatial encoding and decoding of the spins. This problem has been alleviated by the addition of an additional delay prior to data acquisition (Gibbs and Johnson 1991), as well as using a pair of bipolar gradients to effectively cancel the eddy currents (Wu et al. 1995). Additionally,
Figure 4.1

Illustration of the principals behind PFG-diffusion using the Stejskal and Tanner sequence (Stejskal and Tanner 1965). The initial 90° pulse creates the transverse magnetization (a). This coherence is destroyed by the first “encoding” gradient, and the spins are labeled spatially (b). The 180° pulse inverts the spins (c). A second “decoding” gradient is applied (d). If the molecule has not experienced significant translational diffusion during the diffusion delay, (Δ) the second gradient exactly reverses the loss of coherence experienced by the first gradient. However, if the molecule diffused appreciably during the delay, the effect of the second gradient is not exactly opposite of the previous, and complete coherence is not regained, resulting in signal attenuation.
Figure 4.2

The BiPolar-LED pulse program used for the diffusion experiments in this chapter. The differences between this experiment and the one in Figure 4.1 are 1) the magnetization is longitudinal during the diffusion delay ($\Delta$) to minimize the signal loss due to $T_2$ relaxation. And 2) this sequence compensates for eddy currents by employing a short delay prior to acquisition, as well as using a pair of opposing gradient pulses for both the encoding and decoding gradient pulse. Additionally, a WATERGATE sequence was added for water suppression. For reference, the equation describing the signal attenuation as a function of gradient strength is also shown.

$$I(G)=I(0)\exp\left[-\left(\gamma G \delta\right)^2 D \left(\Delta - \frac{\delta}{3} - \frac{\tau}{2}\right)\right]$$
because our samples of interest are in 90% H2O, a WAERGATE sequence (Sklenar et al. 1993) was added for water suppression.

The diffusion experiment is simply this 1D sequence applied repeatedly, gradually incrementing the gradient strength. The resulting signal intensity can be fit to equation 4-5 (Wu et al. 1995) to solve for the diffusion coefficient (D);

$$I = I_0 e^{-D\gamma^2G^2\delta^2(\Delta - \delta/3-\tau/2)}$$

where $I_0$ is the initial intensity, $\gamma$ is the gyromagnetic ratio, $G$ is the gradient strength, $\delta$ is the length of the gradient pulse, $\Delta$ is the length of the diffusion delay, and $\tau$ is length of the gradient recovery delay.

The research described in this chapter was to implement this powerful NMR technique at UCSD, on a well behaved, yet biologically relevant problem. An additional goal of this work was to employ every technique possible to optimize the precision with which diffusion coefficients could be measured for proteins, ultimately to determine the extent to which one could estimate the size (or hydrodynamic radius), in these systems. PFG diffusion NMR has been previously been applied to study the native and non native states of multiple proteins (Jones et al. 1997; Pan et al. 1997; Wilkins et al. 1999; Weljie et al. 2003) These studies provide the biological community with insight into the changes in hydrodynamic radius of proteins under a variety of differing conditions. However, the differences in the conformation of the proteins in these previous studies are quite large compared to other functionally relevant changes in proteins which may be of interest. Therefore knowledge of the limits to which changes hydrodynamic radius which can be accurately determined is important.

Interleukin1-β (IL-1β) is a pleiotropic cytokine with multiple biological functions. The size, shape and solubility of IL-1β are constantly changing. IL-1β is initially synthesized as a precursor protein (Pro-IL-1β) (Black et al. 1988), which is cleaved by Interleukin Converting Enzyme (ICE), also known as Caspase1 (Dinarello 1998). The mature IL-1β is then secreted by a non-classical mechanism in which it is hypothesized to be partially unfolded (Rubartelli et al. 1990). Once secreted, IL-1β has numerous necessary roles in the immune response however, when over expressed or under regulated, IL-1β is also associated with numerous
diseases (reviewed in (Arend 2002)). Interestingly, IL-1β has been found in the fibrils of atherosclerosis (Dinarello and Wolff 1993), as well as the amyloid plaques of Alzheimer’s Disease (Griffin et al. 1995; Griffin et al. 2000). So in addition to serving as an excellent theoretical model the various conformations of IL-1β are biologically relevant.

The native, folded conformation of IL-1β is well documented both by X-ray crystallography and NMR (Finzel et al. 1989; Clore et al. 1990; Clore et al. 1991). It represents the typical β-trefoil fold (Murzin et al. 1992). The structure consists of twelve β-strands, six of which form a barrel like structure, and the remaining six forming a cap. These secondary structural elements are arranged such that the protein exhibits a pseudo 3-fold axis of symmetry. Each trefoil consists of two hairpins, one which is part of the barrel, and the other part of the cap. A somewhat unique aspect of the IL-1β structure is that these hairpins are held together by structural water molecules (Clore et al. 1990). These native structural features are of IL-1β are illustrated in Figure 4.3

Previous equilibrium unfolding studies of IL-1β monitored with intrinsic tryptophan fluorescence have shown a semi-stable hyperfluorescent species present at low denaturant levels (Finke and Jennings 2002). Attempts to study this partially destabilized conformation have elucidated that the overall structure remains intact, that only a handful of residues experience changes in $^{13}$Cα chemical shifts, and that there is a global de-stabilization as measured by an increase in HD exchange (Roy et al. 2005). Kinetic, residue specific H-D exchange NMR studies have shown that the residues to disappear first in the unfolding of IL-1β are those which coordinate the structural waters (Roy and Jennings 2003). These studies suggest that the structural waters could act as lynchpins stabilizing the native structure, and that initial destabilization of the protein may be due to the removal of these interactions.

We chose to monitor the GdnHCl induced unfolding of IL-1β with via PFG-diffusion NMR to determine if there was any detectable change in the hydrodynamic radius of IL-1β to correlate with the structural waters being expelled from the protein, allowing it to expand in this hyper-fluorescent state. Because the change in hydrodynamic radius of this hyper-fluorescent state compared to the native state was predicted to be minor, several measures were taken to
Figure 4.3

The native structure is a well characterized member of the β-trefoil family. The overall structure (A) is that of 12 β-strands, in which 6 form a barrel (shown in red) and the other 6 form a cap (shown in blue). These elements of secondary structure exhibit pseudo 3D symmetry, as shown in (B), the color coding is the same as (A). Each individual trefoil is labeled (I-III). An additional structural feature of IL-1β are the structural waters, shown as yellow dots in (B) and shown in spheres in (C) where it is clearly evident how they serve to hold the β-strands together stabilizing the tertiary structure.
ensure the highest quality and accuracy of data possible were obtained, and are discussed in the following sections. Because such care was taken with the data acquisition and analysis, reproducible results with relatively small error bars were obtained. Therefore it was possible to determine the relative error associated with this type of hydrodynamic radius calculations, and it was demonstrated that there is no detectable difference in the hydrodynamic radius of the hyper-fluorescent species and the native state outside these newly established limits. Furthermore, it was determined that internal small molecule controls for viscosity can be problematic, and that it is possible to use a protein as an external control. Additionally, these PFG-diffusion experiments provide the first known structural information on the unfolded species of IL-1β which is not only proposed to be important for cellular secretion, but also implicated in plaque formation of the Alzheimer's disease.

**Results and Discussion**

**Data collection**

As described above, the diffusion experiment consists of acquiring a series of 1D spectra at increasing gradient strengths, and fitting the resulting intensities to the diffusion equation. The first step to obtaining accurate and reliable diffusion coefficients depends on the quality of this data acquired. There are essentially three main sources of artifacts encountered in PFG-diffusion NMR experiments; 1) the previously mentioned eddy currents, 2) non-uniform gradients, and 3) convection currents caused by temperature gradients (Antalek 2002). The following steps were taken to minimize each of these sources of error. As described in the introduction, contributions from eddy currents were accounted for and eliminated by the choice of pulse sequence. The signal decay in equation 4-5 assumes a linear gradient is applied across the sample, and this is often not the case with commercially available probes, especially at the edges of the gradient coils. The simplest way to overcome this problem is to restrict the sample volume to the central region of the coil which is most linear (Tillett et al. 1998), something easily accomplished with a Shigemi tube. The third source of artifacts, convection currents, is arguably one of the largest sources of errors, and often the most
difficult to control. The temperature control on most spectrometers is accomplished by flowing variable temperature (VT) air through the bottom of the probe past the sample tube ultimately exiting through a port at the top. This easily creates a situation in which the sample near the bottom of the tube may experience a slightly different temperature than the sample near the top of the tube. This temperature gradient can cause convection currents which will artificially modulate the apparent translational diffusion coefficient of interest. Like the eddy currents, there are pulse sequences available for the compensation of convection currents (Jerschow and Muller 1997). However, the additional length of this pulse sequence leads to greater loss of signal due to relaxation. This is especially problematic when working with protein samples which are both initially low in concentration, and highly prone to signal loss due to relaxation. Multiple alternative steps were taken to account for this potential source of error. First, all experiments were conducted at room temperature to minimize the amount of temperature variability necessary from the VT air source. Second, the Shigemi tube has been found to minimize the effects from convection currents (Tillett et al. 1998). More importantly though, if present, the convection currents would have been more noticeable along the longitudinal Z axis than the transverse X and Y axis (Jerschow and Muller 1997; Dehner and Kessler 2005), therefore, diffusion experiments were run using each gradient axis (X,Y & Z) independently. Because all three values were in strong agreement, it can safely be said that the convection currents were adequately compensated for. A representative series of spectra of IL-1β acquired with all these considerations in mind is given in Figure 4.4.

Obtaining diffusion coefficients

As outlined in the introduction, the diffusion coefficient is obtained by fitting the attenuated signal intensity as a function of gradient strength to equation 4-5. The accuracy of the resulting diffusion coefficient is obviously dependent on the quality of the integrated data, and multiple factors were taken into consideration to ensure the highest quality data possible.

To avoid contributions from exchange effects with water (Tillett et al. 1998), the signals from -1.0 – 1.5 ppm were integrated, and fit to equation 4-5 to obtain the diffusion coefficient.
Figure 4.4

Representative data from one of the diffusion experiments. Clearly seen is the signal attenuation as the gradient strength is incremented from 5-95% of its maximum strength, which for the Z axis in this data was determined to be 42.17 gauss/cm. The indicated region from -1.0-1.5 ppm was integrated and resulting normalized volumes were fit to equation 4-5 to solve for the diffusion coefficient.
The importance of this choice of integral region is illustrated in Figure 4-5. The average diffusion coefficient from three independent experiments on each of the three gradient axis obtained from integration of the amide and aliphatic region on a sample of Lysozyme is plotted on a bar graph. Due to chemical exchange with water, the average diffusion coefficient of the amide region \(1.61 \times 10^{-6} \pm 0.05 \times 10^{-6} \) cm\(^2\)/sec is significantly greater than the average diffusion coefficient obtained for the aliphatic region \(1.32 \times 10^{-6} \pm 0.03 \times 10^{-6} \) cm\(^2\)/sec. Adding to the motivation to integrate the aliphatic region of the spectrum rather than the amide is the presence of the strong broad signal from the GdnHCl which was used to denature the protein resonating around 6.7 ppm which overlapped with the amid signals.

In addition to contributing to the exchange effects in the amide region, the residual water signal, especially at the higher GdnHCl concentrations where water suppression was less effective, resulted in a baseline falling with a rate consistent with the attenuation of the water signal. To account for this, a region of noise devoid of any signal (-1.0 to -2.0 ppm) was integrated and the corresponding integral was subtracted from the aliphatic signal prior to normalizing and fitting for the diffusion coefficient.

Examples of the fitted data and residuals from the Z axis for one of the titrations at pH 5.2 are shown in Figure 4.6. The titrations were done at two different pH's because previous work (Finke and Jennings 2002) has shown that the hyperfluorescence is more pronounced at pH 6.5. Additionally, each titration was repeated at two different protein concentrations (0.5, and 0.3 mM) to determine if any contributions from aggregation were present. For error bars, the diffusion experiment was run at least twice on each sample, and the points at 0.0 and 0.8 M GdnHCl where the hyperfluorescence is at its maximum, were run five additional times. When all these replicates were repeated on all three gradient axis, a total of 12-30 independent diffusion coefficients were determined for each point of the two titration curves.
Average Diffusion Coefficients

Amide: $1.61 \times 10^{-6} \pm 0.05 \times 10^{-6} \text{ (cm}^2\text{/sec)}$
Aliphatic: $1.32 \times 10^{-6} \pm 0.03 \times 10^{-6} \text{ (cm}^2\text{/sec)}$

**Figure 4.5**

Diffusion coefficients for three different replicate diffusion experiments on a sample of Lysozyme performed on the three different gradient axes. As clearly seen, the diffusion coefficients for the amide region are artificially enhanced compared to the aliphatic region due to contributions from exchange with the bulk solvent.
Figure 4.6

Representative fitted data from one of the replicates of each GdnHCl point of one of the titrations at pH 5.2. Data shown is from an experiment when the gradients were applied along the Z axis. The normalized integral volumes from each diffusion experiment were fit to equation 4-5. These plots are drawn with the best-fit curve to the data shown in red. The residuals for each curve are shown as well as the blue line in the graphs to the right of the fitted data. The GdnHCl concentration for each point is indicated in the upper right had corner of the fitted data plot.

As described in the text, for error bars, similar curves were obtained in replicate for each sample, and to examine for the possible contributions of convection currents, the diffusion experiments were also repeated on the transverse axis (X, and Y) in addition to the Z axis shown here, for a total of at least 6 curves per sample.

To examine for aggregation effects, the complete titration at each pH (5.2 and 6.5) was repeated at two different protein concentrations (0.5 and 0.3 mM), thus giving at least 12 diffusion coefficients per point in each titration curve.
Correcting for viscosity

Because the diffusion coefficients are dependent on viscosity (eqns 4-1, 4-2), the effects from GdnHCl induced changes on the solution viscosity need to be taken into account to accurately compare diffusion coefficients for the various samples. One previous method to account for viscosity changes (Chen et al. 1995; Jones et al. 1997) has been to use a small molecule as an internal standard. Assuming the hydrodynamic radius of the small molecule reference does not change, the ratio between the diffusion coefficient of the protein, and that of the small molecule can be used as a probe for how the hydrodynamic radius of the protein changes as it unfolds. The reference material used in this work was a small amount (<0.5%) of protonated acetate buffer, which is normally fully deuterated. The buffer is in such excess of the protein that any contributions from protein-buffer interactions would be minimal. Additionally, this minimized the number of chemical species that were added to the samples to interfere with the unfolding reaction. Because the peak from the small molecule is not in an isolated region of the spectrum as shown in Figure 4.7 its decay had to be fit to a double exponential (eqn 4-6) to which the diffusion coefficient for the protein ($D_p$), and initial intensity ($I_p$) were constrained to the values obtained from fitting the aliphatic data.

\[
I = I_p e^{-D_p \gamma^2 G^2 \delta^2 (\Delta - \delta / 3 - \tau / 2)} + I_s e^{-D_s \gamma^2 G^2 \delta^2 (\Delta - \delta / 3 - \tau / 2)}
\]

A representative fit is also shown in Figure 4.7.

Unfolding curves

The plots of the average protein diffusion coefficients relative to the small molecule internal reference vs. GdnHCl are shown in Figure 4.8. The averages plotted were calculated from each of the replicate experiments as describe above, excluding only those outliers falling outside the 99% confidence level of a Q test, the error bars represent the standard deviation of the population. These unfolding curves are close representations of previous unfolding curves of IL-1β with mid-points around 1.2-1.5 M GdnHCl (Finke and Jennings 2002). What is striking is that in the unfolding curve at pH 6.5, there is an initial decrease in relative diffusion.
Figure 4.7

Expansion of the data shown in Figure 4.4 to highlight the acetate peak which was initially used as a small molecule internal control for viscosity. Due to its relatively small size, the acetate peak decays much more rapidly than the protein. Because it is not in an isolated region of the spectrum, the signal was fit to a double exponential equation (4-6) to remove the contribution from the background protein decay. The fitted plot to this data is also shown.
\[ I(G) = I(P) \exp \left[ - (\gamma G \delta)^2 D_p \left( \Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right] + I(A) \exp \left[ - (\gamma G \delta)^2 D_A \left( \Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right] \]
coefficient consistent with the hyperfluorescence. However, there is a similar downward trend in the same region of the Bovine Pancreatic Trypsin Inhibitor (BPTI) unfolding curve which shouldn’t unfold with these concentrations of GdnHCl (Creighton 1977b; a).

**Small molecule internal standards can be problematic for folding studies**

Shown in Figure 4.9 is the plot of the small molecule diffusion coefficient vs. GdnHCl. As clearly visible in the blue series which is the BisTris (pH 6.5), and slightly apparent in the red series, the acetate (pH 5.4), there is a strange non-linear behavior at the same low GdnHCl concentrations in question. To investigate whether this was due to contributions from the protein; either the BisTris interacting with the protein, or the double exponential fit not adequately removing the protein contribution from the diffusion coefficient, the diffusion coefficient of the BisTris was determined in samples of GdnHCl without the protein. This titration is included in Figure 4.9 as well. The chemical shifts of the BisTris peaks were the same with and without the protein, (data not shown) indicating that the strange trend to the diffusion coefficients was not due to the improbable interactions between the BisTris and the protein. Additionally, the chemical shift difference between the BisTris peak at 0 M GdnHCl and 2.0 M GdnHCl is not more than would be expected due to the increasing ionic strength of the sample indicating that the BisTris is not directly interacting with the GdnHCl either. However, the graph of the diffusion coefficient vs. GdnHCl concentration of this sample without the protein (Figure 4.9) reveals the same non-linear behavior, indicating that this is not due to the protein background. While the exact mechanism of how GdnHCl denatures proteins remains uncertain, it is believed to alter the solvation properties of water (Plaxco et al. 1997; Vanzi et al. 1998). The ratio of the diffusion coefficients of two species is approximately equal to the reciprocal of the cube roots of their molecular weights (Waldeack et al. 1997). If the GdnHCl is altering the hydration of the BisTris even by one water molecule, this would have a dramatic effect on the diffusion coefficient, due to the relative molecular weight difference. So
Figure 4.8

Plot of the diffusion coefficients for IL-1β relative to the small molecule internal standard at pH 5.2, (red), pH 6.5 (Blue) and BPTI (Yellow). Curiously, in the pH 6.5 titration, there is an initial decrease in relative diffusion coefficient in the low levels of GdnHCl consistent with the hyperfluorescence seen in the previous unfolding experiments. However, there is also a slight downward trend to the BPTI data as well, and given that BPTI should not unfold at these GdnHCl concentrations, this should be a straight line, and was investigated further.
Figure 4.9

Plot of the diffusion coefficients for the small molecule viscosity standards. Even without the background of the protein, the BisTris (yellow) trend still displays the non-linear trend in the low GdnHCl concentrations as was seen in the samples with the protein (blue). Albeit to a much lower extent, a similar trend can be seen in the acetate data (red).
even when the small molecule isn’t interacting with the protein (Jones et al. 1997), the assumption that its hydrodynamic radius does not change is not always valid, and therefore is not always a suitable standard for viscosity control.

\[
BPTI \text{ as a viscosity standard and to calculate the hydrodynamic radius of IL-1β}
\]

Because of this observed effect of changing hydrodynamic radius due to the GdnHCl, a standard which more closely represented the protein of interest, yet wouldn’t unfold under these conditions was desired. As mentioned previously, BPTI does not unfold at these concentrations of GdnHCl (Creighton 1977b; a). However, due to the extreme resonance overlap between the two proteins, it was necessary to use this as an external control. When the diffusion coefficients of IL-1β were compared relative to those of the BPTI, the unfolding curves are artifact free, and shown in Figure 4.10. This the first time the successful use of an external control for viscosity has been reported for diffusion studies. When fit for thermodynamic parameters (data not shown) the results obtained with this data is in strong agreement with previous studies monitoring the unfolding with CD or fluorescence (Finke and Jennings 2002). Additionally, by using the well established hydrodynamic radius of 15.2 Å for BPTI (Pan et al. 1997) the relative diffusion coefficients can easily be used to calculate the hydrodynamic radius of IL-1β at each titration point using equation 4-7 (Jones et al. 1997)

\[
4-7 \quad r_h^{\text{protein}} = \frac{D_{\text{BPTI}}}{D_{\text{protein}}}(r_h^{\text{BPTI}})
\]

Table 4.1 contains the average diffusion coefficients relative to BPTI, again only excluding the outliers which are outside the limits of a 99% confidence level of the Q test as well as their standard deviations. The calculated hydrodynamic radii for each the various GdnHCl concentrations are also given.
<table>
<thead>
<tr>
<th>[GdnHCl]</th>
<th>Relative D</th>
<th>StdDev</th>
<th>$r_h$ (Å)</th>
<th>StdDev</th>
<th>Relative D</th>
<th>StdDev</th>
<th>$r_h$ (Å)</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0 M</td>
<td>1.45</td>
<td>0.02</td>
<td>22.04</td>
<td>0.30</td>
<td>1.44</td>
<td>0.02</td>
<td>21.89</td>
<td>0.30</td>
</tr>
<tr>
<td>0.2 M</td>
<td>1.46</td>
<td>0.02</td>
<td>22.19</td>
<td>0.30</td>
<td>1.42</td>
<td>0.04</td>
<td>21.58</td>
<td>0.61</td>
</tr>
<tr>
<td>0.4 M</td>
<td>1.45</td>
<td>0.03</td>
<td>22.04</td>
<td>0.46</td>
<td>1.44</td>
<td>0.04</td>
<td>21.89</td>
<td>0.61</td>
</tr>
<tr>
<td>0.6 M</td>
<td>1.45</td>
<td>0.04</td>
<td>22.04</td>
<td>0.61</td>
<td>1.45</td>
<td>0.02</td>
<td>22.04</td>
<td>0.30</td>
</tr>
<tr>
<td>0.8 M</td>
<td>1.47</td>
<td>0.03</td>
<td>22.34</td>
<td>0.46</td>
<td>1.47</td>
<td>0.04</td>
<td>22.34</td>
<td>0.61</td>
</tr>
<tr>
<td>1.0 M</td>
<td>1.49</td>
<td>0.04</td>
<td>22.65</td>
<td>0.61</td>
<td>1.56</td>
<td>0.03</td>
<td>23.56</td>
<td>0.46</td>
</tr>
<tr>
<td>1.2 M</td>
<td>1.59</td>
<td>0.08</td>
<td>24.17</td>
<td>1.22</td>
<td>1.98</td>
<td>0.05</td>
<td>30.10</td>
<td>0.76</td>
</tr>
<tr>
<td>1.4 M</td>
<td>1.82</td>
<td>0.06</td>
<td>27.66</td>
<td>0.91</td>
<td>2.46</td>
<td>0.05</td>
<td>37.39</td>
<td>0.76</td>
</tr>
<tr>
<td>1.6 M</td>
<td>2.28</td>
<td>0.06</td>
<td>34.66</td>
<td>0.91</td>
<td>2.65</td>
<td>0.05</td>
<td>40.28</td>
<td>0.76</td>
</tr>
<tr>
<td>1.8 M</td>
<td>2.53</td>
<td>0.05</td>
<td>38.46</td>
<td>0.76</td>
<td>2.71</td>
<td>0.06</td>
<td>41.19</td>
<td>0.91</td>
</tr>
<tr>
<td>2.0 M</td>
<td>2.61</td>
<td>0.07</td>
<td>39.67</td>
<td>1.06</td>
<td>2.79</td>
<td>0.03</td>
<td>42.41</td>
<td>0.46</td>
</tr>
<tr>
<td>2.5 M</td>
<td>2.71</td>
<td>0.06</td>
<td>41.19</td>
<td>0.91</td>
<td>2.88</td>
<td>0.07</td>
<td>43.78</td>
<td>1.06</td>
</tr>
</tbody>
</table>
Figure 4.10

Plot of the IL-1β diffusion coefficients relative to BPTI. Both of these unfolding curves at pH 5.2 (red) and pH 6.5 are artifact free, and there is no statistical difference between the natively folded (0 M GdnHCl) and hyperfluorescent species (0.8 M GdnHCl). Exact values for each point are given in table 4.1 above.
Establishing the lower limits of changes detectable by PFG-Diffusion NMR

The lowest standard deviation obtained in either titration was 0.3 Å or 1.4%, and is therefore proposed to be the lower limit for change in hydrodynamic radius which can be detected by PFG-NMR under somewhat ideal conditions. However, as seen in the table above, when the samples are perturbed, as in this case by the addition of a denaturant, the errors increase. This can be attributed to multiple factors. First, the translational diffusion coefficient is representative of the average hydrodynamic radius for a given sample. It is well established that proteins do not adopt a single static structure, but rather are present in an ensemble of species (Sadqi et al. 1999). When perturbed by the denaturant, it is reasonable to assume that the range of structures present in this ensemble increases leading to a broader range of diffusion coefficients. Additionally or alternatively the error could be from technical aspects of the experiments given the high salt concentrations. Increasing the ionic strength of the solvent has multiple effects, first it necessitates longer 90° pulses which in turn can lead to sample heating, which would affect the translational diffusion. Additionally, by changing the ionic strength, the viscosity is ultimately affected, and while care was taken to account for changes in viscosity, it is possible that its effects were not completely eliminated.

Characterizing the hyperfluorescent state

Using this data, there is no detectable difference in the hydrodynamic radius of IL-1β between the natively folded species, and the hyperfluorescent species at 0.8 M GdnHCl. The difference between the two is 0.3 Å, which is just at the lower error limit and within the experimentally determined error of 0.46 Å for the 0.8 M GdnHCl concentration. The same average hydrodynamic radius, with just an increase in the error associated with the hyperfluorescent species would be consistent with an increase in dynamic “breathing” motions in this ensemble. This would also be consistent with the previous work which reported no overall structural change but an observed increase in HD exchange rates (Roy et al. 2005).
Characterizing the unfolded state

Previously, Wilkins et all (Wilkins et al. 1999) have shown a correlation between the number of residues and the hydrodynamic radius of a protein, for both the folded and unfolded states. Given that IL-1β is 153 residues, the predicted radii are $20.43 \pm 4.7 \, \text{Å}$ and $38.87 \pm 17.6 \, \text{Å}$ for the folded and unfolded species respectively. In this study, the hydrodynamic radius of the native folded state of IL-1β was determined to be $22.04$ and $21.89 \pm 0.3 \, \text{Å}$ at pH 5.2 and 6.5 respectively. While slightly larger than the predicted 20.43 Å, it is still within the error, and given the large hydrophobic cavity in the β-barrel of IL-1β (Figure 4.3), one would not expect it to have a hydrodynamic radius of a completely collapsed protein. The radii determined for the unfolded state of $41.19 \pm .91 \, \text{Å}$ and $43.78 \pm 1.06 \, \text{Å}$ for pH 5.2, and 6.5 respectively, are also well within the predicted error for the unfolded species. The difference in the values obtained for the two different pH’s could be attributed to the fact that the isoelectric point of IL-1β is around 6.0 so the overall charge of the two states would be different, lending to different electrostatic interactions which could either serve to hold residual structure in the pH 5.2 sample, or lead to electrostatic repulsions in the pH 6.5 sample. These values of the hydrodynamic radius of unfolded IL-1β indicate that upon unfolding, the overall size of IL-1β increases anywhere from 88-100%, a parameter which was previously unknown. This study represents the first structural information of this unfolded species which is found to aggregate in vitro (Finke et al. 2000) and in the amaloid plaques of Alzheimer’s Disease (Griffin et al. 1995; Griffin et al. 2000).

Conclusions

The PFG-Diffusion NMR method of determining translational diffusion coefficients has been implemented in our laboratory, and the experimental limitations have been well defined. These experiments have successfully been used to follow the GdnHCl induced unfolding of IL-1β and are consistent with previous studies (Finke and Jennings 2002). Furthermore, the native hydrodynamic radius of IL-1β was determined to be $22.04 \pm 0.3 \, \text{Å}$, a
value which was previously unknown. More noteworthy, it was determined that the difference in the hydrodynamic radius of IL-1β between the native folded, and hyperfluorescent species is on the edge of our limit of detection. It is more likely that the increase in fluorescence be due to a local change in the environment of the tryptophan, which taken in context of the increase in HD exchange rates (Roy et al. 2005), and this diffusion data must be due to an increase in the dynamic fluctuations of the ensemble, not a global structural change. Furthermore, the hydrodynamic radius of IL-1β was shown to increase by 88%-100% depending on the pH, and these values are quite consistent with other unfolded proteins.

In addition, because these studies were conducted in such a manner to eliminate as many sources of error as possible, with many replications, it was possible to determine a statistical limit for the change in hydrodynamic radius necessary to be detected by PFG-Diffusion NMR. It was determined that under normal solution NMR conditions, the minimum change in hydrodynamic radius must be greater than 1.4%. It is important to point out at this point that this is similar to the error that is associated with such measurements by ultracentrifugation (Kirschner and Schachman 1971; Cantor and Schimmel 1980).

Additionally, these studies suggest that the small molecule standard may not be the ideal choice for a viscosity control. While it was determined that the BisTris was not interacting directly with the GdnHCl or the protein, it did experience changes in hydration with increasing GdnHCl concentrations. Given the proposed model of denaturation by GdnHCl, of altering the solvation properties of water (Plaxco et al. 1997; Vanzi et al. 1998), a molecule in which the relative change in hydration is minimal compared to the molecular weight, such as a protein is a preferred standard. Importantly, this work was the first known demonstration that it is possible to use an external control for viscosity.

Appendix I contains a protocol for setting up these diffusion studies as well as information regarding data analysis. Included in this appendix, are the necessary pulse-programs for the experiments and gradient calibrations, as well as gradient pulse limitations. Sample Matlab scripts use to fit the data are also included.
The text of this chapter in part will be submitted for publication as “The Unfolding of IL-1β Monitored by PFG-Diffusion NMR: Looking for a Molehill Next to the Mountain”, manuscript in preparation. I was the primary researcher and author for this publication, and Richard K. Shoemaker and Patricia A. Jennings the co-authors listed directed and supervised the research.
Chapter 5

Conclusion and Future Directions
The work presented in this dissertation illustrates the application of a diverse set of solution NMR spectroscopy tools, applied in the investigation of structural characterization and conformational changes in biological systems. A variety of experimental methods were applied with the unifying goal of gaining insight into a wide range of conformational changes important to the function of signaling proteins.

**What did we learn about PKC?**

The research in chapters 2 and 3 were directed at understanding the solution structure and conformational changes of the carboxyl tail of PKC. The initial discovery in the early 1980’s that PKC serves as the major receptor for the potent carcinogenic phorbol esters (Castagna et al. 1982), as well as its implication in numerous essential cellular events including differentiation, cellular activation, modulation of secretion, apoptosis, cytoskeletal rearrangements, and the immune response, (reviewed in (Blobe et al. 1996; Toker 1998)), has prompted extensive research into the factors governing the activity and regulation of PKC, and has been extensively reviewed (Nishizuka 1986; Newton 1995; Mellor and Parker 1998; Parekh et al. 2000; Newton 2001; 2003).

Despite this wealth of biological information, there is relatively little structural information regarding PKC. The last 48 residues of PKC serve as the docking site for the upstream kinase PDK-1 (Gao et al. 2001), and HSP70 (Gao and Newton 2002), two protein-protein interactions which are key to regulating PKC activity (Figure 1.5). Recognizing the importance of this carboxyl tail region for regulating the activity of PKC, we set out on a reductionist approach to obtain insight into the regulation of PKC through conformational changes in the carboxyl tail. The initial goal of this project was to determine the structural basis for the phosphorylation dependent binding of these two regulatory proteins to the carboxyl tail of PKC.

By employing solution NMR techniques as described in chapter 2, it was determined that regardless of phosphorylation state, the carboxyl tail of PKC is intrinsically unstructured. There are regions however, which display a propensity to populate α-helical regions of Φ/Ψ
space (Figure 2.6). By comparing these results with the recently published crystal structure of PKC ι (Messerschmidt et al. 2005), it was seen that these helical regions correspond to α-helices which are stable in the full length phosphorylated protein. Importantly, in this crystal structure, it is mainly the phosphate groups which mediate the interaction between the carboxyl tail and the upper lobe of the kinase core. When in the context of the full length protein, and phosphorylated, it is quite possible that the regions displaying α-helical character in PKC-CT from PKC-βII will be stabilized by these tertiary interactions. However, in the absence of these interactions as shown in this work, the carboxyl tail of PKC is intrinsically unstructured. Given the importance of the non-phosphorylated carboxyl tail serving as the docking site for PDK-1 and HSP70, this finding is particularly relevant. As discussed in chapter 2, this inherent disorder provides the flexibility necessary to accommodate the different binding sites of multiple proteins, with appropriate specificity and affinity.

While characterizing the structure of PKC-CT, continual time dependent changes were observed for all constructs (Figure 3.1). As described in chapter 3, through investigating these changes, it was discovered that the peptidyl-prolyl isomerase Pin1 catalyzes the proline isomerization of the carboxyl tail of PKC. Despite growing predominance in regulating signaling proteins through phosphorylation–directed proline isomerization (Zhou et al. 1999; Lu et al. 2002), Pin1 had never been implicated in regulating PKC activity prior to the investigation of the observed time dependent structural changes. More importantly, due to the residue specific information which can be obtained by utilizing NMR spectroscopy, it was determined that the actions of Pin1 towards PKC are more complex than initially predicted, with Pin1 catalyzing an unexpected isomerization of a His-Pro site (Figure 3.8). Through the combination of the structural work presented here and the in vivo work conducted concurrently (Abrahamsen et al. 2006) a new model for PKC regulation was proposed (Figure 3.12 vs. Figure 1.5). By including the discovery of a Pin1 catalyzed isomerization of the carboxyl tail, there is now an explanation for why PDK1 and HSP70 display preferences for the chemically identical, nascent un-phosphorylated vs. de-phosphorylated PKC respectively. Also, in addition to more insight into the processes controlling the down regulation of PKC, the
research described here provides more information into the structural basis for how HSP70 rescues PKC from degradation.

**Where can we go from here?**

The research presented here has certainly provided novel insight into the structure of the carboxyl tail of PKC and conformational changes which are important for regulating activity. However, as with most scientific research, these discoveries can be expanded upon. First, this research was on the functionally relevant (Gao et al. 2001; Gao and Newton 2002), yet isolated tail; it would be interesting to compare the results obtained here with data from full length PKC. With the recent advances in NMR instrumentation, studies on larger proteins are becoming feasible. Despite these advances, one of the remaining hurdles of such studies is the spectral complexity due to the incredible number of resonances. However, this can be alleviated by employing specific labeling strategies. One such strategy is that of segmental labeling through the intein technology (Xu and Perler 1996). If sufficient amounts of the full length kinase could be obtained, by using this method one would be able to isotopically label the same 48 residues of the carboxyl tail used here, but spliced into the context of the kinase, which would be unlabeled. This would allow for heteronuclear NMR investigations of the carboxyl tail, without the spectral complexity of the remainder of the kinase. The same CSI method of analysis could be employed to assess phosphorylation dependence of the structure of the tail. Through this research, we know that the phosphorylations alone do not alter the structure of the carboxyl tail, but, is their functional role to provide the necessary tertiary contacts with the kinase domain, and thus stabilize the $\alpha$-helical structures as proposed here, and if so, to what extent do they become stabilized? While the homonuclear methods employed here to address isomerization states would not be possible with such a construct, there are methods which take advantage of $^{13}$C chemical shifts to determine isomerization states (Schubert et al. 2002)

Additionally, this work revealed complex interactions and activity between PKC and Pin1. Traditionally, Pin1 catalyzes the isomerization of a phosphoThr-Pro peptide bond (Yaffe et al. 1997). However, as presented in chapter 3, Pin1 was discovered to catalyze an
unexpected isomerization of a His-Pro peptide bond (Figures 3.7 & 3.8) which led to fibrilar aggregates of the PKC-CT constructs. As discussed in the proposed model (Figure 3.12) these fibrils are potentially biologically relevant. While the solution NMR methods employed here were not able to address the conformation of PKC in these fibrils, solid state NMR is uniquely suited for such investigations (Tycko 2001; Castellani et al. 2002). It would be interesting to compare the structure and isomerization state of these fibrils to what is found in solution in order to glean more insight into the unexpected catalytic actions of Pin1 towards PKC, and the rescue actions of HSP70.

Additionally, it would be interesting to study the enzyme-substrate complex in this instance to gain further details of the mechanism of Pin1 action. While this was not possible in these investigations due to the tendency of the PKC-CT to aggregate after isomerization, if one would switch to smaller constructs it might be possible to still retain the necessary PKC-Pin1 interactions, while removing those residues which nucleate the fibril formation.

Also, these studies were all conducted with the carboxyl tail sequence of PKC-βIII. The in vivo work (Abrahamsen et al. 2006) demonstrated that the Pin1-PKC interaction was conserved among most of the PKC isozymes. However, as seen in Figure 2.2, the conservation of both of the prolines involved in the isomerizations presented here (Pro637 and Pro642) are not strictly conserved. Does Pin1 still have the multiple catalytic effects on these other isozymes, or is this something specific to the regulation of PKC-βIII?

An additional possible direction would be to incorporate fluorinated proline residues into the full length PKC and assess the activity. It has been shown previously (Renner et al. 2001), that depending on the exact position of the fluorine in the proline residue (4R)- vs. (4S)-F-Pro, the isomerization preferences are altered, with the 4(S) favoring the cis isomer. This could help support the claim made here that the active form of PKC is one in which Ca\(^{2+}\) has stabilized the cis isomer.

Despite these remaining questions, the work described in chapters 2 and 3 not only shed light into the intrinsic disorder of the carboxyl tail of PKC, but also led to the discovery of the Pin1 catalyzed proline isomerization of the carboxyl tail, a conformational change key to
regulating activity of Protein Kinase C. In addition to prompting the discovery of the Pin1 catalyzed isomerization, the residue specific nature of the NMR experiments utilized here demonstrate an unexpected activity of Pin1, which is supported by, and helps explain some of the results from the in vivo investigations.

**PFG-Diffusion as a method to study conformational changes in proteins.**

In addition to providing residue specific information of local conformational changes as described above, NMR can readily provide structural information regarding global conformational changes. NMR methods have also been developed to implement precisely calibrated pulsed magnetic-field gradients (PFGs) to spatially encode the nuclear precession frequencies based upon their physical location in the sample. In addition to forming the basis for Magnetic Resonance Imaging (MRI) tools, which have revolutionized clinical medicine, the ability to spatially encode and decode the nuclear spins using PFG methods allows the NMR spectroscopist to accurately measure the macroscopic molecular motion of the molecules within the samples (Stejskal and Tanner 1965). This motion, commonly referred to as molecular diffusion, depends on the size and shape of the molecule. PFG-diffusion NMR offers a versatile method in which to determine the hydrodynamic radius of a molecule through measurement of the diffusion coefficient. This experimental technique has been used to examine both native and non-native conformations of a variety of proteins (Jones et al. 1997; Pan et al. 1997; Wilkins et al. 1999; Weljie et al. 2003) However, the extent to which the precision of these measurements can be made is not well documented.

With the eventual goal of being able to use the PFG-diffusion methods to study conformational changes in proteins, chapter 4 of this dissertation was aimed at implementing the PFG-diffusion methods on a well behaved yet biologically relevant system, and in the process, optimize the precision with which diffusion coefficients could be measured for proteins, ultimately to determine the extent to which one could estimate the hydrodynamic radius.
In chapter 4, the GdnHCl induced unfolding of IL-1\(\beta\) was monitored by PFG-diffusion NMR, and the hydrodynamic radius of IL-1\(\beta\) was determined for the range of biologically relevant conformations from fully folded to unfolded (Table 1.4). For this work, it was determined that the change in hydrodynamic radius must be greater than 1.4% in order to be detected by PFG-diffusion methods. Because such care was taken to eliminate and control for every known source of error, it is proposed that this represents the limit of precision which can be expected for other PFG-diffusion investigations. Importantly, this precision is similar to what one can expect when utilizing highly optimized detection systems from analytical ultracentrifugation, a biophysical technique commonly employed for such problems (Kirschner and Schachman 1971; Cantor and Schimmel 1980). Additionally, as described in the text, these studies suggest that the previous method of utilizing a small molecule as an internal standard for viscosity (Chen et al. 1995; Jones et al. 1997) may be problematic, and an acceptable alternative is to use an external standard which more closely mimics the protein under investigation.

**How can these PFG-diffusion experiments be used in the future?**

In addition to providing structural information on the variety of biologically relevant conformations of IL-1\(\beta\), the goal of this work was to implement these powerful methods to enable others to study conformational changes in proteins with the highest precision possible. While there are countless examples to which these techniques could be applied, some are particularly relevant to the body of work described in this dissertation.

Because only unlabeled samples are required, the variety of systems this method can be applied to is broader than other NMR methods, and useful information can be obtained on proteins typically inaccessible by other multidimensional NMR techniques. Take for example the many conformations of PKC, it is known that phosphorylation of PKC on the carboxyl tail results in an enzyme which is in a more thermally stable, protease and phosphatase resistant conformation, (Edwards and Newton 1997). Additionally, it is known that PKC adopts a conformation upon activation by Ca\(^{2+}\), that is different from the conformation of just being
recruited to the membrane (Keranen and Newton 1997). Chapter 3, (Figure 3.10) has provided evidence that at least part of this is due to proline isomerization within the carboxyl tail. However, the extent of how this isomerization affects the rest of the kinase is not known. PFG-diffusion NMR provides a method to assess how the global structure of PKC changes under these varying conditions.

An additional application of PFG-diffusion NMR as pointed out by Altieri et al. (Altieri et al. 1995), is the use of PFG-Diffusion to find optimal NMR conditions for new samples. There are a variety of methods available for determining the oligomeric state of a protein such as dynamic light scattering, and ultra-centrifugation. However, in none of these methods are the sample conditions similar to what would be necessary for high resolution NMR studies. PFG-diffusion NMR is a method for determining the oligomeric state of a protein in conditions identical to those which would be used for further NMR investigations. PFG-diffusion experiments do not require labeled samples, so it is quite feasible to optimize solution conditions on a variety of samples prior to obtaining and preparing an expensive labeled one.

**19F NMR to study conformational changes in proteins**

One of the hurdles that must be overcome when attempting to study disordered proteins, or conformational changes in larger proteins by NMR spectroscopy, is resonance overlap caused by the characteristic lack of dispersion among the individual resonances or sheer number of resonances respectively. A technique which can be applied to overcome this hurdle is specific labeling (Mcintosh and Dahlquist 1990; Kelly et al. 1999). Traditional specific labeling for biomolecular NMR involves incorporating $^{15}$N or $^{13}$C only at selected residues. While these nuclei are important to the biomolecular NMR community, another nucleus, $^{19}$F is commonly overlooked. However, there are several useful properties of the $^{19}$F nucleus (Gerig 1989; Danielson and Falke 1996). First, $^{19}$F is a spin ½ nucleus which occurs at 100% natural abundance, and is about 83% as sensitive to NMR detection as $^1$H. For comparison, $^{13}$C is only 1.6% as sensitive. Second, the $^{19}$F chemical shift range is about 100x that of the $^1$H. The $^{19}$F chemical shift is controlled primarily by the lone pair electrons which provide a large
paramagnetic term in the shielding formula. This property also renders the chemical shift exceptionally sensitive to changes in local van der Waals environment as well as local electrostatic fields, making it highly attractive when studying subtle conformational changes. $^{19}$F NMR has successfully been used to study conformational changes in proteins, especially those too large to study by traditional NMR methods (Luck and Falke 1991b; a; Luck et al. 2000; Salopek-Sondi and Luck 2002).

However, $^{19}$F NMR spectra using modern solution-state NMR probes are often complicated by very broad signals in the baseline, which are NMR signals attributable to fluorinated polymers used in probe construction. This broad probe-background is a well known problem with many NMR probes, and the baseline can become an unmanageable, sinusoidal, rolling baseline. In some cases, this can render the NMR spectrum intractable. Since the broad signal from the solid material in the probe will not refocus under a Spin-Echo ($90-\tau-180-\tau$-detect), and the narrow signals from the molecule of interest which is in solution does refocus, the spin-echo technique is a well known and commonly used method to eliminate the broad background signals in $^{19}$F NMR of small molecules (Dr. Richard K. Shoemaker, University of Colorado - Boulder, personal communication). This spin echo method was successfully applied to proteins of varying sizes and labeling patterns and the results are illustrated in the following figures.

**Initial implementation on well behaved proteins**

Just as with the diffusion experiments, a well behaved model system was desired for initial implementation of the $^{19}$F NMR methods. Illustrated in Figure 5.1 are $^{19}$F spectra acquired on a 0.5 mM sample of IL-1$\beta$ specifically labeled with pF-Phenylalanine. Spectrum (A), a simple $90^\circ$ acquire experiment clearly demonstrates the sinusoidal rolling baseline as described earlier. The later spectra (B-C) represent spin echo sequences with increasing $\tau$ delays. As clearly evident, once the echo delay is long enough (C), the spin echo sequence adequately removes the broad background signals from the probe, resulting in spectra with an acceptable baseline, with approximately (~70-75%) the same signal to noise level as the
Figure 5.1

$^{19}$F NMR spectra of IL-1β specifically labeled with pF-Phenylalanine. Spectrum (A) a simple 90° acquire experiment. Clearly visible is the sinusoidal rolling baseline due to the broad background signals from the fluorocarbon polymers used in the construction of the probe. Spectrum (B) was acquired with the Spin-Echo sequence, with a $\tau$ delay of 50 $\mu$s. The background baseline roll is not quite attenuated. Spectrum (C) was acquired with a $\tau$ delay of 300 $\mu$s. Note the flat baseline with little attenuation of the desired protein signals, larger $\tau$ delays will lead to further attenuation of the desired signals in addition to the broad signals of the probe background.
90° acquire sequence. When the τ delay is increased beyond this, the intensity of the protein signal does begin to decrease (data not shown).

This spin-echo sequence was also applied to the SH2 domain of the C-terminal Src kinase (CSK) which was specifically labeled with 5-F-Tryptophan, and spectra without (a) and with the spin-echo (b) are given in Figure 5.2

**A possible cis/trans equilibrium in the Super Folder Green Fluorescent Protein (sfGFP)**

Folding studies in our lab have indicated the folding reaction of the “super folder” variant of the Green Fluorescent Protein is limited by a proline isomerization (data not shown). Because one of the proline residues follows the lone Trp in GFP, $^{19}$F NMR is uniquely suited for probing such a conformational change (Seifert et al. 2002). Figure 5.3 represent the $^{19}$F spectra acquired on a ~0.1 mM GFP sample which is specifically labeled with 5-F-Tryptophan. For comparison, Figure 5.3 (A) is a normal 90° acquire spectra, and 5.2 (B) is the spin echo sequence (300 µs τ delay). The peak at -38 ppm is the pF-Phenylalanine that was added as an internal reference compound, and the broad peak at -45.8 ppm is the resonance from Trp57 in sfGFP. The peak in between the two is an unidentifiable artifact, assigned as such due to the sharp nature, and lack of splitting due to J-coupling. While the signal to noise is too low to definitively determine the presence or lack of multiple populations, this data illustrate the usefulness of the spin echo for even a large (246 residue) protein. If these experiments were repeated on a more concentrated sample, it will be possible to determine if there are multiple conformations of this Trp-Pro bond, and such experiments are in progress.

**Combining the $^{19}$F with the PFG diffusion experiments**

As stated previously, IL-1β is initially synthesized as a precursor protein (Pro-IL-1β) (Black et al. 1988), which is cleaved by ICE, prior to secretion (Dinarello 1998). While it is known that once cleaved, the mature IL-1β proceeds to fold into its well known native structure (Figure 4.3).
Figure 5.2

$^{19}$F NMR spectra of the SH2 domain from CSK specifically labeled with 5-F-Tryptophan, further illustrating the application of a Spin-Echo sequence for removal of the baseline artifacts resulting from the broad signals of the probe polymers. The signal at -38 ppm is pF-Phe which was used as an internal reference (Luck and Falke 1991a) Spectrum (A), was a simple $90^\circ$ acquire experiment, while spectrum (B) was acquired with the Spin-Echo sequence with a $\tau$ delay of 300 $\mu$s.
Figure 5.3

$^{19}$F NMR spectra of sfGFP specifically labeled with 5-F-Trp. Spectrum (A) a simple 90° acquire experiment. Spectrum (B) the Spin-Echo sequence acquired with a $\tau$ delay of 300 $\mu$s, and 4x the number of scans (16384 scans) as spectrum (A). The internal pF-Phe is expanded in (C) to highlight the $^1$H-$^{19}$F coupling present. And the peak from Trp57 in sfGFP is expanded in (D), not visible is a second peak to correspond to a minor population of the Trp-Pro bond.
However, what happens to the 14 kD precursor region is unknown. The sequence of Pro-IL-1β contains two Trp residues, one in the Mature IL-1β region, and one in the precursor region. If one were to specifically label Pro-IL-1β with $^{19}$F Trp, and then cleave it with ICE, it would be possible to monitor the diffusion coefficients of each piece independently. The possibility of such experiment is illustrated in Figure 5.4, which illustrates the PFG-echo-decay series vs. gradient strength from a diffusion experiment performed on a sample of 5-F-Trp.
Figure 5.4

$^{19}$F NMR diffusion decay series of 5-F-Trp, illustrating the possibility of combining the diffusion experiments with nuclei other than $^1$H for the hydrodynamic radius of multiple components in a single solution, by simplifying the number of resonances in the spectra.
Final Conclusion

The work described in this dissertation illustrates the application of a wide variety of solution NMR techniques for studying conformational changes in proteins on a variety of scales, from local residue specific fluctuations to global unfolding transitions.

It was discovered that regardless of phosphorylation state, in the absence of tertiary contacts with the kinase domain, the carboxyl tail of PKC demonstrates a helical propensity, but overall is intrinsically unstructured. By investigating time dependent changes observed in the spectra, an intrinsic proline isomerization in PKC-CT was revealed. More importantly, it was discovered that this isomerization can be catalyzed by Pin1. Because of the residue specific nature of the NMR techniques employed here, an unexpected activity of Pin1 was discovered, which is supported by, and helps explain some of the results from the in vivo investigations. A new model for the maturation and regulation of PKC, addressing some of the previously unknown aspects of the process was proposed.

Additionally, the PFG-diffusion experimental methods were implemented to following the unfolding of Interleukin 1-β, providing some of the first structural information on the biologically relevant non native conformations. Because all known sources of error were addressed and the measurements were acquired in such replicate, it was possible to determine the lower limit to conformational changes which can be detected. Furthermore, this work provides the first demonstration that it is possible to use an external sample as a control for viscosity in diffusion measurements.

Finally, methods for simplifying spectra through specific labeling with $^{19}$F were explored, and experimental protocols for acquiring artifact free spectra were implemented for the future investigations of other systems.
Chapter 6

Materials and Methods
In addition to the main experimental technique of NMR spectroscopy, a large number of methods were necessary to examine how local and global conformational changes regulate the activity of the signaling proteins as studied in this dissertation. The techniques used to complete the research of this dissertation are outlined in this chapter, including the molecular biology, protein expression and purification, sample preparation and NMR experiments. The doctoral degree in part is representative of an aptitude or proficiency with a multitude of techniques. But more importantly, it represents the ability to modify existing “standard procedures” to suit the particular task at hand and overcome the numerous problems associated with non-ideally behaved samples. The theoretical and experimental training derived throughout the course of this work will serve as an excellent background from which future NMR investigations of any system can build upon.

**Protein Expression, Purification and Sample Prep**

**General**

Unlabeled samples were generated by growing the *E. coli* cells in Luria-Bertani broth (LB). Isotopically enriched samples were generated by growing *E. coli* cells in M9 minimal media (Maniatis 1986) supplemented with 2.0 g/L $^{13}$C glucose (Cambridge Isotope Labs) as the sole carbon source and/or 2.0 g/L $^{15}$N ammonium sulfate (Cambridge Isotope Labs) as the sole nitrogen source.

**PKC-CT**

WT and phoso-mimetic mutants of the carboxyl tail of PKC were all expressed and purified according to the following protocol. The pGEX plasmid containing the glutathione S-transferase (GST) tagged carboxyl terminal tail of PKC (PKC-CT) was transformed into *E. coli* BL21-Codon Plus cells (Stratagene). This particular cell line was chosen to enhance the expression levels of this mammalian based sequence in *E. coli* because of the inclusion of rare *E. coli* codons which normally reduces the expression levels in regular *E. coli* BL21 expression systems. Depending on the application, the cells were grown in either LB or
isotopically enriched minimal media containing carbanacylin (100 mg/ L cell culture) at 37°C until the OD$_{600}$ reached ~ 0.5. Protein expression was induced with 0.5 mM IPTG for four-five hours, and harvested by centrifugation. The cells were resuspended in PBS + 20 mM DTT + 20 mM PMSF, and cell lysis was performed by sonication on ice for a total of 30 minutes, 30 seconds on, 30 seconds off (total pulsing time of 15 minutes). The soluble extract was incubated with glutathione immobilized on sepharose beads (Sigma). The beads were washed with five column volumes of cleavage buffer (25 mM NaPO$_4$, 20 mM NaCl, pH 6.8) to remove unbound proteins. The GST-PKC-CT construct was eluted with 100 mM Glutathione in cleavage Buffer, and concentrated to a final volume of 5-7 mL. The GST tag was removed by thrombin cleavage (1 unit thrombin/ mg protein) at 4°C overnight. The PKC-CT was separated from the GST tag, and depleted of thrombin, by size exclusion chromatography (pharmacia S-200 column) equilibrated in column buffer (25 mM NaPO$_4$, 500 mM NaCl, pH 6.8), with the detector set at 254 nm. Fractions containing PKC-CT were pooled and concentrated with a Centricon 3 (Amicon) Purity was judged to be greater than 95% as assessed by SDS-PAGE. Approximate protein concentrations were calculated by using the theoretically determined extinction coefficient for a protein containing 7 phenylalanines (Cantor and Schimmel 1980; Gill and von Hippel 1989), of 1400 M$^{-1}$cm$^{-1}$ at 250 nm. Purified protein was buffer exchanged into NMR buffer (25mM NaPO4, 30 mM NaCl, pH 6.8 for heteronuclear work and 25 mM Bis-TRIS, 5mM NaPO$_4$ pH 6.8 for homonuclear work) using a PD-10 column and concentrated with a Centricon-3 (Amicon) to a final concentration of ~1mM.

**HSP70**

The substrate binding domain (corresponding to residues 387-552) of the *E. coli* Hsp70 chaperone, DnaK was expressed and purified according to the following protocol (Swain et al. 2006). *E. coli* BL21(DE3) cells containing the plasmid for N-terminaly His$_6$-tagged Hsp70 were grown at 34°C in 750 mL LB containing Ampicillin (100 mg / L cell culture) to an OD$_{600}$ of 1.0-1.5. Protein synthesis was initiated through heat induction. Once the cell culture had reached the appropriate OD, 250 mL of LB which had been pre-heated to 60°C
was added to the growing cell cultures, and the temperature of the incubator/shaker was increased to 42° C. Five to six hours after induction, cells were harvested by centrifugation at 10,000 g for 30 min. Cells were resuspended in Buffer A (50 mM NaPO₄, 0.5 M NaCl, 0.02% NaN₃, pH 8.0) plus protease inhibitors (1 mM PMSF, aprotinin at 2.5 mg/L, pepstatin A at 1.0 mg/L, and leupeptin at 5 mg/L). Lysozyme was added to a final concentration of 1 mg/mL and incubated while stirring on ice for 20 minutes. DNAsel was subsequently added to a final concentration of .15 mg/mL and incubated for an additional 20 minutes. Further cell lysis was accomplished via sonication on ice using five 20 second bursts, each separated by 2 minutes cooling on ice. The sonicated cell lysate was centrifuged, and the soluble supernatant was filtered and loaded onto a HiTrap Chelating HP column (Amersham) pre-charged with 50 mM NiSO₄, and equilibrated in Buffer A. Unbound proteins were washed with approximately 3 column volumes of Buffer A, and 5 column volumes of 90% Buffer A, 10% Buffer B (Buffer A + 200 mM imidazole) before a linear gradient (10-100% over 100 mLs) of Buffer B was applied to elute Hsp70. Fractions containing Hsp70 were pooled and concentrated to a volume of approximately 5-7 mL. Bound substrates which co-purified with Hsp70 were removed by denaturing and re-folding. Urea was added to the purified protein to a final concentration of 7 M, and equilibrated for approximately 1 hour. Protein was re-folded by diluting 20 fold drop-wise to 20 mM Sodium Acetate (NaAc) buffer pH 5.0. Additional protein was recovered from the insoluble fraction of the cell lysis by re-suspending the cell pellet in buffer A + 7 M urea. The remainder of the cell pellet purification procedure was as outlined above except all buffers contained 7 M urea until the final refolding into 20 mM NaAc. Protein was concentrated to a final concentration of ~1 mM using a Centricon-10 (Amersham) and buffer exchanged into NMR buffer (25mM NaPO₄, 30 mM NaCl, pH 6.8) using a PD-10 column (Amersham) prior to use.

Pin1

N-terminally His₆ tagged Pin1 expressed and purified with minor modifications to the procedure in (Ranganathan et al. 1997). Depending on the application, the E. coli BL21(DE3)
cells were grown in either LB or isotopically enriched minimal media containing Kanamycin (30 mg/ L cell culture) at 30 °C until the OD_{600} reached ~ 0.8. Protein expression was induced with 0.4 mM IPTG for four hours at 22 °C, and harvested by centrifugation. Cells were resuspended in Lysis buffer (25 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol, and 1% (v/v) Tween 20, pH 8.0), and sonicated on ice for a total of 30 minutes, 30 seconds on, 30 seconds off (total pulsing time of 15 minutes). The sonicated cell lysate was centrifuged, and the soluble supernatant was filtered and loaded onto a HiTrap Chelating HP column (Amersham) charged with 50 mM NiSO_{4}, and equilibrated in Buffer A (Lysis buffer minus the Tween 20). Unbound proteins were washed with approximately 5 column volumes of Buffer A before a linear gradient (40-100% over 125 mLs) of Buffer B (25 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole, 10 mM β-mercaptoethanol) was applied to elute Pin1. Fractions containing Pin1 were pooled and concentrated to a volume of approximately 5-7 mL. His_{6}Pin1 was digested with thrombin (~1 unit thrombin / mg protein) during dialysis against dialysis buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM MgCl_{2}, 2.5 mM CaCl_{2}, 1 mM DTT and 10% (v/v) glycerol pH 8.0) overnight at 4 °C. The dialysate was centrifuged and filtered to remove any precipitated protein, and subjected to further purification by size exclusion chromatography (pharmacia S-200 column) equilibrated in column buffer (10 mM HEPES, 100 mM NaCl, 1 mM DTT pH 7.5). Fractions containing Pin1 were pooled and concentrated to a final concentration of ~1 mM and buffer exchanged into NMR buffer using a PD-10 column (Amersham) prior to use. Purity was judged to be greater than 95% as assessed by SDS-PAGE.

**IL-1β**

IL-1β was grown and purified according to the previously published protocols (Meyers et al. 1987; Finke et al. 2000) with minor modifications. E. coli BL21(DE3) cells containing the plasmid for WT IL-1β were grown at 37 °C in either LB or isotopically enriched Minimal Media containing Kanamycin (30 mg / L cell culture) to an OD_{600} of 0.7 or 0.8 respectively to maximize the induction of the 1-Ala IL-1β isoform and minimize N-terminal processing (Covalt
Protein expression was induced with IPTG at a final concentration of 1 mM. Three to four hours after induction, cells were harvested by centrifugation at 10,000 g for 30 min. Cells were resuspended in 10 mM KPO$_4$, 0.2 mM EDTA, 5 mM DTT, and 1 mM PMSF at pH 8.0 and then lysed by sonication on ice a total of 30 minutes, 30 seconds on, 30 seconds off (total pulsing time of 15 minutes). The supernatant containing soluble IL-1β was initially made 40% saturated in ammonium sulfate; the precipitate was removed via centrifugation. Additional ammonium sulfate was added to the remaining soluble supernatant to a final concentration of 80%. The precipitated pellet containing IL-1β was dissolved in buffer A (25 mM NH$_4$OAc, 2 mM EDTA, and 1 mM β-mercaptoethanol at pH 5.4) and dialyzed overnight in buffer A at 4°C. The dialysate was centrifuged, filtered, and applied to a HiTrap SP cation-exchange column (Pharmacia) equilibrated in buffer A. IL-1β was eluted in a 40 column volume linear gradient of 25-240 mM NH$_4$OAc, pH 5.4, at a flow rate of 3 mL/min. Fractions containing IL-1β were pooled and concentrated using a Centricon-10 (Amicon). Purity was judged to be greater than 95% as assessed by SDS-PAGE. Protein concentrations were calculated by using the previously experimentally determined extinction coefficient of an A$_{280}$ of .6194 corresponding to a protein concentration of 1mg/mL. The purified protein was buffer exchanged into the application appropriate buffer using PD-10 columns (Amerasham), and concentrated to a final concentration between 1.5-2 mM and frozen in approximately 1mL aliquots.

**Specific labeling with $^{19}$F**

$^{19}$F incorporation was accomplished by supplementing the traditional minimal media with 200 mg/L pF-Phe (Sigma) or 5-F-Trp (Sigma). For higher incorporation of 5-F-Trp, 40 mg/L of 3-β-indoleacrylic acid (IAA) (Sigma) was included in the minimal media to inhibit endogenous tryptophan biosynthesis (Leone et al. 2003) Proteins were grown and purified as normal.
NMR Experiments and Data Analysis

Unless otherwise indicated, NMR experiments were performed on a Bruker DMX spectrometer with an operating frequency of 500.13 MHz equipped with a standard 5mm triple-resonance gradient TXI probe at 288 K. Unless specifically stated otherwise, samples were 90% H₂O / 10% D₂O, and were approximately 500 µL, in a standard 5mm NMR tube (Wilmad, Inc). The ¹H dimension was referenced directly to external 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), while the ¹⁵N and ¹³C dimensions were referenced indirectly to DSS as established by (Wishart et al. 1995)

Heteronuclear experiments

The two-dimensional ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectra (Mori et al. 1995) were acquired with broadband ¹⁵N decoupling during acquisition accomplished using a WALTZ16 decoupling scheme (Shaka et al. 1983), and water suppression was achieved using a WATERGATE sequence for gradient tailored binomial water suppression (Sklenar et al. 1993). Spectra were routinely acquired with 128 real t₁ (¹⁵N) points, and 1024 complex t₂ (¹H), the spectral widths were 1520.515 Hz (30 ppm), centered at 120 ppm t₁ (¹⁵N); and 7002.801Hz, (14.00 ppm) centered at the water frequency of 4.7 ppm t₂ (¹H). Quadrature detection in the indirect dimension (t₁, ¹⁵N) was obtained through the use of Echo/Anti-Echo detection (States et al. 1982).

Several triple resonance experiments were employed for the backbone resonance assignments and CSI analysis of the PKC-CT constructs, namely the HNCACB, (H)CC(CO)NH-TOCSY, and HNCA (Sattler et al. 1999). These experiments were conducted on a Bruker DMX spectrometer operating at 600.13 MHz. For all experiments, the ¹H carrier was placed on the water frequency (4.70 ppm), and the ¹⁵N carrier was placed at 120 ppm throughout the duration of the experiments. For the HNCACB and (H)CC(CO)NH-TOCSY, the ¹³C carrier was placed at 43 ppm for aliphatic excitation, for the HNCA, the ¹³C carrier was set to 53 ppm. In all experiments, when necessary the ¹³C carrier was switched to the carbonyl frequency of 173 ppm through the use of Gaussian shaped pulses. For all experiments, ¹⁵N
broadband decoupling was achieved with the WALTZ16 decoupling scheme (Shaka et al. 1983), and $^{13}$C broadband decoupling was achieved with the DIPSI2 decoupling scheme (Shaka et al. 1988). Quadrature detection schemes for all experiments were Echo/Anti-Echo in $^{15}$N and States-TPPI in $^{13}$C (States et al. 1982).

The acquisition parameters for these triple resonance experiments were as follows: The HNCACB was acquired with 128 complex points in $t_1$ ($^{13}$C), 128 complex points in $t_2$ ($^{15}$N) and 1024 real points in $t_3$ ($^1$H). The spectral widths for the HNCACB experiment were; 10261.25 Hz, (68ppm) in $t_1$ ($^{13}$C), 1824.0 (30 ppm) in $t_2$ ($^{15}$N), 8389.26 Hz (13.97 ppm) in $t_3$ ($^1$H). The (H)CC(CO)NH-TOCSY with a mixing time of 14.7 ms was acquired with 128 complex points in $t_1$ ($^{13}$C), 128 complex points in $t_2$ ($^{15}$N) and 1024 real points in $t_3$ ($^1$H). The spectral widths for the (H)CC(CO)NH-TOCSY experiment were; 12104.047 Hz, (80.45 ppm) in $t_1$ ($^{13}$C), 1824.0 (30 ppm) in $t_2$ ($^{15}$N), 8389.26 Hz (13.97 ppm) in $t_3$ ($^1$H). The HNCA was acquired with 220 complex points in $t_1$ ($^{13}$C), 110 complex points in $t_2$ ($^{15}$N) and 1024 real points in $t_3$ ($^1$H). The spectral widths for the HNCA experiment were; 3929.492 Hz, (26 ppm) in $t_1$ ($^{13}$C), 1581.08 (26 ppm) in $t_2$ ($^{15}$N), 8389.26 Hz (13.97 ppm) in $t_3$ ($^1$H).

**CSI analysis**

The chemical shift index method of analysis was use to investigate the presence of secondary structural in the PKC-CT constructs. The secondary shifts ($\Delta \delta$) were determined from the $^{13}$C$^\alpha$ chemical shifts experimentally obtained from the HNCA (Sattler et al. 1999) and comparing them to the published random coil values (Wishart and Sykes 1994a) to which the sequence dependent correction factors for the immediately proceeding and following residues were applied (Schwarzinger et al. 2001).

**Homonuclear experiments**

The two-dimensional homonuclear ($^1$H) TOCSY were acquired using a MLEV-17 mixing sequence for Hartman-Hahn transfer (Bax and Davis 1985). Initially, a variety of mixing times were used ranging from 50-65 mM, the optimum mixing time was determined to be
55.65 ms, and was used for all remaining experiments. Data were acquired with 2048 complex \( t_1 \) points, and 1024 real \( t_2 \) points, employing the TPPI quadrature detection scheme (Marion and Wuthrich 1983). The spectral widths were 5252.101 Hz (10.5 ppm) in both dimensions. Water suppression was achieved using the WATERGATE sequence for gradient tailored binomial water suppression (Sklenar et al. 1993).

The two-dimensional homonuclear \( (^1H) \) NOESY (Piotto et al. 1992) were acquired with mixing times ranging from 300-600 ms, with 600 ms being the most frequently used. Water suppression was again accomplished via the WATERGATE sequence. Data were acquired with 2048 complex \( t_1 \) points, and 512 complex \( t_2 \) points employing the States-TPPI quadrature detection scheme (States et al. 1982). The spectral widths were 5252.101 Hz (10.5 ppm) in both dimensions.

For the 13-residue PKC-CT peptide, when the correlation time of the sample was near the cross-over point for the NOE cross-relaxation pathway, ROESY experiments with a mixing time of 350 ms were acquired for the through space correlations. Data were acquired with 2048 complex \( t_1 \) points, and 512 complex \( t_2 \) points employing the States-TPPI quadrature detection scheme (States et al. 1982). The spectral widths were 5252.101 Hz (10.5 ppm) in both dimensions.

For the studies with the high Ca\(^{2+}\) concentrations, the TOCSY spinlock pulse was changed from the normal 29\(\mu s\) length to 47\(\mu s\) to minimize the amount of power that was necessary in order to protect the probe.

**Data processing**

Except for the diffusion study, which is described below, all NMR data were processed using Felix 2000 software (MSI, San Diego, CA), and analyzed in Sparky (Goddard). Prior to Fourier transformation, each dimension was apodized by a 90° shifted sinbell squared function before zero filling to at least twice the number of collected points. For the homonuclear data, a facelift baseline correction (Chylla and Markley 1993) was applied to both dimensions prior to analysis. When necessary in the indirect dimensions of the heteronuclear experiments, linear
prediction to an additional one third of the number of collected points was applied before apodization.

**PFG Diffusion experiments**

A more detailed step-by-step explanation on how to set up and analyze the data from the PFG diffusion experiments is included in Appendix I of this dissertation. However, the basic procedure, as well as the specific parameters used in this study are outlined below.

*Data collection -*

To minimize effects from convection currents, and to localize the sample to the most linear region of the gradient coil, samples for all diffusion experiments were restrained to 300 µL in a Shigemi tube (Shigemi Inc.). If present, convection current effects would have been most noticeable along the longitudinal (Z) axis (Jerschow and Muller 1997; Dehner and Kessler 2005). Therefore to exclude the possibility of such errors, diffusion coefficients were measured along all three (X, Y and Z) axes.

Gradient strengths were calibrated on a sample of 99.9% D2O (Isotec) for which the diffusion coefficient has been well established \(1.90 \times 10^{-5}\) cm\(^2\)/sec (Longsworth 1960). As described below, the attenuation of the residual water signal was followed and the data fit to the diffusion equation and solved for the gradient strength. For gradient calibration, the standard BP-LED sequence (Wu et al. 1995) was used. The experimental parameters were the following. Sine shaped gradient pulses were used with total length \(\delta\) (pulse length x2) of 3400 µs along the Z axis, and 4500 µs along the X and Y axes. The diffusion delay \(\Delta\) was .029 seconds for all axes. The data were acquired in a pseudo 2D fashion in which the gradient strength was automatically increased in 5% increments from 5%-95% of the maximum gradient power for a total of 19 1D data sets. Maximum gradient strengths for Sine shaped pulses were found to be 41.17, 30.15, and 30.20 gauss/cm on the Z, X and Y gradient coils respectively.
Diffusion experiments on the protein samples employed the BP-LED sequence (Wu et al. 1995) to which a WATERGATE sequence (Sklenar et al. 1993) was added for water suppression. The diffusion parameters δ and Δ were optimized for each sample so that the total attenuation at 95% maximum gradient strength was greater than 98% of the starting signal at 5% maximum gradient strength. Typical values for δ were 6000 - 7500 µs along the Z axis, and 7500 - 9000 µs along the X and Y axes. The values for Δ ranged from 135 - 200 ms. The data were acquired in a pseudo 2D fashion in which the gradient strength was automatically increased in 1% increments from 5-20% to and then in 5% increments up to 95% of the total gradient strength for a total of 31 1D data sets. Spectra were collected with 2048 real points and covered a spectral width of 7002.8 Hz (14.00 ppm).

Buffer conditions for the equilibrium unfolding titrations were either 50 mM Sodium Acetate-d₃ pH 5.2, or 35 mM ²H-Bis-Tris, pH 6.5. GdnHCl concentrations ranged from 0 M – 2.5 M. As an internal control for viscosity (Chen et al. 1995; Jones et al. 1997), 2 mM of respective protonated (¹H) buffer was included in each sample. The unfolding titration at each pH was repeated twice at two different protein concentrations (0.5 mM and 0.3 mM) to examine for possible artifacts arising from sample aggregation. Samples were equilibrated in the appropriate GdnHCl concentration for 24 hours prior to data acquisition.

The external BPTI control samples were 0.5 M BPTI in 50 mM NaAc pH 5.2, and were equilibrated in the GdnHCl 24 hours prior to data acquisition.

Data analysis –

The diffusion data for the gradient calibration was processed in NMR Pipe (Delaglio et al. 1995) using a 90° degree shifted sinbell squared window function after extensive zero filling. The DOSYView.tcl macro included with the NMR Pipe software was used to automatically pick the intensity of the water peak at all incremented gradient strengths. These intensities were then fit in excel to a linearized diffusion equation to solve for G:
where $D^* = D \gamma^2 \delta^2 (\Delta - \delta/3 - \tau/2)$; all known constants for each set of experiments. The resulting gradient values were then used to fit in Matlab using in house scripts, to the following function describing the diffusion attenuation (equation 4-5);

$$I = I_0 e^{-D \gamma^2 \delta^2 (\Delta - \delta/3 - \tau/2)}$$

The gradient values were adjusted according to the following:

$$G(\text{new}) = \left( \frac{D_{\text{exp}}}{D_{\text{lit}}} \right) G(\text{old})$$

until the appropriate diffusion coefficient of $1.90 \times 10^{-5} \text{ cm}^2/\text{sec}$ (Longsworth, 1960) was obtained.

The diffusion data for the protein samples was processed and integrated in SpinWorks (Marat, 2004). Data were extensively zero-filled prior to apodization with a 90° shifted sinbell squared window function. A 51 point sine function was applied for additional post acquisition water suppression. Baseline corrections to account for a falling baseline with increasing gradient strength were applied to the individual gradient increments by integrating a region of the spectrum devoid of signal and subtracting the corresponding area from the integral of the protein signal. To avoid contributions from the GdnHCl, and artifacts arising from exchange processes, the aliphatic region of the protein spectrum was integrated. Integral volumes were normalized prior to fitting in Matlab. Once the diffusion coefficient for the protein was obtained ($D_p$), these values were constrained in a double exponential equation;

$$I(G) = I(P) \exp \left( - (\gamma G \delta)^2 D_p \left( \Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right) + I(A) \exp \left( - (\gamma G \delta)^2 D_A \left( \Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right)$$

which was fit to obtain the diffusion coefficient of the small molecule internal reference standard.
Appendix I

Measuring Diffusion by Pulsed Field Gradient NMR:

A Practical Guide for UCSD
This appendix has been written to aid future researchers at UCSD in not only setting up and running diffusion experiments successfully and safely, but also processing and fitting the data. It is a collection of information from several different sources, including Bruker technical support, other people who have extensive experience with PFG diffusion, as well as published literature. Much effort was spent determining these details, and they are worth documenting to avoid the necessity of someone in the future having to re-discover what has already been done.

**Introduction to Diffusion**

**Molecular Motion**

Molecules in solution are dynamic in nature, and experience a variety of motional degrees of freedom including vibrational, rotational, and translational motion. Translational Brownian molecular motion is often simply called diffusion. Diffusion is dependent on the physical parameters of both the molecule and its solution environment. The diffusion coefficient, $D$, is described by the Stokes Einstein equation,

$$D = \frac{kT}{f}$$

where $k$ is the Boltzmann constant, $T$ is the temperature, and $f$ is the frictional coefficient. For the simple case of a spherical particle,

$$f = 6 \pi \eta r_s$$

where $\eta$ is the viscosity of the solvent and $r_s$ is the hydrodynamic radius of the particle. Therefore, the diffusion coefficient can be used as a probe of size and shape of individual molecules as well as aggregates.

Pulsed-field gradient (PFG) NMR is one method used to measure translational diffusion, by utilizing gradient pulses to spatially encode the position of each nuclear spin.
Some NMR Theory

In a magnetic field, NMR active spins precess at their Larmor frequency according to the following equation, where $\omega_0$ is the Larmor frequency, $\gamma$ is the gyromagnetic ratio, and $B_0$ is the strength of the magnetic field.

$$\omega_0 = -\gamma B_0$$

When an external gradient is applied, the resulting magnetic field at any given coordinate can be described as the sum of the main field and the field due to the applied gradient, where $r$ describes the position of a specific spin

$$B = B_0 + g \cdot r$$

By combining these two equations, one sees that the application of a gradient field alters the Larmor precession of the NMR active nuclei in a spatially dependent and controlled manner. While the direction of the applied gradient can be in any of the three dimensions, it will be restricted to the z-direction for the course of this discussion.

The Diffusion Experiment

The simplest diffusion experiment is shown and described below.
The initial 90° pulse (a) creates the transverse magnetization, which precesses about the magnetic field at a rate proportional to the magnitude of the magnetic field. When the first “encoding” gradient pulse is applied, the coherence created by the 90° pulse is lost due to the individual spins experiencing a slightly different magnetic field from the gradient pulse. After a delay period \( \Delta \), a second “decoding” gradient is applied. Due to the inversion of the spins from the 180° pulse, this gradient pulse has the opposite effect as the first gradient pulse. If the molecule has not experienced significant translational motion during the diffusion delay, the second gradient pulse will exactly reverse the loss of coherence. However, if the molecule diffused appreciably during the delay, the effect of the applied gradient will no longer be exactly opposite of the previous and full coherence will not be retained, and will result in attenuation to the signal. The extent of the molecular motion during this diffusion delay will be apparent in the degree of signal attenuation, which is a function of the diffusion coefficient \( D \) (cm\(^2\)/sec), the gyromagnetic ratio of the observed nucleus \( \gamma \) (gauss*sec\(^{-1}\)), the gradient strength \( g \) (gauss/cm), the length of the gradient pulse \( \delta \) (of sec), and the length of the diffusion delay \( \Delta \) (of sec).

\[
I = I_0 e^{-D\gamma^2g^2\Delta^2\delta^2/(\Delta-\delta/3)}
\]

The basic diffusion experiment is simply acquiring this 1D \(^1\)H spectrum at varying gradient strengths and fitting the resulting intensities to the diffusion equation above. It is now standard practice to increment the gradient strength rather than the diffusion delay (\( \Delta \)) or gradient pulse length (\( \delta \)) so the contributions to signal attenuation due to relaxation can be ignored as they are reflected in the initial \( I_0 \) value and kept constant throughout the experiment.

Rather than measuring absolute diffusion coefficients, which are sensitive to not only molecular size and shape but both temperature and viscosity as well, which can be difficult to control and adequately determine between samples, it is possible to report relative diffusion coefficients. This is where the diffusion coefficient of the protein under investigation is compared to a small molecule reference which should not change size under the varying sample conditions. When selecting a reference molecule, keep in mind that this molecule should neither interact with the protein as the protein changes nor should the solvent have any
effect on the hydrodynamic radius of this standard. If appropriate care is taken to control for solution conditions, it is possible to use an external standard which more closely mimics the sample under investigation.

Further Reading:

The following list of references contain further information on both diffusion theory and experimental details.

1) Johnson, C.S. Prog NMR Spect 34 (1999) 203-256

Measuring Diffusion

Sample Preparation

In order to obtain the best results, when measuring diffusion by PFG-NMR, Shigemi tubes should be used to ensure the sample is positioned in the center of the gradient coil where the gradient response is most linear, and to minimize the effects from convection currents. It is also recommended to check for the contribution of remaining convection currents by measuring the diffusion in all three dimensions, as convection currents will primarily be a problem in the z-axis.

Experimental Parameters

The following sections and directions will assume some familiarity with the operation of the Bruker Spectrometers, and XwinNMR software.

The basic diffusion experiment that is implemented on the DMX500 at UCSD is the Bi-Polar LED sequence, which has been modified to contain a Watergate sequence for water suppression, and the following instructions will reflect such. This pulse sequence reduces the
effects of eddy currents in the probe with minimal chance of obtaining contribution from convection currents. The pulse program diagram is shown in the figure below, but for a more detailed explanation of this sequence see reference #4 Wu et al above. Note the difference in the diffusion equation due to the extra \( \tau \) delays.

\[
I(G) = I(0) \exp \left[ - \left( \gamma G \delta \right)^2 D \left( \Delta - \frac{\delta}{3} - \frac{\tau}{2} \right) \right]
\]

As described in the introduction, the basic diffusion experiment is a set of 1D spectra acquired at a range of gradient strengths. The signal intensity is plotted as a function of gradient strength, and the resulting curve is fit for the diffusion coefficient. The quality of this curve determines the accuracy of the diffusion coefficient.

There are two main parameters which affect the signal attenuation and thus quality of the curve, and need to be set and optimized for each sample prior to collecting the actual diffusion experiment: the gradient length (\( \delta \)) and the diffusion delay (\( \Delta \)). As illustrated in the figure below, the correct values for \( \delta \) and \( \Delta \) are essential to obtaining suitable curves.

In order to reliably fit a curve, at least nine points are required. For optimal diffusion measurements, these nine points should span the entire range of gradient strengths from 5-95% of the maximal value. As shown in the last panel of the figure below with optimal delta values, the signal intensity decayed to < 2% residual intensity (i.e. greater than 98% attenuation due to diffusion) over the entire range of gradient strengths.

Another important experimental consideration is the initial signal intensity. The sample should give sufficient signal to noise such that attenuation over the entire gradient
range is possible. If the initial signal is too weak, this might not be possible, and the number of scans need to be increased.

Optimizing the Delta values:

After initial spectrometer set-up as normal (i.e. shimming, tuning/matching and pulse calibration) the following procedure is used to find the optimal delta values for each sample.

1) Run the 1D version of the diffusion experiment with a gradient strength ( gpz6) of 5%. This is the reference spectrum, and the receiver gain should be set with this experiment. Use the tools in XwinNMR to pick a particular point, and make note of the intensity.

   a. **Parameter set: Diff_Par_Opt**
   
   b. **Pulse Program: bpled_watergate_1D**

2) Then run the same experiment with a gradient strength of 95%. If the signal is still too intense (i.e. not just at the noise level), increase either the diffusion delay $\Delta$ (d20) or the gradient length $\delta$ (p30) until it is indistinguishable from the noise.

   a. Increasing $\delta$ is usually favorable because it results in a bigger effect; $\delta^2$ is determining the signal attenuation while $\Delta$ is only affecting the attenuation linearly. Also, if $\Delta$ is too long, relaxation becomes an issue.

   b. However, be aware of the following duty cycle limits for gradient pulses:

      - When using the bi-polar gradient sequence,
• The maximum gradient pulse length using SINE.100 shaped pulses is 4.5 ms.

• If needed, a square pulse (RECT.10) has ~1.4 times the power of a SINE.100 pulse

• The recommended limit for a square pulse is 3.0 ms – This has been pushed further on other spectrometers, but not tested on ours.

• Please note that Bruker specs recommend staying under 3 ms for SINE.100 pulses and 2.5 ms for square (RECT.10) pulses. While the 4.5 and 3.0 ms pulses have been used under necessity and advisement of qualified individuals – pushing it beyond this limit is not recommended.

3) Again using the tools in XwinNMR – note the intensity of the same point from the reference spectrum (gradient strength of 5%). Make sure that this value is ~2% of the starting value. If not, the initial signal to noise is not sufficient and more scans will be necessary.

4) Once you have parameters which sufficiently attenuate – back off the gradient strength to ~80% and see that there is still signal at this level and that the signal has not decayed too quickly (see figure 3).

5) If major deviations to Δ were made – double check the initial reference spectrum to ensure relaxation has not drastically affected the signal and that there is still sufficient signal at the lower gradient strengths. – If this is the case, increase the number of scans until there is sufficient signal and then double check the higher gradient values again.

The values used in the parameter set are for IL-1β, a 17.5 kD protein, are using SINE.100 gradient pulses, and Δ = 135 ms and δ=2750 µs, for the Z gradients and Δ = 135 ms and δ=3250 µs for the X and Y gradients. Note the larger the molecule, the larger these values will have to be. It is recommended to start with these values and adjust as necessary, staying within the gradient limits defined above and adjusting Δ as necessary. If the diffusion of small
molecules is to be measured, Bruker recommended lower limits of 50 ms and 1000 µs for $\Delta$ and $\delta$ respectively.

**The Diffusion Experiment**

As stated previously in the introduction, the basic diffusion experiment consists of acquiring the 1D spectrum at incremented gradient strengths and fitting the intensities or volumes to the diffusion equation to obtain the Diffusion Coefficient.

While this can be accomplished by running individual 1D experiments, this does get tedious, and if appropriate spectral analysis software are available (i.e. SpinWorks) it is possible to acquire the data in a pseudo-2D fashion where the gradient strength is incremented for each transient rather than a $t_1$ delay. The software has to allow processing of the directly detected dimension and not get confused when the indirect dimension is not processed. Note that NMRPipe can accomplish this, but it does not have the ability to integrate a region. If it is sufficient to use the intensity of a given peak vs. the area of several, NMRPipe is a processing option, and the macro “DOSYview.tcl” (included with the NMRPipe package) can be used to obtain the intensities automatically.

1) Using the parameter set **BPLED_Watergate**, (pulse program **bpled_watergate_2D**) create a new 2D experiment, and input the correct experimental parameters ($^1$H-90, $\delta$, $\Delta$, etc.). For reference, this pulse program is included at the end of this document.

- This pulse program utilizes an additional list (Difframp1) to automatically increment the gradient strength with the command “*diff” in conjunction with the p30:gp6 to define the gradient pulse. This list (Difframp1) uses 31 gradient steps from 5-20% in 1% increments to monitor the decay of the small molecule reference compound and then 5% increments from 25-95% for the remainder of the protein decay. If some other gradient ramp profile is necessary, you will need to create a new list and then change the filename that the pulse program calls up. This command multiplies the gradient strength value defined in the ASED window by the numerical values listed in this file. Therefore, the gradient strength defined in the ASED window for gpz6 should be 100%. If diffusion along another axis is to be
measured, gpz6 should be set to 0% and the desired (gpx6 or gpy6) should be set
to 100%.

2) The necessary parameters for the second dimension are only the correct number of points
(i.e. # of gradient steps included in the Difframp1 file), and the FnMODE be QF.

3) This experiment can now be run as any other 2D experiment, including as a multi-zg,
which is particularly useful when measuring the diffusion along all three gradient axis.

Data Analysis

As stated previously, the ability to use SpinWorks to process the diffusion data
simplifies the process by allowing the collection of data in a pseudo-2D fashion rather than as
individual experiments. The program was developed by Kirk Marat at the University of
Manitoba, Canada. It runs on the Windows platform, and can be downloaded free from
ftp://davinci.chem.umanitoba.ca/pub/marat/SpinWorks. A screen shot highlighting some of the
necessary buttons, and menus is shown below
Processing and integrating the data

1) Once the data folder has been transferred to your computer, use SpinWorks to open the ser file.

2) Open the Processing display with the yellow “Proc” button in the tool bar at the top.

3) From this new menu, edit the processing parameters with the “edit pars” button.

4) In this first tab (F2 detection), set the appropriate processing parameters for the direct dimension of your data. (i.e. number of points, zero filling, window function). When the data is first brought up, use a phase correction of “none”. Interactive phasing will be completed in a later step.

5) In the second tab (F1 evolution), set the size to the number of gradient increments, and the “Detection mode” to anything but “single (COSY type)”. This dimension isn’t actually processed, but if set to single, the program will not allow for interactive phasing.

6) Save the processing parameters with the “apply” button and get back to the original processing menu with the “OK” button.

7) From here, process/transform only the F2 dimension.

8) When the transform is complete, the spectrum will appear on the screen, and interactive phasing is now possible. Click the first yellow button (which looks like a poorly phased peak) to bring up the phasing display. Phase the spectra and make note of the values.

   a. Note – at any time, the data can be magnified with the + or – buttons in the top tool bar. Or a region can be defined with left clicks of the tracking cursor, and then zoomed in on using the “Zo” button. The up and down arrow buttons can be used to move the spectra around in the window.
9) To now apply these phase corrections to the rest of the gradient increments, go back into the processing parameters window and enter the values into the phase correction section, and re-transform the data as before.

10) In the processing window, there is a section “1D Block”. These two buttons are used to scroll through the 1D spectra from the different gradient increments.

11) The spectra can be integrated with the “integrate” button. Bring this window up, and using the tracking cursor, define the region to be integrated. The relative integral value is shown below the data.

12) With both the processing dialog window and integrate window open, scroll through the individual 1D spectra and make note of the integral volumes, These are the values which will be fit to the diffusion equation for the diffusion coefficient.

13) Spinworks occasionally has issues with Bruker’s digitally filtered data and will result in processed data with a sinc wiggle to the baseline centered around the water. This massively distorts the integral values, and will change over the course of the diffusion experiment because the water diffuses faster than the protein, resulting in quicker signal attenuation. To correct for this, process the data using a Sine shaped solvent filter with a size of 51. This removes the Sinc wiggles, but depending on the quality of initial water suppression can lead to a falling baseline throughout the course of the experiment. Spinworks does not have an adequate mechanism for baseline corrections for such pseudo 2-D data. However, this falling baseline can be accounted for by integrating a region of noise, (making note of the exact size of the region compared to the region of interest), and subtracting the corresponding volume from the region of interest prior to data fitting.

**Data Fitting**

Once the data have been integrated, and these integrals imputed into Excel, Matlab can be used to fit the data to the following equation to solve for the diffusion coefficients \( (D_p) \).

The initial intensity \( (I_p) \) is also a dependent variable which is left for Matlab to solve for as well.
\[ I = I_0 e^{-D_{p} \gamma^2 G^2 \delta^2 (\Delta - \delta / 3 - \tau / 2)} \]

The instructions contained here are not intended to serve as a substitute to the Matlab manual, but rather, provide example scripts used to fit diffusion data, and brief instructions in how to use them. Sample scripts are included at the end of the document.

The Excel-95 file called up by Matlab needs to be in the same directory as the Matlab script, and defined as the working directory in Matlab. The file should be only data, (no column headers, or identifiers), only numbers (no equations) and no extra formatting. The file should be laid out in the following way; the first column needs to be the gradient strength in gauss/cm. The remaining columns are the relative integral values. These scripts are expecting an initial value of close to 100, and will error out if this is not the situation. The scripts can be used to fit multiple data sets in batch format, as long as all they all have the same gradient values.

The values which need to be checked and changed when necessary are in red and include:

1) xlsread - the name of the excel file with the data to be fit
2) del – \( \delta \), the gradient pulse (p30) x2 in units of seconds
3) delta - \( \Delta \), the gradient delay (d20) in units of seconds
4) The name of the params file at the end.
5) This script also includes an “if” statement (in blue) to be used when different columns of data use different delta values. Modify the “n” value as necessary – this parameter is the column which contains the data for which the different values are necessary. If this change in parameter definitions is not necessary, comment out or delete this portion of the script.
6) The commands at the end of the script after the “%Plots” are used for creating figures of the fitted data. The red values in the “subplot” line will need to be changed to reflect the number of data sets that are being fit. These numbers represent the # of columns and rows of subplots of individual data sets. The values in the blue are the axis values, and might need to be adjusted depending
on the gradient strengths used, and how well the data fits. The first plot will contain the data, and the best fit, the second will contain the residuals to the fit. Additionally, this script calls up “Diff_rmsd” which is also included at the end. This file needs to be in the folder with the data and other Matlab script.

**Double exponential fitting**

If using a small molecule as an internal control for viscosity, it will be necessary to fit this data to a double exponential in which the protein parameters are constrained to the values obtained from fitting the protein data. The equation for this double exponential fit is the following;

\[ I = I_p e^{-D_p \gamma^2 G^2 \delta^2 (\Delta - \delta /3 - \tau/2)} + I_s e^{-D_s \gamma^2 G^2 \delta^2 (\Delta - \delta /3 - \tau/2)} \]

The first term of this equation contains the values for the slowly diffusing species (i.e. the protein), and the second for the faster diffusing, small molecule. The first term needs to be constrained to values for the protein, and this requires a couple different steps.

1) When integrating the data, make note of which block (gradient strength) the small molecule signal disappears completely behind the background protein signal.

2) Use a separate Excel sheet that contains the data for the small molecule. This sheet needs to be laid out in the same fashion as described previously for the protein.

3) The first step is to get the \(I_p\) values from this data – this is the starting integral for the protein in this particular integrated region. This value is closely approximated by fitting the later portion of the curve which is only protein and constraining the \(D_p\) value to that obtained from fitting the protein data. The script for accomplishing this is included at the end as “Matlab Script #2.”

4) The same values need to be modified as described above as well as a couple additional ones which are highlighted in bold in the sample script

   a. The diffusion coefficients from the protein data need to be added as \(D_{\text{slow}}\), formatted as shown with brackets.
b. The line $Y=A(20:31,n)$ which defines the $Y$ values for the fit, needs to be modified so that only the points in the later part of the curve which contains the protein data are included.

5) Just as the previous script called up Diff_rmsd, this script calls up the “Diff_rmsd_pro” script.

6) Once the $I_p$ values are known, these values along with the $D_p$ values are constrained in the double exponential equation to fit for the small molecule diffusion coefficient with “Matlab Script #3”.

   a. The same values need to be changed as before, as well as adding the $I_p$ values obtained from the previous script.

7) The additional script “Diff_rmsd_pro_ace” is the second script called up by this one.

**Gradient Calibration**

In order to obtain the diffusion coefficient in units of units of cm$^2$/sec, the gradient strength ($g$) in the diffusion equation needs to be in units of gauss/cm. Therefore, the gradients need to be calibrated, to convert the user define gradient strength of % to required gauss/cm.

As of Fall 2005, for the SINE.100 pulses, the gradient strengths at 100% power are around 42.17, 30.15, and 30.20 gauss/cm for the Z, X and Y gradients respectively. For the square pulses, the gradient strengths at 100% power are 65.70, 45.68, and 47.42 gauss/cm for the Z, X and Y gradients respectively. Unless there are drastic hardware changes, these should not vary much, but it is recommended to check them somewhat annually (i.e. once or twice a year).

The method used at UCSD is to calibrate the gradient strength using a sample of a known diffusion coefficient – in our case 99.9% D$_2$O, at 25°C or 298 K for which the diffusion coefficient is $1.90 \times 10^{-5}$ (cm$^2$/sec). (Longsworth L. G. *J Chem Phys* 1960, 64 1914)
This allows for the direct determination of gradient strength for a shaped pulse without having to know the area relationship of the shaped pulse to that of the square pulse, and takes any hardware imperfections into account. Both Square and SINE.100 shaped pulses can be used for the diffusion experiments. Bruker recommends using the SINE.100 shape pulse to minimize eddy current artifacts; however a sine gradient pulse is only about 65% as strong as a rectangular pulse. If measuring the diffusion of a larger protein (>20 kD), the square pulses are recommended in order to maximize the applied gradient strength (g) so the diffusion delay (Δ) can be as short as possible to minimize relaxation.

**Setting Up and Running the Experiment**

The sample of 99.9% D₂O can be treated just as one would treat a “normal” sample in regards to tuning, matching, and calibration of a 90° ¹H pulse. However, it is necessary to shim manually, as gradient shimming does not adequately work on a sample that is 99.9% D₂O.

1) The 1D pulse program used for gradient calibration is called `bpled_1D`, and is the basic BiPolarLED sequence without any solvent suppression. Load this sequence into the Diff_ParOpt experiment and optimize the δ and Δ values as described for the protein sample which should be around 1mS and 50 ms respectively.

2) The 2D pulse program used for gradient calibration is called `bpled_2d`. and just like the 1D version used to establish appropriate δ/Δ values is the basic BiPolarLED sequence without, but has the “*diff” command to automatically increment the gradient strength according to the Diframp1 list.

**Data Processing and Analysis**

The diffusion data for the gradient calibration can be processed in NMR Pipe. The benefit of this being that the DOSYView.tcl macro included with the NMR Pipe software was used to automatically pick the intensity of the water peak at all incremented gradient strengths. These intensities need to be fit in excel to a linearized diffusion equation to solve for G:
where \( D^* = D\gamma^2\delta^2(\Delta-\delta/3-\tau/2) \); all known constants for each set of experiments. The resulting gradient values can then be used in Matlab, to fit to the following function describing the diffusion attenuation for the bi-polar sequence:

\[
I = I_0 e^{-D\gamma^2G^2\delta^2(\Delta-\delta/3-\tau/2)}
\]

The gradient values then need to be adjusted according to the following until the appropriate diffusion coefficient to \(1.90e^{-5}\) cm\(^2\)/sec is obtained:

\[
G(\text{new}) = \left( \frac{D_{\text{exp}}}{D_{\text{lit}}} \right) G(\text{old})
\]
BRUKER Pulse Program for Pseudo 2D Diffusion Experiment w/ Water suppression

;bpled_watergate_2D
;avance-version (03/04/24)
;Adapted by NKruse April 2005
;Pseudo 2D sequence for diffusion measurement using stimulated echo and LED
;using bipolar gradient pulses for diffusion
;using 2 spoil gradients
;USING A WATERGATE FOR WATER SUPPRESSION
;D. Wu, A. Chen & C.S. Johnson Jr.,

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

define list<gradient> diff=<Difframp1>

"p2=p1*2"

"DELTA1=d20-p1*2-p2-p30*2-d16*2-p19-d16"
"DELTA2=d21-p19-d16-4u"

1 ze
2 d1
50u UNBLKGRAD p11:f1
p1 ph1
p30:gp6*diff ;Diffusion Encoding Gradient
d16
p2 ph1
p30:gp6*-1*diff
d16
p1 ph2
p19:gp7
d16
DELTA1
p1 ph3
p30:gp6*diff ;Diffusion Decoding Gradient
d16
p2 ph1
p30:gp6*-1*diff
d16
p1 ph4
p19:gp8
d16
DELTA2
p1 ph5
d12 ;Begin watergate
p16:gp1
d16 pl18:f1
p1*0.231 ph6
d19*2
p1*0.692 ph6
d19*2
p1*1.462 ph6
d19*2
p1*1.462 ph7
d19*2
p1*0.692 ph7
d19*2
p1*0.231 ph7
4u
p16:gp1
d16
4u BLKGRAD ;end watergate

go=2 ph31
30m mc #0 to 2 FLQF (igrad diff)
exi

ph1= 0
ph2= 0 0 2 2
ph3= 0 0 0 2 2 2 2 1 1 1 1 3 3 3 3
ph4= 0 2 0 2 0 2 0 1 3 1 3 3 1 3 1
ph5= 0 0 0 2 2 2 2 1 1 1 1 3 3 3 3
ph6= 0
ph7= 2
ph31=0 2 2 0 2 0 2 1 3 3 1 3 1 1 3

;pl1 : f1 channel - power level for pulse (default)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16 : gradient pulse for watergate
;p19 : gradient pulse 2 (spoil gradient)
;p27 : f1 channel - 90 degree pulse at p118
;p30 : Diffusion gradient pulse (little DELTA * 0.5)
;d1 : relaxation delay; 1-5 * T1
;d16 : delay for gradient recovery
;d19 : delay for binomial water suppression
;d20 : diffusion time (big DELTA)
;d21 : eddy current delay (Te)
;NS : 16 * n, total number of scans: NS * TD0
;DS : 2 * NS
;FnMODE: QF

;use gradient ratio:    gp 6 : gp 7 : gp 8
;                           var : -17.13 : -13.17

;gpz1: 20%
;gpz1: 20%
;gpz1: 20%
;gpz6: 1-100%
;gpz7: -17.13% (spoil)
;gpz8: -13.17% (spoil)

;use gradient files:
;gpnam6: SINE.100
;gpnam7: SINE.100
;gpnam8: SINE.100

;$Id: ledbpgp2sid,v 1.3.2.1 2003/05/06 12:40:00 ber Exp $
Difframp1 file for Automatic Gradient Incrementation

##TITLE= /opt/xwinmr/exp/stan/nmr/lists/gp/Difframp1
##JCAMP-DX= 5.00 Bruker JCAMP library
##DATA TYPE= Shape Data
##ORIGIN= Bruker Analytik GmbH
##OWNER= <nkruse>
##DATE= 05/05/02
##TIME= 11:00:10
##MINX= 0.000000e+00
##MAXX= 0.000000e+00
##MINY= 5.000000e-02
##MAXY= 9.500000e-01
##$SHAPE_EXMODE= Gradient
##$SHAPE_TOTROT= 0.000000e+00
##$SHAPE_BWFAC= 0.000000e+00
##$SHAPE_INTEGFAC= 5.000000e-01
##$SHAPE_MODE= 0
##NPOINTS= 31
##XYDATA= (X++(Y..Y))
5.000000e-02
6.000000e-02
7.000000e-02
8.000000e-02
9.000000e-02
1.000000e-01
1.100000e-01
1.200000e-01
1.300000e-01
1.400000e-01
1.500000e-01
1.600000e-01
1.700000e-01
1.800000e-01
1.900000e-01
2.000000e-01
2.500000e-01
3.000000e-01
3.500000e-01
4.000000e-01
4.500000e-01
5.000000e-01
5.500000e-01
6.000000e-01
6.500000e-01
7.000000e-01
7.500000e-01
8.000000e-01
8.500000e-01
9.000000e-01
9.500000e-01
##END=
Matlab Script #1 – Used for normal Data fitting

clear all     % clears all matlab data (matrices)
cf            % clear figures
global G Y M g del delta tau     % variables for another program, defined here

% _____________
% Define Independent Variable
% _________________
A=xlsread('TitrationA_Aliphatic_Z');
[M,N]=size(A);
G=A(:,1);           % define "X" Points 1-10 M in the first N
G_cont=linspace(0,45,500);
OPTIONS=optimset('Display','Off','TolX', 1e-6, 'TolFun',1e-6,
'MaxFunEvals', 1000, 'MaxIter',1000);
g=26752;           % (gauss*cm)-1
del=5.5e-3;      % (sec)
tau=0.3e-3;        % (sec)
delta=0.135;    % (sec)

% define data set and fit using Diff_rmsd.m program
%for n=2:N
 z_min=1e1;
     Y=A(:,n);
     if n>=22
         del=6.5e-3
         delta=.145
     end
for tries=1:100
     I0_start=100;
     D_start=10*rand*10e-7;
     start=[I0_start,D_start];       % starting values for fminsearch
     [final_values, z0]=fminsearch('Diff_rmsd', start, OPTIONS);
     if z0<z_min
         I0=final_values(1);
         D=final_values(2);
         z_min=z0;
     end
end
params(n-1,:)=[I0,D];

Y_calc=I0*exp(-D*G_cont.^2*del^2*(delta-del/3-tau/2));     % equation for line
Y_calc2=I0*exp(-D*g^2*G.^2*del^2*(delta-del/3-tau/2));
%equation for points
res=Y_calc2-Y;

%Plots

n
figure(1)
subplot(6,6,n-1)
plot(G,Y,'.b',G_cont,Y_calc,'-r')
axis([0 45 -5 130])
xlabel('gauss/cm')
ylabel('Integral Volume')
hold on

figure(2)
subplot(6,6,n-1)
plot(G,res)
axis([0 45 -5 5])
xlabel('gauss/cm')
ylabel('Residual to Fit')
hold on

end

params
save A_Aliphatic_Z.txt params -ascii
Diff_rmsd – To be used with Matlab Script #1

function z=Diff_rmsd(x)
global G Y M g del delta tau

I0=x(1);
D=x(2);

% Data Equation
Y_calc2=I0*exp(-D*G.^2*G.^2*del.^2*(delta-del/3-tau/2));
diff=Y-Y_calc2;
z=sqrt(diff'*diff/M);

% if I0<85 | I0>150
% z=1e30
% end
Matlab Script #2 – Used for Obtaining I_p for the double exponential fit

clear all % clears all matlab data (matrices)
clf % clear figures
global G Y M g del delta tau D % variables for another program, defined here

% _____________
% Define Independent Variable
% _____________
A=xlsread('TitrationA_Acetate_Z');
[M,N]=size(A);
G=A(20:31,1); % define "X" Points 1-10 M in the first N
G_cont=linspace(0,45,500);
OPTIONS=optimset('Display','Off','TolX', 1e-6, 'TolFun',1e-6,
      'MaxFunEvals', 1000, 'MaxIter',1000);
g=26752; % (gauss*cm)-1
del=5.5e-3; % (sec)
tau=0.3e-3; % (sec)
delta=0.135; % (sec)
D_slow=[1.14E-06 1.16E-06 1.16E-06 1.16E-06 1.15E-06 1.16E-06 1.16E-06 1.21E-06 1.22E-06 1.14E-06 1.15E-06 1.12E-06 1.17E-06 1.08E-06 1.10E-06 1.12E-06 1.12E-06 1.15E-06 1.19E-06 1.06E-06 1.09E-06 8.28E-07 8.30E-07 8.11E-07 8.01E-07 5.99E-07 6.28E-07 5.85E-07 5.92E-07 5.76E-07 5.89E-07 5.33E-07 5.45E-07]';

% define data set and fit using Diff_rmsd_pro.m program

for n=2:N
    z_min=1e1;
    Y=A(20:31,n);
    D=D_slow(n-1);
    if n>=22
        G=A(20:31,1)
        Y=A(20:31,n)
        del=6.5e-3
        delta=.145
    end

    for tries=1:100
        I0_start=100;
        D_start=10*rand*10e-7;

        start=[I0_start,]; % starting values for fminsearch
        [final_values, z0]=fminsearch('Diff_rmsd_pro', start, OPTIONS);
        if z0<z_min
            I0=final_values(1);
            %D=final_values(2);
            z_min=z0;
        end
    end
params(n-1,:)=[I0,D];

Y_calc=I0*exp(-D*g^2*G_cont.^2*del^2*(delta-del/3-tau/2));
%equation for line

Y_calc2=I0*exp(-D*g^2*G.^2*del^2*(delta-del/3-tau/2));
%equation for points
res=Y_calc2-Y;

%Plots
n
figure(1)
subplot(6,6,n-1)
plot(G,Y,'.b',G_cont,Y_calc,'-r')
axis([0 45 -5 130])
xlabel('gauss/cm')
ylabel('Integral Volume')
hold on

figure(2)
subplot(6,6,n-1)
plot(G,res)
axis([0 45 -5 5])
xlabel('gauss/cm')
ylabel('Residual to Fit')
hold on

end

params
save A_Acetate_Protein_Z.txt params -ascii
function z=Diff_rmsd(x)
global G Y M g del delta tau D
I0=x(1);

% Data Equation
Y_calc2=I0*exp(-D*g^2*G.^2*del^2*(delta-del/3-tau/2));
diff=Y-Y_calc2;
z=sqrt(diff'*diff/M);
% if I0<85 | I0>150
% z=1e30
% end
Matlab Script #3 – Used for the double exponential fit for $D_s$

clear all %clears all matlab data (matrices)
clf  %clear figures
global G Y M g del delta tau Dp Ip %variables for another 
program, defined here
%
% _______________
%Define Independent Variable %________________
%________________
A=xlsread('TitrationA_Acetate_Z');  
[M,N]=size(A);  % define "X" Points 1-10 M in the first N
G=A(:,1);         % define "X" Points 1-10 M in the first N
G_cont=linspace(0,45,500);
OPTIONS=optimset('Display','Off','TolX', 1e-6, 'TolFun',1e-6, 
'MaxFunEvals', 1000, 'MaxIter',1000);
g=26752;     %(gauss*cm)-1
del=5.5e-3;      %(sec)
tau=0.3e-3;      %(sec)
delta=0.135    %(sec)
D_slow=[1.14E-06 1.16E-06 1.21E-06 1.22E-06 1.14E-06 1.15E-06 
1.12E-06 1.17E-06 1.10E-06 1.06E-06 1.09E-06 8.28E-07 8.30E- 
07 8.11E-07 8.01E-07 5.99E-07 6.28E-07 5.85E-07 5.92E-07 5.76E-07 
5.89E-07 5.33E-07 5.45E-07]';
I_slow=[5.37E+01 4.12E+01 6.29E+01 6.05E+01 6.26E+01 6.26E+01 
6.51E+01 6.36E+01 6.28E+01 5.80E+01 6.55E+01 5.73E+01 4.07E+01 
4.10E+01 1.98E+01 2.06E+01 2.39E+01 2.28E+01 2.39E+01 2.41E+01 
5.10E+01 5.15E+01 9.30E+01 9.33E+01]';
%
%define data set and fit using Diff_rmsd_pro_ace.m program 
%
for n=2:N
  z_min=lel;
  Y=A(:,n);
  Dp=D_slow(n-1);
  Ip=I_slow(n-1);
  if n>=22
    G=A(:,1)
    Y=A(:,n)
    del=6.5e-3
    delta=.145
  end
  for tries=1:100
    Ia_start=80;
    Da_start=10*rand*10e-7;
    start=[Ia_start,Da_start];  %starting values for 
fminsearch
    [final_values, z0]=fminsearch('Diff_rmsd_pro_ace', start, 
OPTIONS);
    if z0<z_min
      Ia=final_values(1);
Da=final_values(2);

z_min=z0;
end

params(n-1,:)=[Ia,Da];

Y_calc=Ia*exp(-Da*g^2*G_cont.^2*del^2*(delta-del/3-tau/2))+Ip*exp(-Dp*g^2*G_cont.^2*del^2*(delta-del/3-tau/2));
%equation for line

Y_calc2=Ia*exp(-Da*g^2*G.^2*del^2*(delta-del/3-tau/2))+Ip*exp(-Dp*g^2*G.^2*del^2*(delta-del/3-tau/2));
%equation for points
res=Y_calc2-Y;

%Plots

figure(1)
sn
subplot(6,6,n-1)
plot(G,Y,'.b',G_cont,Y_calc,'-r')
axis([0 45 -5 130])
xlabel('gauss/cm')
ylabel('Integral Volume')
hold on

figure(2)
sn
subplot(6,6,n-1)
plot(G,res)
axis([0 45 -5 5])
xlabel('gauss/cm')
ylabel('Residual to Fit')
hold on

end

params
save A_Acetate_DoubleExpoFit_Z.txt params -ascii
function z=Diff_rmsd(x)

global G Y M g del delta tau Dp Ip

Ia=x(1);
Da=x(2);

% Data Equation
Y_calc2=Ia*exp(-Da*g^2*G.^2*del^2*(delta-del/3-tau/2))+Ip*exp(-Dp*g^2*G.^2*del^2*(delta-del/3-tau/2));

diff=Y-Y_calc2;
z=sqrt(diff'*diff/M);

% if I0<85 | I0>150%
% z=1e30%
% end
Appendix II

Acquiring High Quality $^{19}$F NMR Spectra:

A Practical Guide for UCSD
Even though the UCSD biomolecular NMR facility does not have a probe with a dedicated \(^{19}\)F channel in the facility, it is still possible to acquire high quality spectra with our current equipment. This appendix has been written to aid future researchers properly configure the DMX500 spectrometer to acquire high quality \(^{19}\)F detected NMR spectra.

**Why \(^{19}\)F NMR?**

There are several useful properties of the \(^{19}\)F nucleus (Gerig 1989; Danielson and Falke 1996) First, \(^{19}\)F is a spin \(\frac{1}{2}\) nucleus which occurs at 100% natural abundance, and is about 83% as sensitive to NMR detection as \(^{1}\)H. (for comparison, \(^{13}\)C is only 1.6% as sensitive). Second, the \(^{19}\)F chemical shift range is about 100x that of the \(^{1}\)H. The \(^{19}\)F chemical shift is controlled primarily by the lone pair electrons which provide a large paramagnetic term in the shielding formula. This property also renders the chemical shift exceptionally sensitive to changes in local van der Waals environment as well as local electrostatic fields, making it highly attractive when studying subtle conformational changes.

**The necessary hardware**

While it is defiantly preferred to have a dedicated \(^{19}\)F channel, such a probe is not an absolute necessity. The Larmor frequencies of \(^{1}\)H and \(^{19}\)F are close enough (\(^{1}\)H = 500.0; \(^{19}\)F = 470.39) to allow a spectroscopist to “de-tune” the \(^{1}\)H channel of a probe to the \(^{19}\)F frequency. The draw backs to this are a loss in sensitivity, because the probe is not designed to detect at this frequency, and the loss of ability to pulse on \(^{1}\)H, thus the inability to decouple the \(^{19}\)F spins from the surrounding \(^{1}\)H spins.

In the UCSD biomolecular NMR facility, most probes are not able to be de-tuned this far off from the proton frequency, however, the inverse Broad Band (BBi) probe for the DMX500 has this capability.
Configuring the Instrument Hardware

Once the BBi probe has been correctly installed, a new shim map will be necessary if one wants to gradient shim. (Note this is not necessary if you only need manual shimming) Each probe has its own shim maps. The normal TXI triple resonance probe is Probe #32, and the BBi probe is #10. Once the X-winNMR software is aware that the probe has been changed (edhead command), a shim map can be created as normal. Bear in mind that the shim map procedure as well as any gradient shimming requires a $^1$H signal, so the instrument and probe need to be configured as such at this point.

Once the shimming is as desired, it is possible to re-cable the instrument to pulse and detect on $^{19}$F through the $^1$H amplifier and probe channel. For traditional $^1$H experiments, the cable running out of the $^1$H amplifier is normally attached to the $^1$H output jack; however, for $^{19}$F this needs to be switched to the X_QNP jack. Due to $^{19}$F filters in the F1 ($^1$H) channel of the preamp, this cable needs to be routed through the F2 ($^{13}$C) channel of the preamp, from which the $^{13}$C band-pass filter has been removed. This cable running from the front of the F2 channel of the preamp is connected to the $^1$H coil of the BBI probe.

There is nothing special about tuning the probe, other than it is probably way off resonance, and the SW of the WOBB window will have to be increased from around 10 Hz to 50 Hz. Use the computer screen to help get it close, then the LED lights on the Preamplifier can be used for fine tuning.

Configuring the Software

When first setting up to acquire a $^{19}$F spectra, it is recommended that the Bruker parameter set “F19” be utilized to set the majority of the parameters including the offsets, and EDASP configuration. This parameter set should set the EDASP window according to the following diagram:
Data Collection –

The $^{19}$F parameter set brings over the standard 90° acquire 1D sequence “zgflqn” which is included at the end of this document for reference. The main and key difference between this pulse program and one which would be used for $^1$H acquisition is the command “10m QNM_X” (highlighted in red) at the beginning of the pulse sequence. This command is what tells the hardware to switch to the X_QNP jack of the $^1$H amplifier. This command can be added to any current $^1$H pulse sequence for $^{19}$F acquisition.

The use of a spin-echo to remove baseline artifacts

$^{19}$F NMR spectra using modern solution-state NMR probes are often complicated by very broad signals in the baseline, which are NMR signals attributable to fluorinated polymers used in probe construction. This broad probe-background is a well known problem with many NMR probes, and the baseline can become an unmanageable, sinusoidal, rolling baseline. In some cases, this can render the NMR spectrum intractable. Since the broad signal from the solid material in the probe will not refocus under a Hahn-Echo (90-τ-180-τ-detect), and the narrow signals from the molecule in solution do refocus, the spin-echo technique is a well know and commonly used method to eliminate the broad background signals in 19F NMR of small molecules (Dr. Richard K. Shoemaker, University of Colorado – Boulder, personal communication). This sequence “1D_19F_SpinEcho” is included at the end of this appendix.
as well. The length of the echo delay depends on the final desired spectral width and number of scans. For example a 300 µs delay is often suitable for a 80 ppm spectral width, but when a 200 ppm spectral width is necessary, this does not completely remove the baseline artifacts and 1 ms delay is necessary.

Estimation of the appropriate τ delay can be approximated by examination of the FID. The rolling baseline is caused by the first few points of the FID. If a simple 90° acquire spectra is recorded, and the FID is examined, one sees the initial points are massively distorted. If the scale at the bottom is in time, one can take the time at which the FID begins to resemble a normal FID as the length of the necessary echo delay.

Referencing the spectra

The choice of reference compound varies among the NMR community for 19F standards. The bimolecular community appears to go with referencing TFA to 0 ppm, which places F-Phe at -38 ppm. (Luck and Falke 1991a; Gerig 1994; Danielson and Falke 1996) The other common standard is CFCl₃ at 0.0 ppm. When referenced this way, TFA resonates at -76.55.

When finished with 19F NMR –

It is recommended to re-tune the BBi probe back to ¹H, so the next user is not caught completely off guard, when they attempt to tune the probe.
“zgflqn”
90° acquire pulse program for $^{19}$F Detection

;zgflqn
;advance-version (02/05/31)
;1D sequence for F-19 observe
;for QNP-operation (F-19 via X-QNP output of switchbox)

#include <Avance.incl>

10m QNP_X
1 ze
2 30m
d1
   p1 ph1
   go=2 ph31
   30m mc #0 to 2 F0(zd)
exit

ph1=0 2 2 0 1 3 3 1
ph31=0 2 2 0 1 3 3 1

;p11 : f1 channel - power level for pulse (default)
;p1 : f1 channel - high power pulse
;d1 : relaxation delay; 1-5 * T1
;NS: 1 * n, total number of scans: NS * TD0

;$Id: zgflqn,v 1.7 2002/06/12 09:05:22 ber Exp $
"1D_19F_SpinEcho"
Spin Echo Pulse Program for $^{19}$F Detection

zgflqn_SpinEcho
; advance-version (02/05/31)
; 1D sequence for F-19 observe with SpinEcho
; for QNP-operation (F-19 via X-QNP output of switchbox)

#include <Avance.incl>

10m QNP_X
1 ze
2 30m
d1
   p1 ph1
d2
   p2 ph1
d2
go=2 ph31
30m mc #0 to 2 F0(zd)
exit

ph1=0 2 2 0 1 3 3 1
ph31=0 2 2 0 1 3 3 1

; pl1 : f1 channel - power level for pulse (default)
; p1 : f1 channel - high power pulse
; p2 : f1 channel - 180 pulse
; d1 : relaxation delay; 1-5 * T1
; d2 : SpinEcho Delay
; NS: 1 * n, total number of scans: NS * TD0

;$Id: zgflqn,v 1.7 2002/06/12 09:05:22 ber Exp$
References


Gao, T., and Newton, A.C. 2002. The turn motif is a phosphorylation switch that regulates the binding of HSP70 to protein kinase C. *J Biol Chem* **277**: 31585-31592.


Goddard, T.D., Kneller, D.G., Sparky3 University of California San Francisco


Spolar, R.S., and Record, M.T., Jr. 1994. Coupling of local folding to site-specific binding of proteins to DNA. Science 263: 777-784.


