INSIGHTS INTO ARF ACTIVATION FROM NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY

by

RONALD D. SEIDEL, III

(Under the Direction of James H. Prestegard)

ABSTRACT

ADP Ribosylation Factors (Arfs) comprise a family of Ras-related GTPases regulating a wide variety of intracellular signaling pathways, specifically impacting bi-directional membrane traffic within Golgi membranes. Arfs perform these functions through cyclic interactions with guanine nucleotide exchange factors (GEFs), GTPase activating proteins (GAPs), membranes, and effectors. Here, novel NMR evidence is presented for a sequence of events within Arf activation leading from membrane association to GTP incorporation.

It is clear from previous models that Arfs have the ability to undergo substantial structural changes resulting from activation. This structural switching hinges on exposure of a myristoylated N-terminal $\alpha$-helix. A previously reported mechanism whereby this N-terminal domain is exposed is based on a comparison of the basal (Arf1•GDP) and activated ($\Delta$17Arf1•GTP, N-terminally truncated) crystal structures. However, it has been assumed that all structural changes leading to activation, observed through these models, are the direct result of GTP incorporation rather than displacement of this N-terminal extension. Separation of conformational changes resulting from nucleotide exchange and N-terminal truncation cannot be addressed using these existing structures because both the bound nucleotide and the

N-terminus differ. Results from NMR experiments presented within, including residual dipolar couplings (RDCs), has allowed the generation of an additional model, Δ17Arf1•GDP, and reveals substantial structural variation resulting solely from N-terminal truncation.

Incorporation of these findings within a mechanistic framework centers on displacement of the N-terminal helix upon membrane interaction, previously mimicked through truncation, but what may cause the needed destabilization of the Arf1 N-terminus in full-length forms? It is known that signaling lipids, specifically phosphatidylinositol (4,5)-bisphosphate ($PI(4,5)P_2$), regulate the activities of a number Arf effectors. The ability of Arf itself to bind these same lipids has also been reported previously, though with minimal structural detail. In what follows, NMR evidence is presented in support of lipid-mediated conformational changes within Arf1 preceding nucleotide release. A patch of positive potential on the protein surface is identified as the lipid-binding site, and structural changes extending to the N-terminal helix are also reported. Taken together, the existing biochemical data and newly presented structural information have allowed the generation of a mechanistic model for Arf1 activation.

INDEX WORDS: ADP Ribosylation Factor 1, Arf, Energy Minimization, Molecular Refinement, NMR, Residual Dipolar Coupling

# INSIGHTS INTO ARF ACTIVATION FROM NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY

by

RONALD D. SEIDEL, III

B.S., University of Georgia, 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2005

# INSIGHTS INTO ARF ACTIVATION FROM NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY

by

RONALD D. SEIDEL, III

Major Professor: James H. Prestegard

Committee: Alan E. Przybyla
Jeffrey Urbauer
Bi-Cheng Wang
William S. York

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2005

# DEDICATION

To my Father, a friend and confidant, without his interminable support this never would have been possible.

In loving memory of my Mother, whose strength gave me the courage to try and the motivation to succeed.

Barbara Seidel
August 12, 1947 – September 26, 2003

# ACKNOWLEDGEMENTS

I would like to thank a number of people without whom this dissertation would never have been possible. First, my major professor and mentor, James Prestegard, for his support and guidance throughout this project; he is a man of exemplary character with whom it has been my distinct privilege to work under. Next, my committee members, Alan Przybyla, B.C. Wang and Will York, for their encouragement and suggestions. James Omichinski, specifically, has been helpful in my scientific development through his rather direct comments and critiques. Collaborators Rick Kahn and Juan-Carlos Amor through constant discussions have pushed my knowledge of Arf biology without which the research presented would not have been possible. I would like to also thank members of the Prestegard lab, not only for their helpful discussions but also for their friendship. Cathy Bougault, Ziad Eletr, Mike Ford, Andrew Fowler, John Glushka, Anita Kishore, Laura Morris, Fang Tian, Greg Wylie, and Homayoun Valafar. I would also like to thank Peter Horanyi and Alison Venable for making the coursework bearable. I would like to recognize the contributions, both personal and professional, of Avinash Sujan, a colleague and friend. Although he passed away this year he leaves behind a legacy through his wife, Sabeen, and daughter, Anya. He is an inspiration to us all through his unremitting positive spirit and he will be deeply missed. I would like to thank my friends and roommates, who have made my tenure at the University unique. Finally, I would like to express my unending gratitude to my wife, Trish, and her family, who through difficult times, have been a pillar of support to which I owe my strength of mind.

TABLE OF CONTENTS

Page

CHAPTER

## Chapter 1

## Introduction: The Arf1 Activation Pathway

Proteins move from one organelle to another within the secretory pathway by means of small transport vesicles, surrounded by one of two types of cytosolic cage proteins, Clathrin and Coatamer. Clathrin-coated vesicles are composed of a fibrous protein taking the form of a three limbed triskelion, each limb containing one heavy and one light chain. These three-legged structures combine to form a cage surrounding the vesicle surface. The purpose of these protein coats is simply to shape the membrane into a bud and to capture molecules for transport. It was once thought that Clathrin-coated vesicles dominated secretory pathways, hauling proteins from the rough ER through the Golgi and out to the cell surface. This was later deemed highly improbable as yeast mutants found deficient in Clathrin retained an ability to transport proteins to a certain extent. At that time it was observed that a second type of coated vesicle, one using a multitude of cytosolic proteins (e.g.CopI/II), existed. These coatamer-coated vesicles are now known to prevail in ER to Golgi and intra-Golgi transport, leaving Clathrin-coated vesicles to carry proteins from the trans-Golgi network to the cell surface and lysosomes or from the cell surface to endosomes. In either case, Clathrin or Coatamer, the coat is lost after budding, allowing for subsequent vesicle fusion.

The goal of much current research has been to identify all the proteins involved in the formation, targeting, and fusion of transport vesicles. The two most widely accepted approaches in the study of coated vesicle transport are biochemical analyses of cell-free

systems and the aforementioned genetic analysis of yeast mutants, having deficiencies at specific stages within the transport pathway. Not surprising, both approaches have shown that a host of macromolecules are required for the formation of coated vesicles. Prominent among these is a class of small GTP binding proteins, termed ADP Ribosylation Factors (Arfs).

Arfs were first identified as cofactors required in the ribosylation of the α subunit of the hetero-trimeric G protein, $G_{s,}$ by Cholera toxin, leading to activation of adenylate cyclase *(1, 2)*. Today, however, this function is of minor importance compared to the role of Arfs in signaling cascades, specifically those impacting organelle structure and cell morphologies. Comprising a family of 21kDa Ras-related GTPases, Arfs are divided into three classes based on size, amino acid sequence, gene structure, and phylogenetic analysis: Arf1, Arf2, and Arf3 form class I, Arf4 and Arf5 are in class II and Arf6 is in class III *(3)*. The role of class I Arfs, Arf1 and 3, are well established; having functions in ER to Golgi and intra-Golgi transport *(4)*. In contrast, Arf6 has been implicated in transport at the plasma membrane and endosomal compartments, and is involved in remodeling of the actin cytoskeleton *(4-6)*. While the functions of the remaining Arfs are not well understood, class I Arfs, particularly Arf1, have been extensively studied and this has led to the generation of a mechanistic model of Arf action. It is my intent to examine in more detail the mode of Arf action leading from membrane association to activation. A comprehensive understanding of Arf behavior may ultimately lead to tailored therapeutics, as deterioration of Arf-mediated vesicle transport has been implicated in enhanced levels of apoptosis in cancer cell lines *(7)*.

Essential components of the action of Arfs *in vivo* are the abilities to reversibly bind to biological membranes, to activate lipid modifying enzyme activities [e.g., phospholipase D or PI(4)P 5-kinase *(8-10)*], and to recruit or stabilize the binding of other proteins to those same

membranes; prominent among the latter are Coatamer-coat proteins, including COPI, GGAs, adaptins, and MINTs *(11-14)*. Membrane binding of Arf1 is controlled by exposure of a myristoylated, N-terminal α-helix upon exchange of GDP for GTP. Though the changes in protein structure that accompany nucleotide exchange are critical aspects of all regulatory GTPases, they are particularly important to Arfs, which directly couple the activation process to membrane translocation *(15-17)*. Indeed, the formation of activated Arf1 on a membrane surface is viewed as the initial event that nucleates the formation of a Coatamer-coated bud and subsequent coated vesicle. In what follows, a structural view describing the mode of Arf1 activation will be given. This will provide a convenient background for subsequent discussions. While the discourse focuses primarily on functions related to Arf1, results also extend to activities that are certainly present in Arf3 (sharing 96% sequence homology) and present to an uncertain extent in the remaining classes of Arfs.

### 1.1 Structural Snapshots of the Nucleotide Binding Cycle

It is important to note that all members of the Ras super-family have the ability to bind nucleotides, however, Arfs remain distinctive in that they preferentially bind GDP over triphosphate analogs in solution *(18)*. What is more, Arfs are unique in having a near absolute requirement for phospholipids in nucleotide exchange processes. Deletion of the N-terminal helix results in a protein devoid of Arf activity, but one that retains the capacity to bind nucleotides, at least the GTP form *(19)*. Interestingly, unlike the basal state, the binding of GTP to this deletion mutant was found to be independent of added phospholipids, the structure of which is now purported to represent an “activated” form of Arf1. Based on these

observations, displacement of this N-terminal domain is deemed crucial for GTP incorporation and membrane association. Current proposed mechanisms whereby both the myristoyl-chain and the N-terminal helix are exposed are based on a detailed comparison of the full-length, non-myristoylated, GDP-bound, x-ray crystal structure (later referred to as Arf1•GDP) *(20)* to other non-myristoylated structures; specifically, the Arf1 truncation mutant, bound to a non-hydrolysable GTP analog, Gpp(NH)p, (later referred to as Δ17Arf1•GTP) *(15)*.

### *1.1.1 The N-terminal Helix*

Structurally Arfs, as represented in the Arf1•GDP crystal structure *(20)*(Figure 1.1A), consist of a mixed parallel/antiparallel 6 β-strand core surrounded by 6 α-helices and 12 connecting loops. Residues 2-17 constitute the aforementioned N-terminal α-helix (αNt) and flexible connecting loop region ($L_{Nt}$, found adjacent to αNt, however not labeled). The peptide sequence for the N terminal domain is myr$G_2$N**IF**AN**LF**KG**LF**G$KKE_{17}$ (hydrophobic residues in bold). This can clearly fold into an amphipathic helix and act as a complement to myristoyl insertion during membrane association *(20, 22-25)*. In Arf1•GDP the N-terminal helix as a whole lies in a cleft parallel to the C-terminal helix (αCt) and is held in place predominantly by hydrophobic interactions *(15, 26, 27)*. It has been shown that removal of the myristoyl chain decreases the association of Arf1 with lipid vesicles; however, deletion, and point mutation (F9A and L8A-F9A), of the N-terminal 17 residues abolishes it *(28)*. Clearly, the interaction of both the myristate and the amphipathic helix are needed in anchoring the protein to the membrane surface.

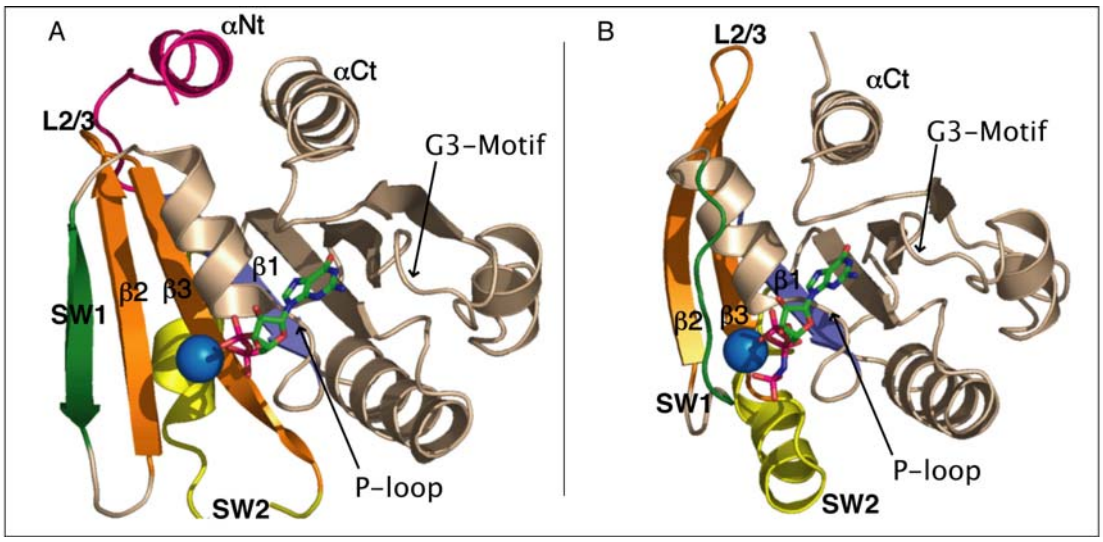


Figure 1.1. Important structural features of (A) Arf1•GDP and (B) Δ17Arf1•GTP. Regions having the lowest degree of structural variation (RMSD of overlayed structures is 0.3A) are shown in neutral tones. Highly variable regions are labeled and color-coded as follows: N-terminal α helix [αNt (residues 2-17), magenta], Switch 1 [SW1 (42-52), green], Inter-Switch β strands [β2-β3 (53-68), orange], and Switch 2 [SW2 (70-85), yellow]. The molecular representations (current and all subsequent) are generated using the PyMol Molecular Graphics System *(21)*.

### *1.1.2 The Nucleotide Binding Site*

The nucleotide-binding pocket is located on the face of the protein opposite of the N-terminal binding cleft. The residues involved in nucleotide association, conserved throughout all six known mammalian Arfs[†], are found in the phosphate binding loop (P-loop) and the guanine base-binding motif (G3Motif) (Figure 1.1). The α and β phosphates of GDP, and γ phosphate of GTP when present, lie across the P-loop and the N-terminus of α helix A (later referred to as the nucleotide binding helix) in a sequence of residues termed the Walker

[†] As determined from the Protein Sequence Database. Accession numbers are as follows: Arf1 (A33283), Arf2 (A36367), Arf3 (A41570), Arf4 (B38622), Arf5 (A23741) and Arf6 (B23741)

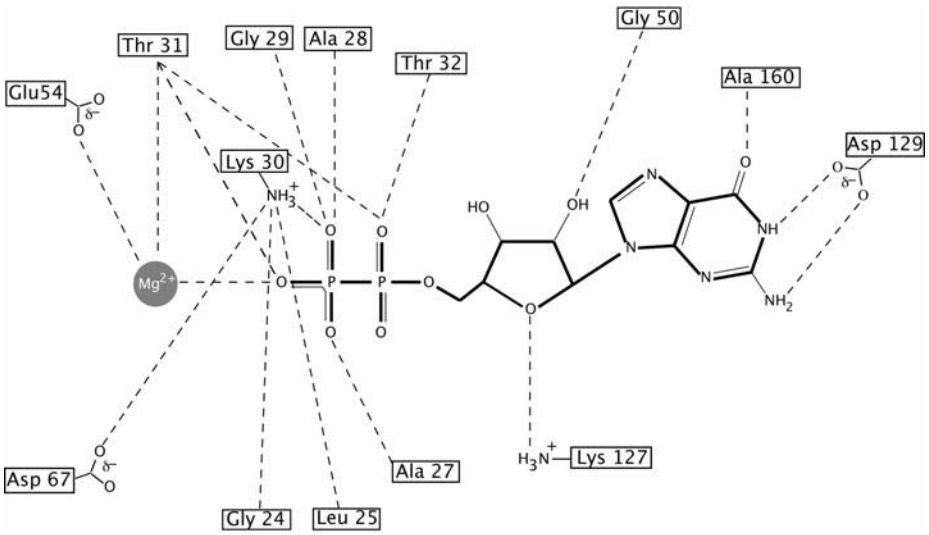


Figure 1.2. Schematic diagram of the nucleotide and $Mg^{2+}$ bound to Arf1. Dotted lines imply a distance of less than or equal to 3Å, showing the intricate network of potential hydrogen bonds with the terminal β phosphate of GDP as well the ligated $Mg^{2+}$ ion. Also notable is the presence of partial negative charge in the binding pocket as denoted by the symbol (δ- ).

sequence by analogy to Ras (residues 24-GLDAAGK-30)*(20)*. The guanine base interacts with the TCAT box (residues 158-TCAT-161) and the highly conserved guanine specificity region (G3Motif (126-NKXD-129). Unlike the proximal α phosphate, the terminal β phosphate of GDP (γ phosphate of GTP) is held in place by an intricate network of hydrogen bonds and $Mg^{2+}$ coordination, conferring upon Arfs a requirement for divalent metals in nucleotide association (Figure 1.2). The positive charge of this ion is very well compensated by charges on the phosphate and carboxylate side chains (largely due to the presence of E54 and D67 near the terminal phosphate of GDP). The presence of negative charge in this region may well destabilize GTP through interactions with the additional negatively charged terminal phosphates and produce the enhanced affinity for GDP in basal forms.

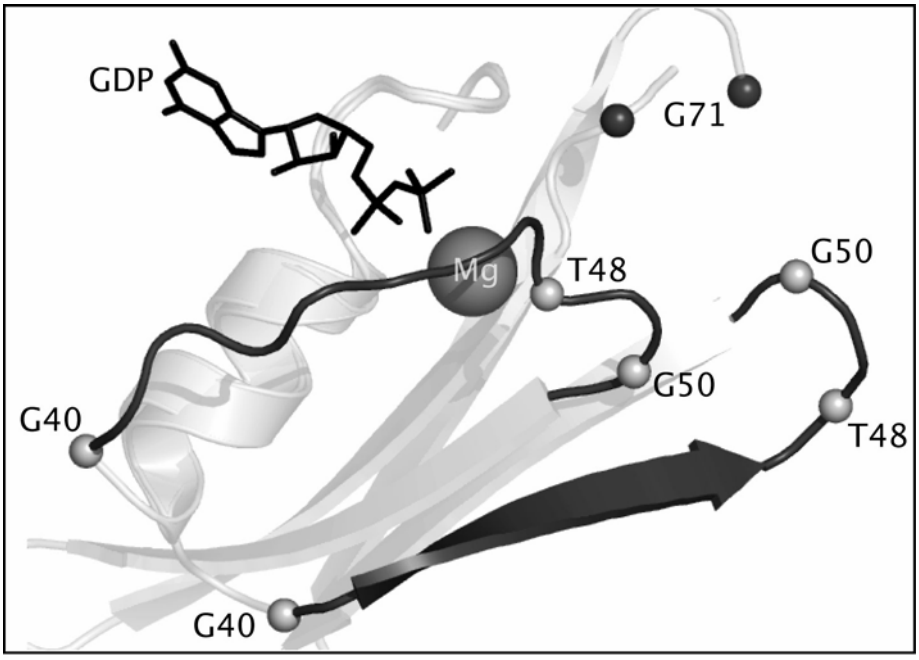


Figure 1.3. Conformational changes observed within switch 1 upon nucleotide exchange. Arf1•GDP has SW1 as a β strand at the periphery of the Arf1 core with Δ17Arf1•GTP undergoing a transition to an extended loop (shown in dark gray). Cα positions of important amino acids are illustrated as spheres and labeled by primary sequence and one letter amino acid code.

### *1.1.3 The Switch Regions*

Connecting the nucleotide binding site and the N-terminal helix are distinctive switch regions [termed SW1, SW2, and the Inter-Switch by analogy to Ras proteins (Figure1.1) *(29-31)*]. These change conformation in response to the type of bound nucleotide. Switch 1 (SW1, residues 42-54), forms a seventh β strand at the periphery of β2 in Arf1•GDP. This is remote from the nucleotide. Upon GTP incorporation, however, SW1 loses its secondary structure and moves in space allowing for interaction between the invariant Threonine (T48) and the $Mg^{2+}$ ion as well as the γ phosphate of GTP (Figure 1.3) *(15)*. This structural

rearrangement is permitted through action of terminal glycine residues preceding (G40) and succeeding (G50) the SW1 β strand, which offer the needed conformational flexibility.

Conversely, residues 70 to 85 referred to as switch 2 (SW2), seem to undergo a transition from an extended loop to a more ordered helical segment as a result of nucleotide exchange, again permitting tighter nucleotide binding (Figure 1.1). The functional significance of the SW2 shift is found not only in GTP incorporation but also in the hydrolysis of GTP to GDP in later steps of the cycle. Upon activation, Glycine 71 moves by 5 Å adopting a position similar to that observed for Ras (Gln 61). This has been found critical for orientation of a hydrolytic water molecule needed for GAP-catalyzed hydrolysis of GTP to GDP *(32)*. Mutation of this catalytically critical Glycine to Leucine in Arf1 is lethal to mammalian and yeast cells alike *(33, 34)*. The relative position of Gly 71 in inactive conformers, unique to Arfs, may be the explanation for the lack of intrinsic GTPase activity of the isolated proteins.

### *1.1.4 The Interswitch Toggle*

Lastly, the Inter-Switch β-strands (ISR, defined as β2, β3, and $L_{2/3}$ [residues 51-68]), undergoing the largest degree of observed conformational change, are displaced by 7Å allowing for occupation of the cleft used for binding the N-terminal helix in the Arf1•GDP complex *(35)*. This displacement corresponds to a two-residue change in register between β3 and β1 in a direction parallel to the β strands. The change in register allows for insertion of the γ phosphate of GTP by positioning the invariant aspartic acid (D67) away from the terminal phosphate (Figure 1.4). This is suggested to dispense with the excess electro-

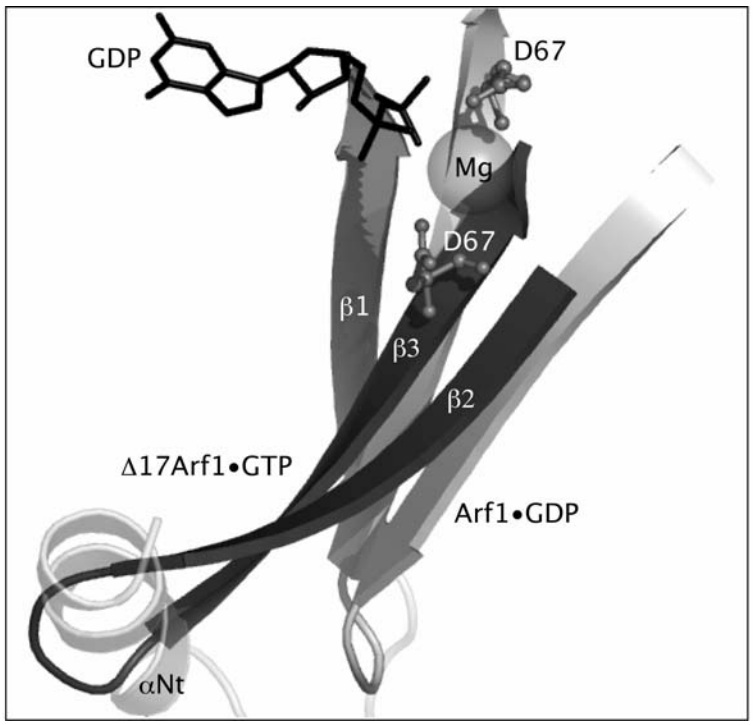


Figure 1.4. Superimposition of Arf1•GDP (light gray) and Δ17Arf1•GTP (dark gray) showing conformation changes of the interswitch β strands upon activation. Most notable is the displacement of the β strands from the protein core into the N-terminal binding cleft orchestrated by shifting of the invariant aspartic acid (D67).

negativity in the nucleotide-binding site of GDP loaded samples. Loss of edge-to-edge contacts with SW1 and rearrangement of the hydrogen bond network within the interswitch strands are likely coupled.

## 1.2 Guanine Nucleotide Exchange Factors (GEFs) and Nucleotide Exchange

As shown above, the trademark of Arf proteins is their ability to undergo structural changes in response to the bound nucleotide. The GDP bound "off" state and the GTP bound

"on" state, being structurally distinct, have been shown to recognize different partner proteins in the cell. This type of behavior allows these small GTPases to perform as molecular switches in signaling cascades. As the degree of conformational change needed for molecular switching remains energetically unfavorable in solution, Arfs do not act in isolation. Activation by GTP requires the aid of Guanine nucleotide Exchange Factors (GEFs) and inactivation requires GTPase Activating Proteins (GAPs). The latter are beyond the scope of the current work and will not be described. Recently, a host of Arf structures showing association with various GEFs have been solved *(15, 35)*, allowing for a model for GEF interaction to be established. In the coming sections I will again take a structural approach in describing the mode in which GEFs act on their target Arfs allowing for this hallmark structural switch.

### *1.2.1 A glimpse of GEF association*

GEFs have been classified into three families based on sequence similarity and the ability to bind specific GTPases. Arf Nucleotide Binding Site Opener (ARNO), specific to Arf1 in mammalian cells, along with Gea2 (yeast) and other low molecular weight class I Arf-GEFs are multi-domain proteins with each domain being responsible for discrete functions in the activation cycle. The 200 residue catalytic Sec7 homology domain, common to all Arf GEFs identified to date, is responsible for the enzyme-catalyzed exchange of GDP for GTP on Arf *in vitro (15, 36, 37)*. The Sec7 region is seen flanked by multiple domains believed responsible for events such as protein-protein interactions, protein-membrane interactions and oligomerization *(38)* (some of which will be described in later sections).

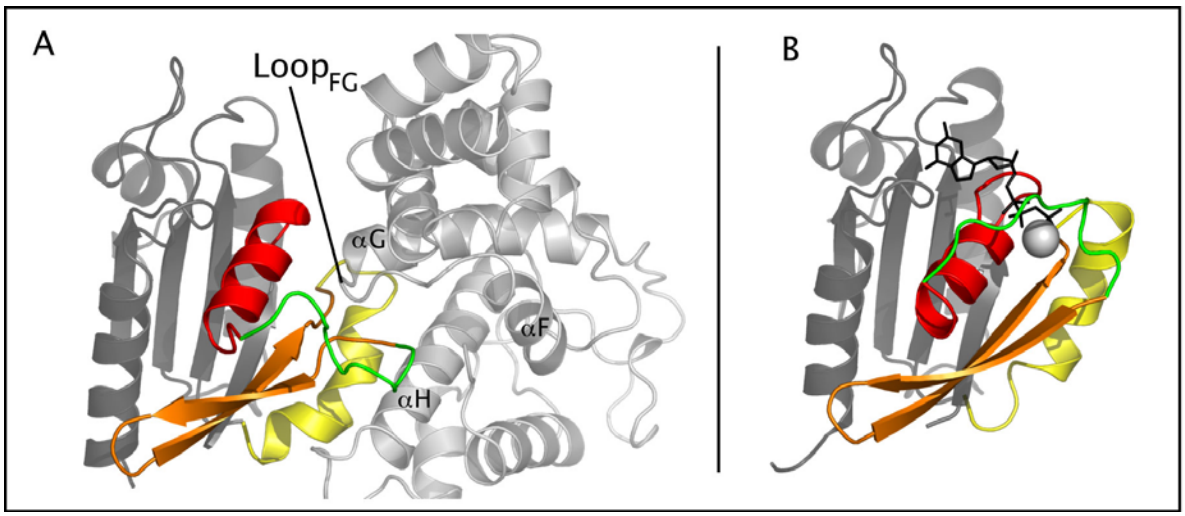


Figure 1.5. Structural comparison of (A) Δ17Arf1•apo–Sec7 and (B) Δ17Arf1•GTP. Regions within Arf1 affected by GEF association are color coded as follows: Nucleotide binding helix and phosphate binding loop (αA/P-loop, Red), Switch 1 (SW1, green), Inter-Switch β strands (β2-β3, orange), and Switch 2 (SW2, yellow); unaffected region shown in gray. Gea2 Sec7 domain, located on the right in panel A, shown in light gray. Alpha helices forming Arf1 binding site labeled as αF/G/H, respectively. Hydrophilic loop implicated in nucleotide displacement labeled as Loop$_{FG}$.

The key to GEF action lies in the ability to associate with, and stabilize, a form of Arf with a reduced affinity for nucleotides. This is best illustrated in the structure of an Arf1 truncation mutant bound to the catalytic Sec7 homology domain of Gea2 which lacks both nucleotide and the N-terminal Helix (termed Δ17Arf1•apo–Sec7, Figure 1.5)*(15)*. It is important to note that truncation of Arfs has proven necessary in the structural study of Arf-GEF interactions as the presence of the N-terminal helix impedes the catalyzed exchange reaction *(39)*. It is immediately apparent that the Apo conformation closely resembles that of the GTP bound activated form discussed previously, with the major changes again being localized to SW1, SW2 and the interswitch region.

The “active site” for GEF interaction is centered on the SW1 face of Arf1 through contact with a series of alpha helices forming a hydrophobic groove in Sec7 (αF/G/H) and insertion of a hydrophilic loop ($Loop_{F/G}$) into the nucleotide-binding pocket of Arf1. Mutagenesis studies also support a model in which SW1 and SW2 provide the primary recognition sites for GEF association *(36, 40)*. The FG loop is seen in Figure 1.6 (A) inserted into the GTPase, leading to interactions with glutamate 97 (Gea2 - Sec7) and the pocket used for terminal phosphate placement and $Mg^{2+}$ coordination in Arf1. Mutation of this invariant glutamate (E97) reduces GEF activity 1200 fold *(40)* and has led to the paradigm of “glutamic finger” assisted GDP release. Interestingly, the Sec7 domain shows no interaction with the Guanine specificity region (G3 Motif, TCAT box). This region is therefore likely to be the site for initial incorporation of nucleotide triphosphates.

SW1 rearrangements lie at the heart of GEF association. Residues 40 – 51 have lost secondary structure, much like the GTP bound active form, and pivoted away from the GEF binding site towards the nucleotide-binding region on Arf1 (Figure 1.6B). The latter half of SW1 (residues 48-53) is found buried in the hydrophobic groove of Sec7 with the remainder found mobile in solution. It is suggested that this action opens Arf1 up for subsequent insertion of the glutamic finger. Mutational analysis of buried residues (I49, F51) revealed a 25-fold decrease in GEF activity, further substantiating a heightened role for the SW1 interfacial region *(41)*.

SW2 has become partially ordered forming a $3_{10}$ helix (residues 71-80). However this is spatially distinct from the activated structure previously described. Notably, the SW2 region interacts intimately with the GEF Sec7 domain, likely serving multiple functions. The change in both conformation and spatial displacement may induce or stabilize an altered

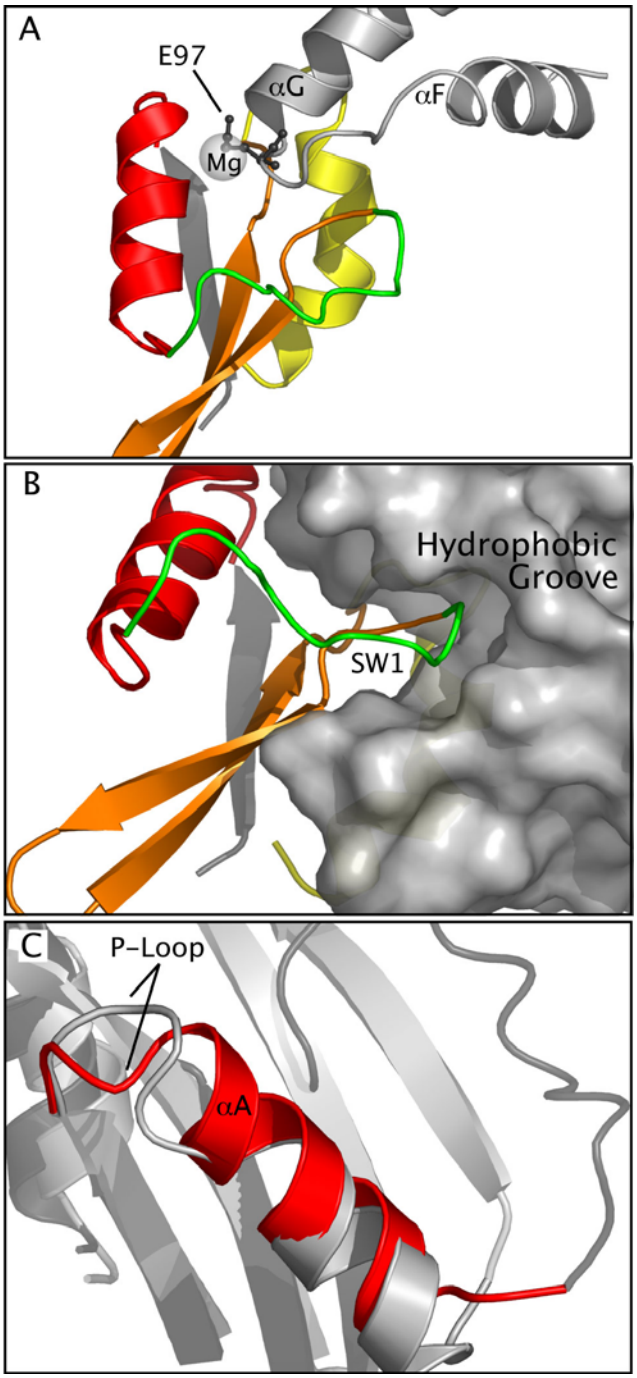


Figure 1.6. GEF catalyzed nucleotide release on Arf1 highlighting (A) insertion of the FG Loop into Arf1, led by interactions with glutamate 97 (Gea2) and the pocket used for terminal phosphate placement and $Mg^{2+}$ coordination, (B) SW1 – hydrophobic groove interactions, and (C) the nucleotide-binding helix (αA, residues 26-30) and the P-loop (K30 and T31) have both altered their conformations in the complex structure. Regions of interest within Arf1 are color coded as follows: Nucleotide binding helix and phosphate binding loop (αA/P-loop, Red), Switch 1 (SW1, green), Inter-Switch β strands (β2-β3, orange), and Switch 2 (SW2, yellow).

conformation of the Arf1 P-loop (described below) as well as forming numerous side chain interactions with the GEF, possibly contributing to Arf-GEF specificity.

The nucleotide-binding helix (αA, residues 26-30) and the P-loop (K30 and T31) have both altered their conformations in the complex structure, likely through direct interaction with E97 (Gea2 - Sec7) (Figure 1.6C). Although the shifting of αA can be viewed as a destabilizing component during nucleotide release, it is not currently considered a key feature and is likely to be a secondary response to nucleotide displacement *(15)*. The interswitch has performed its trademark toggle, thus permitting the needed shifting of G70/E71 and D67 as a result of glutamic finger insertion allowing for a destabilization of the ligated $Mg^{2+}$ ion and insertion of GTP.

Interestingly, the insertion of the glutamic finger does not impede GDP association; having a 3Å approach with the terminal β phosphate, it does, however, place a strong steric and electrostatic force on GDP in the binding pocket. The exchange reaction itself is reversible, also promoting the exchange of GTP for GDP *(26)*. It is thought that excess GTP in the cytosol drives the reaction in the forward direction. From this, it has been suggested that a ternary structure must first form; containing Arf1-Nucleotide-GEF, through interactions with the hydrophobic groove (Sec7) and SW1 (Arf1), consequently necessitating a multi-step process in GEF catalyzed nucleotide release. The GEF is later dissociated through incorporation of GTP, dispensing with the glutamic finger interactions and thus the need for GEF association.

### *1.2.2 The trapped intermediates: Opening the door for membrane induced change*

The normal mode of nucleotide exchange can be inhibited using specific Arf constructs and Arf complexes; the study of which can result in a detailed view of the intermediate changes leading to Arf1 activation. Binding of the small molecule Brefeldin A (BFA) can inhibit the activity of a number of class I Arf GEFs (e.g. Gea2). The mode of inhibition by BFA is unique, involving only the Arf1•GDP-Sec7 ternary complex and leaving the Arf1•GDP and Sec7 isolates untouched. This quaternary complex (Arf1•GDP - BFA - Sec7) remains cytosolic in biochemical assays, removing the previous requirement for membrane anchoring upon GEF association *(35)*. Alternatively, nucleotide dissociation can be inhibited by specific point mutations within the GEF itself. Mutation of the invariant glutamic acid (Gea2 E97, ARNO E156) to lysine yields an abortive complex halting nucleotide exchange. In this case, however, there remains a strong need for a hydrophobic environment, suggesting that this complex lies further down the path to nucleotide dissociation. It is understood that these two complexes represent distinct intermediate stages along the activation pathway and until very recently (December 2003, *(35)*) the structures remained elusive in crystal trials. It had been previously thought that all structural changes upon Arf1 activation were the direct result of GTP incorporation, and that glutamic finger insertion in Apo forms mimicked these changes. New models, however, allow for the possibility of structural rearrangements prior to nucleotide release or GTP incorporation.

In the presence of BFA, shown in Figure 1.7(A), the Sec7 domain interacts with Δ17Arf1 in a similar fashion to the nucleotide free complex (shown again in Figure 1.7C). SW1 has lost its secondary structure and is found partially inserted into the hydrophobic groove on ARNO Sec7. SW2, in this instance, has only partially ordered as a result of GEF

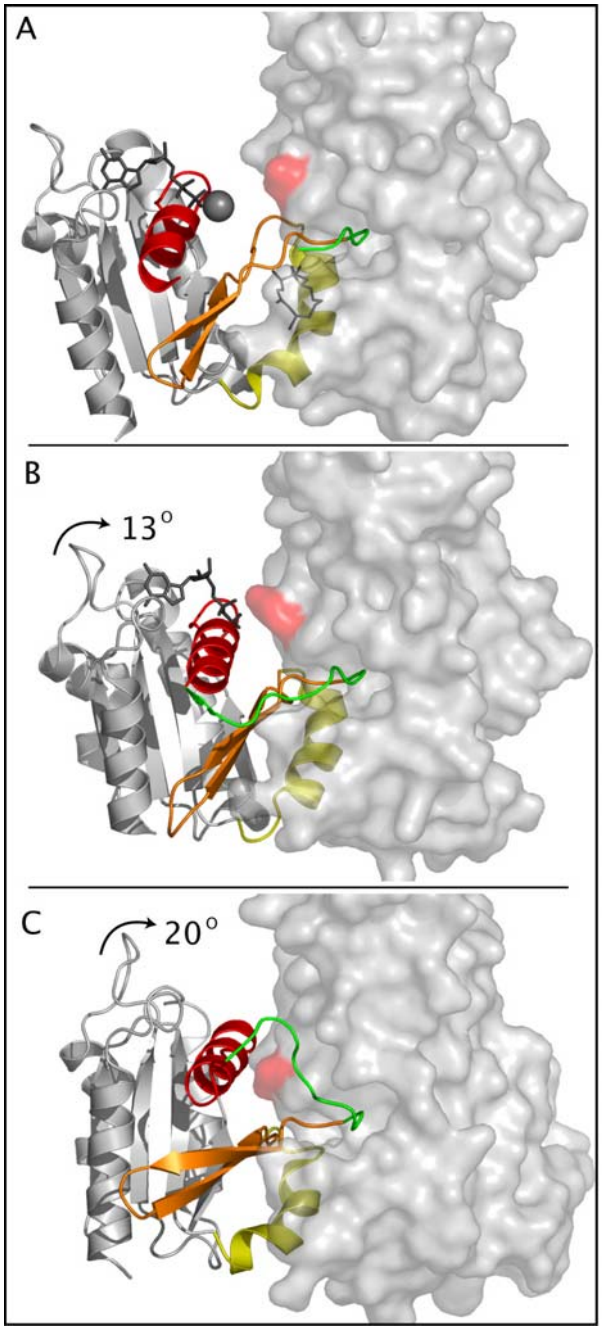


Figure 1.7. Intermediate stages of GEF catalyzed nucleotide release on Arf1 as illustrated in the X-ray crystal structures of (A) Δ17Arf1•GDP - BFA - Sec7, (B) Δ17Arf1•GDP - Sec7 (E156L) and (C) Δ17Arf1•apo- Sec7. In all cases, the GEF is represented as a semi-transparent surface plot (light gray). The catalytically crucial glutamic acid highlighted in red. Regions within Arf1 affected by GEF association are color coded as follows: Nucleotide binding helix and phosphate binding loop (αA/P-loop, Red), Switch 1 (SW1, green), Inter-Switch β strands (β2-β3, orange), and Switch 2 (SW2, yellow). Rotation of the Arf core (shown in gray) towards GEF denoted by arrows and degree of change.

interaction allowing the Arf1 P-loop to remain unfettered. The incorporation of BFA, through interaction with the interswitch β strands, has yielded a structure in which the interswitch has yet to toggle forcing a rotational transformation of the Arf1 core (shown in gray) of 20° when compared to the nucleotide free complex (comparing Figure 1.7A to C). The center of rotation is around the interfacial region, SW1 and SW2. As seen from the perspective of the GEF, the hydrophobic groove shows a clamping of 9° around SW1 of Arf1 immediately upon interaction. Since it is known that BFA has the ability to trap the Arf-GEF complex in a cytosolic fashion, as described above, one may assume that N-terminal truncation of Arf1 is unwarranted in this case. However, the truncated model presented (Δ17Arf1•GDP - BFA - Sec7) shows little difference to that of the full-length complex (Arf1•GDP - BFA - Sec7). The Arf1 N-terminal domain is not seen in these crystal models due to high thermal factors. This suggests that the N-terminal helix of Arf1 is destabilized in complex structures preceding GDP release. However, it is not currently known whether this destabilization occurs concurrent with or prior to GEF association.

Inhibition by the glutamic finger mutation allows for a similar association with the switch regions, although as a direct result of “glutamic finger” insertion, the interswitch has translated from the protein core allowing for a 13° rotation of Arf1 towards the exchange factor (Figure 1.7B). The insertion of the mutated “glutamic finger” domain has disrupted the interaction with the $Mg^{2+}$ ion while leaving the nucleotide intact. The GEF in this instance shows no further change in conformation. The final rotation of the Arf1 core towards the GEF seen in the nucleotide free complex (1.7C) is thought to be instigated by attractive forces of the intact glutamic finger and the P-loop, placing loop residues in a position to mimic the displaced phosphate oxygens.

## 1.3 Membrane Localization and the Role of Phosphoinositides (PIs)

Arf in its inactive state adopts a conformation that is unable to fit into the binding pocket of its respective GEF in solution. It is therefore believed that Arf1 first undergoes a lipid – mediated conformational change before it can form a productive complex with a membrane associated GEF. In this regard, Arf1 activation takes place after both the GEF and its target Arf have been localized to the membrane. However, what is responsible for the spatial localization of Arf1 and its associated Arf –GEF? Arf1•GDP might bind to the appropriate membrane site through interactions with proteins and/or lipids and then attract the GEF. Alternatively, Arf1•GDP might interact rather indiscriminately with membranes, but be activated by its GEF only at the correct location using the GEF as the membrane discriminating sensor, or some permutation of the two. If this is believed, there then must be a mechanism for controlling the association of Arf-GEFs and possibly Arf1 itself to specific membranes within a cell.

### *1.3.1 PI mediated membrane translocation of Arf-GEFs*

The mechanism for exchange factor recruitment is easily explained upon examination of the structure of Arf-GEF proteins. All low molecular weight class I Arf-GEFs (ARNO, cytohesin, Gea2, etc) contain pleckstrin homology (PH) domains *(37),* which have been shown to provide specific binding sites for the polar head group of membrane embedded phospholipids, specifically phosphatidylinositol (4,5)-bisphosphate [Ptdins(4,5)$P_2$] *(37, 42)* and an adjacent carboxy-terminal polybasic domain that cooperates with the PH domain to

enhance membrane binding . The PH domain extends nearly 110 amino acids forming a highly conserved β sandwich comprised of two nearly orthogonal β sheets having 3 to 4 antiparallel β strands each *(43)*. The binding site for inositol polyphosphates is found in the loops connecting these strands that are rich in lysines, arginines, and histidines forming a solvent exposed positively charged patch adjacent to an amphipathic alpha helix (C-terminal) *(43, 44)*. This helix is believed to be used in anchoring the domain to the vesicle surface subsequent to poly-phosphate interaction. It has been shown that this PH domain is responsible for membrane recruitment of these soluble Arf-GEFs. However, the ability of Arf1 to bind these same phosphoinositides is now also under investigation *(45)*.

### *1.3.2 PIs Enhance GTP binding to Arf1*

The ability of Arf1 itself to bind phosphoinositides was first predicted from the presence of a "positive patch" on the surface of the protein *(20)* found adjacent to the amphipathic N-terminus; this is reminiscent of the afore-mentioned PH domain. The importance of this patch was confirmed by observations that mutation of the lysines and arginines in this polybasic domain resulted in alterations in lipid interactions and the lipid requirements for GAP-stimulated hydrolysis of Arf-bound GDP *(46)*. In contrast to the role of phosphoinositides as required cofactors, e.g., in the Arf-stimulated activation of phospholipase D *(8, 47)* or Arf-GAP activities *(48)*, PI(4,5)$P_2$ has the ability to independently promote the release of GDP from Arf1 *in vitro* and stabilize the apo-protein *(49, 50)*, thus potentially acting as an Arf-GEF. Additional evidence supporting a direct functional role for the interaction of phosphoinositides with Arf1 comes from the observation that nucleotide

exchange stimulated by phosphoinositide-dependent GEFs is still present upon removal of the membrane recruitment domain of the exchange factor *(51)*. Hence, there is evidence for the functional importance of phosphoinositide interactions with Arf1, but few details on the structural perturbations that accompany the interactions between these lipids and Arfs (to be addressed in Chapter 4).

## 1.4 Current Model for Arf1 Activation

Combining biochemical information with the crystal structures described allows for the generation of a mechanistic view of Arf1 action leading to activation. This is summarized in Figure 1.8. (1) Arf1 is found in rapid equilibrium between cytosolic and membrane associated forms through interaction with its attached myristate. It is thought that SW1, SW2 and the interswitch remain intact, placing the invariant aspartic acid (D67) in a position to prohibit GTP incorporation. The N-terminal helix is believed to remain protected in the protein core as a result of the interswitch placement. It is important to note, however, that no structural information currently exists for this structural intermediate (to be addressed in Chapter 3 and 4). (2) The exchange factor, being cytosolic in nature, is recruited to the membrane surface through specific interaction with its membrane recruitment domain and membrane imbedded $PI(4,5)P_2$. Arf1 subsequently interacts with its GEF at the membrane surface through SW1 – hydrophobic groove interactions, however it is not currently known whether conformational changes within Arf1 are needed prior to GEF association (to be addressed in Chapter 3). It is clear, however, that upon interaction, SW1 has lost its secondary

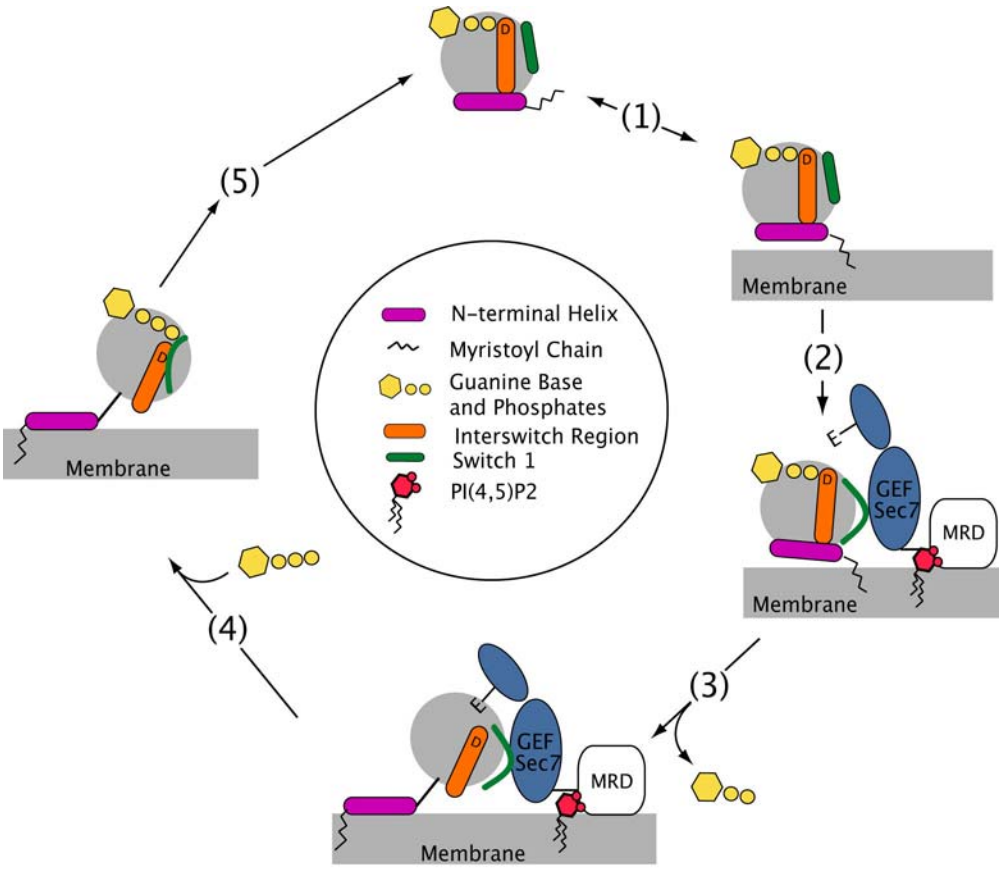


Figure 1.8. Mechanistic model of Arf1 activation. See text for details on independent steps within the nucleotide binding cycle (1-5).

structure and is involved in clamping of the GEF hydrophobic groove. (3) Rotation of the Arf1 core towards the GEF allows for insertion of the glutamic finger into the nucleotide-binding pocket of Arf1. This act displaces both the bound nucleotide and coordinated $Mg^{2+}$ ion by disrupting the hydrogen bond network surrounding the terminal phosphate. The interswitch has shifted out of the protein core as a result of Arf1 rotation forcing a release of the N-terminal helix as well as dispensing with the excess electro-negativity found in the nucleotide-binding pocket through a 2 residue shifting of D67. Interaction with the Glutamic finger allows for displacement of the P-loop, compensating for the loss of phosphate oxygens,

stabilizing the apo form. (4) Incorporation of GTP, likely interacting first with the Guanine base, allows for release of the associated GEF and locks Arf1 tightly onto the membrane surface. (5) Hydrolysis of GTP to GDP is needed for release of Arf from the membrane and subsequent vesicle fusion. As Arf1 has no innate GTPase activity, it requires the use GTPase activating proteins (GAPs) for this step.

While it seems the mechanism leading to Arf1 activation is well established, some key questions remain. First, because structures of GDP and GTP loaded complexes also differ due to N-terminal helix removal, it is not clear how much of the conformational change currently attributed to nucleotide exchange is actually associated with N-terminal helix deletion, a process that may mimic membrane association. Second, it is not currently clear whether the interaction between Arf1 and its GEF occurs in a conformation resembling the basal (GDP) state of Arf1 or only after some membrane induced structural change occurs (Figure 1.8, step 2). Third, while there is biochemical evidence speaking to the significance of phosphoinositide interactions with both Arf1 and its associated Arf-GEF, details of the conformational rearrangements accompanying the interactions between PIs and Arfs remain elusive. In the coming chapters, NMR evidence will be presented that may provide answers to these key questions.

**References**


1. Cassel, D. and T. Pfeuffer. (1978) Mechanism of cholera toxin action: covalent modification of the guanalyl nucleotide binding protein of the adenylate cyclase system. *Proc. Natl. Acad. Sci. USA*, **75**: p. 2669-2673.
2. Kahn, R.A. and A.G. Gilman. (1984) Purification of a protein cofactor required for ADP-ribosylation of the stimulatory regulatory component of adenylate cyclase by cholera toxin. *J Biol Chem*, **259**(10): p. 6228-34.
3. Lebeda, R.A., S.K. Johnson, M.I. Stewart, and R.S. Haun. (2003) Sequence, Genomic Organization, and Expression of the Human ARF6 Gene: A Class III Arf. *DNA an Cell Biology*, **22**(11): p. 737-741.
4. Chavrier, P. and B. Goud. (1999) The role of Arf and rab GTPases in membrane transport. *Curr. Opin. Cell Biol*, **11**: p. 466-475.
5. D'Souza-Schorey, C., G. Li, M.I. Columbo, and P.D. Stahl. (1995) A regulatory role for Arf6 in receptor-mediated endocytosis. *Science*, **267**: p. 1175-1178.
6. Cavenagh, M.M., J.A. Whitney, K. Carroll, C. Zhang, A.L. Boman, A.G. Rosenwald, I. Mellman, and R.A. Kahn. (1996) Intracellular distribution of Arf proteins in mammalian cells. Arf6 is uniquely localized to the plasma membrane. *J Biol Chem*, **271**(36): p. 21767-74.
7. zhu, J.-W., H. Nagasawa, F. Nagura, S.B. Mohamad, Y. Uto, K. Ohkura, and H. Hori. (2000) Elucidation of Strict Structural Requirements of Brefeldin A as an inducer of Differentiation and Apoptosis. *Bioorg. Med. Chem*, **8**(2): p. 455-463.
8. Kahn, R.A., J.K. Yucel, and V. Malhotra. (1993) ARF signaling: a potential role for phospholipase D in membrane traffic. *Cell*, **75**(6): p. 1045-8.
9. Shome, K., C. Vasudevan, and G. Romero. (1997) ARF proteins mediate insulin-dependent activation of phospholipase D. *Curr Biol*, **7**(6): p. 387-96.
10. Godi, A., P. Pertile, R. Meyers, P. Marra, G. Di Tullio, C. Iurisci, A. Luini, D. Corda, and M.A. De Matteis. (1999) ARF mediates recruitment of PtdIns-4-OH kinase-beta and stimulates synthesis of PtdIns(4,5)P2 on the Golgi complex. *Nat Cell Biol*, **1**(5): p. 280-7.
11. Hill, K., Y. Li, M. Bennet, M. McKay, X. Zhu, J.F. Shern, E. Torre, J. Lah, A. Levey, and R. Kahn. (2003) Munc18 interacting proteins- ADP-ribosylation factor-dependent coat proteins that regulate the traffic of beta-Alzheimers precursor protein. *J Biol Chem*, **278**(38): p. 36032-36040.
12. Takatsu, H., K. Yoshino, K. Toda, and K. Nakayama. (2002) GGA proteins associate with Golgi membranes through interaction between their GGAH domains and ADP-ribosylation factors. *Biochem J*, **365**(Pt 2): p. 369-78.
13. Zhdankina, O., N.L. Strand, J.M. Redmond, and A.L. Boman. (2001) Yeast GGA proteins interact with GTP-bound Arf and facilitate transport through the Golgi. *Yeast*, **18**(1): p. 1-18.
14. Donaldson, J.G., D. Cassel, R.A. Kahn, and R.D. Klausner. (1992) ADP-ribosylation factor, a small GTP-binding protein, is required for binding of the coatomer protein beta-COP to Golgi membranes. *Proc Natl Acad Sci U S A*, **89**(14): p. 6408-12.

15. Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95**(2): p. 237-48.
16. Pasqualato, S., J. Menetrey, M. Franco, and J. Cherfils. (2001) The structural GDP/GTP cycle of human Arf6. *EMBO Rep*, **2**(3): p. 234-8.
17. Losonczi, J.A., F. Tian, and J.H. Prestegard. (2000) Nuclear magnetic resonance studies of the N-terminal fragment of adenosine diphosphate ribosylation factor 1 in micelles and bicelles: influence of N-myristoylation. *Biochemistry*, **39**(13): p. 3804-16.
18. Kahn, R.A. and A.G. Gilman. (1986) The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein. *J Biol Chem*, **261**(17): p. 7906-11.
19. Kahn, R.A., P. Randazzo, T. Serafini, O. Weiss, C. Rulka, J. Clark, M. Amherdt, P. Roller, L. Orci, and J.E. Rothman. (1992) The amino terminus of ADP-ribosylation factor (ARF) is a critical determinant of ARF activities and is a potent and specific inhibitor of protein transport. *J Biol Chem*, **267**(18): p. 13039-46.
20. Amor, J.C., D.H. Harrison, R.A. Kahn, and D. Ringe. (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature*, **372**(6507): p. 704-8.
21. DeLano, W.L., *The PyMol Molecular Graphics System*. 2002, DeLano Scientific: San Carlos.
22. Amor, J.C., J.R. Horton, X. Zhu, Y. Wang, C. Sullards, D. Ringe, X. Cheng, and R.A. Kahn. (2001) Structures of yeast ARF2 and ARL1: distinct roles for the N terminus in the structure and function of ARF family GTPases. *J Biol Chem*, **276**(45): p. 42477-84.
23. Greasley, S.E., H. Jhoti, C. Teahan, R. Solari, A. Fensome, G.M. Thomas, S. Cockcroft, and B. Bax. (1995) The structure of rat ADP-ribosylation factor-1 (ARF-1) complexed to GDP determined from two different crystal forms. *Nat Struct Biol*, **2**(9): p. 797-806.
24. Kahn, R.A., C. Goddard, and M. Newkirk. (1988) Chemical and immunological characterization of the 21-kDa ADP- ribosylation factor of adenylate cyclase. *J Biol Chem*, **263**(17): p. 8282-7.
25. Franco, M., P. Chardin, M. Chabre, and S. Paris. (1993) Myristoylation is not required for GTP-dependent binding of ADP- ribosylation factor ARF1 to phospholipids. *J Biol Chem*, **268**(33): p. 24531-4.
26. Franco, M., P. Chardin, M. Chabre, and S. Paris. (1996) Myristoylation-facilitated binding of the G protein ARF1GDP to membrane phospholipids is required for its activation by a soluble nucleotide exchange factor. *J Biol Chem*, **271**(3): p. 1573-8.
27. Zhu, X., A.L. Boman, J. Kuai, W. Cieplak, and R.A. Kahn. (2000) Effectors increase the affinity of ADP-ribosylation factor for GTP to increase binding. *J Biol Chem*, **275**(18): p. 13465-75.
28. Antonny, B., S. Beraud-Dufour, P. Chardin, and M. Chabre. (1997) N-terminal Hydrophobic Residues of the G-Protein ADP Ribosylation Factor-1 Insert into Membrane Phospholipids upon GDP to GTP Exchange. *Biochemistry*, **36**: p. 4675-4684.
29. Reinstein, J., I. Schlichting, and A. Wittinghofer. (1990) Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of Escherichia coli. *Biochemistry*, **29**(32): p. 7451-9.

30. Lacal, J. and F. McCormick, *The Ras Superfamily of GTPases*. 1993, Boca Raton, FL: CRC Press.
31. Vetter, I.R. and A. Wittinghofer. (2001) The guanine nucleotide-binding switch in three dimensions. *Science*, **294**(5545): p. 1299-304.
32. Scheffzek, K., M.R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, and A. Wittinghofer. (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, **277**: p. 333-338.
33. Zhang, C.J., A.G. Rosenwald, M.C. Willingham, S. Skuntz, J. Clark, and R.A. Kahn. (1994) Expression of a dominant allele of human ARF1 inhibits membrane traffic in vivo. *J Cell Biol*, **124**(3): p. 289-300.
34. Kahn, R.A., J. Clark, C. Rulka, T. Stearns, C. Zhang, P. Randazzo, T. Terui, and M. Cavenagh. (1995) Mutational Analysis of Saccharomyces cerevisiae ARF1. *J Biol Chem*, **270**(1): p. 143-150.
35. Renault, L., B. Guibert, and J. Cherfils. (2003) Structural Snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. *Nature*, **426**(6966): p. 525-530.
36. Mossessova, E., J.M. Gulbis, and J. Goldberg. (1998) Structure of the guanine nucleotide exchange factor Sec7 domain of human arno and analysis of the interaction with ARF GTPase. *Cell*, **92**(3): p. 415-23.
37. Donaldson, J.G. and C.L. Jackson. (2000) Regulators and Effectors of the ARF GTPases. *Current Opinion in Cell Biology*, **12**: p. 475-482.
38. Macia, E., S. Paris, and M. Chabre. (2000) Binding of the PH and polybasic C-terminal domains of ARNO to phosphoinositides and to acidic lipids. *Biochemistry*, **39**(19): p. 5893-901.
39. Franco, M., P. Chardin, M. Chabre, and S. Paris. (1996) Myristoylation facilitated Binding of the G Protein ARF1-GDP to membrane Phospholipids is Required for Its Activation by a Soluble Nucleotide Exchange Factor. *J Biol Chem*, **271**(3): p. 1573-1578.
40. Beraud - Dufour, S., S. Robineau, P. Chardin, S. Paris, M. Chabre, J. Cherfils, and B. Antonny. (1998) A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg2+ and the beta-phosphate to destabilize GDP on Arf1. *EMBO J.*, **17**: p. 3651-3659.
41. Beraud-Dufor, S., S. Robineau, P. Chardin, S. Paris, M. Chabre, J. Cherfils, and B. Antonny. (1998) A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg2+ and the beta-phosphate to destabalize GDP on Arf1. *EMBO J.*, **17**(3): p. 3651-3659.
42. Roth, M.G. (1999) Lipid Regulators of Membrane Traffic through the Golgi Complex. *Trends in Cell Biology*, **v.9**(5): p. 174-179.
43. Yoon, H.S., P.J. Hajduk, A.M. Petros, E.T. Olejniczak, R.P. Meadows, and S.W. Fesik. (1994) Solution Structure of a pleckstrin-homology domain. *Nature*, **369**(23): p. 672-675.
44. Hyvonen, M., M.J. Macias, M. Nilges, H. Oschkinat, M. Saraste, and M. Wilmanns. (1995) Structure of the binding site for inositol phosphates in a PH domain. *The EMBO Journal*, **14**(19): p. 4676-4685.

45. Seidel III, R.D., J.C. Amor, R.A. Kahn, and J.H. Prestegard. (2004) Structural Perturbatins in Human ADP Ribosylation Factor-1 Accompanying the Binding of Phosphatidylinositides. *Biochemistry*, **43**: p. 15393-15403.
46. Randazzo, P. (1997) Functional interaction of ADP-ribosyaltion factor 1 with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem*, **272**(12): p. 7688-7692.
47. Martin, A., F.D. Brown, M.N. Hodgkin, A.J. Bradwell, S.J. Cook, M. Hart, and M.J. Wakelam. (1996) Activation of phospholipase D and phosphatidylinositol 4-phosphate 5- kinase in HL60 membranes is mediated by endogenous Arf but not Rho. *J Biol Chem*, **271**(29): p. 17397-403.
48. Randazzo, P.A. and R.A. Kahn. (1994) GTP hydrolysis by ADP-ribosylation factor is dependent on both an ADP- ribosylation factor GTPase-activating protein and acid phospholipids. *J Biol Chem*, **269**(14): p. 10758-63.
49. Terui, T., R.A. Kahn, and P.A. Randazzo. (1994) Effects of acid phospholipids on nucleotide exchange properties of ADP- ribosylation factor 1. Evidence for specific interaction with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem*, **269**(45): p. 28130-5.
50. Kahn, R.A., T. Terui, and P.A. Randazzo. (1996) Effects of acid phospholipids on ARF activities: potential roles in membrane traffic. *J Lipid Mediat Cell Signal*, **14**(1-3): p. 209-14.
51. Paris, S., S. Beraud-Dufor, S. Robineau, J. Bigay, B. Antonny, M. Chabre, and P. Chardin. (1997) Role of Protein-Phospholipid Interactions in the Activation of ARF1 by the Guanine Nucleotide Exchange Factor Arno. *J Biol Chem*, **272**(35): p. 22221-22226.

## Chapter 2

## Residual Dipolar Couplings and Protein Structure Determination

Within its relatively short history, Nuclear Magnetic Resonance (NMR) spectroscopy has managed to play a significant role in the study of macromolecular structure and dynamics. The methods employed in structure determination, however, have relied almost exclusively on the measurement of significant amounts of semi-quantitative local restraints. Prominent among these are homonuclear $^{1}H$-$^{1}H$ nuclear overhauser effects (NOEs). In many cases, the NOE is quantitated and reported as a percentage of peak intensity observed with and without NOE enhancement. The resulting change in intensity is binned into strong, medium, and weak effects providing distance information for atoms separated by approximately 2.5Å, 3.7Å, and 5Å, respectively. Also, measurement of three-bond scalar coupling constants ($^{3}J$), either homonuclear ($^{1}H$-$^{1}H$, $^{13}C$-$^{13}C$) or heteronuclear ($^{13}C$-$^{1}H$, $^{13}C$-$^{15}N$, $^{15}N$-$^{1}H$), has provided information on the intervening dihedral angles ($\phi$, $\varphi$) through comparison with a semi-empirical curve described by the Karplus equation. For couplings between the amide proton and the $\alpha$ proton of an amino acid in a protein this is given by $^{3}J = (6.4\cos^{2}\theta - 1.4\cos\theta + 1.9)$ Hz; where $\theta = (\phi - (\pi/3))$ *(4)*. These Karplus relationships are used to provide information on the secondary structure of proteins under question, and numerous methods have been developed over the years for the measurement of such couplings *(6)*. Lastly, chemical shifts of backbone atoms ($^{1}H$, $^{13}C_{\alpha}$, $^{13}C_{o}$, and $^{15}N$) have proven to be a valuable source of structural

information. In this case, as in the last, a distinct relationship to backbone torsion angles has been seen and chemical shift values can therefore be used to improve the quality of NMR derived structures.

While these methods are widely accepted, they are not without their inherent limitations. The steep distance dependence ($1/r^6$) and resultant short-range interaction of the NOE makes extensive assignment of protein backbone and, more importantly, side chain resonances a must. Also, the scalar coupling and chemical shift values, while useful, provide strictly local information. Propagation of errors obtained from these short-range constraints can leave remote parts of a molecule poorly defined. What is more, the time consuming aspects of assignments needed for NOE based structures may not be justified in all situations. For many problems we begin with well-defined structures from X-ray, NMR, or molecular modeling and are primarily interested in changes associated with steps along a biochemical process. Are there more efficient approaches applicable in these situations? In what follows, I will focus on one approach to alleviate some of these limitations, the measurement and application of residual dipolar couplings (RDCs).

## 2.1 An Overview of High Resolution Liquid Crystal NMR

Before I begin, I would like to briefly describe the properties of the liquid crystals to be discussed below, as it will provide a convenient background for future discussions. Molecules dissolved in liquid phases retain a high level of molecular mobility. This mobility has two aspects: (1) Translational molecular mobility and (2) Rotational molecular mobility. Translational mobility refers to the molecules capacity to move past one another and travel a

range of distances. Rotational mobility refers to the molecules propensity to spin around its center of mass. Both are very important in NMR spectroscopy, as they tend to average out many nuclear spin interactions.

When molecules are dissolved in isotropic liquids, the translational and rotational mobilities are the same in all directions. Molecular motion in these liquids effectively averages out most spin-spin and spin-anisotropic magnetic field interactions resulting in much simplified NMR spectra. Conversely, anisotropic liquids (referred to as liquid crystals) are those in which the solvent molecules sample translational and rotational degrees of freedom in a non-isotropic fashion. This motional anisotropy leads to incomplete averaging of nuclear spin interactions resulting in more complicated, but information rich, NMR spectra.

It is important to note that the study of molecules in oriented phases is not new. In fact, the NMR spectra of nematic liquid crystals, and solutes in nematic liquids, dates back to 1963. The term nematic, in this case, refers to liquid crystals ordered in only one direction. In a classic example, Saupe and Englert had shown a change in the proton NMR spectra of benzene dissolved in PAA (*p*-azoxyanisole) from the single sharp resonance commonly observed in isotropic liquids to complex spectra having more than 30 peaks *(12)*. Interestingly, the benzene molecule retained a high level of mobility and relatively narrow lines. The inter-molecular interactions were averaged out as in isotropic liquids. Moreover, the reorientation of benzene was also sufficiently rapid to reduce the intra-molecular interactions to the same average value for every molecule in solution. Later, Emsley and Lindon (1975) reported a similar ability to dissolve small molecules (norbornadiene) in ordered nematic solvents to obtain residual dipolar information while retaining a high level of spectral resolution *(15)*. However, the nematic phases used were not suitable for extension to

biological macromolecules as vast amounts of long-range homonuclear $^{1}H$-$^{1}H$ coupling remained in these well-ordered systems, making spectra problematic. Moreover, the systems used were based primarily on organic solvents. The ability to use nematic liquid crystals in macromolecular systems would remain undeveloped for more than 20 years awaiting isotopic labeling schemes and tunable aqueous alignment media.

Increases in magnetic field strength allowed exploration of alternatives to nematic liquids for molecular ordering. In the late 70's, molecules with anisotropic susceptibilities were shown to preferentially align along the magnetic field *(18)*. However the overall alignment was small leaving dipolar effects hard to measure. Quadrupolar splittings, which show the same orientational dependence but have larger couplings, were more commonly reported. While application of quadrupolar splittings to small organics showed promise, poor chemical shift resolution made extension to larger biological systems difficult. Measurement of the weaker dipolar interaction, although vanishingly small as reported, was decidedly more advantageous.

In 1981, Bothner-By and co-workers showed that small molecules containing paramagnetic ions could also undergo field-induced ordering. In this case, the level of alignment was an order of magnitude higher than the previously reported diamagnetic systems *(18)*, allowing for direct assessment of residual dipolar splittings *(21)*. Magnetic susceptibilities in these systems, both diamagnetic and paramagnetic, were known to be additive *(22)* and ultimately allowed for extension to more biologically relevant samples *(23)*.

The first application to a protein system was reported in 1995, again relying on direct field dependent orientation of a molecule with large anisotropic magnetic susceptibility. This pioneering work demonstrated an ability to measure dipolar contributions to splittings of

amide pairs in a $^{15}N$ labeled sample of cyanometmyoglobin *(28)*. The measured splittings in this case spanned 2.5 Hz, substantially larger than the fraction of a Hertz previously reported. The ability to isotopically label the protein, the use of paramagnetic ions for molecular alignment and the increased field strengths available (750 MHz) allowed for the needed compromise between spectral simplicity and the level of molecular order. Although this proved to be a viable means of extracting dipolar couplings in protein systems, it remained only applicable to systems with metal centers. A more robust method for macromolecular alignment was needed for more widespread applicability.

Rapid advances in aqueous liquid crystal media quickly bridged this gap *(31-34)* and their application to proteins was finally illustrated in the work of Tjandra and Bax *(35)*. They demonstrated an ability to measure both homonuclear ($^{13}C_{\alpha}$-$^{13}C_{o}$) and heteronuclear ($^{15}N$-$^{1}H$, $^{13}C_{\alpha}$-$^{1}H_{\alpha}$) dipolar couplings in an isotopically labeled sample of ubiquitin. The observed couplings spanned more than 20 Hz ($^{15}N$-$^{1}H$), an order of magnitude larger than that reported with field-induced alignment alone *(28)*. The solute molecules retained a high level of mobility and the resulting line-widths were only slightly perturbed, resulting in reported errors of only a fraction of a Hertz. The liquid crystals employed were a form of discotic phospholipid micelles, termed “bicelles”, consisting of mixtures of DHPC and DMPC *(32)*. The degree of protein alignment obtained at low bicelle concentrations was large enough to accurately measure a wide array of dipolar couplings, yet small enough to prevent spectral crowding and the other deleterious effects previously observed in liquid crystal studies *(12, 15)*. Moreover, the dipolar couplings yielded structural information with accuracy unprecedented in macromolecular NMR. It would seem that the final obstacle to routine use of residual dipolar couplings as a structural tool had been breeched.

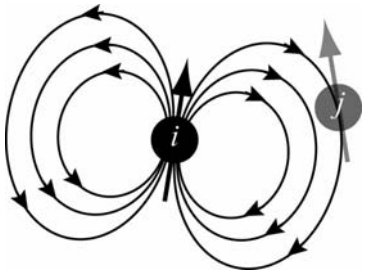


Figure 2.1. The magnetic field interaction of nuclear spin *i* with nuclear spin *j*. This figure was adapted from the literature *(36)*.

## 2.2 Residual Dipolar Coupling: Theory

The presence of the dipolar interaction is quite simple to imagine. As an example, consider the case of a two-spin system, having spin-1/2 nuclei, *i* and *j*. As each spin in the system is magnetic, it generates a magnetic field. The consequent field lines of spin *i* encircle the surrounding space in a manner dictated by the direction of the magnetic moment. A nearby nuclear spin, in this case *j*, interacts with the dipolar field (Figure 2.1). This through space interaction is referred to as direct dipole-dipole coupling, or simply dipolar coupling, as the interaction between the two nuclei transmit through the intervening space.

The nature of the dipolar interaction, on the other hand, is more intricate, requiring a mathematical description in the form of the time dependant dipolar Hamiltonian shown below:

$$H_{i,j}^{DD}(t) = D_{ij}^{\text{Static}} [3 (\hat{I}_i \bullet e_{ij})( \hat{I}_j \bullet e_{ij}) - \hat{I}_i \bullet \hat{I}_j], \qquad [1]$$

where $e_{ij}$ represents a unit vector joining the centers of the two nuclei, *i* and *j*, and $\hat{I}_{i,j}$ refers to the spin angular momentum operator for spin *i* or *j*. $D_{ij}^{\text{Static}}$ defines the magnitude of the

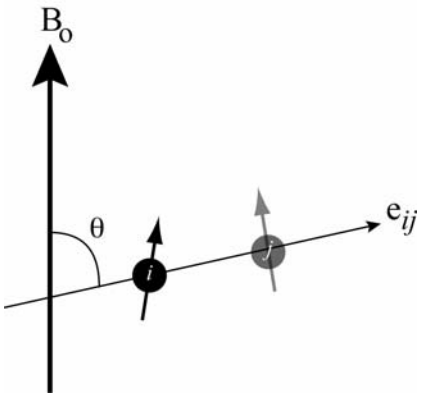


Figure 2.2. The angle (θ) between an internuclear vector ($e_{ij}$) and the magnetic field ($B_o$). This figure was adapted from the literature *(36)*.

dipolar interaction and is referred to as the dipolar coupling constant, illustrated in equation 2 below:

$$D_{ij}^{Static} = - \mu_0(h/2\pi)\,\gamma_i\,\gamma_j/(4\pi^2 r^3_{ij}), \qquad [2]$$

with $\mu_0$ representing the magnetic permittivity of a vacuum, h being Planck's constant, $\gamma_{i,j}$ is the magnetogyric ratio of spin *i* or *j*, and $r_{ij}$ is the internuclear distance between nuclei *i* and *j*. It is important to note that the dipolar coupling constant is not orientation dependant, it scales based on the type of nuclei being observed and the distance between them. The dipolar Hamiltonian, however, remains time dependant as the direction of the vector, $e_{ij}$, changes as the molecule rotates. It is not surprising, then, that residual dipolar couplings would vanish in isotropic systems as random Brownian motion generates a uniform distribution of vector orientations.

At the high field limit, the non- secular terms of the dipolar Hamiltonian may be discarded. Therefore, only the z component of the local field of one nuclear dipole at the position of the second remains relevant. This truncated Hamiltonian is represented in equation 3.

$$H_{i,j}^{DD}(t) = D_{ij}^{\text{Static}} \langle \hat{I}_{i,z} \hat{I}_{j,z} (3\cos^2\theta - 1) \rangle, \qquad [3]$$

where the angular brackets denote an average over allowed molecular orientations, $\hat{I}_{i,j\,(z)}$ being the z component of the spin angular momentum operator for spin *i* or *j*, and $\theta$ being the angle between the internuclear vector ($e_{ij}$) and the magnetic field ($B_o$)(Figure 2.2). It is clear that this description yields both distance and angular information. However, when the distance is fixed, as in a directly bonded pair, the splitting of the NMR resonances for the two nuclei due to residual dipolar coupling yields solely angular information. This resonance splitting is defined by the following equation:

$$D_{ij}^{\text{RES}} = D_{ij}^{\text{Static}} \langle P_2(\cos\theta) \rangle, \qquad [4]$$

with $P_2(x) = 1/2(3x^2 - 1)$. Here $D_{ij}^{\text{Static}}$ is the dipolar coupling between the nuclei in a perfectly ordered system with $\theta = 0$. Averaging over allowed orientations, as denoted by the brackets, greatly reduces the observed values, but the residual values can be analyzed to provide both structural information and the nature of molecular order when sufficient measurements can be made *(37, 38)*.

### *2.2.1 Rigid Body Approximation*

In solution, the time dependence of the dipolar Hamiltonian arises from reorientation of the internuclear vector, $e_{ij}$, with respect to the magnetic field, as well as small fluctuations in internuclear distances and angles (i.e. local bond librations) *(37)*. This problem can be simplified when measuring couplings from directly bonded pairs, as the internuclear distance can be assumed fixed. While it is understood that small fluctuations in this distance may still persist, the use of an effective internuclear distance allows the elimination of explicit consideration of these motions *(39)*. In general a description of average orientation for a single vector requires 2 angles, more than a single measurement can provide.

The problem can be further simplified if one assumes the molecule (or molecular fragment) to be rigid. In proteins, secondary structural elements such as $\alpha$ helices and $\beta$ strands may represent suitable rigid bodies. However, in a more general case, the peptide plane itself can provide the needed rigid entity. In this case, the orientation of the internuclear vector, $r_{ij}$, in an arbitrarily chosen molecular coordinate system, fixed with respect to the molecule, can be described in Cartesian space by the angles $\alpha_x$, $\alpha_y$, and $\alpha_z$ (defined in spherical space by the Polar angles $\theta$, $\phi$), as illustrated in Figure 2.3. The time dependent orientation of the axes of this coordinate system with respect to the magnetic field can then be defined by the angles $\beta_x$, $\beta_y$, and $\beta_z$ ($\zeta_{(t)}$, $\xi_{(t)}$). While the Cartesian approach illustrated remains more suitable for physical interpretation, a similar representation in spherical space is also employed [Figure 2.3(B)], as it is convenient in separating the rotational properties of vectors. The discussion to follow therefore relies on the Polar description of vector orientation; the Cartesian representation is employed in the following section, “The Saupe Order Matrix”.

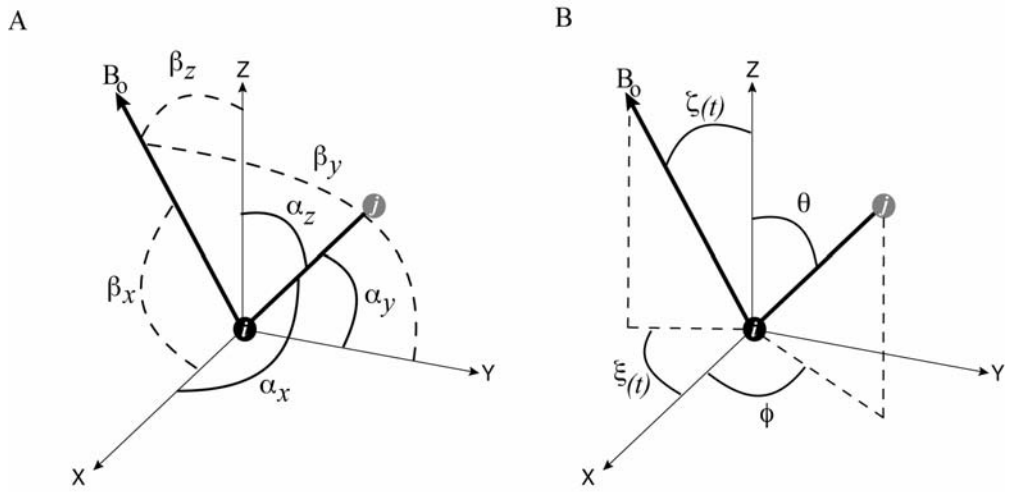


Figure 2.3. Definition of the angles relating an internuclear vector ($r_{ij}$) and the molecular coordinate frame in (A) Cartesian space using the angles $\alpha_x$, $\alpha_y$, and $\alpha_z$ and (B) in Spherical space using the angles θ, ϕ. The time dependent angles, (A) $\beta_x$, $\beta_y$, and $\beta_z$ and (B) $\zeta_{(t)}$, $\xi_{(t)}$ similarly define the orientation of the magnetic field relative to the x, y, and z axes of the coordinate system.

Defining the value cos (θ) (equation 4) as the product between a unit vector along the magnetic field and one in the direction of the internuclear vector, the second rank Legendre function, $P_2(\cos\theta)$, can be expanded in spherical space using the Legendre addition theorem *(40)*. The resulting Legendre polynomial is shown in equation 5 below.

$$P_2(\cos\theta) = 4\pi/5 \; \Sigma^2_{m=-2} \; Y^m{}_2(\theta, \phi) \; Y^m{}_2(\zeta_{(t)}, \xi_{(t)}), \qquad [5]$$

The expression Y in this case relates to the surface spherical harmonics of the system and the superscript *m* represents the magnetic quantum number. After comparison to equation 4 above, $D_{ij}^{RES} = D_{ij}^{Static} \langle P_2(\cos\theta)\rangle$, it is easily seen that the rigid molecule approximation has

generated a description of dipolar coupling in which the angular time dependence is factored into only one set of angles, ($\zeta_{(t)}$, $\xi_{(t)}$).

### *2.2.2 The Saupe Order Matrix*

Defining, again, the value cos (θ) (equation 4) as the scalar product between a unit vector in the internuclear direction and one along the magnetic field ($B_o$), $P_2(\cos\theta)$ can be expanded in Cartesian space to:

$$\langle P_2(\cos\theta)\rangle = 3/2\ \langle(\cos\beta_x \cos\alpha_x + \cos\beta_y \cos\alpha_y + \cos\beta_z \cos\alpha_z)^2\rangle - 1/2, \qquad [6]$$

Defining $A_i = \cos(\alpha_i)$ and $B_{i,(t)} = \cos(\beta_i)$, (the symbol ($t$) in this case corresponds to the time dependant angles) equation 6 can be further expanded to:

$$\langle P_2(\cos\theta)\rangle = 3/2\ [B^2_{x,(t)}A_x^2 + B^2_{y,(t)}A_y^2 + B^2_{z,(t)}A_z^2 + 2(B_{x,(t)}\,B_{y,(t)}A_xA_y) + 2(B_{x,(t)}\,B_{z,(t)}A_xA_z) + 2(B_{y,(t)}\,B_{z,(t)}A_yA_z)] - 1/2, \qquad [7]$$

Upon setting the expression $S_{kl}$ equal to $3/2(B_{k,(t)}\,B_{l,(t)}) - 1/2\delta_{kl}$, where $\delta_{kl}$ represents the Kronecker delta function[1], the above equation reduces to:

$$\langle P_2(\cos\theta)\rangle = \Sigma_{kl\,=\,[x,\,y,\,z]}\ S_{kl} \cos(\alpha_k) \cos(\alpha_l), \qquad [8]$$

[1] The Kronecker Delta function has the following meaning: $\delta_{kl}$ evaluates to 1 if k = l, otherwise $\delta_{kl} = 0$.

$$\begin{pmatrix} S_{xz} \\ S_{yz} \\ S_{xy} \\ S_{xx\,-\,yy} \\ S_{zz} \end{pmatrix}$$

Figure 2.4. Vector description of the five remaining elements in the Saupe order tensor obtained after matrix diagonalization.

The internal geometry is now represented by direction cosines ($\cos\alpha_k$, $\cos\alpha_l$) relating the orientation of the internuclear vector with that of the arbitrarily chosen fragment frame (axes x, y, z). The time dependant angles, those related to the orientation of the alignment frame, are encompassed by the variables $S_{kl}$. This 3 x 3 matrix is commonly referred to as the Saupe order matrix *(41)*, the order tensor, or simply the order matrix. As the diagonal elements ($S_{xx}$, $S_{yy}$, and $S_{zz}$) sum to zero, the matrix is traceless, and since the elements $S_{kl}$ and $S_{lk}$ are equivalent, it is also symmetric, yielding only five independent elements (consistent with the five spherical harmonic sums obtained in equation 5 above when expanding in Polar space). The five remaining elements of the order tensor are illustrated in vector notation above (Figure 2.4).

If the structure of the molecule (or molecular fragment) is known (represented in the above equation by direction cosines), the five independent elements of the order matrix can be solved provided that dipolar couplings for at least five internuclear vectors are available. If any of the five internuclear vectors are found collinear, however, more measured couplings are required. This is generally not a problem in macromolecular NMR as many internuclear

vectors are commonly observed, resulting in an over-determined evaluation of the order matrix *(38)*.

Extraction of order information from measured couplings is achieved by diagonalization of the order tensor. This is performed by sequential rotation of the experimentally determined axis system (called the principal alignment frame, PAF) into that of the arbitrarily chosen axis system (x, y, z axis from Figure 2.3). In this case, the elements of the order tensor (Figure 2.4) can be reduced to two variables, $S_{zz}$, which defines the magnitude of order, and $\eta$ [where $\eta = (S_{xx} - S_{yy})/(S_{zz})$] defining the deviation from axial symmetry. This rotation is generally represented in terms of three Euler angles relating the orientation of the PAF to the molecular frame, thus defining the average orientation of the molecule, or molecular fragment. As will become clear in later sections, it is advantageous to work in this principal alignment frame in a number of applications. In this instance, the dipolar interaction defined in equation 7 can be rewritten as:

$$D_{ij}^{RES} = D_a (3\cos^2\theta - 1) + 3/2\ R (\sin^2\theta \cos 2\phi), \qquad [9]$$

where $D_a$, defining the magnitude of anisotropy, is equal to $[1/2(S_{zz})( D_{ij}^{Static})]$ and the rhombicity (R) is equal to $2/3(\eta)$.

### *2.2.3 Singular Value Decomposition and Dipolar Analysis*

Performing an order matrix analysis in practice means setting up and solving a series of linear equations of the form $Ax = b$. Where A is a matrix composed of direction cosines

$$\begin{pmatrix} (1)\,[\cos^2(\alpha)y - \cos^2(\alpha)x] & [\cos^2(\alpha)z - \cos^2(\alpha)x] & [2\cos(\alpha)x\cos(\alpha)y] & [2\cos(\alpha)x\cos(\alpha)z] & [2\cos(\alpha)y\cos(\alpha)z] \\ (2)\,[\cos^2(\alpha)y - \cos^2(\alpha)x] & [\cos^2(\alpha)z - \cos^2(\alpha)x] & [2\cos(\alpha)x\cos(\alpha)y] & [2\cos(\alpha)x\cos(\alpha)z] & [2\cos(\alpha)y\cos(\alpha)z] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (n)\,[\cos^2(\alpha)y - \cos^2(\alpha)x] & [\cos^2(\alpha)z - \cos^2(\alpha)x] & [2\cos(\alpha)x\cos(\alpha)y] & [2\cos(\alpha)x\cos(\alpha)z] & [2\cos(\alpha)y\cos(\alpha)z] \end{pmatrix} \begin{pmatrix} S_{xz} \\ S_{yz} \\ S_{xy} \\ S_{xx-yy} \\ S_{zz} \end{pmatrix} = \begin{pmatrix} (1), D_{ij} \\ (2), D_{ij} \\ \cdot \\ \cdot \\ (n), D_{ij} \end{pmatrix}$$

Figure 2.5. System of linear equations used in order matrix determination of residual dipolar couplings. Equations take the form Ax=b, where A is a matrix composed of direction cosines relating various internuclear vectors to the magnetic field, and x and b are vector descriptions of the Saupe order matrix and measured dipolar couplings. This figure has been adapted from the literature *(42)*.

relating various internuclear vectors to the fragment frame, and x and b are vector descriptions of the molecular order and measured dipolar couplings, respectively *(42)*, illustrated in Figure 2.5. In the past, solutions to these equations were performed by grid or random search approaches *(43, 44)*. These brute force methods, while sufficient for simple cases, prove inconvenient when systems of linear equations are large, as seen in biomolecular NMR. Singular Vale Decomposition (SVD) is a powerful algebraic technique for solving large sets of linear equations, allowing for more expeditious extension to macromolecular systems *(42)*. This decomposition is not trivial, however, and has necessitated its incorporation into a more robust and user-friendly analysis tool *(45)*. It is clear that this type of analysis provides a simultaneous description of the average molecular orientation (or molecular fragment orientation) and its overall level of order, having unparalleled utility for the NMR spectroscopist.

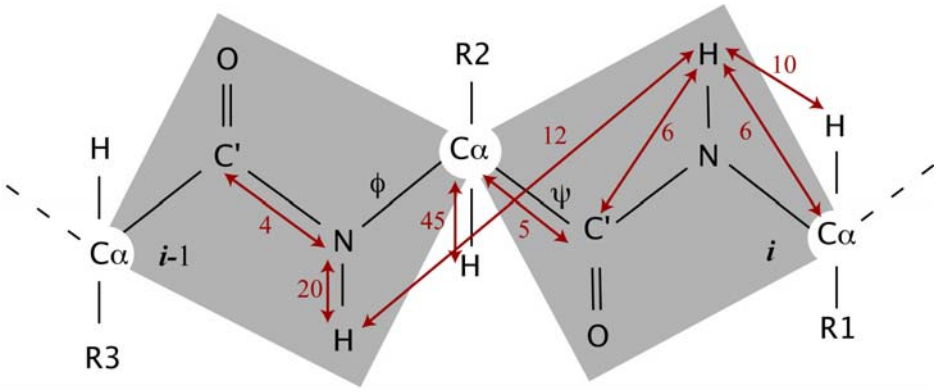


Figure 2.7. Magnitude of observable 1, 2, and multiple bond tensorial interactions within adjoining peptide planes (illustrated above as *i* and *i*-1). Rx signifies side chain atoms associated with residue X. The numerical values depicted represent estimated dipolar couplings under typical levels of molecular ordering ($10^{-3}$).

## 2.3 Measurement of RDCs

As a result of their known internuclear distance, measurement of residual dipolar couplings has focused primarily on one-bond $^{1}H$-$^{15}N$ and $^{1}H$-$^{13}C$ ($^{1}D_{HN\text{-}N}$ and $^{1}D_{C\text{-}H}$) interactions. These are also generally the easiest to measure. Moreover, as the bond lengths are short and reasonably invariant, they tend to lead to substantial dipolar interactions, thus increasing the experimental accuracy *(46)*. Two-bond interactions ($^{2}D_{HN\text{-}C\alpha}$ and $^{2}D_{HN\text{-}Co}$) also benefit from a fixed internuclear distance. However, due to the greater degree of separation, and corresponding diminished RDC value, they are more difficult to measure at the same level of relative accuracy. Measurable one, two and multiple bond dipolar interactions are illustrated in Figure 2.7. The numerical values depicted represent maximum dipolar couplings for each vector that might be observable in a molecule ordered to a level of $10^{-3}$ (levels commonly achieved in liquid crystalline phases).

Table 2.1. Common experiments used in measuring $^{1}D$, $^{2}D$, and $^{m}D$ dipolar couplings*

| Type | RDC | Experiment | Method | Ref. |
|---|---|---|---|---|
| | $H_N$-N | Coupled-HSQC | *f* | *(1)* |
| | | SCE-HSQC | *f* | *(2)* |
| | | IPAP-HSQC | *f* | *(5)* |
| | | E.COSY HSQC | *f* | *(7)* |
| | | $S^3$E-HSQC | *f* | *(8)* |
| | | J-HSQC | *I* | *(9)* |
| $^{1}D$ | | phase encoded-HSQC | *I* | *(10)* |
| | $C_\alpha$-$H_\alpha$ | HNCO | *f* | *(11)* |
| | | IPAP-(HA)CANH | *f* | *(13)* |
| | | CT-J-HSQC | *I* | *(14)* |
| | N-$C_o$, N-$C_\alpha$ | TROSY-HNCO | *I* | *(16)* |
| | | J-correlated HNC | *I* | *(17)* |
| | $C_o$-$C_\alpha$ | CT-HSQC | *f* | *(19)* |
| | $H_N$-$C_\alpha$ | soft-HNCA-E.COSY** | *f* | *(20)* |
| $^{2}D$ | | CT-J-HSQC | *I* | *(14)* |
| | $H_N$-$C_o$ | $S^3$E/IPAP-HNCO | *f* | *(24)* |
| | | semi-CT-HSQC | *f* | *(25)* |
| | H-$H_\alpha$ | HNHA | *I* | *(26)* |
| $^{m}D$ | H-H | COSY-HMQC | *f* | *(27)* |
| | | SS-HMQC | *f* | *(29, 30)* |

*This table has been adapted from the literature with permission from the primary authors *(46)*. The symbols *f* and *I* correspond to Frequency and Intensity based experiments, respectively.

The NMR experiments employed for the measurement of dipolar couplings can be divided into two categories, frequency based and intensity based methods. The frequency resolved methods are those in which separation of peak centers is measured in the frequency domain *(47)*; while intensity based experiments extract couplings directly from resonance intensity rather than from experimental splittings *(48, 49)*. In either case, the coupling

observed is the sum of scalar and dipolar contributions, necessitating measurement in both isotropic and aligned conditions. Although this may be considered unfortunate as it doubles the number of needed experiments, variations in the backbone scalar (*J*) couplings can yield useful information on local geometry in certain cases, as described previously *(4)*. In what follows, I focus my discussion primarily on frequency resolved methods for extracting dipolar couplings arising from amide resonances ($^1D_{HN-N}$) in isotopically labeled samples. Splittings of these resonances are easily observed with high sensitivity in HSQC based experiments with only minimal modification to existing pulse sequences. Also, amide pairs are well distributed throughout proteins providing global structural information. The reader is directed to Table 2.1 for a more comprehensive review of experiments commonly used for extraction of the remaining measurable couplings.

### *2.3.1 Coupled HSQC*

The pulse sequence for a standard decoupled $^1H$-$^{15}N$ Heteronuclear Single Quantum Coherence (HSQC) is shown in Figure 2.8 below *(50, 51)*. In this example, thin bars represent 90° pulses with 180° pulses being illustrated as thick bars. The phase of the pulse is written above the bar. It is important to note that the sensitivity of this experiment is proportional to the cube of the magnetogyric ratio for the nucleus used both as the origin of magnetization and means of detection at the end. As a result, protons are used as the source of magnetization and detection whenever possible.

In the experiment shown, proton spin magnetization is initially transferred to its attached $^{15}N$ nucleus, chemical shift evolution is observed during $t_1$ and the magnetization is finally transferred back to the proton spin for detection ($t_2$). Fourier transforming in $t_1$ and $t_2$

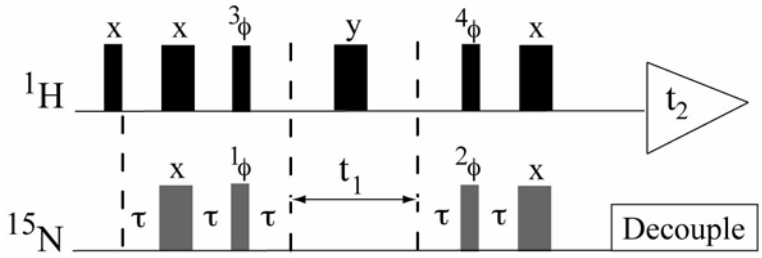


Figure 2.8. Pulse sequence for the $^{1}$H-$^{15}$N-heteronuclear single quantum coherence (HSQC) spectrum. Thin bars represent 90 pulses and thick bars represent 180 pulses. The phases of the pulse is written above the bar. The phase cycles represented by the symbol $\phi$ take the following form; $^{1}\phi$, $^{2}\phi$, $^{4}\phi$ = x, -x and $^{3}\phi$ = y, -y.

gives the corresponding frequencies $\upsilon_1$ and $\upsilon_2$. Plotting $\upsilon_2$ on the x-axis and $\upsilon_1$ on the y-axis yields the spectrum shown in Figure 2.9 (A/B). As every amino acid in the protein contains one amide site (save proline), each dot (or cross-peak) illustrated represents a single amino acid, resolved through chemical shift dispersion. Removal of the 180°$_y$ pulse found in the middle of the $t_1$ delay, allows for evolution of both scalar couplings as well as chemical shift. This experiment is referred to as a coupled HSQC *(1)*. The resulting spectrum is illustrated in Figure 2.9(C). It is clear that removal of the $t_1$ 180 allows for splitting of resonances in the indirect dimension.

The doublet splittings are directly proportional to the scalar (*J*) coupling in isotropic systems. However, upon incorporation into liquid crystalline media, the *J*-coupling is replaced by a term including both the scalar and the residual dipolar coupling. The change in resonance splitting is illustrated in Figure 2.9(D). The anisotropic doublet splitting can be either larger or smaller than the isotropic splitting, depending on the average orientation of the $^{15}$N-$^{1}$H vector with respect to the magnetic field. Subtraction of splittings arising from

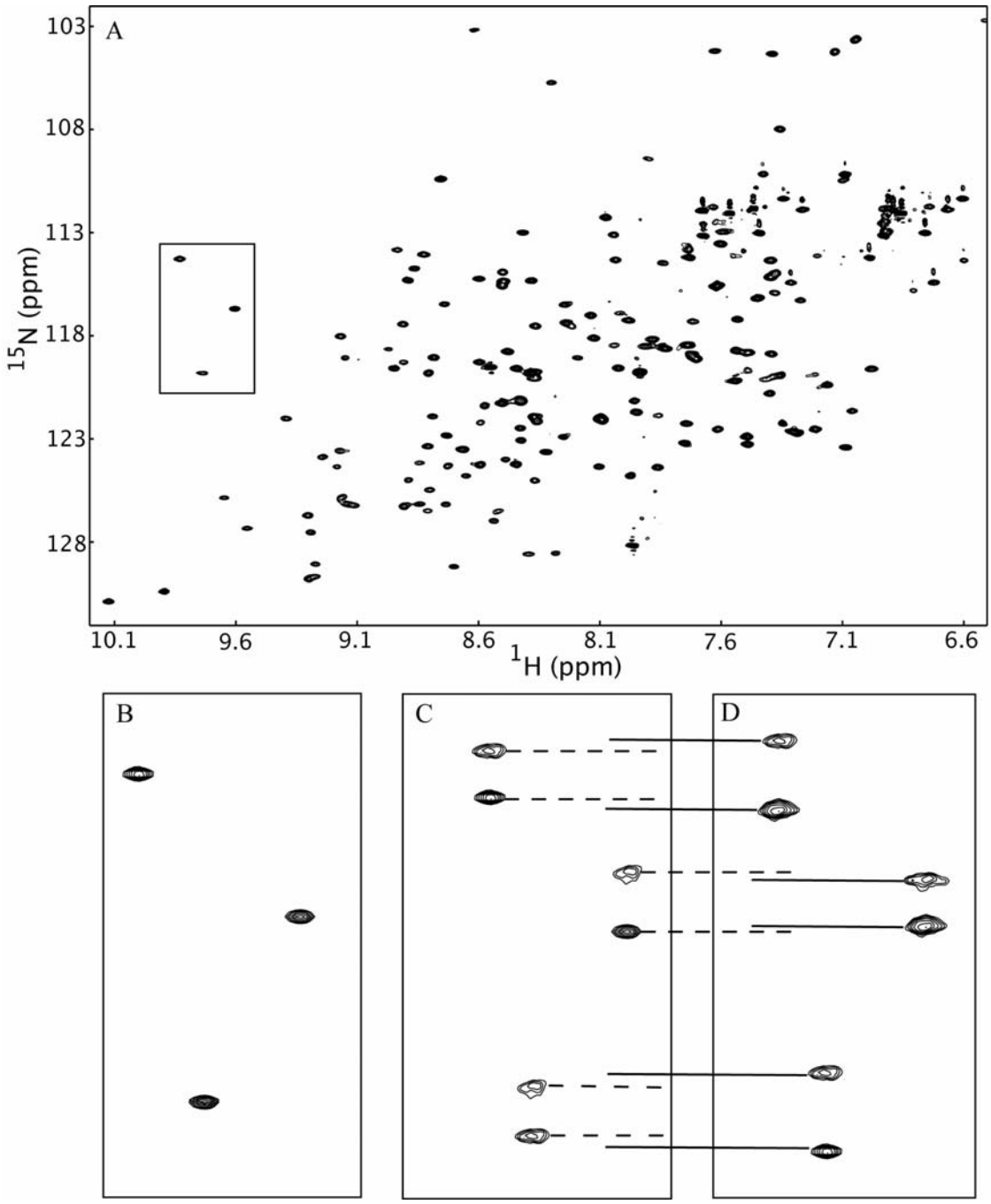


Figure 2.9. Representative spectra of (A) $^{1}$H-$^{15}$N HSQC on a 0.5mM protein sample (Arf1) in 10mM phosphate buffer (pH 7.0). Spectra recorded at 298K with a $^{1}$H resonance frequency of 800 MHz. Resonances from all amino acids, save proline, (180 in total) are represented. (B) Expansion of selected resonances observed in (A). (C) Isotropic resonance splitting in the indirect ($^{15}$N) dimension due to removal of the $t_1$ 180° decoupling pulse. The splitting is proportional the scalar (*J*) coupling as measured by the distance between pairs of dotted lines. (D) Anisotropic resonance splitting, now proportional to the sum of scalar and dipolar couplings, illustrated as pairs of solid lines.

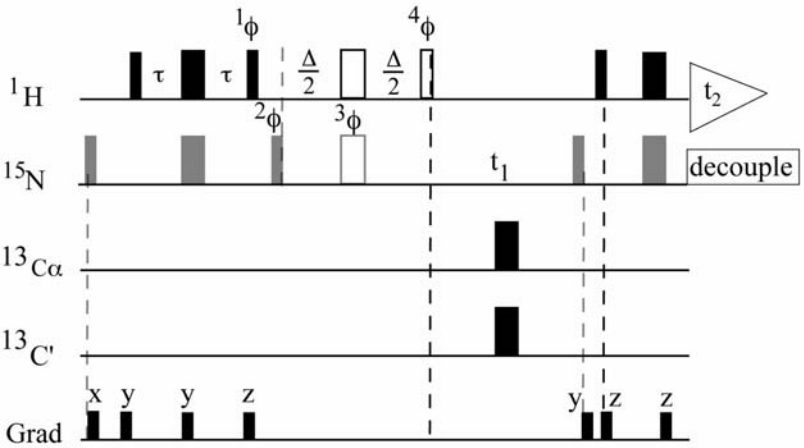


Figure 2.10. Pulse sequence of the $^{1}$H-$^{15}$N IPAP-HSQC experiment. Narrow and wide bars correspond to 90° and 180° pulses, respectively. The open pulses (non-colored bars) and optional delay (Δ/2) are only used in generating the antiphase (AP) signal and are omitted during collection of inphase (IP) spectra. $^{13}C_{\alpha}$ and $^{13}C$' pulses are applied to decouple $^{13}C$ from $^{15}N$ during $t_1$. Values of phases used in phase cycling (ϕ) are dependant on IP or AP collection. Delay durations are as follows: τ = 2.5 ms, and Δ = 5.3 ms. The reader is directed to the primary literature for a more complete description of gradient shapes and durations, and proper phases*(52)*.

isotropic and aligned conditions yields the needed residual dipolar values. It would seem that this technique, being highly sensitive and simple to employ, would be widely used. However, with each added amino acid, the resulting HSQC becomes more crowded and complex. Also, removal of the $t_1$ 180, needed for measurement of couplings, effectively doubles the number of observed peaks and therefore limits the size of proteins amenable to coupled HSQC schemes to less than 20 kD.

### *2.3.2 Inphase – Antiphase (IPAP) HSQC: decreasing spectral complexity*

An interesting solution to solving this overcrowding problem is to record the HSQC spectrum such that doublet components are stored separately during data acquisition. A

conceptually simple way of obtaining such a spectrum is illustrated in Figure 2.10. This method separates the $t_1$-coupled HSQC into two distinct sub-spectra; one utilizing the optional de-phasing delays at the beginning of $t_1$ (shown as open pulses), and one without. These experiments are collected in an interleaved manner, and results in an anti-phase and in-phase component of each HSQC doublet. It is important to note that the anti-phase spectrum is typically 10 percent weaker in intensity than its in-phase counterpart. This is due in part to the presence of the additional delay. A scaling factor is therefore used during data processing to account for this discrepancy. After scaling the anti-phase element, the difference of in-phase and anti-phase spectra only yields the downfield component, whereas the sum yields the up-field component. This approach is referred to as an in-phase, anti-phase (IPAP) HSQC *(5)* and is illustrated in Figure 2.11. This allows for extension to proteins up to 30 kD in size. For larger proteins, even the IPAP HSQC may show too much resonance overlap. It then becomes advantageous to obtain $^{1}D_{HN\text{-}N}$ couplings from 3-dimensional experiments, such as the HNCO, in which the $^{1}$H-decoupling pulse ($t_1$ 180) has similarly been removed from the $^{15}$N dimension *(52).*

### *2.3.3 Extraction of accurate dipolar contributions*

In the absence of line-shape or phase distortions, the accuracy of peak position in frequency resolved spectra is directly dependent on its signal-to-noise ratio and inversely dependent on its line-width *(53).* Once data are collected, however, the precision with which a coupling can be measured will depend highly on the procedure used for extracting the couplings. Numerous methods have been developed over the years in an attempt to maximize the experimental precision of measured couplings *(54-58).* Of notable importance to the

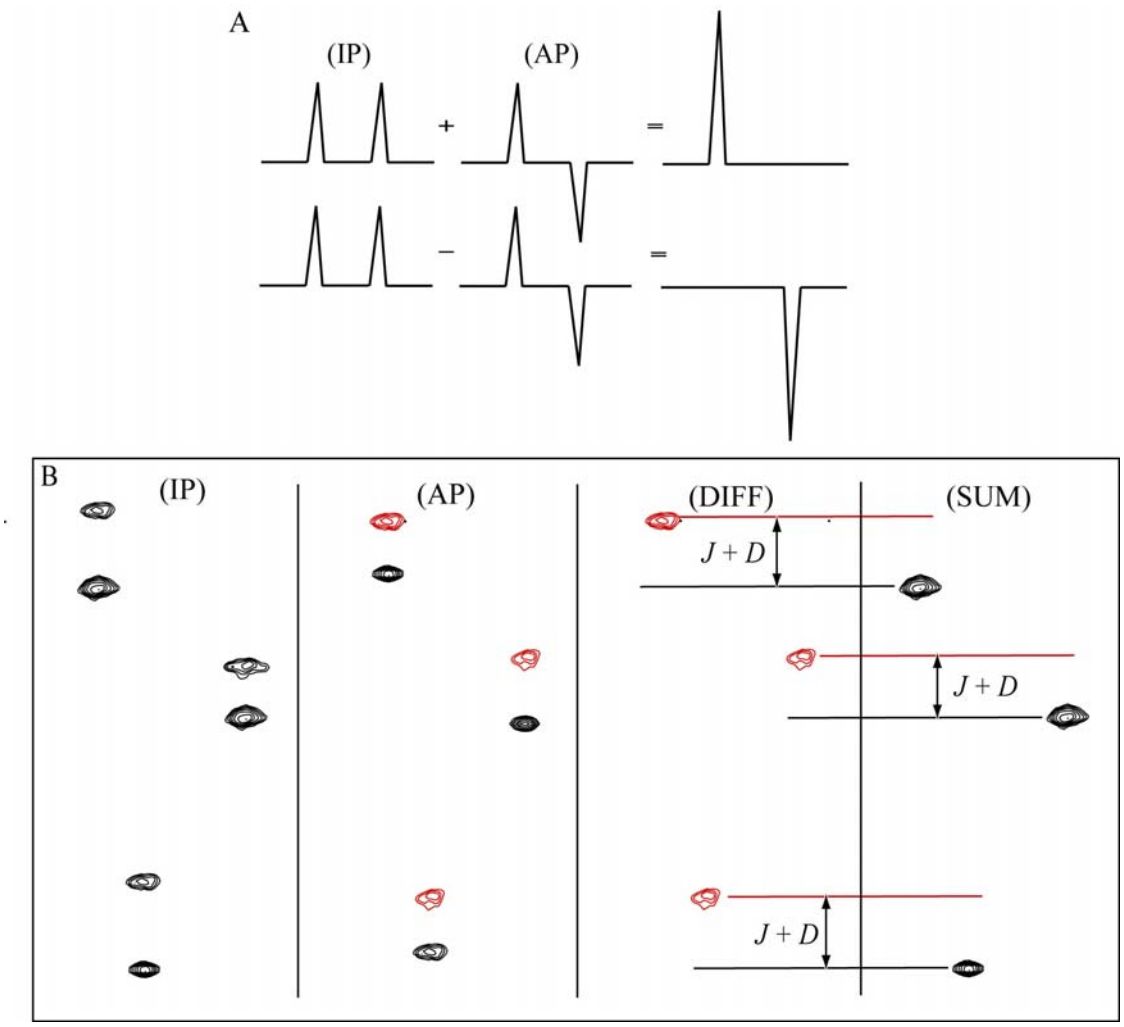


Figure 2.11. (A) Schematic representation of the Inphase (IP) and Antiphase (AP) method for separation of doublet components. The difference with the in-phase spectrum yields the downfield component, whereas the sum yields the up-field component. (B) Section of a $^{1}H$-$^{15}N$ IPAP HSQC showing the Inphase (IP), Antiphase (AP), Summation (SUM), and Difference (DIFF) spectra.

current work is the frequency-domain least squares module (NLinLS) found within NMRDraw *(59)*. NLinLS attempts to minimize the least-squares error between regions in the 2-dimensional target spectrum and the selected line-shape model by iteratively adjusting the model parameters. It also allows simultaneous modeling of several peaks in cases where resonance overlap exists. Importantly, confidence limits used in estimating experimental

error are derived from the standard deviations of the optimized model parameters. Although these methods are routinely employed to improve experimental precision, they remain highly dependent on the line-shape and line-width of the observed resonance. One could use alternative software packages, including PIPP *(55)* and XRAMBO *(58)*, in overcoming some of these obstacles. As a detailed description of each is beyond the scope of the present work, the reader is directed to the primary literature for further explanation.

## 2.4 Molecular Alignment

As mentioned previously, a prerequisite to measurement and analysis of residual tensorial interactions is that a suitable level of molecular order must first be achieved. The nature of this ordering is rather straightforward. Incorporation of most molecules into a magnetic field induces the formation of a magnetic dipole moment, with the strength of this induced dipole being proportional to its magnetic susceptibility for a particular direction. The magnetic moment will then interact with the external magnetic field and, if the molecule has an anisotropic susceptibility tensor, the orientation with the lowest interaction energy will be more populated. Macromolecular applications of direct field alignment are a simple extension of this principle and were therefore the first to be tested, as described previously *(28, 30)*.

The liquid crystal phases currently employed in generating the requisite level of order ($10^{-3}$) rely on an indirect interaction with the magnetic field. In this case the anisotropic distribution of electron density, and high magnetic susceptibility anisotropy, of the medium itself is used in a cooperative fashion to order the liquid crystal. This alignment is then imparted to dissolved solute particles by collisional or electrostatic mechanisms. It is

Table 2.2. Commonly used alignment media for the collection of dipolar couplings*

| Medium type | Director Orientation | Temp range (°C) | Properties | Ref |
|---|---|---|---|---|
| DMPC:DHPC | Perpendicular | 27-45 | can be doped to alter surface charge | *(26)* |
| DLPC:DHPC | Perpendicular | 7-50 | lower temperature of alignment | *(60)* |
| DIODPC:DIOHPC | Perpendicular | 10-55 | resistant to acid-base hydrolysis | *(61)* |
| DMPC:DHPC($Ln^{3+}$) | Parallel | 35-90 | $Ln^{3+}$ ion alters director orientation | *(62)* |
| Viral Particles | Parallel | 5-45 | wide temperature and concentration range, highly negatively charged | *(63)* |
| Cellulose Crystallites | Perpendicular | 37-?? | stable, moderately negatively charged | *(64)* |
| Cetylpyridinium halide/ n-hexanol/sodium halide | Parallel | 0-70 | sensitive to salt | *(65)* |
| Polyacrylamide gels | Mechanical | 5-45 | can be positive,negative, or neutral | *(66)* |
| n-alklyl-poly(ehtylene-glycol)/n-alkyl alcohol | Perpendicular | 0-40 | nonionic, stable | *(67)* |

*This table has been reproduced with permission from the primary authors *(68)*.

important to note that there is not yet a single, ideal liquid crystal and a host of solvent systems have recently been introduced, summarized in Table 2.2. In what follows, several of the most widely used systems are briefly described.

### *2.4.1 Bicelles*

Bicelles were the first liquid crystal media used for weak alignment of proteins and nucleic acids *(32, 35)*. They are most commonly composed of mixtures of

dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcoline (DHPC). In water, these lipid molecules aggregate spontaneously into bilayers, in which the more hydrophilic head groups shield the hydrophobic tails. Under certain conditions, the bilayers no longer form immobile extended sheets, but adopt a disc-shaped aggregate, termed bicelle, which can move freely in solution *(34, 69)*. These bicelles are thought to be circular in shape, several hundred angstroms in diameter (200-250Å) and 40Å thick *(69)*. DMPC is believed to constitute the bulk of the plane or surface with the shorter chain lipids protecting the edges. They tend to orient spontaneously in the presence of a strong external magnetic field and, in many cases, orient such that the bicelle normal, or director, is perpendicular to the field as shown in Figure 2.12.

At the low lipid concentrations used in most NMR applications, the distance between bicelle disks is thought to exceed 400Å *(35)*, allowing for only a small level of macromolecular order when compared to that of the liquid crystal. In this case, the rotational and translational mobility of the biomolecule is retained allowing for collection of high-resolution NMR spectra.

The bicelle system is tunable. At room temperature, this system remains isotropic, presumably composed of smaller, thicker and less anisotropic bicellar disks *(35)*. As temperatures rise above 30°C, the system switches from an isotropic phase to a nematic liquid crystalline phase *(14)*. This is important as an isotropic reference spectrum and its aligned counterpart can then be collected under near identical conditions. Moreover, the relative level of solute alignment increases as temperature increases allowing for optimization of molecular ordering prior to data acquisition.

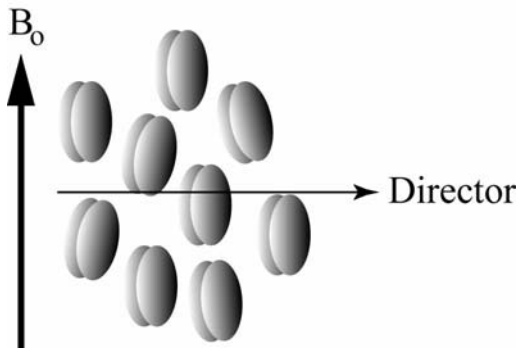


Figure 2.12. Cooperative alignment of bicelle disks along the external magnetic field. Steric interference between the disks causes them to stack on one another and allows the bicelle to orient with respect to the magnetic field as well as to each other. The average direction of alignment (Director) is perpendicular to the field ($B_o$).

It is important to note that the regular phospholipids used in bicelle preparations are subject to acid-base catalyzed hydrolysis of the ester linkage between the fatty acid chains and the glycerol backbone. The rate of hydrolysis is at a minimum at pH 6.5, but rapidly increases as the pH varies *(70)*. Use of ether linked lipids such as 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine (DIODPC) and 1,2-di-O-hexyl-sn-glycero-3-phosphocholine (DIOHPC) can prevent this hydrolysis. Samples such as these have proven stable at a range of pH values for a period of years *(61)*. Also, the transition temperature can be modified using alternate lipids. Substituting with lipids containing shorter alkyl chains, such as dilauroyl-phosphatidylcoline (DLPC), for DMPC in bicelle preparations decreases the transition temperature by as much as 12˚ *(60)*. However, this improvement in alignment temperature is not without its drawbacks. Samples prepared with shorter chain lipids tend to be less stable and less predictable in their behavior. As a result, DMPC/DHPC bicelles remain the most commonly used. Other innovations in bicelle design include doping with lanthanide ions causing bicelles to orient with their normal parallel to the magnetic field *(62)*.

### *2.4.2 Bacteriophage Pf1*

Soon after the introduction of bicelles, it was demonstrated that phage or viral particles could also form an effective alignment media. Prominent among tested particles is Bacteriophage Pf1, developed independently and simultaneously by Clore et al. *(71)* and Pardi and coworkers *(63)*. While it has risen in popularity in part due to its commercial availability, it is also remarkably robust and extraordinarily stable over a host of conditions, inlcuding a wide temperature range (5˚C - 45˚C). Bacteriophage Pf1 consists of a single stranded circular DNA genome packaged into a coat protein at a protein-nucleotide ratio of 1:1. This forms an extended rod approximately 20,000Å in length and 60Å in diameter. Pf1 has a tendency to orient with its long axis parallel to the external magnetic field and, due to its size, has the ability to achieve liquid crystallinity over a very wide range of phage concentrations. Owing to its nucleotide content, it is not surprising that Pf1 carries a substantial amount of net surface charge. Therefore, unlike the bicelles previously described, electrostatic forces dominate protein alignment. This makes Pf1 a perfect complement to bicelle preparations. However, due to its net charge, Pf1 is only suitable for use with negatively charged proteins.

### *2.4.3 Alkyl polyethylene glycol ethers*

The need for alternate alignment frames and media more compatible with protein samples has facilitated the design of alternative nonionic liquid crystals, specifically those composed of *n*-alkyl-poly (ethylene glycol)/*n*-alkyl alcohol mixtures *(67)*. These media are

uncharged, insensitive to ionic strength and pH, and can be used over a wide temperature range (0°C to 40°C). Different alkyl-poly (ethylene glycol) molecules are denoted as C*m*E*n*, where *m* is the number of carbons in the alkyl chain and *n* is the number of glycol units in the poly (ethylene glycol) moiety. Of particular interest to the current work is the reported phase diagram for $C_8E_5$/octanol in $H_20$. Under certain conditions, these systems form a liquid crystalline phase, where the hydrophobic alkyl chains aggregate into planar bilayers with the hydrophilic poly (ethylene glycol) head group facing the solvent. These bilayers order in a way reminiscent of bicelles, forming a Lamellar like structure. The dissolved macromolecule again occupies the inter-bilayer region. Unlike bicelles, however, these are believed to form extended sheets rather than discoidal structures *(67)*.

Upon incorporation into a magnetic field, these bilayers spontaneously align with their directors perpendicular to the field, presumably forming a super-structure in which the bilayers bend into a set of concentric tubes of different radii *(38)*. The addition of octanol is necessary to stabilize the liquid crystalline phase at the low surfactant concentrations needed for weak solute alignment. $C_8E_5$/octanol systems are now used routinely, in the concentration range of 3%-8%, to obtain residual dipolar couplings in proteins.

## 2.5 Application to Structure Determination

The application of residual dipolar couplings has grown exponentially in popularity since their introduction to the macromolecular scene in 1995. Residual dipolar couplings are now being used to address a number of different problems pertaining to structural biology,

ranging from structure validation and homology searches to refinement of predetermined structures to *de novo* structure determination. In what follows, I will give a brief overview of recent advances made in the area of protein structure determination using residual dipolar couplings.

### *2.5.1 Protein Structure Validation*

Preliminary three-dimensional atomic coordinates of atoms as provided by NMR or X-ray crystallography are often evaluated and improved through a refinement process. In crystallography, the progress of this refinement is monitored by cross-validation with a residual index known as the R-value *(72)*. This index is therefore the primary measure of the accuracy and precision of output structures. Numerous methods have been proposed for the incorporation of an NMR R-factor for cross-validating NOE-derived structures *(73-75)*, however none have become widely accepted. This is likely due to the semi-quantitative way in which NOE restraints are analyzed and an inability to back-calculate observables without the use of a number of simplifying assumptions.

Residual dipolar couplings do not carry the same limitation, and an exact value for couplings can be more easily back calculated from structural models. For NMR, a common validation parameter is the quality factor, *Q*, and is defined as *(76)*:

$$Q = \mathrm{rms}(D^{calc} - D^{obs})/\mathrm{rms}(D^{obs}), \qquad [10]$$

Where $D^{obs}$ and $D^{calc}$ are observed and calculated one-bond dipolar couplings respectively with the term *rms* referring to the root mean square. Ideally, dipolar couplings used in the

comparison are not used in the structure determination protocol, and an independent assessment of quality can be obtained.

The value *Q* is related to the Pearson's correlation coefficient, R, with R=0.9 equivalent to *Q*=42% and R=0.99 equivalent to *Q*=14% *(76)*. It is important to note that even for a perfect structure, *Q* will not approach zero. This is likely due to the variation in tensor orientations from site to site with respect to the reported average protein orientation, resulting in a lower limit for *Q* of approximately 10% *(77)*. This approach has been used in a number of cases in validating NMR derived structures post refinement *(78, 79)*.

### *2.5.2 Protein domain orientation*

It has been previously mentioned that the measurement of residual dipolar couplings can provide a complement to the more traditional NOE-based distance constraints; as they can restrict the relative orientation of proteins or protein domains regardless of their spatial separation. This is potentially important for multi-subunit or multi-domain proteins as NOEs across protein interfacial regions, found in protein-protein and protein-ligand complexes, can be sparse and difficult to observe with any measure of relative accuracy *(3, 46)*. X-ray crystallography has been the experimental method of choice in determining macromolecular complexes, however crystal-packing forces can affect relative domain orientation.

One of the earliest applications to the orientation of protein domains was in the case of a two-domain fragment from barley lectin, BLBC *(3)*. In this instance, the structures of the isolated domains were previously known. However due to the highly mobile unstructured region connecting the two domains, the complex structure remained in question. Orienting the

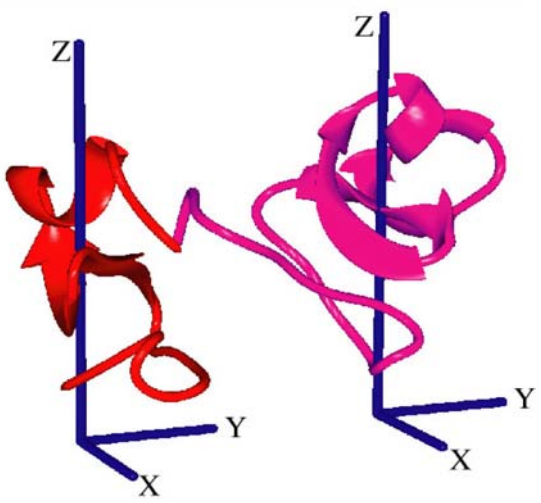


Figure 2.14. Domain orientation of B and C domains of Barley lectin, BLBC. $^{1}H$-$^{15}N$ dipolar couplings were used to solve for the order tensor for each domain. Orienting the principal alignment frames such that they coincide generated the depicted structure. This figure has been reproduced with permission from the primary authors *(3)*.

experimentally determined principal alignment frames from each protein domain such that they coincided generated a model for the protein structure, as illustrated in Figure 2.14. It is important to note that relative domain orientations cannot be uniquely determined from a single set of measured dipolar couplings, no matter how numerous *(3, 42, 80)*, and necessitates the use of multiple non-linearly dependent alignment frames in removing this degeneracy *(81)*.

A more functionally relevant application includes studies involving T4 Lysozyme in which domain closure is seen upon ligand binding *(82)*. Finally, and of notable importance, was the recent use of RDCs to restrict the domain orientations of hemoglobin *(83)*. In this instance, it was found that the structure observed in solution, determined by RDCs, was in agreement only with the average position of backbone atoms when comparing two previously determined crystal structures (known as the R and $R_2$ states), suggesting that the

physiologically relevant structure of hemoglobin is the dynamic average between existing crystal models.

### *2.5.3 Direct Structure Determination*

It has also been shown possible to build molecular models using residual dipolar couplings alone, even in cases where the initial structure is not known. In an early application, six residual dipolar couplings were measured per peptide plane from a sample of rubredoxin *(84)*. The phi and psi torsion angles connecting the peptide planes were iteratively modified and the subsequently predicted couplings were compared to experimental values, providing for the best peptide geometry. Additional peptides were added sequentially in cases where direct resonance connectivity could be made, resulting in the generation of 6 molecular fragments each separated by a single proline. Solving for the order matrix in each case and orienting the respective PAFs such that they coincided determined the relative fragment orientation, illustrated in Figure 2.15 (A). Any remaining orientational ambiguity in fragment geometry was removed upon incorporation of RDCs from a second alignment media. Covalent-bonding restrictions imposed by the connecting prolines provided the needed translational positioning. This procedure led to the generation of a backbone structure of rubredoxin (Figure 2.15(B)) that agreed with a homologous crystal structure to within 1.8Å *(84)*. Interestingly, sequence specific resonance assignment is not a pre-requisite to structure determination using this procedure. Sets of dipolar couplings involving particular resonances need simply be connected. This connectivity can be determined using a minimal number of triple resonance experiments when compared to standard protocols. Procedures such as these are currently being incorporated into structural genomics initiatives.

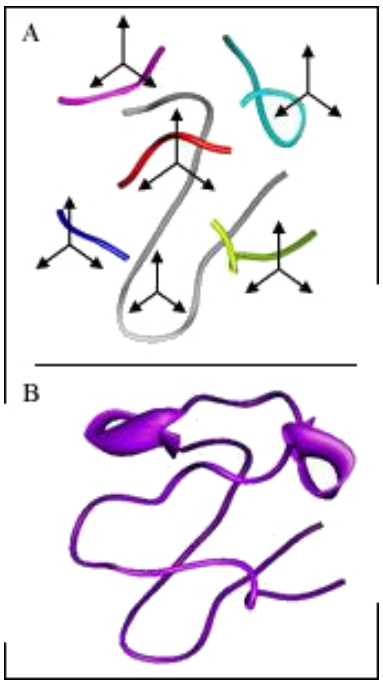


Figure 2.15. Direct protein structure determination using residual dipolar couplings. (A) In cases where direct resonance connectivity could be made, molecular fragments were generated, each separated by a single proline. Solving for the order matrix in each case and orienting the PAFs such that they coincided determined the relative fragment orientation shown. (B) Unique translational positioning was achieved through covalent-bonding restrictions generating the complete structure. This figure has been reproduced with permission from the primary authors *(84)*.

### *2.5.4 Protein Structure Refinement*

Common protocols used in three-dimensional refinement of macromolecules simultaneously attempt to preserve the covalent geometry of the protein and satisfy the constraints derived from NMR measurements. These protocols typically start from extended distance geometry structures but are equally amenable to proteins with known fold. The penalty function to be optimized is generally a sum of terms representing the energetics of the system and the experimental restraints. When starting from extended structures, standard procedures, such as conjugate gradient or rigid body minimization may be employed to

minimize these penalty functions; however, these typically get trapped in nearby local minima. It has therefore become more common to use molecular dynamics or simulated annealing type protocols to obtain more robust optimization.

A universal method for incorporating dipolar couplings into simulated annealing style structure determination has been developed for use with the XPLOR-NIH program *(85)*. In order to effectively incorporate RDC constraints into these refinement schemes, pseudo-atoms *X, Y, Z,* and *O* representing the tensor orientation in each alignment frame is first added to the starting coordinates. These pseudo-atoms are held rigid with respect to one another but allowed rotational degrees of freedom centered on *O* during refinement. An energy penalty function is then added for RDC data ($E_{SANI}$) that accounts for the difference between an observed coupling and one predicted if the orientation of the alignment tensor were given by the *OXYZ* axis system.

As the axis system is free to rotate, it aligns itself in a best-fit fashion to the principal alignment frame during initial rounds of simulated annealing. It is therefore advantageous to define the dipolar interaction in terms of this principal alignment frame. The resulting dipolar equation, as described above in section 2.2.2 (equation 9), is rewritten below. In this instance, the dipolar interaction can be summarized as:

$$D_{ij}^{RES} = D_a (3\cos^2\theta - 1) + 3/2\ R (\sin^2\theta \cos 2\phi), \qquad [9]$$

where $D_a$, defining the magnitude of anisotropy, is equal to $[1/2(S_{zz})( D_{ij}^{Static})]$ and the rhombicity (R) is equal to $2/3(\eta)$. Thus, the relationship of the measured couplings with its associated PAF, utilized in refinement schemes for the purpose of back-calculation, is defined

by the values $D_a$ and R. These must be experimentally determined and along with the weighting factor, *k* (as described below), are incorporated as coefficients in simulated annealing approaches.

For a dipolar coupling between a pair of atoms *i* and *j*, the energy penalty function, $E_{SANI}$, is defined as:

$$E_{SANI} = k(D_{ij}^{calc} - D_{ij}^{obs})^2, \qquad [11]$$

where $D_{ij}^{calc}$ and $D_{ij}^{obs}$ represent the calculated and observed dipolar coupling, respectively, and *k* is the aforementioned weighting factor. When $E_{SANI}$ is included in the simulated annealing protocol, the weighting factor is increased exponentially during the cooling stages, generally ranging from 0.0001 to 1 kcal/Hz$^2$ for N-H couplings *(38)*. It is important to note that the final value of *k* is chosen such that the correlation between measured and predicted couplings does not become tighter than experimental precision. Extensive cross-validation may be used to determine the appropriate value for *k (53)*.

Also, the energy landscape employing a term for dipolar couplings tend to have a very large number of sharp local minima. As a result, in the absence of distance constraints, simulated annealing processes are not efficient ways of finding a global minimum starting from an extended structure. These schemes are generally employed where the initial fold is determined through more traditional approaches, generating a model that is already close to the true structure using NOE data. They rely solely on the use of RDCs in late-stage refinement to increase the precision and accuracy of output structures. This approach has

been used extensively in the literature *(53, 79, 86)* and remains the most established use of RDCs to date.

In cases where structural homology is considered high and a suitable starting model exists, incorporation of RDC constraints as penalty functions in more standard energy minimization processes can also result in the generation of model structures from experimental data. This removes the previous dependency on extended distance geometry structures and is particularly useful in cases where one expects only minor conformational perturbations in response to processes such as ligand binding or protein-protein interaction. The macromolecule, in this instance, is not permitted excessive deviation from the starting model. As a result, this procedure should be amenable for model generation in the absence of distance constraints. The design and implementation of this type of refinement schedule is described in detail in the following chapter.

Above is just a sampling of current applications of residual dipolar couplings in biomolecular NMR, ranging from structure validation to direct structure determination. The use of these orientational parameters has seen an exponential growth in the past 10 years and new experimental techniques and applications are continually expanding. In the coming chapters, new approaches will be introduced and old ones exploited in determining a mechanistic model for protein activation, further illustrating the power of residual dipolar interactions in the area of protein structure determination.

**References**


1. Tolman, J.R. and J.H. Prestegard. (1996) Measurement of amide 15N-1H one-bond couplings in proteins using accordion heteronuclear-shift-correlation experiments. *J Magn Reson B*, **112**(3): p. 269-74.
2. Jain, N.U., S. Noble, and J.H. Prestegard. (2003) Structural characterization of a mannose binding protein-trimannoside complex using residual dipolar couplings. *J. Biol. NMR.*, **328**: p. 451-462.
3. Fischer, M.W., J.A. Losonczi, J.L. Weaver, and J.H. Prestegard. (1999) Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. *Biochemistry*, **38**(28): p. 9013-22.
4. Pardi, A., M. Billeter, and K. Wuthrich. (1984) Calibration of the Angular Dependence of the Amide Proton - C-alpha proton Coupling Constants, 3JHN-Alpha, in a globular protein - Use of 3JHN-alpha for Identification of Helical Secondary Structure. *J. Mol. Biol*, **180**(3): p. 741-751.
5. Wang, Y.X., J.L. Marquardt, P. Wingfield, S. Stahl, S. Lee-Huang, D. Torchia, and A. Bax. (1998) Simultaneous measurement of 1H-15N, 1H-13C, and 15N-13C dipollar couplings in a perdeuterated 30kDa protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, **120**: p. 7385-7386.
6. Bax, A., W. Vuister, S. Grzesiek, F. Delaglio, A.C. Wang, R. Tschudin, and G. Zhu. (1994) Measurement of homonuclear and heteronuclear J-couplings from quantitative J-correlation. *Meth. Enzym.*, **239**: p. 79-105.
7. Ding, K. and A.M. Gronenborn. (2003) Sensitivity enhanced 2D IPAP, TROSY-anti-TROSY, and E.COSY experiemtns: alternatives for measuring dipolar 15N-1HN couplings. *J. Mag. Res.*, **163**(2): p. 208-214.
8. Vander Kooi, C., E. Kupce, E.R.P. Zuiderweg, and M. Pellecchia. (1999) Line narrowing in spectra of proteins dissolved in a dilute liquid crystalline phase by band-selective adiabatic decoulpings: Application to HN-N residual dipolar coupling measurements. *J. Biomol. NMR*, **15**(4): p. 335-338.
9. Tjandra, N., S. Grzesiek, and A. Bax. (1996) Magnetic field dependence of nitrogen-proton J splittings in 15N-enriched human Ubiquitin resulting from relaxation interference and residual dipolar coupling. *J. Am.chem. Soc.*, **118**: p. 6264-6272.
10. Tolman, J.R. and J.H. Prestegard. (1996) A quantitative J-correlation experiment for the accurate measurement of one-bond amide 15N-1H couplings in proteins. *J Magn Reson B*, **112**(3): p. 245-52.
11. Yang, D.W., J.R. Tolman, N.K. Goto, and L.E. Kay. (1998) An HNCO-based pulse scheme for the measurement of13C-1H one bond dipolar couplings in 15N, 13C labeled proteins. *J. Biomol. NMR*, **12**(2): p. 325-332.
12. Saupe, A. and G. Englert. (1963) High-resolution Nuclear Magnetic Resonance Spectra of Oriented Molecules. *Phys. Rev. Lett.*, **11**(10): p. 462-464.
13. Zweckstetter, M. and A. Bax. (2001) Single-step determination of protein substructures using dipolar couplings: Aid to structural genomics. *J. Am.chem. Soc.*, **123**(38): p. 9490-9491.

14. Tjandra, N., J. Omichinski, A. Gronenborn, G. Clore, and A. Bax. (1997) Use of dipolar H-N and H-C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nature Structural Biology*, **4**(9): p. 732-738.
15. Emsley, J.W. and J.C. Lindon. (1975) The proton N.M.R. spectrum of partially oriented norbornadiene. Molecular structure and the effect of vibraitonal averaging. *Mol. Phys.*, **29**(2): p. 531-538.
16. Chou, J.J., F. Delaglio, and A. Bax. (2000) Measurement of one-bond 15N-13Co dipolar couplings in medium sized proteins. *J. Biomol. NMR*, **18**(2): p. 101-105.
17. Wienk, H.L., M.M. Martnez, G.N. Yalloway, J.M. Schmidt, C. Prez, H. Ruterjans, and F. Lohr. (2003) Simultaneous measrement of protein one-bond and two-bond nitrogen-carbon couling constants using an internally referenced quantitatve J-correlated [15N,1H] TROSY-HNC experiment. *J. Biomol. NMR*, **25**(2): p. 133-145.
18. Lohman, J.A.B. and C. MaClean. (1978) Alignment Effects on High Resolution NMR Spectra, Induced by the Magnetic Field. *Chem. Phys.*, **35**: p. 269-274.
19. Mittermaier, A. and L.E. Kay. (1999) Measurement of methyl 2H quadrupolar couplings in oriented proeins. How uniform is the quadrupolar couplng constant. *J. Am.chem. Soc.*, **121**(45): p. 10608-10613.
20. Tian, F., C.A. Fowler, E.R. Zartler, F.A. Jenney, Jr., M.W. Adams, and J.H. Prestegard. (2000) Direct measurement of 1H-1H dipolar couplings in proteins: a complement to traditional NOE measurements. *J Biomol NMR*, **18**(1): p. 23-31.
21. Bothner-By, A.A., P.J. Domaille, and C. Gayathri. (1981) Ultr-High field NMR Spectroscopy: Observation of Proton-Proton Dipolar Couplings in Paramagnetic Bis[tolyltris(pyrazoyl)borato]cobalt(II). *J. Am. Chem. Soc.*, **103**: p. 5602-5603.
22. Flygare, W.H. (1974) Magnetic-Interactions in Molecules and an Analysis of Molecular Electronic Charge Distribution from Magnetic Parameters. *Chem. Rev.*, **74**(6): p. 653-687.
23. Lisiki, M.A., P.K. Mishra, A.A. Bothner-By, and J.S. Lindsey. (1988) Solution Conformation of a Porphyrin-Quinone Cage Molecule Determined by Dipolar Magnetic Field Effects in Ultra-High-Field NMR. *J. Phys. Chem.*, **92**: p. 3400-3403.
24. Wang, Y.X., J.L. Marquardt, P. Wingfield, S.J. Stahl, S. Lee-Huang, D.A. Torchia, and A. Bax. (1998) Simulataneous measurement of 1H-15N, 1H-13C, and 15N-13C' dipolar couplings in a perdeuterated 30kDa protein dissolved in a dilute liquid crystalline phase. *J. Am.chem. Soc.*, **120**(29): p. 7385-7386.
25. Ottiger, M. and A. Bax. (1998) Determination of relative N-HN, N-C', Ca-Ha effective bond lengths in a protein by NMR in a dilute liquid crystalline phase. *J. Am.chem. Soc.*, **120**(47): p. 12334-12341.
26. Tjandra, N. and A. Bax. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science.*, **278**: p. 1111-1114.
27. Wu, Z. and A. Bax. (2002) Measurement of long-range H1-H1 dipolar couplings in weakly aligned systems. *J. Am.chem. Soc.*, **124**(33): p. 9672-9673.
28. Tolman, J.R., J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc Natl Acad Sci U S A*, **92**(20): p. 9279-83.

29. Meier, S., D. Haussinger, P. Jensen, M. Rogowski, and S. Grzesiek. (2003) High-accuracy residual 1H(N)-13C and 1H(N)-1H(N) dipolar couplings in perdeuterated proteins. *J. Am.chem. Soc.*, **125**(1): p. 44-45.
30. Tolman, J.R., J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard. (1997) NMR evidence for slow collective motions in cyanometmyoglobin. *Nat Struct Biol*, **4**(4): p. 292-7.
31. Tyrell, P.M. and J.H. Prestegard. (1986) Structural Studies of Carbohydrates by Deuterium NMR: Sucrose. *J. Am. Chem. Soc.*, **108**(14): p. 3990-3995.
32. Ram, P. and J.H. Prestegard. (1988) Magnetic field induced ordering of bile salt/phospholipid micelles: new media for NMR structural investigations. *Biochim Biophys Acta*, **940**(2): p. 289-94.
33. Sanders, C.R., 2nd and J.H. Prestegard. (1990) Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophys J*, **58**(2): p. 447-60.
34. Sanders, C.R. and J.P. Schwonek. (1992) Characterization of Magnetically Orientable Bilayers in Mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphasphatidylcholine by Solid-State NMR. *Biochemistry*, **31**: p. 8898-8905.
35. Bax, A. and N. Tjandra. (1997) High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium. *J. Biomol. NMR*, **10**: p. 289-292.
36. Levitt, M., *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. 2001, West Sussex, England: John Wiley and Sons.
37. Prestegard, J.H., H.M. Al-Hashimi, and J.R. Tolman. (2000) NMR Structures of biomolecules using field oriented media and residual dipolar couplings. *Q Rev Biophys*, **33**(4): p. 371-424.
38. Bax, A., G. Kontaxis, and N. Tjandra. (2001) Dipolar Couplings in Macromolecular Structure Determination. *Methods Enzymol*, **339**: p. 127-174.
39. Case, D.A. (1999) Calculations of NMR dipolar coupling strengths in model peptides. *J. Bio. NMR*, **15**: p. 95-102.
40. Weisstein, E.W., *Spherical Harmonic Addition Theorem*. 2000, Mathworld--A Wolfram Web Resource.
41. Saupe, A. (1964) *Z. Naturfors.*, **19a**: p. 161.
42. Losonczi, J.A., M. Andrec, M.W. Fischer, and J.H. Prestegard. (1999) Order matrix analysis of residual dipolar couplings using singular value decomposition. *J Magn Reson*, **138**(2): p. 334-42.
43. Sanders, C.R., 2nd and J.H. Prestegard. (1992) Headgroup orientations of alkyl glycosides at a lipid interface. *J. Am. Chem. Soc.*, **114**: p. 7096-7106.
44. Losonczi, J.A. and J.H. Prestegard. (1998) Nuclear magnetic resonance characterization of the myristoylated, N- terminal fragment of ADP-ribosylation factor 1 in a magnetically oriented membrane array. *Biochemistry*, **37**(2): p. 706-16.
45. Valafar, H. and J.H. Prestegard. (2004) REDCAT: A Residual Dipolar Coupling Analysis Tool. *J Magn Reson*, **167**: p. 228-241.
46. Prestegard, J.H., C.M. Bougault, and A.I. Kishore. (2004) Residual Dipolar Couplings in Structure Determination of Biomolecules. *Chem. Rev.*, **104**(8): p. 3519-3540.
47. Montelione, G.T. and G. Wagner. (1989) Accurate measurements of homonuclear HN-Ha coupling constants in polypeptides using heteronuclear 2D NMR experiments. *J. Am.chem. Soc.*, **111**: p. 5474-5475.

48. Vuister, G.W. and A. Bax. (1993) Quantitative J-correlation: a new approach for measuring homonuclear three bond J(HNHa) coupling constants in 15N-enriched proteins. *J. Am.chem. Soc.*, **115**: p. 7772-7777.
49. Bax, A., G.W. Vuister, S. Grzesiek, F. Delaglio, A.C. Wang, R. Tschudin, and G. Zhu. (1994) Measurement of homonuclear and heteronuclear J-couplings from quantitative J correlation. *Meth. Enzym.*, **239**: p. 79-105.
50. Bax, A., M. Ikura, L.E. Kay, D.A. Torchia, and R. Tschudin. (1990) Comparison of different modes of 2-dimensional reverse-correlation NMR for the study of proteins. *J. Mag. Res.*, **86**(2): p. 304-318.
51. Kay, L.E., P. Keifer, and T. Saarinen. (1992) Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Coherence Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.*, **114**: p. 10663-10665.
52. Yang, D.W., R.A. Venters, G.A. Meuller, W.Y. Choy, and L.E. Kay. (1999) TROSY-based HNCO pulse sequences for the measurement of (HN)-1H-15N-(Co)-13C, (HN) -1H -(Co)-13C, (Co)-13C-13Ca and (HN)-1H-13Ca dipolar couplings in 15N, 13C, and 2H labeled proteins. *J. Biomol. NMR*, **14**(4): p. 333-343.
53. Bax, A., G. Kontaxis, and N. Tjandra. (2001) Dipolar couplings in macromolecular structure determination. *Meth. Enzym.*, **339**: p. 127-174.
54. Bretthorst, G.L. (1990) Bayesian analysis. 3. applications to NMR signal detection, model selection, and parameter estimation. *J. Mag. Res.*, **88**: p. 571-595.
55. Garrett, D.S., R. Powers, A.M. Gronenborn, and G.M. Clore. (1991) A common sense approach to peak picking in two-, three-m and four- dimensional spectra using automatic computer analysis of contour diagrams. *J. Mag. Res.*, **95**: p. 214-220.
56. Miller, M.I., S.C. Chen, D.A. Keufler, and D.A. Davignon. (1993) Maximum-likelihood and the em algorithm for 2D NMR spectroscopy. *J. mag. Res. Ser. A*, **104**: p. 247-257.
57. Martin, Y.L. (1994) A global approach to accurate and automatic quantitative analysis of NMR spectra by complex least squares curve fitting. *J. mag. Res. Ser. A*, **111**: p. 1-10.
58. Andrec, M. and J.H. Prestegard. (1998) A Metropolis Monte Carlo implementation of bayesian time-domain parameter estimation: application to coupling constant estimation from antiphase multiplets. *J Magn Reson*, **130**(2): p. 217-32.
59. Delaglio, F., S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer, and A. Bax. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**: p. 277-293.
60. Wang, H., M.O. Eberstadt, E. T., R.P. Meadows, and S.W. Fesik. (1998) A liquid crystalline medium for measuring residual dipolar couplings over a wide range of temperatures. *J. Bio. NMR*, **12**: p. 443-446.
61. Cavagnero, S., H. Jane Dyson, and P.E. Wright. (1999) Improved low pH bicelle system for orienting macromolecules over a wide temperature range. *J. Bio. NMR*, **13**(4): p. 387-391.
62. Prosser, R.S., S.A. Hunt, J.A. DiNatale, and J.A. Vold. (1996) Magnetically aligned membrnae model systems with positive order parameters: switching the sign of Szz with paramagnetic ions. *J. Am.chem. Soc.*, **118**: p. 269-270.

63. Hansen, M.R., M. Rance, and A. Pardi. (1998) Observation of long range 1H-1H distances in solution by dipolar coupling interactions. *J. Am.chem. Soc.*, **120**(43): p. 11210-11211.
64. Fleming, K., D. Gray, S. Prasannan, and S. Matthews. (2000) Cellulose crystallites: a new and robust liquid crystalline medium for the measurement of residual dipolar couplings. *J. Am.chem. Soc.*, **122**: p. 5224-5225.
65. Prosser, R.S., J.A. Losonczi, and I.V. Shiyanovskaya. (1998) Use of a novel aqueous liquid crystalline medium for high-resolution NMR of macromolecules in solution. *J. Am.chem. Soc.*, **120**: p. 11010-11011.
66. Sass, H.J., G. Musco, S.J. Stahl, P. Wingfield, and S. Grzesiek. (2000) Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J. Biomol. NMR*, **18**: p. 303-309.
67. Ruckert, M. and G. Otting. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am.chem. Soc.*, **122**(32): p. 7793-7797.
68. Prestegard, J.H. and A.I. Kishore. (2001) Partial alignment of biomolecules: an aid to NMR characterization. *Curr Opin Chem Biol*, **5**(5): p. 584-90.
69. Hare, B.J., J.H. Prestegard, and D.M. Engelman. (1995) Small angle x-ray scattering studies of magnetically oriented lipid bilayers. *Biophys J*, **69**(5): p. 1891-6.
70. Ottiger, M. and A. Bax. (1998) Characterization of magnetically oriented phospholipid micelles for measurement of dipolar coulplings in macromolecules. *J. Bio. NMR*, **12**: p. 361-372.
71. Clore, G.M., M.R. Strarich, and A.M. Gronenborn. (1998) Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am.chem. Soc.*, **120**(40): p. 10571-10572.
72. Glusker, J.P., M. Lewis, and M. Rossi, *Crystal Structure Analysis for Chemists and Biologists.* First ed. Methods in Stereochemical Analysis, ed. A.P. Marchand. 1994: Wiley-VCH.
73. Gonzalez, C., J.A.C. Rullman, A. Bonvin, R. Boelens, and R. Kaptein. (1991) Towards an NMR R factor. *J. Mag. Res.*, **91**(659-664).
74. Withka, J.M., J. Srinavasan, and P.H. Bolton. (1992) Problems with, and alternatives to, the NMR R factor. *J. Mag. Res.*, **98**(611-617).
75. Brunger, A.T., G.M. Clore, A.M. Gronenborn, R. Saffrich, and M. Nilges. (1993) Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation. *Science.*, **261**: p. 328-331.
76. Cornilescu, G., J.L. Marquardt, M. Otiiger, and A. Bax. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute crystalline phase. *J Am Chem Soc*, **120**: p. 6836-6837.
77. Cornilescu, G. and A. Bax. (2000) Measurement of proton, nitrogen, and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline media. *J. Am.chem. Soc.*, **122**: p. 10143-10154.
78. Clore, G.M. and J. Kuszewski. (2003) Improving the accuracy of NMR derived structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am.chem. Soc.*, **125**(6): p. 1518-1525.

79. Clore, G.M. and C.D. Schweiters. (2004) How much backbone motion is required to account for dipolar coupling data measured iin multiple alignment media as assessed by independent cross-validation? *J. Am.chem. Soc.*, **126**(9): p. 2923-2938.
80. Bolon, P.J., H.M. Al-Hashimi, and J.H. Prestegard. (1999) Residual dipolar coupling derived orientational constraints on ligand geometry in a 53 kDa protein-ligand complex. *J Mol Biol*, **293**(1): p. 107-15.
81. Al-Hashimi, H.M., H. Valafar, M. Terrell, E.R. Zartler, M.K. Eidsness, and J.H. Prestegard. (2000) Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J Magn Reson*, **143**(2): p. 402-6.
82. Goto, N.K., N.R. Skrynnikov, F.W. Dahlquist, and L.E. Kay. (2001) What is the average conformation of bacteriophage T4 lysozyme in solution? a domain orientation study using dipolar couplings measured by solution NMR. *J. Mol. Biol.*, **308**(4): p. 745-764.
83. Lukin, J.A., G. Kontaxis, V. Simplaceanu, Y. Yuan, and A. Bax. (2003) Quarternary structure of hemolglobin in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **100**(2): p. 517-520.
84. Tian, F., H. Valafar, and J.H. Prestegard. (2001) A dipolar coupling based strategy for simultaneous resonance assignment and structure determination of protein backbones. *J Am Chem Soc*, **123**(47): p. 11791-6.
85. Schwieters, C.D., J.J. Kuszewski, N. Tjandra, and G.M. Clore. (2003) The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Res.*, **160**: p. 66-74.
86. Legler, P.M., M.L. Cai, A. Peterkofsky, and G.M. Clore. (2004) Three-Dimensional Structure of the cytoplasmic B domain of the mannitol transporter II Mannitol of the Escherichia coli phosphotransferase system. *J. Biol. Chem.*, **279**(37): p. 39115-39121.

# Chapter 3

## Conformational Changes in Human Arf1 on Nucleotide Exchange and Deletion of Membrane Binding Elements

Membrane binding of Arf, needed for Coatamer attachment and subsequent vesicle budding, is controlled by exposure of a myristoylated, N-terminal α-helix upon exchange of GDP for GTP. As previously described, the molecular pathway that leads to membrane anchoring and GDP dissociation within Arf most likely involves multiple steps, including [1] transient membrane docking, [2] conformational rearrangements at the SW1/Sec7 interfacial region, either prior to or concomitant with GEF association, and [3] nucleotide release. These steps have been seen in biochemical studies [BFA and membrane interactions *(1, 2)*] but most have remained elusive in crystal structure attempts. Existing information, however, provides evidence for the pivotal role of the N-terminal helix and attached acyl-chain pairing the nucleotide binding cycle to membrane interactions. We present here structural data relevant to this process.

Previous studies, comparing crystal structures for Arf1•GDP and Δ17Arf1•GTP, identified several regions of structural variation and suggested that these be associated with nucleotide exchange rather than removal of the N-terminal helix. However, separation of conformational changes due to nucleotide binding and N-terminal truncation cannot be addressed in comparing these structures, because both the bound nucleotide and the N-terminus differ. Resolving the two effects is important as any structural changes involving the

N-terminus may represent membrane-mediated conformational adjustments that precede GTP binding. In solution it is possible to examine a third form of Arf1, Δ17Arf1•GDP, which allows for separation of these associated changes by direct comparison to crystal structures differing in nucleotide (Δ17Arf1•GTP) and differing by truncation (Arf1•GDP). We have been able to assign backbone resonances using triple resonance assignment strategies *(3)*, and have subsequently carried out a structural examination using chemical shift perturbations of heteronuclear single quantum coherence (HSQC) spectra, changes in amide proton exchange rates, and residual dipolar couplings (RDCs), all collected from NMR data on a $^{15}N$ labeled form of the protein. In particular, measured residual dipolar couplings from Δ17Arf1•GDP, along with calculated values from the crystal structures of Arf1•GDP and Δ17Arf1•GTP, have made it possible to discern which changes are likely to be a direct result of nucleotide exchange as opposed to those derived from truncation.

In what follows we propose a solution structure of Δ17Arf1•GDP derived by incorporation of RDC associated penalty functions into an energy minimization protocol. We use this structure to assess regions of Δ17Arf1•GDP that deviate when compared to both the Arf1•GDP and Δ17Arf1•GTP crystal structures. The former comparison identifies structural differences that can be associated with N-terminal truncation, and the latter comparison identifies differences that can be associated with nucleotide exchange. The differences are supported by a limited set of NOE data in some critical aspects of the structure. Additional data showing differences in amide exchange rates between full length Arf1•GDP and Δ17Arf1•GDP also highlight subtle changes in the nucleotide binding pocket that may be responsible for the alteration in relative affinities for GDP and GTP on truncation.

## 3.1 Experimental Procedures

### *3.1.1 Sample Preparation*

BL21(DE3) cells were transformed with the expression plasmid for Δ17Arf1 (pOW14) or Arf1 (pOW12) respectively in media containing 60ug/ml ampicillin incubated at 37°C. Both sets of transformed cells were grown in 3ml LB/Amp to saturation, collected and placed in 50ml LB/Amp and grown to an $OD_{600}$ of 0.6. Cells were collected, resuspended and added to 1L MOPS/Amp media *(4)*, supplemented with 1x of GIBCO Vitamin mix (Cat# 11120-052), adding 1g/L $^{15}NH_4Cl$ and 2g/L $^{13}C_6$-Glucose (Cambridge Isotope Laboratories) (Δ17Arf1) or 4g/L unlabeled glucose (Arf1). Protein expression was initiated by the addition of IPTG to 0.5mM (at $OD_{600}$ of 1.0) and cells were incubated for an additional 3-5hrs. Cells were resuspended in 10ml TM buffer (20mM Tris, 2mM $MgCl_2$ pH7.6 at room temperature (RT)), broken using a French Press at 20-25k psi, and subsequently clarified by centrifugation (100k x g, 1hr).

Proteins were purified in two chromatographic steps. First using a 50ml Q resin column (BioRad, Q-macro, high substitution) equilibrated in TM buffer (20mM Tris, 2mM $MgCl_2$ pH7.6 at room temperature (RT)) and resolved with a 20% gradient of TMN buffer (containing 1M NaCl). Protein fractions were pooled and concentrated to 0.5ml and loaded onto a 60cm, 120ml Superdex75 column (Amersham) and eluted at 1.0ml/min with phosphate buffer (10mM $KPO_4$, 50mM NaCl, 1mM $MgCl_2$ and 5mM $NaN_3$ pH 7.0 at RT).

Nucleotide exchange on Δ17Arf1 was achieved by incubating the protein with a 50 or 10 fold molar excess of GDP or GTP (not an analog) respectively, including EDTA at 10mM and incubating the reaction either overnight at RT for GDP or for 60min at RT for GTP.

Exchange reactions were quenched by addition of $MgCl_2$ to 30mM. Excess nucleotide, EDTA and $MgCl_2$ were removed by chromatography on a Superdex75 column using phosphate buffer. Protein containing fractions were pooled and concentrated to a final concentration of 1mM.

Samples used for triple resonance assignment (Δ17Arf1•GDP) and HSQC-NOESY experiments (Δ17Arf1•GDP and Arf1•GDP), contained 0.5mM protein in potassium phosphate buffer (mentioned above) with 10 % $^2H_2O$ added to provide a lock signal. The samples for measurement of residual dipolar couplings were prepared as follows: Arf1•GDP was prepared to be 0.5mM protein in the aforementioned phosphate buffer with 10 % $^2H_2O$. The 7% (w/v) bicelle solution contained dimyristoyl-phosphatidylcholine (DMPC)/ dihexanoylphosphatidylcholine (DHPC) at a 2.8/1 molar ratio doped with dimyristoylphosphoglycerol (DMPG) (15:1 DMPC: DMPG). This gave a 13.8 Hz splitting of the $^2H$ NMR signal from 10% $^2H_2O$ in the sample at 305 K. The sample of Δ17Arf1•GDP was prepared to be 0.5mM protein, again in the aforementioned phosphate buffer. The 7.5% (w/v) ether bicelle solution contained 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine (DIODPC)/ 1,2-di-O-hexyl-sn-glycero-3-phosphocholine (DIOHPC) (Avanti Polar Lipids, Alabaster, AL) at a 3/1 molar ratio. This gave a 15 Hz splitting of the $^2H$ NMR signal from 10% $^2H_2O$ in the sample at 303 K. A similar sample using 4% pentaethylene glycol octylether ($C_8E_5$): octanol (Sigma-Aldrich Corp., St. Louis, MO), having a molar ratio (r) of 1.01, as an alignment medium was used for additional measurements of dipolar couplings on Δ17Arf1•GDP. The $C_8E_5$:octanol system yielded a 14 Hz splitting of the $^2H$ NMR signal from 10% $^2H_2O$ present in the sample buffer at 301 K.

Table 3.1. $^{31}$P Chemical Shift Values from Arf Samples†

| | α | β | γ | Pi |
|---|---|---|---|---|
| Δ17•GDP | -9.561 | -3.612 | * | 2.15 |
| Δ17•GDP -(Mg) | -9.63 | -4.34 | * | ** |
| Free GDP | -10.7 | -6.3 | * | 2.2 |
| Δ17•GTP | -9.4 | -15.1 | -4.8 | 2.12 |

† *, no gamma phosphate present; **, succinate buffer used, therefore no inorganic phosphate detected from buffer.

Due to the heightened affinity for nucleotide triphosphates in truncated Arf1 samples, an assessment of possible contamination from GTP incorporation into the target protein (Δ17Arf1•GDP) was performed. GDP and GTP have distinctive $^{31}$P chemical shift values, also reflecting the coordination of $Mg^{2+}$ when present, allowing for a direct detection of bound nucleotide by observation of phosphorous atoms in the nucleotide itself. The results show $Mg^{2+}$-coordinated GDP as the bound nucleotide with GDP and trace amounts of inorganic phosphate and GMP found in solution (Table 3.1). Comparing integrated peak intensities of $^{31}$P peaks observed in the spectra from GDP to internal standards (inorganic phosphate from sample buffer) indicates greater than 85% occupation of the nucleotide-binding pocket.

### *3.1.2 Removal of Common Bacterially Expressed Protein Contaminants*

It has been observed, and previously reported *(5)*, that E. *coli* strains commonly used for heterologous protein expression contain a polysaccharide termed cyclic enterobacterial

common antigen ($ECA_{CYC}$). More recently it has been reported, in collaboration with our group *(6)* that the N-acetyl signals of $ECA_{CYC}$ can be observed in $^{15}N$-$^{1}H$ HSQC spectra, and have arisen in various bacterially expressed protein samples. This may suggest that this water-soluble carbohydrate is a common contaminant. Given that the $^{15}N$-$^{1}H$ signals from the N-acetyl moiety of $ECA_{CYC}$ all fall within the spectral region associated with protein amides, the presence of this contaminant can account for unexpected increases in spectral complexity and resonance ambiguity. The characterization and removal of $ECA_{CYC}$ is briefly reported below.

$ECA_{CYC}$ was found in preparations of HIFd (MW 13.2 kDa, $pI_{CALC}$=6.2), two samples of GABPα, GABPα$^{138\text{-}254}$ (residues 138-254, MW 13.7 kDa, $pI_{CALC}$=4.9) and GABPα$^{168\text{-}254}$ (residues 168-254, MW 10.3 kDa, $pI_{CALC}$=5.2), along with Δ17Arf1•GDP (MW 18.8 kDa, $pI_{CALC}$=5.6). The purification schemes for all proteins followed closely with that described above for Δ17Arf1•GDP, including anion exchange and size exclusion chromatography. Unfortunately, this has proven inadequate in removing $ECA_{CYC}$ contamination; separation using anion exchange procedures are complicated by the charge of this carbohydrate, giving it a modest binding affinity for Q-type resins. Moreover, the heterogeneity in shape produces a broad elution profile and correspondingly poor separation using gel filtration chromatography.

Two methods have therefore been developed to eliminate undesirable $ECA_{CYC}$ contamination. The first of which is to eliminate biosynthesis of this carbohydrate, specifically using BL21 based mutant bacterial strains found deficient in ECA biosynthesis. Three such strains, containing mutations in the *wecA*, *wecF*, and *wecG* genes have been constructed. As an alternative to this strategy for strains already in common use among NMR laboratories, one can also reduce the amount of $ECA_{CYC}$ present by altering growth conditions

used for protein over-expression. Spectra of samples purified after 3, 7 or 14 hours post induction show that the amount of $ECA_{CYC}$ appears to roughly correlate with length of induction, although not linearly. Moreover, comparison of spectra obtained from samples expressed in either M9 or MOPS media show that sufficiently higher levels of $ECA_{CYC}$ are present in samples grown in M9. Decreasing induction length to less than or equal to 3 hours using MOPS media has allowed for adequate removal of $ECA_{CYC}$ contamination of bacterially expressed samples.

### *3.1.3 NMR Spectroscopy*

Spectra for the heteronuclear triple resonance assignment experiments, as well as the coupled IPAP $^{1}H$-$^{15}N$ HSQC experiments used to measure residual dipolar couplings, were collected on Varian 800MHz and 600MHz spectrometers respectively using Z-gradient triple resonance probes (Varian Inc, Palo Alto, CA). Heteronuclear 3D experiments (including HNCA, HNCACB, CBCACONH, HNCO, and HN(CO)CA) were run at 298 K using standard gradient assisted pulse sequences as supplied as part of the Varian protein pack. In addition a $^{15}N$ edited TOCSY experiment *(7, 8)* provided assignment of most alpha proton resonances.

$^{15}N$ Edited $^{1}H$-$^{1}H$ HSQC-NOESY experiments were collected on a Varian 600MHz spectrometer using a Z-gradient triple resonance probe (Varian Inc, Palo Alto, CA). Experiments were run at 298K using standard pulse sequences as supplied as part of the Varian protein pack. Data collection included 1024 $t_3$ points, 24 $t_2$ points and 64 $t_1$ points,

subsequently extended with linear prediction and zero filling to 2048, 64, and 128 points respectively.

The $^{31}P$ spectra used for analysis of nucleotide content were collected on a Mercury 300 MHz spectrometer, with a Nalorac 4 nucleus probe, using standard 1D experiments having proton decoupling only during acquisition. Relatively long recycle times of 2s were selected to minimize variations in peak intensities due to differential spin relaxation. $^{31}P$ chemical shifts were externally referenced to 85% phosphoric acid and internally referenced to 10mM inorganic phosphate from sample buffer (pH 7.0). The IPAP $^{1}H$-$^{15}N$ HSQC experiments used to measure RDCs were adapted from a previously described pulse sequence *(9)* and run under both isotropic (298 K) and partially aligned (303 K, 301 K) conditions on both ether bicelle and $C_8E_5$ samples. Data collection for the IPAP $^{1}H$-$^{15}N$ HSQC experiments typically included 1024 $t_2$ points and 256 $t_1$ points, subsequently extended with linear prediction and zero filling to 2048 and 1024 points respectively. Differences in the splittings from both aligned and isotropic conditions [303 K (ether bicelle) / 301 K ($C_8E_5$) and 298 K (both)] yielded residual dipolar couplings.

$^{1}H$-$^{15}N$ Cleanex HSQC experiments were used to identify amide protons that rapidly exchange with protons from water in both full length and Δ17Arf1•GDP samples. Experiments were adapted from a previously described pulse sequence *(10)* and conducted on samples made to be 0.5mM protein in the aforementioned potassium phosphate buffer. Data collection for the Cleanex $^{1}H$-$^{15}N$ HSQC experiments included 2048 $t_2$ points and 128 $t_1$ points subsequently extended with linear prediction and zero filling to 4096 and 512 points, respectively. Data processing, including extraction of residual dipolar couplings, was performed with the nmrPipe/nmrDraw package *(11)*, spectral figures were generated with the

NMRview graphics package *(12)*, and molecular graphics were generated using the Pymol Molecular Graphics package *(13)*

### *3.1.4 RDC assisted Chemical Shift Assignment of Arf samples*

Backbone sequential assignments were obtained for 90% of the residues using the following heteronuclear 3D spectra: HNCA, HNCACB, CBCACONH, HNCO, and HN(CO)CA. The small number of remaining residues were assigned with the aid of residual dipolar couplings, as published *(3)* and summarized below. For brevity, the assignment strategy for only Arf1•GDP is described; assignment of Δ17Arf1, loaded with either GDP or GTP, was a simple extension of this method.

Splittings were extracted for every set of peaks in the coupled $^{1}H$-$^{15}N$ HSQC experiment by visually determining peak centers using columns that ran through peak maxima. Differences in the splittings from both aligned and isotropic conditions (308 K and 298K) yielded residual dipolar couplings. Elements of the alignment tensor were found by analysis of all assigned peaks using the crystal structure *(14)* and a refined version of a previously published program *(15)*. Residual dipolar couplings, ranging from –31 to 32 Hz, were then back calculated for all $^{15}N$-$^{1}H$ vectors using this alignment tensor .

Dipolar couplings were measured for 5 unassigned peaks. In 2 cases, unique matches to predicted couplings were found. In 3 additional cases, prediction of chemical shift allowed an unambiguous assignment. For these cases backbone chemical shifts were calculated using the ShiftX program *(16)*.

### *3.1.5 RDC Analysis and Back-calculation*

Experimental dipolar couplings were initially analyzed using the REDCAT analysis package to allow an evaluation of a principal order frame and order parameters *(17)*. Upon obtaining the assignments of the $^{15}$N-$^{1}$H crosspeaks for Δ17Arf1•GDP, elements of the alignment tensor were found by analysis of couplings from 40 assigned peaks in the structurally conserved region. Principal order parameters using Arf1•GDP (or Δ17Arf1•GTP) crystal coordinates obtained from the above analysis are as follows for ether bicelles: $S'_{xx}$, $S'_{yy}$, $S'_{zz}$ = 0.00012, 0.00004, -0.00016 (0.00009, 0.00003, -0.000093); Euler angles relating the principal order frame to the molecular frame of the crystal structure are: α, β, γ = -83, 86, 50 (-56, 233, 34). The corresponding parameters for the $C_8E_5$ medium were found to be: $S'_{xx}$, $S'_{yy}$, $S'_{zz}$ = 0.00042, 0.00006, -0.00048 (0.00026, 0.00010, -0.00036) and α, β, γ = -44, 85, 119 (-29, 54, 122). Residual dipolar couplings were back calculated for all $^{15}$N-$^{1}$H vectors in both crystal structures (Arf1•GDP and Δ17Arf1•GTP) using alignment and derived order tensor elements. These couplings, which ranged from –6 to 6 Hz (ether bicelle) and –13 to 15 Hz ($C_8E_5$), were compared to experimental couplings to provide an initial evaluation of the similarity of Δ17Arf1•GDP to these structures.

### *3.1.6 Structure Refinement against Dipolar Restraints*

Refinement of existing crystal structures for Arf1•GDP, modified by removing N-Terminal helix coordinates, and Δ17Arf1•GTP was performed using the standard energy functions (bond, angle, improper, Van der Waals, electrostatic) and RDC constraint functions

(sani) in the XPLOR-NIH package *(18)*. Both starting structures included the bound nucleotide, the coordinated magnesium ion, and explicit water molecules. Refinement made use of a combination of torsion angle minimization (IVM Module *(19)*) and Powell minimization in Cartesian space as described in the literature *(20)*. The latter was used to allow full adjustment in position of all atoms, represented as small departures from idealized covalent geometry, in response to the RDC constraints as described below (Protocol Validation). There has been concern that undesired distortion of peptide plane geometry might occur in response to application of RDC constraints on H-N bond vectors using solely Cartesian methods, making torsion or rigid body minimization a preferred procedure. However, it is widely recognized that modest (≤ 6°) departures from ideal peptide plane geometry can be the result of the local environment within proteins *(21)* and suppression of these expected variations might force compensation in backbone torsion angles upon refinement. As this would lead to a greater degree of $\phi$ and $\psi$ angle variation, refinement first in torsion angle space followed by short rounds of Cartesian minimization was chosen. The force constants for bonds and angular components (angles and improper torsions) were set to 1000 kcal $mol^{-1}$ $Å^2$ and 500 kcal $mol^{-1}$ $rad^2$, respectively *(20)*, resulting in peptide bond distortions well within the expected range (average deviation from planarity of 1.7° and maximum deviation of 6°). Forces used with RDC constraints were adjusted in order to bring agreement between experimental and back-calculated RDCs only within experimental error.

Harmonic well potentials were then added for RDC data ($E_{SANI}$) along with values of DFS (Dynamic Frequency Shift), Da ($(1/2)S_{zz}$), and Rhombicity of 0.0 (0.0), -19.55 (-17.45), 0.51 (0.33), for $C_8E_5$ and ether bicelle () data respectively. Minimization progressed using the following schedule: Initial Powell minimization was first carried out in Cartesian space (no

dipolar restraints) until convergence. Powell minimization in torsion angle space was then implemented followed by short rounds of Cartesian space minimization with RDC force constants increasing linearly over 500 steps to their final values (2.0 [ether bicelle] and 1.0 [$C_8E_5$]) using 1000 rounds of minimization at each step. Finally, Cartesian minimization (500 steps; no RDC restraints) was carried out to allow for return to ideal covalent geometry and Powell minimization in torsion angle space, with RDC forces set to their final values for 750 rounds, was used to allow for RDC convergence without disrupting peptide plane geometry. A typical minimization schedule, shown in Appendix A, required 2 hours on a dual 2.8 GHz Intel Xenon system with 2Gbytes of RAM.

The degree of correlation between refined structures and back-calculated data was expressed as both the root mean square deviation (RMSD in Hz) of data from the best-fit correlation line and the Quality factor ($Q$) *(22)*, defined below.

$$Q = \mathrm{RMS}(D^{calc} - D^{obs})/\mathrm{RMS}(D^{obs}), \qquad [1]$$

Where $D^{obs}$ and $D^{calc}$ are observed and calculated one-bond dipolar couplings respectively with the term RMS referring to the root mean square.

The precision of the refinement scheme was examined by conducting ten separate calculations having left out separate 10% fractions of the dipolar data. This was done as an attempt to generate alternate minimization trajectories during refinement. All structures produced agree to within 1.0Å RMS deviation of $C_\alpha$ positions when compared to structures generated using all available dipolar data. These structures are being separately deposited

(PDB ID, 1U81). However, the small variation in overall structures adds confidence to observed deviations from the starting model to be discussed below.

### *3.1.7 Protocol Validation*

Residual Dipolar Couplings were simulated from three coordinate files [1HUR (Arf1), 1BQ8 (Rubredoxin), and 1D3Z (Ubiquitin)] obtained from the Brookhaven Protein Databank. The structures were allowed to relax using XPLOR based constant temperature (CT) dynamics simulations for 20 picoseconds at 500 K, 1000 K, 1500 K, and 2000 K (10 ps); which yielded structures with $C_\alpha$ atoms that had RMS deviations from starting models of 0.8Å, 1.2Å, 2.2Å, and 3.1Å, respectively. Energy minimization refinement was attempted on these model structures using solely Cartesian based, solely torsion based, or combined methods as previously described (Initial Minimization with no RDC constraints, Minimization with linearly increasing RDC forces, Final Minimization) using H-N residual dipolar couplings simulated from two independent alignment frames ($S_{zz}= 6.9e^{-4}$ and $\eta=0.31$ [RDC 1], $S_{zz}= -1e^{-3}$ and $\eta=0.19$ [RDC 2]).

The effectiveness of the refinement protocol was determined by comparing $C_\alpha$ RMS deviations and RDC correlations between starting (1HUR, 1D3Z, 1BQ8) and minimized models. Powell minimization relying solely on torsional degrees of freedom only returned backbone coordinates to within 1Å RMS deviation in models relaxed to 2.2Å (CT 1500K, 20ps) with 3.2 Hz ($Q$ = 21%) agreement of simulated RDCs. Deviations in backbone atoms are in this case largely the result of translational differences in secondary structural elements and loop regions. Powell Minimization in Cartesian space (with modifications to the

topology file converting dihedrals defining planarity to impropers) allowed for 2.3 Hz RDC correlation ($Q$ = 15%) and an ability to return models relaxed to 3.1Å (CT 2000K, 10ps) to within 1Å of the starting structure in all cases attempted. It is important to note that the major deviations in backbone positions observed in Cartesian minimization are found in loops, those that correspond to secondary structural elements show deviations less than 0.6Å. The translational deviations observed in torsion angle minimization are absent when minimizing in Cartesian space. It is likely that the smaller set of variables available in torsion angle minimization leads to a more restricted path over the energy landscape and more difficulty in returning to the initial structure. Using a combined torsion angle, Cartesian approach as described above allows for an improved level of convergence over both approaches based on Cartesian minimization and torsion minimization alone. This approach showed an ability to refine structures relaxed to 3.1Å to *less than* 1Å RMS deviation when compared to starting structures (1HUR, 1D3Z, 1BQ8) and allowed for 1.98 Hz ($Q$ = 13%) agreement between dipolar couplings. Hence, the combined approach was chosen for application to Arf1 structures.

## 3.2 Results

### *3.2.1 Qualitative Assessment of Conformational Change as a Result of Truncation*

If most structural alterations in the Arf1 molecule were dependent solely on nucleotide exchange we would anticipate the closest structural homolog of Δ17Arf1•GDP to be Arf1•GDP. We therefore begin our presentation with a qualitative comparison of these two molecules. As previously mentioned, the N-terminal helix of Arf1•GDP lies in a cleft

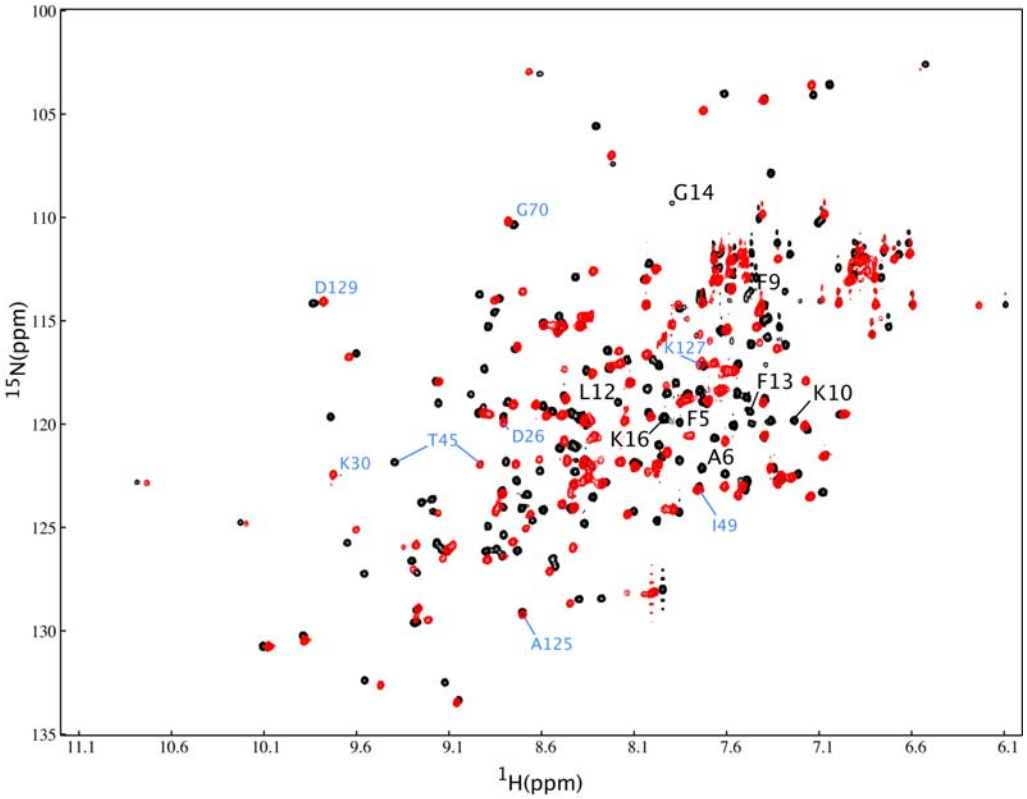


Figure 3.1. Overlay of the $^{1}$H, $^{15}$N-HSQC spectra of 0.5mM Arf1•GDP (black) on spectra of Δ17Arf1•GDP (red) both in 10mM phosphate buffer (pH 7.0). Spectra recorded at 298K with a $^{1}$H resonance frequency of 800 MHz. Residues of the N-terminal α helix found in Arf1•GDP not present in Δ17Arf1•GDP are labeled in black. Residues found within the nucleotide-binding pocket [P-loop (D26, K30), G3Motif (A125, K127, D129)] are labeled in light blue and considered representative peaks with little expected variation between models. Interestingly, those of SW1 (T45) show a large degree of chemical shift variance.

parallel to the C-terminal helix (αCt) located opposite of the nucleotide-binding site. One might expect that the N-terminal helix could be removed with minimal perturbation to most parts of the molecule, except for possible perturbations to the nearby αCt-helix and portions of the ISR found near the N-terminal helix, and that this structure, with the helix removed, would serve as a good model for Δ17Arf1•GDP (in which the nucleotide is unchanged).

Table 3.2 Chemical Shift Differences between Arf1•GDP and Δ17Arf1•GDP*

| | Chemical Shift Assignments (ppm) | | | | |
|---|---|---|---|---|---|
| | | Arf1•GDP | | Δ17Arf1•GDP | |
| | | $^{1}H$ | $^{15}N$ | $^{1}H$ | $^{15}N$ |
| Nt | Phe5 | 7.85 | 119.9 | *NP | *NP |
| Nt | Ala6 | 7.67 | 120.6 | *NP | *NP |
| Nt | Phe9 | 7.49 | 113.9 | *NP | *NP |
| **SW1** | **Val43** | **9.12** | **132.5** | ***NP** | ***NP** |
| **SW1** | **Thr44** | **8.90** | **126.1** | ***NP** | ***NP** |
| **SW1** | **Thr45** | **9.39** | **121.8** | **8.94** | **120.1** |
| **SW1** | **Ile46** | **8.42** | **122.3** | **8.34** | **120.8** |
| **ISR** | **Asn52** | **7.73** | **122.1** | **7.46** | **116.1** |
| **ISR** | **Val53** | **8.89** | **124.9** | **8.68** | **118.3** |
| **ISR** | **Glu54** | **8.81** | **126.4** | **8.29** | **120.4** |
| **ISR** | **Val56** | **8.61** | **122.2** | ***NP** | ***NP** |
| **ISR** | **Lys59** | **8.74** | **122.7** | **8.02** | **115.2** |
| **SW2** | **Gly70** | **8.75** | **110.3** | **8.77** | **108.8** |
| **SW2** | **Gln71** | **8.54** | **119.4** | **8.56** | **118.1** |
| Ct | Leu170 | 8.42 | 121.1 | 8.30 | 119.2 |
| Ct | Asn179 | 7.47 | 116.1 | 7.42 | 113.9 |
| Ct | Gln180 | 7.56 | 120.0 | 7.47 | 117.2 |

* Abbreviations used: Nt, N-terminal helix; Ct, C-terminal helix; SW1, Switch 1 region; SW2, Switch 2 region; ISR, Inter-Switch Region. Residues shown are representatives found in secondary structural regions having either no currently identified partner peak (labeled NP) or a significant degree of shift. Unexpected shifts are in bold.

Indeed, it has been reported that a crystal structure for Δ17Arf1•GDP had been determined with coordinates little changed from those of Arf1•GDP *(23)*.

$^{1}H$-$^{15}N$ Heteronuclear Single Quantum Coherence (HSQC) spectra are often used as a fingerprint for structural characteristics of soluble proteins. Positions of cross-peaks are sensitive to a combination of amino acid type and local structural alterations *(24, 25)*. If

structure was preserved between Arf1•GDP and Δ17Arf1•GDP except for N-terminal helix deletion, HSQC spectra should superimpose with primary differences being the absence of peaks for the N-terminal 17 residues and shifting of peaks for residues belonging to the C-terminal helix. Overlaying the $^{1}H$-$^{15}N$ HSQC spectra for these molecules shows differences that are far more extensive (Figure 3.1). Analysis of the differences that do exist requires peak assignment for both spectra. Assignments for Arf1•GDP have been reported previously *(3)*. Assignments for Δ17Arf1•GDP were carried out using triple resonance methods as described in experimental procedures. These assignments are reported in Appendix B and have been deposited in the Protein data bank (1U81). Representative cross peaks for Arf1•GDP that have shifted or lack obvious counterparts in Δ17Arf1•GDP are summarized in Table 3.2. Residues found within the nucleotide-binding pocket [P-loop (D26, K30) and the G3Motif (A125, K127, D129)] are also labeled and considered representative peaks with little expected variation between models. Peaks that exhibit significant shifts are in fact distributed throughout the structure (illustrated in Figure 3.1 and summarized in Table 3.2).

### *3.2.2 Qualitative Assessment of Conformational Change as a Result of Nucleotide Exchange*

Given these differences, we must also examine the possibility that most differences seen between Arf1•GDP and Δ17Arf1•GTP crystal structures are actually triggered by truncation and that the Δ17Arf1•GTP crystal structure may be a better model for Δ17Arf1•GDP. This would result in a greatly expanded list of peaks that deviate in comparing HSQC spectra of Δ17Arf1•GDP and Arf1•GDP. Peaks might include residues near SW1, SW2, and the interswitch region, as well as residues in contact with the nucleotide-binding site. As discussed previously, conformations in these regions deviate significantly

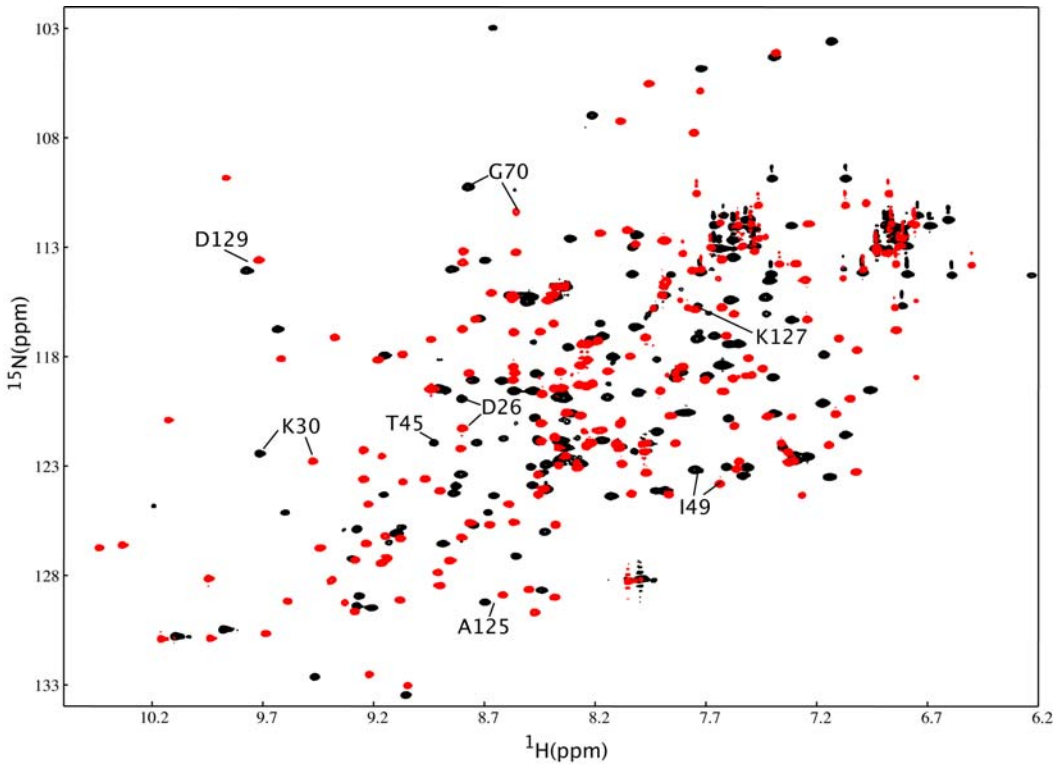


Figure 3.2. Overlay of the $^{1}$H, $^{15}$N-HSQC spectra of 0.5mM Δ17Arf1•GTP (black) on spectrum of Δ17Arf1•GDP (red) each in 10mM phosphate, pH 7.0. Spectra were recorded at 298K with a $^{1}$H resonance frequency of 800 MHz. Residues found within the nucleotide-binding pocket [P-loop (D26, K30), G3Motif (A125, K127, D129)] are labeled in black and shown as representative peaks with a higher degree of expected variation between models. Those of both SW1 (T45, I49) and SW2 (G70) also show a large degree of chemical shift deviation.

between structures. Several residues in the suggested regions do shift as labeled on the HSQC overlay of Figure 3.1 and in the list of Table 3.2. However an equal number do not, specifically members of the phosphate-binding loop and a few members of the guanine base binding motif labeled in Figure 3.2 [P-loop (residues 26-30) and G3Motif (residues 126, 127, 129), respectively].

If one were to propose that Δ17Arf1•GTP might provide a better structural homolog to Δ17Arf1•GDP, one might also expect their HSQC spectra to superimpose with possible

perturbations within residues found solely in the nucleotide-binding pocket. However, upon overlaying HSQC spectra obtained for Δ17Arf1•GDP and Δ17Arf1•GTP (Figure 3.2), variations in chemical shifts are far more numerous than the number that might be expected to surround the binding pocket, specifically those of both SW1 (T45, I49) and SW2 (G70) also show a large degree of chemical shift deviation. This suggests that neither of the two crystal structures are a good model for Δ17Arf1•GDP, allowing for the possibility that Δ17Arf1•GDP is a structural intermediate between the two states (basal and activated).

### *3.2.3 Residual Dipolar Coupling Comparison and Structure Refinement*

A more direct measure of structural change lies in residual dipolar coupling measurements. Here we compare experimentally measured N-H RDCs to predictions based on directions of N-H bond vectors seen in the crystal structures of Δ17Arf1•GTP and Arf1•GDP. Calculation of predicted couplings for comparison to structural models requires an initial evaluation of order parameters from experimental data. This in turn requires a conserved structural segment within molecules to be identified.

Comparing crystal structures for Arf1•GDP and Δ17Arf1•GTP reveals residues 90 through 165 as structurally conserved. Major structural differences are restricted to β strands 1 through 3, including both switch regions and the connecting loops as described (Goldberg, 1998). We have used data from the structurally conserved segment to solve for the order matrix and used these order parameters to back-calculate other expected couplings.

Because there are always minor structural variations between proteins in solution and those in crystals, it is important to calibrate this back-calculation procedure on a system where structural homology is expected to be high. Comparing the measured residual dipolar

couplings for Arf1•GDP to those predicted from the crystal structure of the same molecule should set the expected level of agreement (baseline) between a well-defined structure in solution and its crystal counterpart. Using 37 dipolar coupling values in the most conserved region, an order tensor for Arf1•GDP in a bicelle medium was calculated. All couplings were then back calculated using these order parameters along with atomic coordinates from the full-length crystal structure *(14)*.

A correlation plot of experimental and back-calculated RDC data for Arf1•GDP is presented in Figure 3.3(A). This shows a strong correlation between measured and experimental values having a root mean square deviation of 5 Hz. This corresponds to deviations that are somewhat larger than experimental precision (2Hz), but the structural variations leading to this deviation can be quite small. An alteration of 10 degrees in the average orientation of an N-H vector, if near a direction of optimal sensitivity in a molecule ordered to the extent observed here, would produce a change in dipolar coupling of 5.2Hz. There are also a few outliers in the plot. These correspond to residues in loops (L25, A27 (P-loop); K129, N135 ($L_{\beta 5/\alpha D}$); Q180, K181 (C-terminus)) whose structure might be expected to deviate more between solution and crystal. The remaining deviations seen above 2 Hz therefore correspond to small adjustments in local N-H vector angles.

Refinement of the structure through an energy minimization regime consisting of Powell minimization in torsion angle space followed by short rounds of Powell minimization in Cartesian space (see methods; hereafter referred to as torsion angle minimization) can make these small adjustments without significant changes in backbone structure. Using an RDC force constant of 2, the RMSD between experimental and back-calculated couplings reduced to below the expected experimental precision (2Hz); the associated Quality factor (*Q*) for this

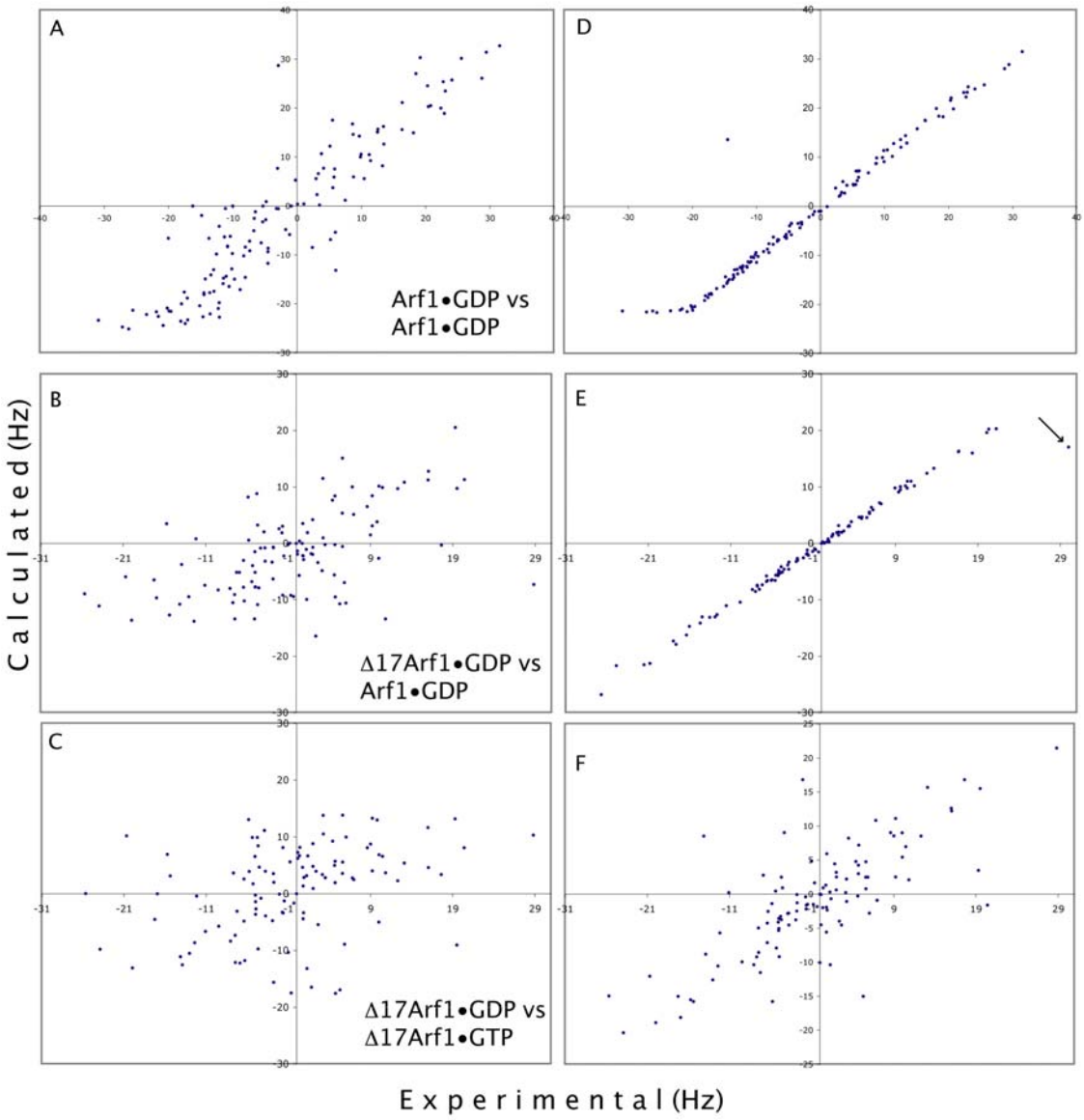


Figure 3.3. Correlation plot of experimentally measured $^{15}N$-$^{1}H$ dipolar couplings versus those calculated from unrefined crystal coordinates for A) Arf1•GDP(experimental) vs. Arf1•GDP(calculated), B) Δ17Arf1•GDP(experimental vs. Arf1•GDP(calculated), and C) Δ17Arf1•GDP(experimental) vs. Δ17Arf1•GTP(calculated)). In all cases, the X-axis denotes experimentally measured couplings and the Y-axis denotes calculated couplings from the specified crystal coordinates. The scatter seen in plots A/B/C indicate strong, modest, and poor levels of correlation between experimental couplings and those predicted from model structures having a resulting RMSD of 5, 9.7, and 12.3 Hz and associated Quality factors of 44%, 86% and 103%, respectively. These measured RDCs also show an ability to refine to convergence using torsion angle Energy Minimization when using Arf1•GDP as the starting model (D, E) having a resulting refined RMSD of 1.75 and 3.4 Hz respectively (*Q*= 13%, 33%). When using Δ17Arf1•GTP coordinates in refinement, the RDC correlation showed no improvement (F) with a corresponding RMSD of 13 Hz (*Q*=107%). The outlier denoted by an arrow is from residue 70, omitted from energy minimization refinement due to a high B-factor in the starting model.

refined structure is 13% (described in Chapter 2 and summarized in experimental procedures). The resulting correlation plot is shown in Figure 3.3(D). The backbone $C_\alpha$ atom positions of the resulting molecular structure have an RMSD from those in the starting structure of 0.8Å, largely due to movement of SW2 loop residues from the starting crystal structure (1HUR); these residues were previously determined to have a high B factor *(14)*. Removing these residues from the RMSD comparison reduces the corresponding $C_\alpha$ deviation to 0.5Å.

Comparison of experimental residual dipolar coupling data for Δ17Arf1•GDP to coupling data predicted from the Arf1•GDP structure can now be undertaken. This should give an experimental view of changes induced by truncation without the complication of simultaneous nucleotide exchange. In this case data from two different media were available, greatly improving an ability to assess structural differences. Elements of the alignment tensor in these two media, ether bicelles and $C_8E_5$:octanol ($C_8E_5$), were found by using residual dipolar coupling data for 40 assigned peaks in the structurally conserved region and the Arf1•GDP crystal coordinates (all measured RDCs are given in Appendix C). It is important to note that the overall range of dipolar couplings found for Δ17Arf1•GDP is smaller than that of Arf1•GDP due to a smaller degree of alignment in the liquid crystal media (-6 – 6 Hz in ether bicelle media and –13 – 15 Hz in $C_8E_5$ media as compared to-30 – 31 Hz for Arf1•GDP in bicelle media). Therefore, in what follows, couplings of Δ17Arf1•GDP have been scaled to match the range of dipolar couplings originally observed for Arf1•GDP to facilitate direct comparison.

A correlation plot of experimentally measured values for Δ17Arf1•GDP in ether bicelle and $C_8E_5$ media versus predicted values based on the Arf1•GDP crystal structure is

shown in Figure 3.3(B). This shows a modest correlation between measured and predicted values having scaled RMSDs of 9.9 Hz and 9.7 Hz respectively (*Q*=86%). However, the agreement is significantly less than in the Arf1•GDP solution to crystal comparison given above. Again, torsion angle refinement of the Arf1•GDP coordinates against RDC constraints using a force constant adjusted to compensate for the change in RDC range can impose a greater degree of correlation between dipolar values and, in fact, reach a value within experimental precision [RMSD of 3.4 Hz (*Q*= 33%), Figure 3.3(E)]. However, in this instance backbone torsion angles were significantly modified as a result of refinement. The RMSD of $C_\alpha$ backbone atom positions relative to the starting structure was 2.5Å.

Comparison of experimental RDC data for Δ17Arf1•GDP to coupling data predicted from the structure of Δ17Arf1•GTP should give a representation of changes induced by nucleotide exchange. A correlation plot of experimentally measured versus predicted values is shown in Figure 3.3(C). This shows a very poor correlation between measured and predicted values, having an overall scaled RMSD of 12.3 Hz (*Q*=103%). In principal a refinement of this structure (Δ17Arf1•GTP) should lead to the same refined structure as that obtained starting from Arf1•GDP. Refinement of the Δ17Arf1•GTP coordinates against RDC constraints using the same force constant given above, in this instance, however, does not improve the correlation between dipolar values [Figure 3.3(F), RMSD 13.0 Hz (*Q*= 107%)]. This type of behavior is not unexpected in a minimization routine when conversion between structures would involve substantial energy barriers. Simulated annealing protocols *(18, 26)* offer a possibility for overcoming these barriers under circumstances where both translational and orientational degrees of freedom are restricted, but with only N-H RDC restraints these methods prove unproductive *(27)*. We can, however, conclude that the structure of

Δ17Arf1•GDP must deviate more from the Δ17Arf1•GTP crystal structure than from that of Arf1•GDP.

### *3.2.4 Additional information relating truncated and non-truncated Arf1 structures*

NOE data can confirm some of the similarity between the Δ17Arf1•GDP structure in solution and the Arf1•GDP crystal structure. $^{15}$N edited HSQC-NOESY experiments involving Δ17Arf1•GDP and Arf1•GDP were collected using a mixing time of 150 ms on samples similar to those used for the RDC experiments except for the absence of alignment media. These data are rather limited because only NOEs involving amide protons are detected, and cross peaks only to assigned resonances along the backbone can be identified. However, assigned NOEs in some critical regions do occur. In particular, the amide proton from methionine 22 (β1) in Δ17Arf1•GDP shows a weak NOE connectivity with the α proton from valine 64 (β3). This bridges beta-strands involved in the interswitch hydrogen bond network. The same contact is seen in Arf1•GDP, suggesting this portion of the interswitch remains intact (data not shown).

Another set of data easily accessed through assigned $^{15}$N-$^{1}$H HSQC spectra is amide proton exchange rates. A commonly used model assumes that the amide group is in rapid equilibrium between occluded and exposed forms with base catalyzed exchange of the amide proton for a water proton in the open form being the rate-limiting step *(28)*. Under these conditions, increases in rates of exchange reflect primarily destabilization of occluded forms and a shift in equilibrium toward conformations that expose amide protons to solvent. Normally rates are measured by monitoring decreases in HSQC peak intensities as amide protons are replaced by deuterium from deuterated water added to the sample. When rates of

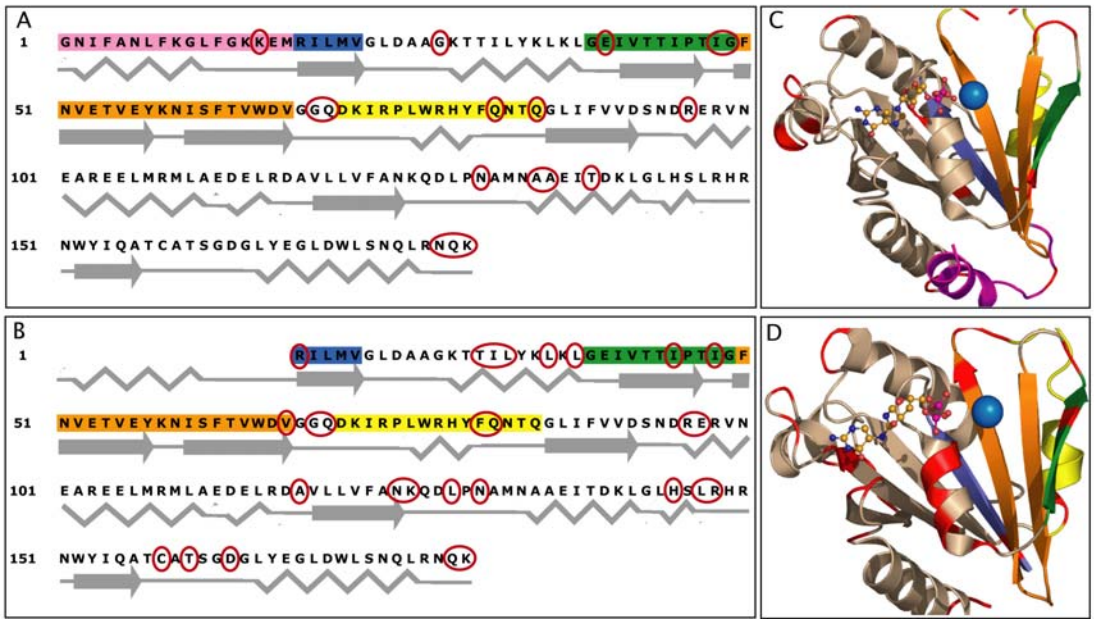


Figure 3.4. $^{1}$H-$^{15}$N Cleanex HSQC comparisons, showing areas of rapid exchange between water and amide protons of Arf1•GDP (A) and Δ17Arf1•GDP (B). Secondary structural elements are shown as a cartoon beneath one letter amino acid code for the protein sequence. Resonances evident in a $^{1}$H-$^{15}$N Cleanex HSQC experiment are circled in red. To relate these regions of enhanced exchange to a three-dimensional model, structures of Arf1•GDP (C) and Δ17Arf1•GDP (D) are shown as a cartoon representation color-coded as follows: N-terminal α helix (magenta), Switch 1 (green), Inter-Switch Beta Strands (orange), and Switch 2 (yellow); with Cleanex peaks represented as red areas in the respective 3-D models.

interest are fast, however, an equivalent experiment can be conducted in which water protons are labeled by saturation or inversion of magnetization. Differences between spectra with water magnetization perturbed and not perturbed highlight cross peaks belonging to amides experiencing rapid proton exchange. The CLEANEX HSQC is an experiment of this class that has been designed to minimize artifacts from inadvertent perturbation of alpha protons underlying the water resonance *(10)*. Here we present data on Arf1 proteins from the CLEANEX experiment.

Results for CLEANEX HSQC data for Arf1•GDP and Δ17Arf1•GDP are presented in Figures 3.4(A) and (B), respectively. There are clearly more peaks in the Δ17Arf1•GDP case indicating an increase in exposed residues or destabilization of secondary structure elements that protect amide protons from exchange in Arf1•GDP. The amides showing rapid exchange have been mapped onto the sequences of the two proteins in Figures 3.4(C) and (D). As expected, the majority of rapidly exchanging amides are in loops between secondary structure elements in both proteins. Major differences occur in the loss of some rapidly exchanging elements in the D helix of Arf1•GDP on truncation and the addition of rapidly exchanging elements in helix A, the Guanine binding Motif (G3Motif), and loop immediately preceding the C-terminal helix in Δ17Arf1•GDP.

### *3.2.5 Quantitative Comparison of Conformational Change as a result of Truncation*

The best way to compare changes in Arf1 structure that might be induced by N-terminal truncation, as opposed to nucleotide exchange, is on the basis of a three dimensional model of Δ17Arf1•GDP. Fortunately, torsion angle minimization of the Arf1•GDP crystal structure under RDC constraints collected on Δ17Arf1•GDP in solution did produce a modified structure. Torsion angle minimization schedules using 2 sets of dipolar restraints have shown an ability to converge RDCs ($Q$= 70%[start], $Q$= 13%[refined against simulated RDCs with no experimental error]) in structures having greater than 3Å backbone RMS deviation from starting models in test cases (see methods; Protocol Validation). These tests also show an ability to return Cα backbone atoms from their initial 3Å RMS deviations to *less than* 1Å RMS deviation from the starting model, with most remaining deviations found in

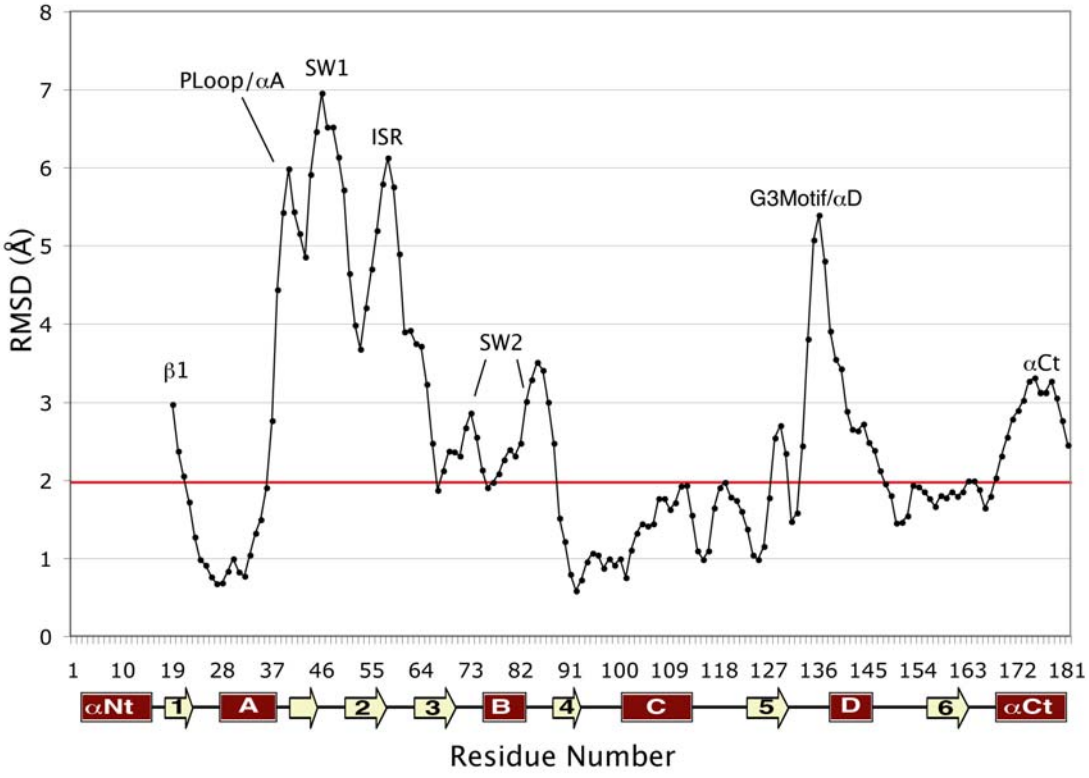


Figure 3.5. Root Mean Square Deviation in Angstroms seen upon overlaying the proposed Δ17Arf1•GDP solution structure with that of the starting model (Arf1•GDP), specifically overlaying β strands 4, 5, and 6 found predominantly unchanged in the protein core. This highlights global deviations throughout the whole of the protein, ranging from 0.6 to 7Å. Those deviations found significantly above these small global rearrangements (as denoted by a horizontal red line at 2.0Å) are labeled to correspond to regions shown in structure overlays of Figure 3.6. The x-Axis denotes residue number with corresponding cartoon representations of secondary structure elements and the y-axis depicts RMSD in angstroms.

loop regions. Given that RDCs back-calculated from N-H bond vector positions in our Δ17Arf1•GDP model agree with both sets of experimental data within estimated precision (RMSD 3.4Hz (*Q*= 33%), this must be considered a valid structure.

The structure appears to be of reasonable quality. Examination of the structure using Procheck *(29, 30)* indicates that 69% of ϕ and ψ backbone angles fall within the most favored region of Ramachandran space and an additional 29% fall within the acceptable region. There

are no unacceptable bond angles or bond lengths. Because of the use of RDC constraints on N-H vectors to drive conformational changes, it was essential to examine N-H bond geometry in more detail. The out-of-plane distortions, as measured by omega angles, are minimal (standard deviation of 1.7° and maximum deviation of 6°). These are comparable to deviations seen after minimization without RDC constraints (1.6° and 5.6°) and do not reflect unreasonable distortions due to RDC constraints. The same can be said for in-plane distortions as measured by HN-N-Ca bond angles (1.2° and 2° for the RDC constrained structure versus 1° and 1.8° for the unconstrained structure).

The model produced does not differ globally from the crystal structure of Arf1•GDP used as a starting point, as discussed below (2.5Å RMSD for backbone alpha carbons). However, there are distinct regions that do differ. These are easily identified by looking at RMS Deviations as a function of residue number for Cα positions in superimposed structures for Δ17Arf1•GDP and the Arf1•GDP crystal structure. It is readily apparent that Cα positions differ not only in and around the switch regions, but also throughout several other regions of the protein (Figure 3.5). Among the regions showing the greatest degree of variation (RMSD > 2Å) are the switch regions, the interswitch strands, the C-terminal helix, and the now N-terminal β-strand (β1). Surprisingly, residues involved in the nucleotide-binding pocket, specifically the nucleotide binding helix (αA), P-loop, and G3 binding motif also show a large degree of Cα deviation. The model itself will be presented in the discussion below.

### 3.3 Discussion

Much like the original starting model (Arf1•GDP), the Δ17Arf1•GDP structure produced by refinement against dipolar couplings is composed of a mixed parallel/antiparallel 6 β strand core surrounded by 5 α helices and 12 connecting loops. However, SW1 now encompasses a much-diminished 7th antiparallel β strand at the periphery of the core. As a direct result of removal of the N-terminal helix and connecting loop region, a multitude of small but global modifications have also occurred. Where these differences in structure are quite large (2.0 – 7Å), they can be readily detected in superimposed structures (illustrated in Figure 3.6).

Most notably, the nucleotide binding helix (αA) has adjusted its tilt allowing for a concerted shift in the coordinated magnesium ion (Figure 3.6A). SW1, losing most of its extended sheet conformation, has both changed its pitch and its spatial location; propagating a spatial displacement within the interswitch β strands, specifically residues 50-53(β3) and 65-68(β4), without a corresponding change in register (Figure 3.6B). The nucleotide-binding pocket, shown in Figure 3.6(C), also shows a modest degree of structural variation. Specifically, displacement of P-loop residues causes closer contacts with the ribose of GDP, forcing a disengagement of the guanine base from specific contacts in the G3Motif. Lastly, the C-terminal helix, owing to a greater degree of conformational freedom, has now shifted its tilt towards the N-terminal binding cleft (Figure 3.6D).

Some of the conformational adjustments indicated above can be further validated using amide exchange data from the $^{1}H$-$^{15}N$ Cleanex HSQC experiment. In comparing the results shown in Figure 3.4(B/D), the segments containing the largest number of rapidly

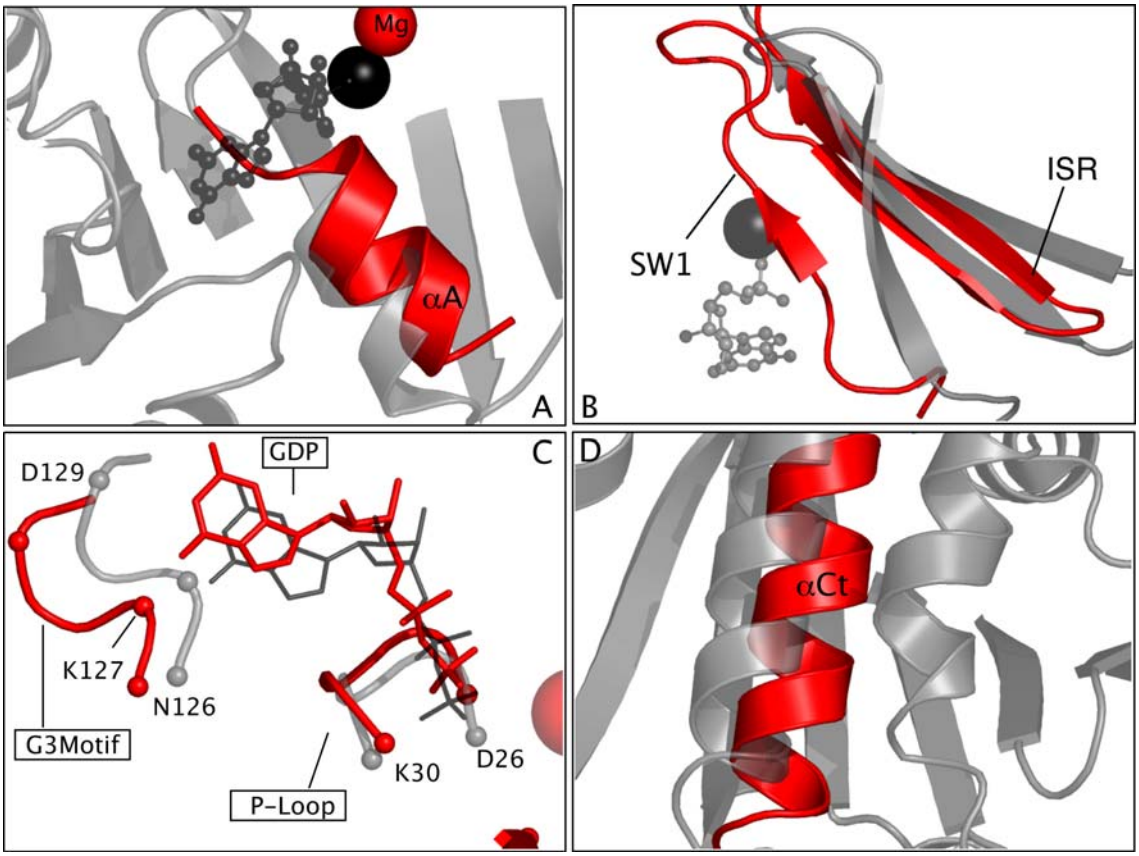


Figure 3.6. Superimposition of the proposed Δ17Arf1•GDP solution structure with Arf1•GDP using core β strands 4, 5, and 6 as anchor points. In all cases darker colors (grey, black) indicate Arf1•GDP with red being representative of the proposed Δ17Arf1•GDP model. Panel A shows a change in tilt of the nucleotide binding helix (αA) allowing for a concerted shift in the coordinated magnesium ion. SW1 (Panel B), losing most of its extended sheet conformation, has changed its pitch and spatial location, propagating a displacement within the interswitch β strands without a correlated change in register. The nucleotide-binding pocket, shown in Panel C shows displacement of P-loop residues causing closer contacts with the ribose of GDP and forcing a disengagement of the guanine base from specific contacts in the G3Motif. Residues directly involved in nucleotide coordination are labeled by amino acid type and residue number with respective Cα positions denoted by spheres. The C-terminal helix (Panel D) has shifted towards the N-terminal binding cleft as a result of N-terminal truncation.

exchanging residues are relegated to the secondary structural element αA which appears to be destabilized (shown in Figure 3.6A), the loop near the C-terminus (158-163) which is also more exposed or destabilized after truncation, and the members of the G3Motif (N126 and

K127), which have changed their association with the guanine base. It is interesting that all of these segments make direct contact with GDP or its associated $Mg^{2+}$ ion in the Arf1•GDP crystal structure.

It has been suggested that all conformational differences between the basal and activated crystal structures result from the binding of GTP and not from truncation of Arf1. In support of this, Goldberg reports, "the crystal structure of Δ17Arf1•GDP at 2.0 Å resolution closely resembles Arf1•GDP" *(23)*. This is largely true in the sense that secondary structural elements seen in Arf1•GDP are preserved in our solution structure of Δ17Arf1•GDP and the difference in positions of common backbone atoms in the two structures have an overall RMSD of only 2.5Å. Also, there is a much more significant deviation from the Δ17Arf1•GTP crystal structure; in fact, the description of structural changes in SW1, SW2 and interswitch regions on comparing our solution structure of Δ17Arf1•GDP to the crystal structure of Δ17Arf1•GTP would remain qualitatively the same as that already in the literature.

However, there are significant shifts of some of the elements and apparently energetically significant rearrangements of residues near the nucleotide-binding site on comparing our truncated Δ17Arf1•GDP structure to full length Arf1•GDP. These shifts and rearrangements are apparently the source of the difference in chemical shift patterns seen in simple HSQC experiments when comparing Arf1•GDP to Δ17Arf1•GDP (Figure 3.1). Moreover, they appear to be responsible for small adjustments in N-H vector directions that show up in residual dipolar coupling deviations distributed throughout the structure.

The structure produced by refinement against the dipolar couplings cannot in general be viewed as adopting some of the changes seen on comparing Δ17Arf1•GTP to Arf1•GDP.

Although, the splaying and partial unraveling of the beta strand of SW1 and the changes in the ISR strands might be seen as movements along the path to complete unraveling of the SW1 beta strand seen in the crystal structure of Δ17Arf1•GTP. The Δ17Arf1•GDP solution structure we have produced is more appropriately viewed as a structural intermediate between the two crystal states.

The functional significance of our observations, including the destabilization observed for residues in contact with GDP in the binding site of Δ17Arf1, as compared to Arf1, and the shift in the position of the $Mg^{2+}$ ion, may be connected with the altered nucleotide binding affinities of the truncated protein. It has been shown that truncated Δ17Arf1 has a higher affinity for GTP while full length Arf1 has a higher affinity for GDP. To the extent that truncation of the N-terminal helix mimics events that may happen on membrane association, these observations may be relevant to the effect of membrane association on nucleotide affinity. Despite early models that have Arf1 associating with membranes only in the GTP form, there is really little experimental evidence to support this; in fact more recent models have begun to include association of the GDP bound form *(31, 32)*. If this is correct, the above data may indicate that membrane association actually promotes exchange of GDP for GTP. Association of guanine nucleotide exchange factors (GEFs) at the membrane surface may also be promoted by some of the changes seen. These factors are known to associate with residues at the surface of SW1. In our model, the extra strand provided by SW1 has moved in space potentially opening the protein to allow for subsequent association with its corresponding GEF. Hence, a model for structural reorganization upon contacting the membrane surface, prior to GEF association and nucleotide exchange, is supported.

The current findings, while significant, hinge on the fact that the N-terminal helix must be released or destabilized on membrane interaction to allow for the level of conformational change observed, having been mimicked in the current work through N-terminal truncation. This concept is not without foundation, as Arf1 is known to interact with its corresponding GEF in a GDP bound state on the membrane surface only after the N-terminal helix has been disrupted. However, what may cause the needed destabilization of the Arf1 N-terminus? These and other issues are at the heart of discussions presented in the following chapters.

**References**


1. Paris, S., S. Beraud-Dufor, S. Robineau, J. Bigay, B. Antonny, M. Chabre, and P. Chardin. (1997) Role of Protein-Phospholipid Interactions in the Activation of ARF1 by the Guanine Nucleotide Exchange Factor Arno. *J Biol Chem*, **272**(35): p. 22221-22226.
2. Renault, L., P. Christova, P. Guibert, S. Pasqualato, and J. Cherfils. (2002) Mechanism of Domain Closure of Sec7 Domains and role in BFA sensitivity. *Biochemistry*, **41**(11): p. 3605-3612.
3. Amor, J., R. Seidel, F. Tian, R. Kahn, and J. Prestegard. (2002) Letter to the Editor: H-1, N-15 and C-13 assignments of full length human ADP Ribosylation Factor 1 (ARF1) using triple resonance connectivities and dipolar couplings. *J Biomol NMR*, **23**(3): p. 253-254.
4. Neidhart, F.C., P.L. Bloch, and D.F. Smith. (1974) Culture Medium for Enterobacteria. *J Bact*, **119**: p. 736.
5. Erbel, P.J., K. Barr, N. Gao, G.J. Gerewig, P.D. Rick, and K.H. Gardner. (2003) Identification and Biosynthesis of Cyclic Enterobacterial Common Antigen in Escherichia Coli. *J. Bacteriol.*, **185**: p. 1995-2004.
6. Erbel, P.J.A., R.D. Seidel III, S.E. Macintosh, L.N. Gentile, J.C. Amore, R.A. Kahn, J.H. Prestegard, L.P. McIntosh, and K.H. Gardner. (2004) Cyclic enterobacterial common antigen: Potential contaminant of bacterially expressed protein preperations. *J. Biomol. NMR.*, **29**: p. 199-204.
7. Griesinger, C., G. Otting, K. Wuthrich, and R. Ernst. (1988) Clean TOCSY for H1 Spin System-Identification in Macromolecules. *J. Am. Chem. Soc.*, **110**(23): p. 7870-7872.
8. Fesik, S. (1988) Isotope-Edited NMR Spectroscopy. *Nature*, **332**(6167): p. 865-866.
9. Wang, Y.X., J.L. Marquardt, P. Wingfield, S. Stahl, S. Lee-Huang, D. Torchia, and A. Bax. (1998) Simultaneous measurement of 1H-15N, 1H-13C, and 15N-13C dipollar couplings in a perdeuterated 30kDa protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, **120**: p. 7385-7386.
10. Hwang, T.-L., S. Mori, A.J. Shaka, and P.C.M.v. Zijl. (1997) Application of Phase Modulated CLEAN Chemical EXchange Spectroscopy (CLEANEX-PM) to Detect Water-Protein Proton Exchange and Intermolecular NOEs. *J. Am. Chem. Soc.*, **119**: p. 6203-6204.
11. Delaglio, F., S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer, and A. Bax. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**: p. 277-293.
12. Johnson, B. and R. Blevins. (1994) NMRView - A Computer-Program for the Visualizaation and Analysis of NMR Data. *J. Biomol. NMR*, **4**: p. 603-614.
13. DeLano, W.L., *The PyMol Molecular Graphics System*. 2002, DeLano Scientific: San Carlos.
14. Amor, J.C., D.H. Harrison, R.A. Kahn, and D. Ringe. (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature*, **372**(6507): p. 704-8.
15. Losonczi, J.A., M. Andrec, M.W. Fischer, and J.H. Prestegard. (1999) Order matrix analysis of residual dipolar couplings using singular value decomposition. *J Magn Reson*, **138**(2): p. 334-42.

16. Wishart, D.S. and D.A. Case. (2001) Use of chemical shifts in macromolecular structure determination. *Meth. Enzym. A.*, **338**: p. 3-34.
17. Valafar, H. and J.H. Prestegard. (2004) REDCAT: A Residual Dipolar Coupling Analysis Tool. *J Magn Reson*, **167**: p. 228-241.
18. Schwieters, C.D., J.J. Kuszewski, N. Tjandra, and G.M. Clore. (2003) The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Res.*, **160**: p. 66-74.
19. Schwieters, C.D. and G.M. Clore. (2001) Internal Coordinates for Molecular Dynamics and Minimization in Structure Determination and Refinement. *J Magn Reson*, **152**: p. 288-302.
20. Clore, G.M. and C.D. Schwieters. (2004) How Much Backbone Motion in Ubiquitin is Required To Account for Dipolar Couplng Data Measured in Multiple Alignment Media as Assessed by Independent Cross-Validation? *J Am Chem Soc*, **126**: p. 2923-2938.
21. MacArthur, M.W. and J.M. Thornton. (1996) Deviations from Planarity of the Peptide Bond in Peptides and Proteins. *Journal of Molecular Biology*, **264**(5): p. 1180-1195.
22. Cornilescu, G., J.L. Marquardt, M. Otiiger, and A. Bax. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute crystalline phase. *J Am Chem Soc*, **120**: p. 6836-6837.
23. Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95**(2): p. 237-48.
24. Kay, L.E., P. Keifer, and T. Saarinen. (1992) Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Coherence Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.*, **114**: p. 10663-10665.
25. Stonehouse, J., G.L. Shaw, J. Keeler, and E.D. Laue. (1994) Minimizing Sensitivity Losses in Gradient-Selected N-15-H-1 HSQC Spectra of Proteins. *J. Magn. Res. A.*, **107**: p. 178-184.
26. Tjandra, N., J. Omichinski, A. Gronenborn, G. Clore, and A. Bax. (1997) Use of dipolar H-N and H-C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nature Structural Biology*, **4**(9): p. 732-738.
27. Chou, J.J., L. Shipeng, and A. Bax. (2000) Study of Conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. *JB NMR*, **18**: p. 217-227.
28. Dempsey, C.E. (2001) Hydrogen Exchange in peptides and proteins using NMR spectroscopy. *Prog. NMR Spec.*, **39**: p. 135-170.
29. Morris, A.L., M.W. MacArthur, E.G. Hutchinson, and J.M. Thornton. (1992) Sterochemical quality of protein structure coordinates. *Proteins*, **12**: p. 345-364.
30. Laskowski, R.A., M.W. MacArthur, D.S. Moss, and J.M. Thornton. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**: p. 283-291.
31. Franco, M., P. Chardin, M. Chabre, and S. Paris. (1996) Myristoylation-facilitated binding of the G protein ARF1GDP to membrane phospholipids is required for its activation by a soluble nucleotide exchange factor. *J Biol Chem*, **271**(3): p. 1573-8.
32. Renault, L., B. Guibert, and J. Cherfils. (2003) Structural Snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. *Nature*, **426**(6966): p. 525-530.

# Chapter 4
## Structural Perturbations in Human Arf1 Accompanying the Binding of Phosphatidylinositides

In performing their role in intracellular signaling pathways, specifically involving the regulation of membrane traffic and organelle morphology, Arfs interact with Arf-specific guanine nucleotide exchange factors (GEFs), GTPase activating proteins (GAPs), and effectors. Signaling phosphatidylinositides, most commonly phosphatidylinositol (4,5)-bisphosphate ($PI(4,5)P_2$) or phosphatidylinositol (3,4,5)-trisphosphate ($PI(3,4,5)P_3$), have been shown previously to regulate the activities of a number of these regulators and effectors of Arf. The ability of Arf itself to bind these same phosphatidylinositides has also been reported previously, though without much structural detail. For instance, it has been established that GTP incorporation is fundamental in membrane association and activation of Arfs. However, the extent of *in vitro* binding of GTP to Arfs is quite low (<10%) *(1)* when compared to other members of the Ras super-family, and this binding is highly dependent on a hydrophobic environment. The amphipathic N-terminus of Arf1, absent in other GTPases, acts as an inhibitory domain that limits GTP binding and confers this dependence on lipids, as previously described. In contrast, vesicles containing lipids, such as phosphatidylinositol (4,5)-bisphosphate ($PI(4,5)P_2$), not only retain the ability to bind to Arfs, but also promote the release of GDP and stabilize the apo form *(2, 3)*. What is more, the subsequent removal of

$PI(4,5)P_2$ allows GTP to bind at a higher stoichiometry than had been previously reported *(2)*. It may be that the binding of charged phospholipids represents an avenue for structural alteration directly upon membrane association.

We have therefore investigated the ability of non-myristoylated Arf1•GDP (Arf1•GDP) to bind soluble $PI(4,5)P_2$ analogs using $^{15}N$-$^{1}H$ Heteronuclear Single Quantum Coherence (HSQC) based NMR experiments. HSQC experiments are ideally suited for resolving local structural variation accompanying ligand binding due to their innate ability to resolve individual amino acid resonances and sense local alterations in their electronic environment. Cross peaks in these spectra appear at positions that correlate amide proton chemical shifts with amide nitrogen chemical shifts. Changes in local environment are then reflected in differences in cross peak positions between spectra of protein and protein-ligand complexes *(4, 5)*. In the current study, we have examined the ability of Arf1•GDP to bind to myo-inositol (1, 4, 5)-trisphosphate ($I(1,4,5)P_3$), the soluble head group found in phosphatidylinositol (4,5)-bisphosphate ($PI(4,5)P_2$), as well as a soluble, short acyl chain $PI(4,5)P_2$ analog, using HSQC based NMR experiments. By using the previously determined $^{15}N$-$^{1}H$ backbone assignments on Arf1•GDP *(6)*, selective changes in chemical shifts for certain resonances of these protein-ligand complexes were used both to map a putative binding pocket within Arf1•GDP and to measure relative binding affinities for some of the ligands.

Chemical shift changes are sometimes not confined to a single region of the protein. When this occurs, making a distinction between effects from ligand binding at one site, ligand binding at multiple sites, and structural changes induced at remotes sites is often difficult. However, distinction can be achieved by combining shift perturbation experiments with a

number of other NMR-based techniques. One class of experiment that can contribute to a distinction relies on the direct protection of surface residues from spin relaxation enhancement by a soluble paramagnetic agent *(7)*. Enhanced relaxation in the presence of such an agent is easily seen via a loss in intensity for HSQC peaks from unprotected residues. TEMPOL provides a neutral, soluble agent that will not be directly affected by electrostatic changes on binding of the highly charged ligands to be investigated. Monitoring changes in amide proton exchange rates upon ligand binding *(8)* can further substantiate the nature of changes induced directly at ligand binding sites or indirectly at remote sites. The more rapidly exchanging amides likely to be involved in these sites are easily monitored through experiments in which peak intensities depend on magnetization from water protons that exchange into the sites [CLEANEX experiments, *(9)*]. Altered rates of magnetization transfer can reflect changes in stability of structural elements at various sites as previously described (Chapter 3). Such stability changes may be particularly important in promoting nucleotide release and Arf-GEF associations needed for Arf activation.

Raising the possibility that phosphoinositide binding may induce more remote conformational changes is not without basis. It has been suggested, for example, that some of the conformational changes within Arf1 needed for activation and GTP binding occur prior to GDP release and upon interaction of Arf1 with its corresponding GEF on the membrane surface *(10)*. Moreover, as described in the previous Chapter, a few key structural changes within Arf1 are suggested to occur upon truncation/membrane association directly *(11)*. These membrane-induced changes could well be mediated by specific interactions with phosphoinositides. While several structures have been described for Arf bound to a GEF or GAP, there is currently no direct structural information on Arf binding to a lipid or a

membrane. We hope to provide a step toward resolving this deficiency with the studies of phosphoinositide binding that follows.

## 4.1 Experimental Procedures

### *4.1.1 Sample Preparation and Effector Selection*

Human Arf1 was expressed in BL21(DE3) as previously described *(12)*but with modifications to allow the incorporation of stable isotopes. Transformed cells were grown to saturation in 3ml LB/Amp, collected and placed in 50ml LB/Amp media and grown to an $OD_{600}$ of 0.6. Cells were then collected and placed into 1L MOPS/Amp media *(13)*, supplemented with 1x GIBCO Vitamin mix (Cat# 11120-052) and 1g/L $^{15}NH_4Cl$ and 4g/L glucose. At an $OD_{600}$ of 1.0, protein expression was induced by the addition of IPTG to 0.5mM. After 3-5 hrs, cells were collected, resuspended in 10ml TM buffer (20mM Tris, 2mM $MgCl_2$ pH7.6 at room temperature (RT)) and lysed using a French press. Cell lysate was clarified by centrifugation at 100k xg for 1hr.

Soluble Arf1 was purified in two chromatographic steps, as previously described *(14)*. A Q-macro resin column (BioRad, high substitution; 50 ml bed volume) was equilibrated in TM buffer (20mM Tris, 2mM $MgCl_2$ pH7.6 at RT) and proteins were resolved with a NaCl gradient. Fractions containing Arf1 were pooled, concentrated to 0.5ml and loaded onto a Superdex75 gel filtration column (Amersham; 60 cm x 1.6 cm) at a flow rate of 1.0ml/min in phosphate buffer (10mM $KPO_4$, 50mM NaCl, 1mM $MgCl_2$ and 5mM $NaN_3$ pH 7.0 at RT);

these conditions have been selected to allow for accurate transfer of $^{15}N$-$^{1}H$ assignments as previously collected *(6)*.

Samples for titration experiments were 0.5mM Arf1•GDP in the phosphate buffer with 10 % $^{2}H_2O$ added to provide a lock signal. Due to low ligand volumes, data collection proceeded using either a 300 µl sample in a 5mm (outer diameter) Shigemi tube (myoI, myoI-(1)P, myoI-(1,4)$P_2$, I(1,4,5)$P_3$) or 200 µl in a 3mm Shigemi tube (myoI-(4,5)$P_2$ and dibutyryl-PI(4,5)$P_2$ (0.3mM Arf1)). When solubility allowed, binding of effectors was monitored in 0.25 mM increments at concentrations ranging from 0 to 10 mM effector. NMR samples for separation of pH-induced shifts were 1.0mM Arf1•GDP in the same buffer with pH values ranging from pH 6.0 to 7.0 in 0.25 unit steps.

The choice of phosphoinositide analogs follows that described previously *(15)* and included the following ligands. D-myo-inositol 1, 4, 5 trisphosphate [I(1,4,5)$P_3$] (Sigma Aldrich, I 9766), the soluble head group for PI(4,5)$P_2$, was used as an easily accessible ligand with established affinity, and D-myo-inositol [myoI] (Sigma Aldrich, I 5125), containing only the inositol ring with no phosphates, were used as a negative control. Specificity was tested by varying the position and number of phosphates located on the inositol ring, as illustrated in Figure 4.1, including D-myo-inositol (1)-monophosphate [myoI-(1)P], D-myo-inositol (1,4)-bisphosphate [myoI-(1,4)$P_2$] and D-myo-inositol (4,5)-bisphosphate [myoI-(4,5)$P_2$](Sigma Aldrich, I 2523, I 0510, I 3264, respectively). A soluble PI(4,5)$P_2$ analog, dibutyryl-phosphatidylinositol (4,5)-bisphosphate [dibutyryl-PI(4,5)$P_2$] (Echelon Biosciences Inc., P 4504) was also used. However, due to its more limited solubility, dibutyryl-PI(4,5)$P_2$ titrations ranged only from 0-3.5mM.

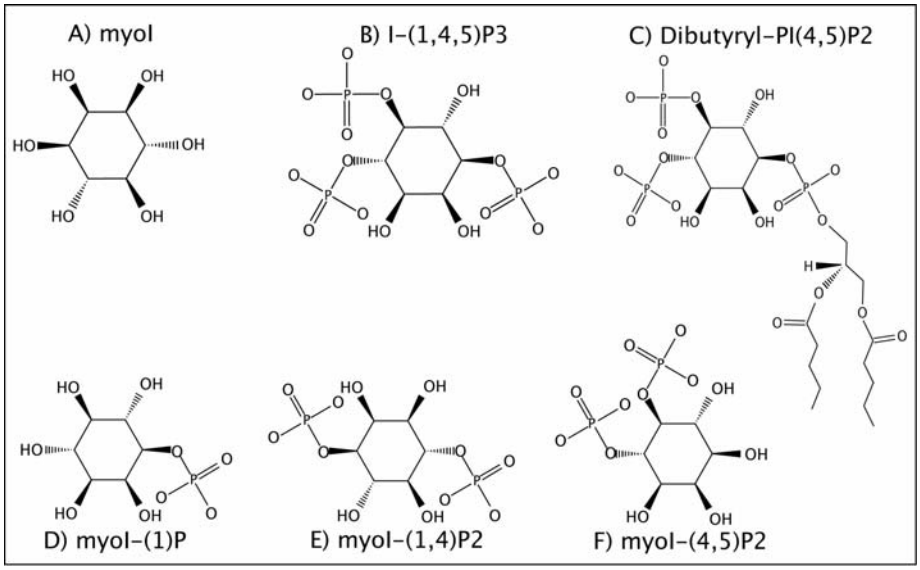


Figure 4.1. Ligands used in phosphoinositide binding studies with Arf1•GDP included (**A**) D-myo-inositol, containing only the inositol ring with no phosphates, (**B**) D-myo-inositol (1, 4, 5)-trisphosphate, the soluble head group for $PI(4,5)P_2$, (**C**) dibutyryl-$PI(4,5)P_2$, a short acyl-chain soluble mimic for $PI(4,5)P_2$, (**D**) D-myo-inositol (1)-monophosphate, (**E**) D-myo-inositol (1,4)-bisphosphate, and (**F**) D-myo-inositol(4,5)-bisphosphate.

### *4.1.2 NMR Experiments*

HSQC-based titration experiments for pH studies and determination of chemical shift perturbations with samples containing myoI, $I(1,4,5)P_3$ and dibutyryl-$PI(4,5)P_2$ were performed at 298K on a Varian 800MHz spectrometer (Varian Inc, Palo Alto, CA). Spectra for samples containing myoI-(1)P, myoI-$(1,4)P_2$ and myoI-$(4,5)P_2$ were collected on a Varian 600 MHz Spectrometer at 298K using a Z-gradient triple resonance cold probe for enhancement of signal to noise (Varian Inc, Palo Alto, CA). Average data collection for the $^{1}H$-$^{15}N$ HSQC experiments included 2048 $t_2$ points and 128 $t_1$ points, which were subsequently modified with linear prediction and zero filling to 4096 and 512 points

respectively. Each spectrum required approximately 1hr. All data processing was performed with the NMRpipe/NMRdraw and NMRview packages *(16, 17)*.

The $^{31}P$ spectra used for analysis of nucleotide content were collected on a Mercury 300 MHz spectrometer, with a Nalorac 4 nucleus probe, using standard 1D experiments having proton decoupling only during acquisition. Relatively long recycle times of 2s were selected to minimize variations in peak intensities due to differential spin relaxation. $^{31}P$ chemical shifts were externally referenced to 85% phosphoric acid and internally referenced to 10mM inorganic phosphate present in the sample buffer (pH 7.0).

$^{1}H$-$^{15}N$ Cleanex HSQC experiments were used to identify amide protons that rapidly exchange with protons from water in Arf1•GDP samples. Experiments were adapted from a previously described pulse sequence *(9)* and conducted on samples made at 0.5mM protein in phosphate buffer. Data collection for the Cleanex $^{1}H$-$^{15}N$ HSQC experiments included 2048 $t_2$ points and 128 $t_1$ points subsequently extended with linear prediction and zero filling to 4096 and 512 points, respectively. Each spectrum required approximately 16 hrs.

$^{1}H$-$^{15}N$ HSQC based surface accessibility experiments were run on Arf1•GDP (0.5mM) with and without the presence of the paramagnetic spin probe TEMPOL (50mM) (4-hydroxy-2,2,6,6-tetramethyl-piperidine-1-oxyl; Sigma Chemical Co.) to topologically map the solvent exposed molecular surface both with and without the putative ligand (dibutyryl-$PI(4,5)P_2$). Data collection for the $^{1}H$-$^{15}N$ HSQC experiments included 2048 $t_2$ points and 128 $t_1$ points subsequently extended with linear prediction and zero filling to 4096 and 512 points, respectively. Each spectrum required approximately 4 hrs.

### *4.1.3 Determination of Dissociation constants*

Quantitation of perturbations seen at each titration point was achieved using equation 1 below, which takes the difference in chemical shifts and scales them with respect to the relative ranges of the chemical shift variation observed at saturation. Dissociation constants were extracted using a Tcl/Tk based NMRview module, available upon request from Kevin Gardner (www.freedom7.swmed.edu).

$$\delta_{1H15N} = \frac{\sum (\delta_i)}{\delta_{i(max)}} \qquad [1]$$

$$\Delta\delta = \frac{\Delta\delta_{max} * \{(L + P_T + K_d) - [(L + P_T + K_d)^2 - (4*L*P_T)]^{1/2}\}}{[2*P_T]} \qquad [2]$$

The extraction of $K_d$ values is based on fitting theoretical curves, generated using equation 2, where the value L is ligand concentration, $P_T$ is the total protein concentration, $K_d$ is the equilibrium dissociation constant, $\Delta\delta$ is the combined chemical shift change ($\delta_{1H15N}$) at a particular ligand concentration, and $\Delta\delta_{max}$ is the chemical shift change at saturation. In order to improve the quality of fit, chemical shift values obtained from 9 residues showing a substantial degree of chemical shift variation as a result of ligand addition (F5, A6, K10, L12, F13, G14, K16, R178, K181) were scaled by chemical shift deviations seen at saturation and combined to produce a single titration curve. This presumes that all shifts arise from a single binding event, and to support this assumption binding curves for individual residues showing large shift deviations from different regions of the protein were examined. They showed no significant deviations from the averaged curve. Moreover, the curves agree with a stoichiometry of 1:1 indicating the absence of multiple sites with similar binding constants.

Ideally, the presence of ligand is the only contributor to chemical shift changes. However, some of the residues monitored also showed slight alterations in chemical shift

owing to small changes in pH as a result of ligand addition. We attempted to remove pH contributions as follows. First, a separate pH titration was conducted. Second, during ligand titrations, pH changes were monitored using the chemical shift variation of H146, a residue shown to be very sensitive to local alterations in pH and being well removed from sites showing ligand binding effects. Finally, a pH contribution to chemical shifts of other residues was calculated by scaling the H146 shift at a ligand titration point by the ratio of the pH shift for each residue to the pH shift for H146 seen in the pH titration. The resulting chemical shift value (in ppm) for each residue was subsequently subtracted from the observed chemical shift deviation seen at each point in the ligand titration. All chemical shift deviations are represented on a normalized scale (0 to 1) as fractional shifts observed.

### *4.1.4 Quantitation of Paramagnetic Attenuation*

As previously reported *(18, 19)*, paramagnetic effects were measured by comparing auto-scaled crosspeak attenuation values ($A_i$), defined as:

$$A_i = [2-(v_{ip}/v_{id})], \qquad [3]$$

with equation 3 being the individual deviations from the average of the crosspeak auto-scaled volumes, $v_{ip,d,}$ the latter being defined as:

$$v_{ip,d} = v_{ip,d}/[(1/n)(\sum_i V_{ip,d})], \qquad [4]$$

where n is the number of measured crosspeak volumes and $V_{id}$ and $V_{ip}$ are the protein individual crosspeak volumes in the absence and presence of the spin probe, respectively. The individual $A_i$'s were then plotted as a function of protein primary sequence. Values lying

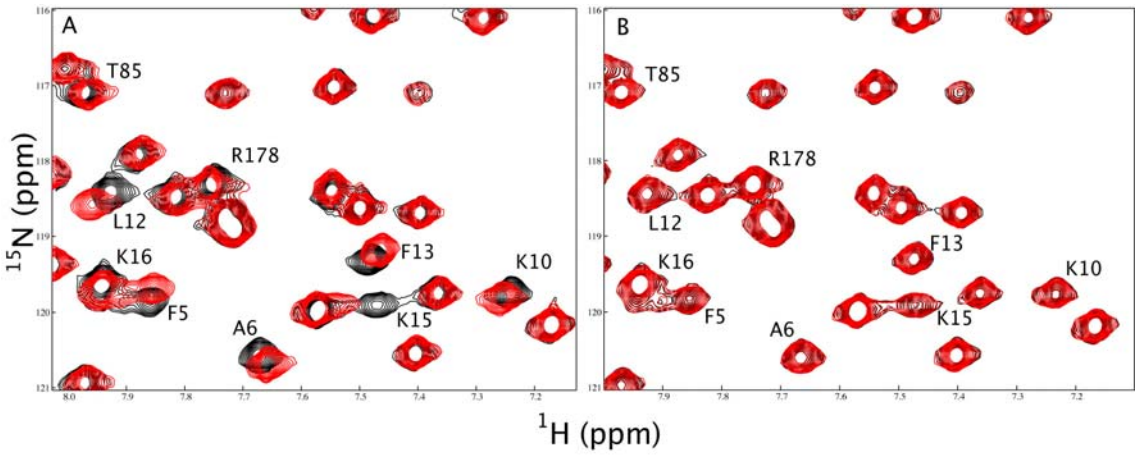


Figure 4.2. Section of an HSQC spectra collected on 0.5mM Arf1•GDP in 10mM potassium phosphate (pH 7.0), 1mM $MgCl_2$, 1mM GDP, and 100mM NaCl. Shown in black are peaks in the absence of ligand and shown in red are peaks in the presence of ligand, (**A**) 8.0 mM I(1,4,5)P3 and (**B**) 8.0 mM myoI. Peaks showing chemical shift variation as a result of ligand addition are labeled with residue number and amino acid type for those residues assigned. All spectra were recorded at 298K with a proton resonance frequency of 800MHz, as described under Methods.

above or below the average attenuation level [set at 1, since $(\sum v_{ip,d}/n)=1$] correspond to high or low spin probe accessibility, respectively. In order to better represent the surface-exposed residues, values lying above 1.2 were color-coded and superimposed on a three-dimensional surface plot of Arf1.

## *4.2 Results*

### *4.2.1 Phosphoinositide Binding*

Upon overlaying the HSQC spectra obtained for Arf1•GDP in the presence (8 mM) and absence of $I(1,4,5)P_3$ , distinct variations between the two states were readily observed, as

shown in Figure 4.2A. These variations were absent from samples of Arf1•GDP containing D-myo-inositol, the core of I(1,4,5)$P_3$ lacking the phosphates, at the same concentration (Fig. 4.2B). This serves as a control that assigns shifts seen with I(1,4,5)$P_3$ to specific interactions with phosphorylated inositols. The selective perturbation of chemical shifts shown in this segment of the HSQC spectra is for residues found in the positive patch on Arf1 or for amino acids comprising the amino terminal helix (F5, A6, L12, F13, K15).

A more complete detailing of resonances perturbed in response to the addition of a number of different phosphoinositides to Arf1 is shown in Figure 4.3. Here, limiting shifts in the $^{1}$H and$^{15}$N dimensions are plotted, weighting shifts by the ratio of observed sweep widths for $^{1}$H and$^{15}$N to remove bias as a result of the greater sensitivity of $^{15}$N shifts (in ppm) to small conformational changes. The horizontal blue line at 0.02 ppm denotes the standard deviation observed over all changes. It is apparent that upon addition of I(1,4,5)$P_3$ to Arf1 (Figure 4.3E), selective perturbations of individual resonances occur and that many of these are repeated for some of the other ligands (Figure 4.3C and 4.3D). Although the number of significant chemical shift perturbations observed might seem larger than expected, most can be assigned to a small number of structural regions. These changes map to three regions, one within the N-terminal helix and loop (residues F5, A6, K10, F13, K15, K16, R17), a second including the positive patch (positive residues K10, K15, K16, K181and their adjacent residues R17, N179, Q180), and a third around the nucleotide binding region including the P-loop (residues D26, A28, G29), helix A (residues Y35, K36), and the G3 binding motif (residues K127, Q128, N132). Outside of these three regions, minor, or more isolated, perturbations are observed in β strands running through the protein core and connecting the N-terminal helix with the nucleotide binding pocket; specifically residues in β1 (19-RIL-21),

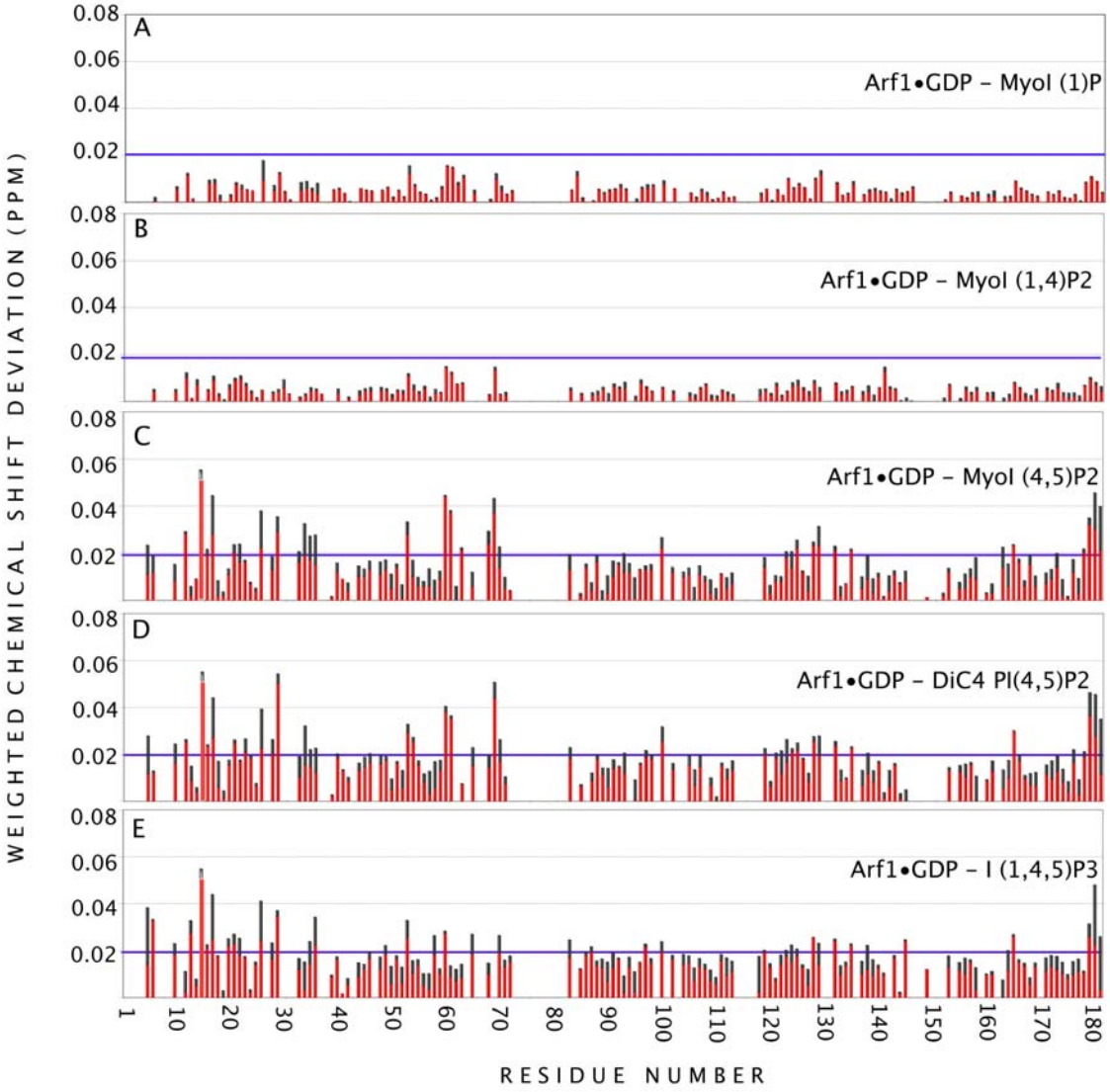


Figure 4.3. $^{15}N/^{1}H$ weighted chemical shift deviations in Arf1•GDP (0.5mM) samples with various phosphoinositides are plotted as a function of residue number. Ligands used are as follows: (**A**) myoI-(1)P (8mM), (**B**) myoI-(1,4)$P_2$ (8mM), (**C**) myoI-(4,5)$P_2$ (8mM), (**D**) dibutyryl- PI(4,5)$P_2$ (3mM) and (**E**) I-(1,4,5)$P_3$ (8mM). Contributions to chemical shift variation from the $^{1}H$ dimension are shown in red with $^{15}N$ specific contributions represented in grey. All values are shown as the overall magnitude of the deviation observed. Values above one average standard deviation are found above the horizontal blue line at 0.02 ppm.

the interswitch region involving β2 and β3 (I49, V53, Y58, N60, V65), and a residue in Switch 2 (G69) are perturbed. Residues R99 and D164, located in surface exposed loops, also

show some perturbation. Note that resonances for residues K30 and T31 of helix A undergo substantial line broadening as a result of I(1,4,5)$P_3$ binding, preventing the measurement of chemical shift offsets; they should also be included in the list of perturbed resonances.

Cell signaling by phosphatidylinositides is a very specific process that is controlled by addition of phosphate groups primarily to the 3, 4, and 5 positions on the inositol ring. Because charge effects make an important contribution to the protein interactions in which these multiply phosphorylated compounds engage, the stereochemistry of the phosphates is an important indicator of specificity. The importance of phosphate placement on the inositol ring was tested by examining chemical shift offsets in Arf1•GDP resonances in response to ligands which varied both in the position and number of phosphates on the inositol ring; these included myoI-(1)P, myoI-(1,4)$P_2$, and myoI-(4,5)$P_2$, as illustrated in Figure 4.1.

Chemical shift deviations observed for myoI-(1)P and myoI-(1,4)$P_2$ are small, even at the limiting concentrations of our studies (Figure 4.3A and 4.3B). This indicates weaker binding and shows that the simple presence of two phosphates on the inositol ring is not sufficient to cause the larger chemical shifts described above. However, myoI-(4,5)$P_2$ (Figure 4.3C) shows selective chemical shift perturbations very similar to those seen for I(1,4,5)$P_3$. These include residues in the surface exposed positive patch on Arf1•GDP, as well as residues assigned to the N-terminal helix and the nucleotide binding loops. A few additional perturbations are found in switch 2 elements (V68, G70) as well as helix A (L34). These observations suggest that the placement of the 5-phosphate or the vicinal (4, 5)-bisphosphate combination is a critical element for binding and that I(1,4,5)$P_3$ and myoI-(4,5)$P_2$ are likely to share a common binding site.

Because the soluble I(1,4,5)$P_3$ and membrane-associated PI(4,5)$P_2$ each have signaling capabilities in cells, we wanted to assess the importance of the hydrophobic portion of the lipid in interactions with Arf1. A water soluble, short acyl-chain analog (dibutyryl-phosphatidylinositol (4,5)-bisphosphate [dibutyryl-PI(4,5)$P_2$]), was examined for its ability to bind Arf1•GDP (molecule diagrammed in Figure 4.1). A binding pattern again reminiscent of I(1,4,5)$P_3$ association was found (Figure 4.3D), suggesting that the polar head group remains involved in direct binding to Arf1 for all these ligands. However, as we illustrate below, the lipid portion appears to enhance binding.

### *4.2.2 Determination of Dissociation Constants*

In select cases it is possible to use the concentration-dependent changes in chemical shifts to determine binding constants for ligands. To do this, chemical shift differences for I(1,4,5)$P_3$ and dibutyryl-PI(4,5)$P_2$ upon addition of ligand were plotted and fit using equations 1 and 2 (see methods) over the titration range 0- 3mM and 0-1.2mM, respectively (Figure 4.4). Normalized deviations in chemical shift are plotted for each titration point, as described in the methods section. The observed chemical shifts for I(1,4,5)$P_3$ clearly shift monotonically as a function of concentration (Figure 4.4 A), indicating a ligand that is in fast exchange on the NMR timescale. Shifts also cease to deviate at high ligand concentrations (above 1.25mM), indicating an ability to saturate the binding site. This behavior allows an accurate determination of the dissociation constant ($K_{D(apparent)}$ = 56 ± 15 μM) for I(1,4,5)$P_3$. Extracting a binding constant from the titration with dibutyryl-PI(4,5)$P_2$, is more difficult (Figure 4.4 B). An ability to saturate the binding site (limiting concentration of 0.4mM) is

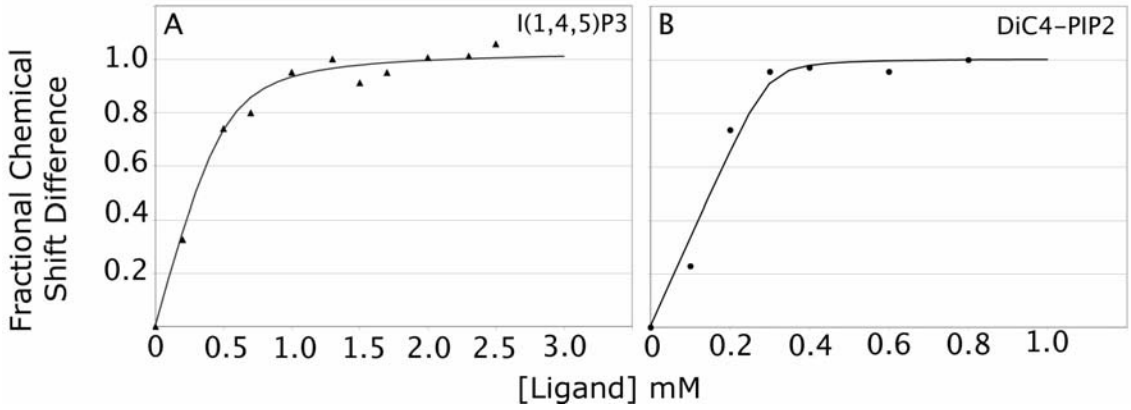


Figure 4.4. Quantitation of perturbations as a function of ligand concentration. HSQC spectra were obtained for samples containing Arf1•GDP (0.5mM) with varying amounts of I(1,4,5)$P_3$ (A, closed triangles) or Arf1•GDP (0.3mM) with varying amounts of dibutyryl-PI(4,5)$P_2$ (B, closed circles). Normalized deviations in chemical shifts observed for each point were obtained using equation 1 (see methods). Dissociation constants were extracted using a Tcl/Tk based NMRview module and equation 2 as described in methods. A line simulated using the derived dissociation constant is included for comparison.

demonstrated and the binding constant is clearly higher. However, one can only place limits on the size of a binding constant. Also, there are indications that the ligand may be in slow exchange on the NMR time scale. Peaks from spectra below 0.3 mM show broadening that can be interpreted as superposition of pairs of poorly resolved peaks. A change in intensity distribution between pairs gives the appearance of a change in chemical shift. While this behavior prohibits the accurate determination of the dissociation constant, an upper limit can be estimated from the saturation point ($K_D$ < 10 μM).

For myoI-(1)P and myoI-(1,4)$P_2$ significant shifts were not observed even at high concentrations. This is indicative of very weak binding ($K_D$ > 10 mM) or the lack of an ability to induce shift perturbations. MyoI-(4,5)$P_2$ produced limiting shifts comparable to I(1,4,5)$P_3$, but due to the small number of titration points collected with the limited quantity of ligand available to us, a quantitative determination of binding was not attempted.

However, the fact that a limiting shift was reached near 2.0mM allows us to set an upper limit for the dissociation constant of 500μM. Hence MyoI-(4,5)$P_2$ binds with an affinity noticeably lower than I(1,4,5)$P_3$ and dibutyryl-PI(4,5)$P_2$ binds with an affinity considerably higher than I(1,4,5)$P_3$. Thus, although neither the glycerol nor the acyl-chains are required for interaction with Arf1, their presence increases the affinity for Arf1, either through direct contacts with the protein or indirectly through effects on the structure or solvation of the inositol and associated phosphates.

### *4.2.3 Binding Site Verification*

In cases where selective perturbations are seen upon ligand addition (myoI-(4,5)$P_2$, dibutyryl-PI(4,5)$P_2$ and I(1,4,5)$P_3$) the perturbations occur in at least three regions of the protein. Because chemical shift perturbations can occur at remote sites due to allosteric propagation, it is difficult to assess which region of the protein is contacting the ligand directly. It is therefore important, whenever possible, to supplement chemical shift data with other information, in order to determine the direct site of ligand interaction on the protein surface. One of the additional experiments that can supplement chemical shift offsets is perturbation of relaxation times for surface exposed amide proton-amide nitrogen crosspeaks by water-soluble paramagnetic species *(7, 18)*. By monitoring crosspeak attenuations in HSQC spectra upon addition of the stable free radical TEMPOL (50mM) to the solvent *(18)*, topological mapping of the Arf1 protein surface was achieved. Crosspeaks most attenuated are those derived from surface exposed residues. An example of the data acquired is illustrated in Figure 4.5. The variation in crosspeak attenuation levels between spectra

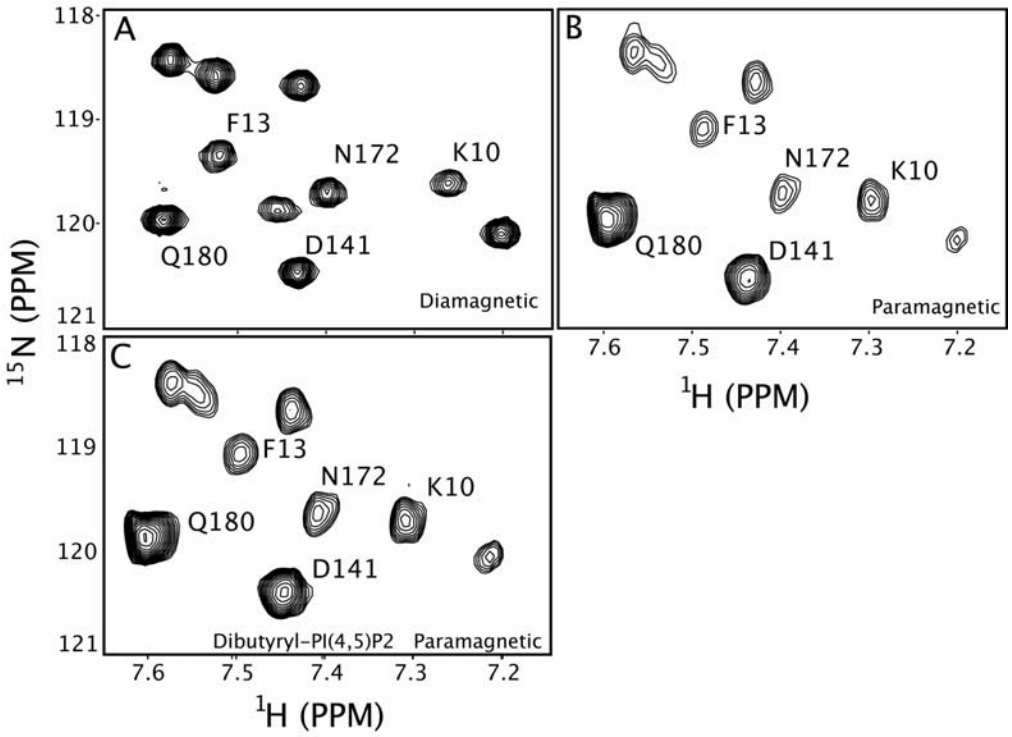


Figure 4.5. Section of an HSQC spectrum collected on 0.5mM Arf1•GDP in (**A**)10mM potassium phosphate (pH 7.0), 100mM NaCl, 1mM $MgCl_2$, and 1mM GDP, (**B**) in the same buffer with the addition of 50mM TEMPOL, and (**C**) upon addition of 3.0mM dibutyryl-PI(4,5)$P_2$ and 50mM TEMPOL. Representative peaks showing paramagnetic attenuation as a result of TEMPOL accessibility are labeled with residue number and amino acid type. All spectra were recorded at 298K with a proton resonance frequency of 600MHz, as described under Methods.

recorded in the presence and absence of small ligand molecules can also show regions of protection afforded to the molecular surface as a result of ligand addition. Because the origin of these differences is largely steric exclusion, and is very different in origin from that of chemical shift perturbations, the data nicely complement the chemical shift information.

Figure 4.6(A) and (B) show the results in the absence of ligand mapped onto a surface representation of the protein (PDB code 1HUR) for both the nucleotide binding face (A) and surface exposed positive patch (B). The crosspeak attenuations observed are scaled, as

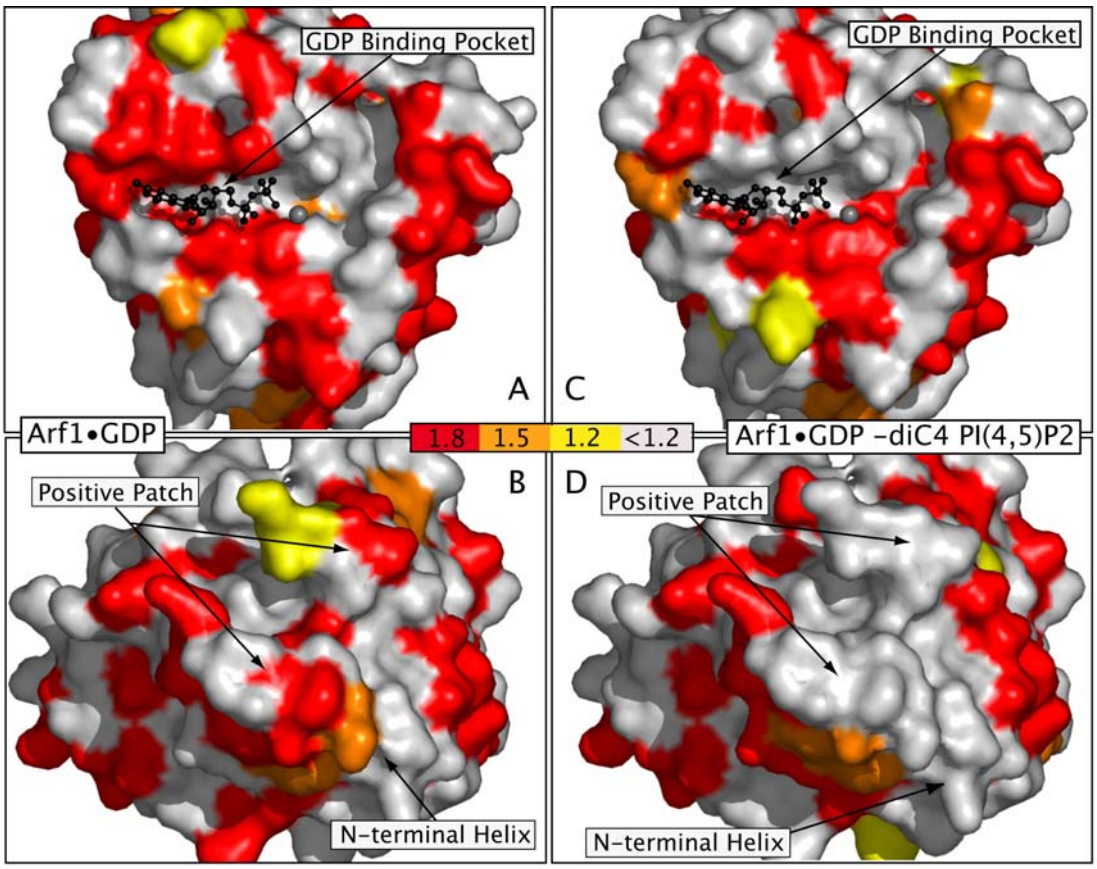


Figure 4.6. Arf1•GDP surface accessibility map derived from TEMPOL spin probe attenuation observed in both the absence (panels **A** and **B**, labeled Arf1•GDP) and presence (panels **C** and **D**, labeled Arf1•GDP – diC4 PI(4,5)$P_2$) of dibutyryl-PI(4,5)$P_2$ (3mM). Crosspeak attenuations in the nucleotide binding face (panels **A** and **C**) and the surface exposed positive patch (panels **B** and **D**) are superimposed onto a three-dimensional surface model of Arf1. The A-B and C-D pairs are related by approximately a 180° rotation. The crosspeak attenuations observed are scaled to a maximum value of 2 with values lying above 1 corresponding to increased spin-probe accessibility levels. The representative surface maps are color-coded to illustrate the degree of paramagnetic attenuation [red (values 1.8 or greater), orange (1.5-1.8), yellow (1.2-1.5), and grey (less than 1.2)].

described in the methods sections, to a maximum value of 2 with values lying above 1 corresponding to increased spin-probe accessibility levels. The representative surface maps are color-coded to illustrate the degree of paramagnetic attenuation [red (having values 1.8 or greater), orange (1.5), yellow (1.2), and grey (<1.2)]. Addition of the ligand (dibutyryl-

$PI(4,5)P_2$) changes the TEMPOL-derived surface accessibility patterns, as shown in Figure 4.6(C) and (D). Extensive changes are seen near the positive patch upon ligand addition; specifically, a decrease in surface accessibility of residues K10, L12, F13, K15, 178-RNQK-181 is seen. Alterations to the nucleotide-binding region were also seen upon ligand addition, specifically in residues V123, K127 and D129, but because these changes are fewer in number and lie along one face of the bound nucleotide they were deemed much less likely to result from direct interaction with the ligand. These data lead us to conclude that the chemical shift perturbations observed for resonances found in and around the positive patch result from direct contact with the ligand and those observed for resonances from the nucleotide binding site, or intervening structural elements, are more likely a secondary effect of ligand binding observed through allosteric propagation.

### *4.2.4 Conformational Destabilization upon Phosphoinositide Binding*

It has been previously suggested that $PI(4,5)P_2$ binding might stimulate the release of bound GDP from membrane associated Arf1 *(2, 20, 21)*. To address the potential for conformational destabilization in Arf1•GDP seen in complex with the phosphoinositides used in our study, we compared the changes in energetic properties of the protein between ligand bound and unbound states, as assessed through measurement of amide proton exchange rates. The amide group is assumed to be in rapid equilibrium between protected and exposed forms with base catalyzed exchange of the amide proton for a water proton in the open form being the rate-limiting step *(8)*. As described in Chapter 3, increases in rates of exchange primarily reflect destabilization of protected forms and a shift in equilibrium toward conformations that

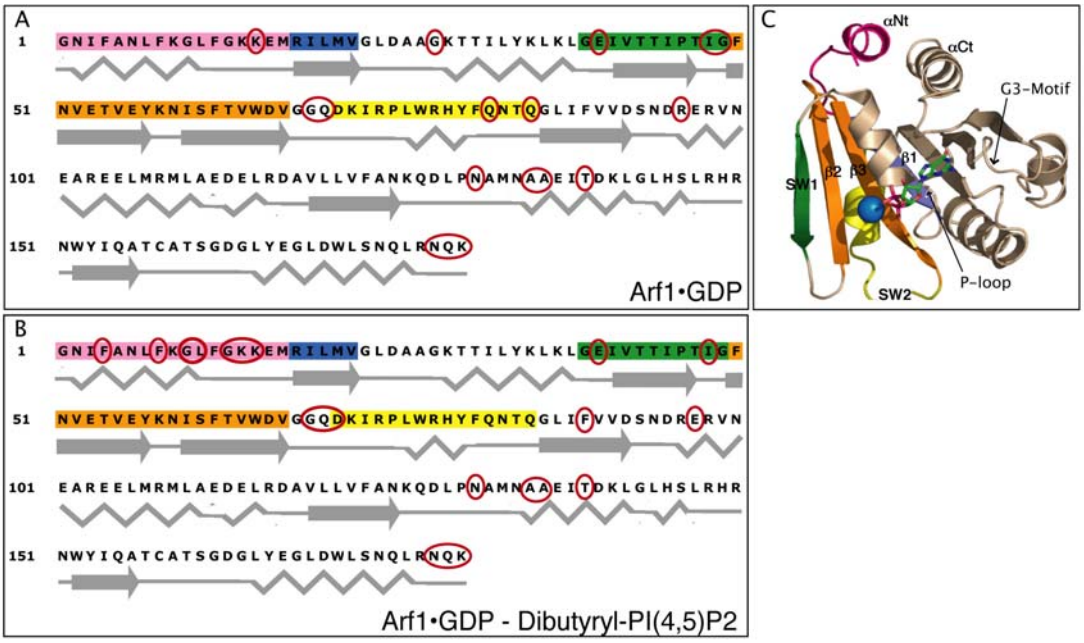


Figure 4.7. $^{1}H$-$^{15}N$ Cleanex HSQC comparisons, showing areas of rapid exchange between water and amide protons of Arf1•GDP (**A**) and Arf1•GDP-dibutyryl-PI(4,5)$P_2$ (**B**). Secondary structural elements are shown as a cartoon beneath the protein sequence, shown using the single letter amino acid code. Resonances evident in a $^{1}H$-$^{15}N$ Cleanex HSQC experiment are circled in red. To relate these regions of enhanced exchange to a three-dimensional model, the structure of Arf1•GDP is shown as a cartoon representation color-coded as follows: N-terminal α helix (magenta), switch 1 (green), inter-switch region (orange), and switch 2 (yellow).

expose amide protons to solvent. Observing changes in CLEANEX HSQC peak intensity can highlight regions where rates of interest are moderately fast *(9)*.

CLEANEX HSQC data for Arf1•GDP in the absence and presence of dibutyryl-PI(4,5)$P_2$ are summarized in Figures 4.7(A) and (B), respectively. Here residues having amides showing exchange sufficiently rapid to acquire magnetization from water protons are circled. There are more such residues in the ligand bound case (Fig. 4.7B), indicating an increase in exposed residues or destabilization of secondary structure elements that protect amide protons from exchange in unbound Arf1•GDP (Fig. 4.7A). As expected, the majority

of rapidly exchanging amides are in loops between secondary structural elements in both ligand-free and –bound states. Major differences seen upon ligand binding occur in the gain of some rapidly exchanging elements in the N-terminal helix of Arf1•GDP (Fig. 4.7B) along with elements in switch 2 (e.g., D72) and the loss of rapidly exchanging elements in the P-loop (G29, Fig. 4.7A), the loop succeeding switch 1 (G50) and the loop from 83-85 at the end of the switch 2 helix.

Because a number of changes in amide exchange rates involve loops near the nucleotide binding site we wanted to examine more closely the specific effects on the bound nucleotide that accompanied ligand binding. This was addressed by observing the naturally occurring $^{31}P$ resonances attached to the guanine base in both free and bound GDP in the Arf1•GDP-dibutyryl-PI(4,5)$P_2$ complex. GDP has distinctive $^{31}P$ chemical shift values for the β phosphorus in the bound and free states, allowing for a direct detection of bound nucleotide. $^{31}P$ spectra were recorded from Arf1•GDP in the absence and presence of dibutyryl-PI(4,5)$P_2$ (Figure 4.8) The spectra show an overall retention of the nucleotide within the binding pocket of Arf1 in solution. However, comparing integrated peak intensities of $^{31}P$ peaks observed in the spectra from GDP to internal standards indicates a change in the binding equilibrium favoring nucleotide release upon ligand addition. The ratio of free to bound β phosphate intensities changes from 0.11 to 0.26. It was seen that upon binding dibutyryl-PI(4,5)$P_2$ there are energetic changes in protein elements near the nucleotide binding site and there appears to be a coupled shift toward a lower binding constant for GDP and consequent release of the nucleotide.

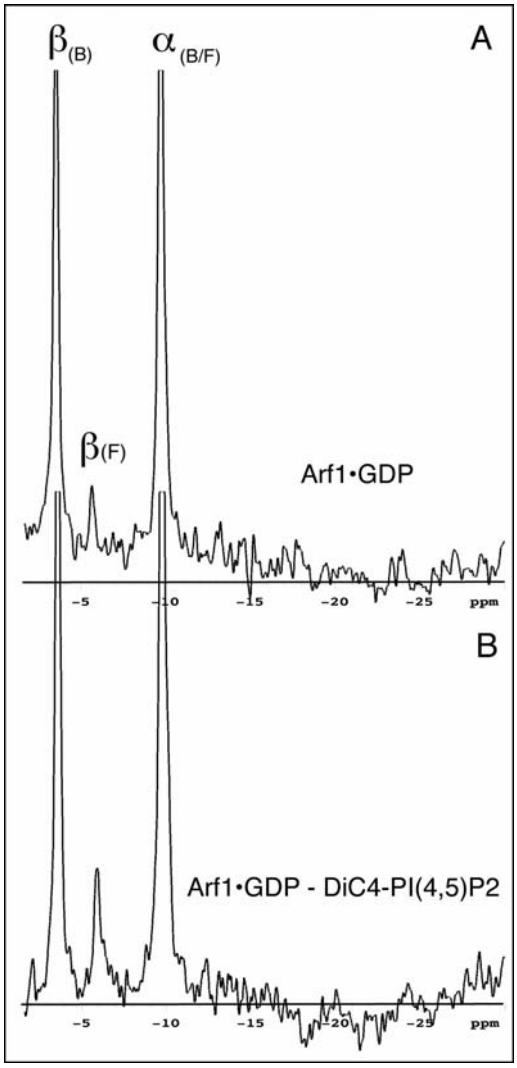


Figure 4.8. $^{31}$P spectra recorded on Arf1•GDP (0.5mM) in the absence (**A**) and presence (**B**) of dibutyryl-PI(4,5)$P_2$ (3mM). Phosphates observed are labeled by phosphate type ($\alpha/\beta$) along with bound state (F=free or B=bound). Note that $\alpha$ resonances are overlapped between free and bound GDP. Integration of peak intensities shows as an increase in the amount of the free $\beta$ phosphate from GDP upon ligand binding (the ratio of $\beta_F/\beta_B$ changes from 0.11 to 0.26). All spectra were recorded with a proton resonance frequency of 300 MHz at 298 K.

## 4.3 Discussion

The localized production and sensing of phosphorylated forms of phosphatidylinositols is an integral part of Arf biology. This is both because activated Arfs

can recruit and directly activate the enzymes (PI 4-kinase and PI(4)P 5-kinase) controlling the formation of these signaling lipids and because the product of these kinases (PI $(4,5)P_2$) is a required cofactor in both Arf-dependent phospholipase D and Arf GAP activities and assists in the recruitment to membranes of a subset of Arf GEFs. A further complication to the development of molecular models of Arf action stems from the observation that $PI(4,5)P_2$ can bind directly to Arf1 and alter its affinity for guanine nucleotides *(2, 3)*. Several studies have suggested structural changes involved in the activation of Arf proteins and in their binding to Arf GEFs. These include substantial changes to the amphipathic N-terminal helix, the nucleotide-binding pocket, interswitch region, and connecting strands *(10, 22)*. We sought additional structural evidence to test the hypothesis that binding to phosphatidylinositols may play a facilitative role in the activation of Arf1. Using a number of NMR techniques, we document here specificity in the binding to different phosphatidylinositides, chemical shift changes in specific positively charged residues on one surface of Arf1 that we propose to result from direct interaction with the ligand, and allosteric changes in the nucleotide binding site that accompany ligand binding, all of which support the conclusion that binding to $PI(4,5)P_2$ can promote the nucleotide exchange process and facilitate activation of Arfs.

Tests of specificity for the ability to induce chemical shift changes in Arf1 revealed a dependence on vicinal phosphates of the inositol ring (see Fig. 4.3). Non-specific charge effects are not viewed as a viable explanation for the changes observed here as $myoI(1,4)P_2$ was without effect while $myoI(4,5)P_2$ produced comparable perturbations to those induced by $I(1,4,5)P_3$ or dibutyryl-$PI(4,5)P_2$. The concentration dependence in chemical shifts induced by $I(1,4,5)P_3$ was then used to determine an apparent binding constant and the value obtained ($K_{D\ (apparent)} = 56 \pm 15\ \mu M$) is comparable to that determined previously for the binding of

Arf1•GTP to PI(4,5)$P_2$ in Triton X-100 micelles (45 ±13 μM; *(23)*). In contrast, we found that the addition of even very short acyl chains (dibutyryl-PI(4,5)$P_2$) increased the affinity to give a dissociation constant below 10 μM. This increase in affinity may result either from an induced alteration in the structure or charge distribution of the inositol and associated phosphates to give a form that binds better to Arf1, or from a direct involvement of the glycerol or one or both butyrates in binding to the protein. While the similarity in binding properties of Arf1 observed in solution and in a detergent micelle *(23)* suggest an ability to extrapolate to a membrane environment, it is important to point out that we still do not know the affinity of any Arf for PI(4,5)$P_2$ in the context of a biological membrane. However, based on the dibutyryl-PI(4,5)$P_2$ data we can speculate that the dissociation constant is below 10 μM and is in the physiological range. Our data were all collected for Arf1 in solution but with the use of membrane mimics (e.g., bicelles) we should be able to extend the use of NMR techniques to an even better approximation of the physiologically important Arf-lipid interactions.

Structural consequences of ligand binding extend far beyond the actual phosphoinositide-binding patch (shown as a surface plot in Figure 4.9A with perturbed resonances highlighted in red). Assigned resonances comprising the N-terminal helix (Figure 4.9B) and the nucleotide-binding pocket (αA/P-loop and G3 Motif, Figure 4.9C) are also perturbed, indicating a change in the local electronic environment around those residues. CLEANEX HSQC experiments show changes in stability of secondary structural elements of several of these segments supporting a propagated change in structure to the N-terminal helix and the nucleotide-binding site. Nucleotide destabilization and subsequent release in native environments that contain phosphoinositides have been described in the literature *(2, 3, 24)*

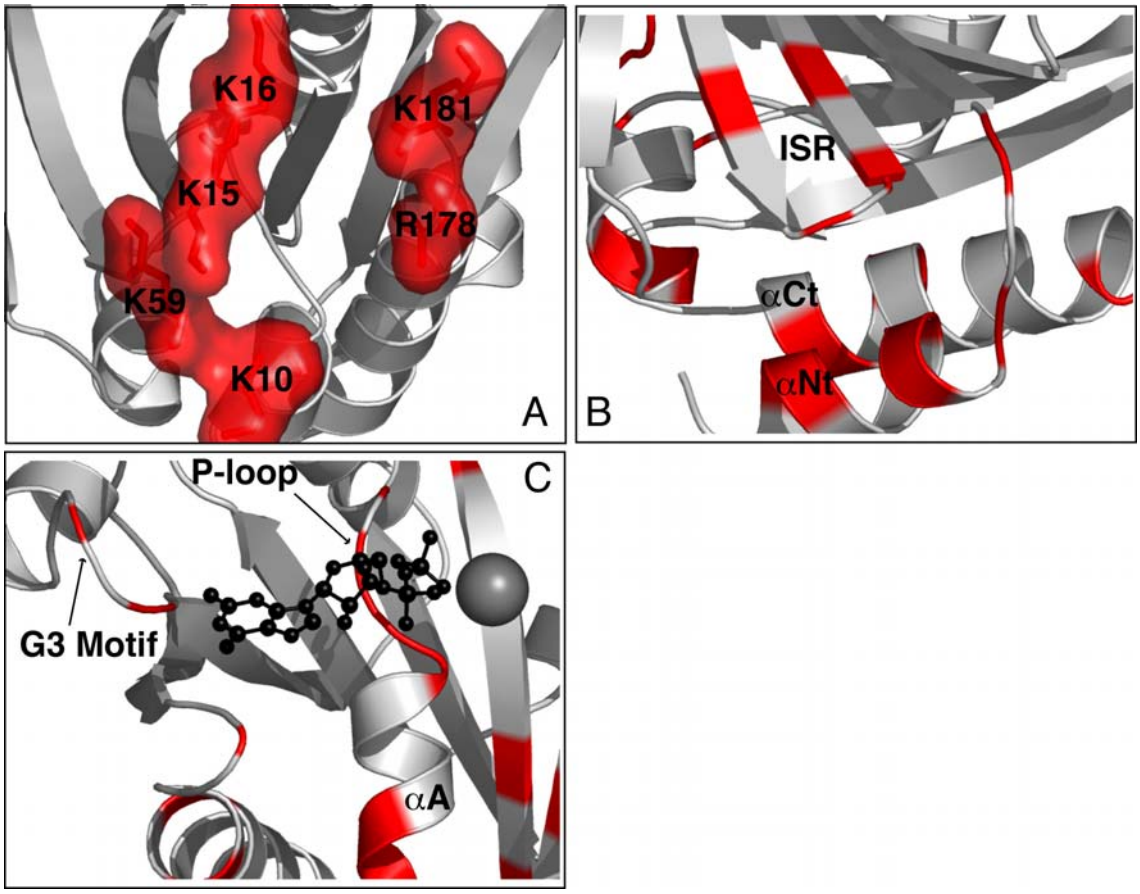


Figure 4.9. Crystal structure of Arf1•GDP highlighting regions affected by phosphoinositide binding. (**A**) the solvent exposed positive patch on Arf1•GDP, (**B**)the interswitch region (ISR), the N-terminal helix (αNt) and the C-terminal helix (αCt), and (**C**) the nucleotide binding pocket (P-loop, αA and G3 Motif). Assigned residues showing chemical shift perturbations in HSQC titrations colored red.

and our study has shown that this extends to solution conditions where a change in the equilibrium of GDP binding occurs as evidenced in $^{31}$P experiments. The stabilization changes in protein structural elements and these changes in nucleotide binding properties are likely to be coupled. On the basis of our observations, Arf1•GDP is proposed to associate with PI(4,5)$P_2$ at the membrane surface, allowing for the local structural modifications

required for nucleotide destabilization and possibly alteration in the position and stability of the N-terminal helix.

## References


1. Zhu, X., A.L. Boman, J. Kuai, W. Cieplak, and R.A. Kahn. (2000) Effectors increase the affinity of ADP-ribosylation factor for GTP to increase binding. *J Biol Chem*, **275**(18): p. 13465-75.
2. Terui, T., R.A. Kahn, and P.A. Randazzo. (1994) Effects of acid phospholipids on nucleotide exchange properties of ADP- ribosylation factor 1. Evidence for specific interaction with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem*, **269**(45): p. 28130-5.
3. Kahn, R.A., T. Terui, and P.A. Randazzo. (1996) Effects of acid phospholipids on ARF activities: potential roles in membrane traffic. *J Lipid Mediat Cell Signal*, **14**(1-3): p. 209-14.
4. Kay, L.E., P. Keifer, and T. Saarinen. (1992) Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Coherence Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.*, **114**: p. 10663-10665.
5. Stonehouse, J., G.L. Shaw, J. Keeler, and E.D. Laue. (1994) Minimizing Sensitivity Losses in Gradient-Selected N-15-H-1 HSQC Spectra of Proteins. *J. Magn. Res. A.*, **107**: p. 178-184.
6. Amor, J., R. Seidel, F. Tian, R. Kahn, and J. Prestegard. (2002) Letter to the Editor: H-1, N-15 and C-13 assignments of full length human ADP Ribosylation Factor 1 (ARF1) using triple resonance connectivities and dipolar couplings. *J Biomol NMR*, **23**(3): p. 253-254.
7. Niccolai, N., A. Ciutti, O. Spiga, M. Scarselli, A. Bernini, L. Bracci, D.D. Maro, C. Dalvit, H. Molinari, G. Esposito, and P. Temussi. (2001) NMR Studies of Protein Surface Accessibility. *J Biol Chem*, **276**(45): p. 42455-42461.
8. Dempsey, C.E. (2001) Hydrogen Exchange in peptides and proteins using NMR spectroscopy. *Prog. NMR Spec.*, **39**: p. 135-170.
9. Hwang, T.-L., S. Mori, A.J. Shaka, and P.C.M.v. Zijl. (1997) Application of Phase Modulated CLEAN Chemical EXchange Spectroscopy (CLEANEX-PM) to Detect Water-Protein Proton Exchange and Intermolecular NOEs. *J. Am. Chem. Soc.*, **119**: p. 6203-6204.
10. Renault, L., B. Guibert, and J. Cherfils. (2003) Structural Snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. *Nature*, **426**(6966): p. 525-530.
11. Seidel III, R.D., J.C. Amor, R.A. Kahn, and J.H. Prestegard. (2004) Conformational Changes in Human Arf1 on Nucleotide Exchange and Deletion of Membrane Binding Elements. *J. Biol. Chem.*, **279**(46): p. 48307-48318.
12. Randazzo, P.A. (1997) Resolution of two ADP-ribosylation factor 1 GTPase-activating proteins from rat liver. *Biochem J*, **324**(Pt 2): p. 413-9.
13. Neidhart, F.C., P.L. Bloch, and D.F. Smith. (1974) Culture Medium for Enterobacteria. *J Bact*, **119**: p. 736.
14. Randazzo, P.A. and R.A. Kahn. (1995) Myristoylation and ADP-ribosylation factor function. *Methods Enzymol*, **250**: p. 394-405.

15. Hyvonen, M., M.J. Macias, M. Nilges, H. Oschkinat, M. Saraste, and M. Wilmanns. (1995) Structure of the binding site for inositol phosphates in a PH domain. *The EMBO Journal*, **14**(19): p. 4676-4685.
16. Delaglio, F., S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer, and A. Bax. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**: p. 277-293.
17. Johnson, B. and R. Blevins. (1994) NMRView - A Computer-Program for the Visualizaation and Analysis of NMR Data. *J. Biomol. NMR*, **4**: p. 603-614.
18. Scarselli, M., A. Bernini, C. Segoni, H. Molinari, G. Esposito, A. lesk, F. Laschi, P. Temussi, and N. Niccolai. (1999) Tendamistat surface accessibilty to the TEMPOL paramagnetic probe. *J Biomol NMR*, **15**: p. 125-133.
19. Molinari, H., G. Esposito, M. Pegna, L. Ragona, N. Niccolai, and Z. L. (1997) Probing Protein structure by solvent Perturbation of NMR Spectra: The Surface Accessibility of bovine Pancreatic Trypsin Inhibitor. *Biophys J*, **73**: p. 382-296.
20. Paris, S., S. Beraud-Dufor, S. Robineau, J. Bigay, B. Antonny, M. Chabre, and P. Chardin. (1997) Role of Protein-Phospholipid Interactions in the Activation of ARF1 by the Guanine Nucleotide Exchange Factor Arno. *J Biol Chem*, **272**(35): p. 22221-22226.
21. Randazzo, P. (1997) Functional interaction of ADP-ribosyaltion factor 1 with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem*, **272**(12): p. 7688-7692.
22. Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95**(2): p. 237-48.
23. Randazzo, P.A. (1997) Functional interaction of ADP-ribosylation factor 1 with phosphatidylinositol 4,5-bisphosphate. *J Biol Chem*, **272**(12): p. 7688-92.
24. Zheng, Y., J.A. Glaven, W.J. Wu, and R.A. Cerione. (1996) Phosphatidylinositol 4,5-bisphosphate provides an alternative to guanine nucleotide exchange factors by stimulating the dissociation of GDP from Cdc42Hs. *J Biol Chem*, **271**(39): p. 23815-9.

# Chapter 5

## Proposed Model for Arf1 Activation and Feedback-Regulation

Eukaryotic cells contain an intricate assortment of organelles, connected with each other by a continuous flow of small carrier vesicles whose purpose is to transport macromolecules between the compartments. These vesicles are initially detached from the donor membrane through a process termed budding and incorporated into the target membrane by a vesicle fusion process termed docking. The vesicles undergoing budding utilize a mass of coat proteins that are instrumental for shaping the vesicle, ultimately allowing for release from the donor membrane. Vesicle formation is a multi-step event, being initiated by the recruitment of GDP bound Arf to the membrane surface. Arfs then assist in the recruitment of coat protein complexes as well as serve as direct activators of lipid modifying enzymes (i.e. phospholipase D or PI(4)P 5-kinase *(1-3)*), directly influencing membrane structure. Taken together, the existing biochemical data and newly presented structural information have allowed for the generation of a mechanistic model for Arf1 activation. Moreover, current findings, specifically the ability of Arf1 to associate with acidic phospholipids (Chapter 4), have permitted the proposal of a possible feedback loop that can enhance basal levels of membrane recruitment.

## 5.1 Activation

As summarized in Figure 5.1, in an inactive state, [1] Arf1 is seen in rapid equilibrium between cytosolic and membrane associated forms through interaction with its N-terminal myristate. Hydrophobic insertion of the fatty acyl chain into the membrane is necessary for stable association with bilayers, although it is not sufficient to tightly anchor the protein to the membrane surface. [2] Upon interaction with negatively charged membrane imbedded lipids, possibly $PI(4,5)P_2$, the amphipathic N-terminal helix may become destabilized allowing for a more viable membrane interaction. In this way association of the amphipathic helix can provide additional binding energies and a more prolonged membrane association. This is not without biochemical foundation. It has been seen that, in the absence of phospholipids, ARNO, an Arf-GEF, can stimulate nucleotide exchange on truncated Arf that lacks its N-terminal helix, but cannot stimulate a similar exchange using full-length forms. Moreover, the phospholipid dependent reaction of full-length ARNO, or its catalytic Sec7 domain alone, with full-length Arf is enhanced by the presence of $PI(4,5)P_2$ *(4)*. It is therefore likely that $PI(4,5)P_2$ acts directly on Arf to make it more accessible to the exchange factor. $PI(4,5)P_2$ mediated displacement of the N-terminal helix from the protein core may facilitate the "opening" of SW1, a process needed for GEF association. This displacement is mimicked in our current work through N-terminal truncation. While SW1 has shifted on membrane association, SW2 and the interswitch remain intact, placing acidic residues in a position to prohibit GTP incorporation. However, the propagated structural and energetic changes observed for residues in contact with GDP upon N-terminal displacement (Chapter 3) or lipid

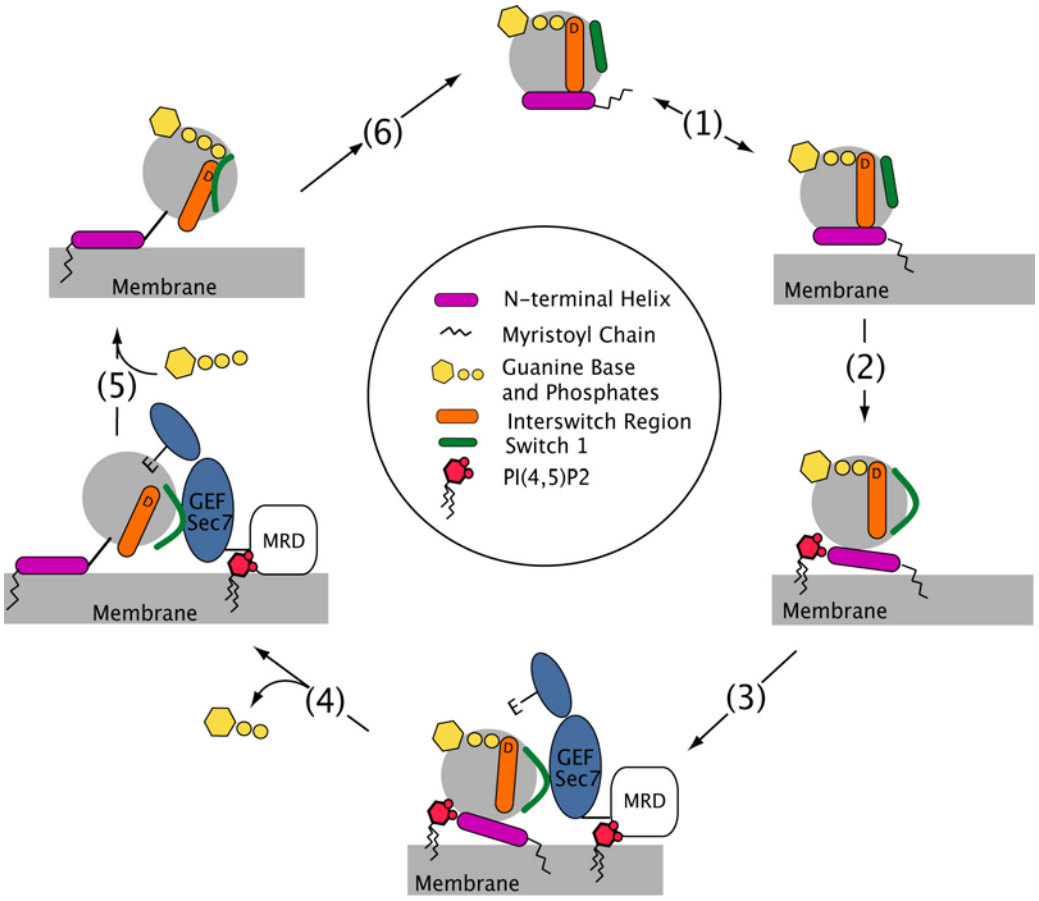


Figure 5.1. Proposed model of Arf1 activation. See text for details on independent steps within the nucleotide binding cycle (1-6). The membrane recruitment domain (MRD) within the Arf-GEF, not present in crystal structures, is shown as an open cartoon.

interaction (Chapter 4) suggest that GDP, while retained, becomes destabilized awaiting nucleotide release.

[3] Exchange factors such as ARNO, being cytosolic in nature, are recruited to the membrane surface through specific interaction with membrane imbedded PI(4,5)$P_2$. Arf1 subsequently interacts with its GEF at the membrane surface through SW1-hydrophobic groove interactions, enhanced by a prior SW1 displacement. [4] Rotation of the Arf1 core

towards the GEF allows for insertion of the glutamic finger into the nucleotide-binding pocket of Arf1, dislodging the previously destabilized bound nucleotide. The interswitch has shifted out of the protein core forcing a full release of the N-terminal helix as well as dispensing with the excess electro-negativity found in the nucleotide-binding pocket. [5] Incorporation of GTP, first interacting with the guanine base, allows for release of the associated GEF and locks Arf1 tightly onto the membrane surface. [6] Hydrolysis of GTP to GDP is needed for membrane release and subsequent vesicle fusion, requiring the association of GTPase activating proteins (GAPs).

## 5.2 Regulation

The mode by which vesicles are formed and consumed in membrane transport is regulated by initial lipid composition, and this composition is suggested to transiently modify during this process *(5)*. Based on our evidence that PI(4,5)$P_2$ binds to Arf1 in a way that can mediate membrane association (Chapter 4), a process that was previously controversial, a feedback loop can be proposed. It is known that the major pathway by which PI(4,5)$P_2$ is synthesized is through a sequential phosphorylation of phosphoinositides (PI) by PI 4- kinase to PI(4)P followed by a PI(4)P 5-kinase. Three PI 4-kinases have been currently identified ($\alpha$, $\beta$, and $\gamma$), separated based on size and organelle localization *(6)*. Interestingly, membrane recruitment and activation of PI 4-kinase $\beta$, the predominant form in Golgi membranes, is known to be regulated by Arf *(3)*. There are also three identified PI(4)P 5-kinases ($\alpha$, $\beta$, and $\gamma$), which again are differentially expressed *(7)*. However, all PI(4)P 5-kinases can be

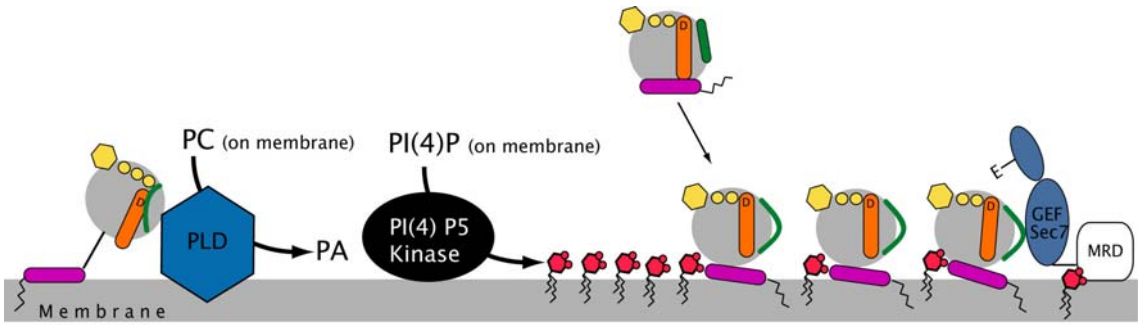


Figure 5.2. Possible model for positive feedback regulation of Arf1. GTP bound Arf interacts with PLD o the membrane converting PC to PA, which in turn, stimulates PI(4)P 5-kinase, increasing local $PI(4,5)P_2$ concentration. This may define the site at which Arf and its GEF interacts. Not shown is the stimulation of PI(4)P kinase by Arf increasing the local concentration of PI(4)P for subsequent conversion by PI(4)P 5-kinase (shown above). Direct stimulation of PI(4)P 5-kinase by Arf1 is also neglected. The reader is directed to the legend on Figure 5.1 for a description of cartoons used. Figure was adapted from the primary literature *(5)*.

stimulated, more than 100 fold *(8)*, by phosphatidic acid (PA) *in vitro (7)*. Interestingly, activation of Arf can cause a conversion of phosphatidylcholine (PC) to PA by stimulating phospholipase D1 (PLD) *(1, 9)*. Moreover, in the absence of PA, Arf1 has been found to directly stimulate PI(4)P 5-kinase $\alpha$ *(10)*. This may well act as a positive feedback loop for Arf recruitment.

As illustrated in Figure 5.2, after initial membrane recruitment and GEF catalyzed nucleotide exchange of Arf (shown in Figure 5.1), Arf is free to interact with membrane bound effectors, such as PLD. In this model, activated Arf interacts with PLD converting PC to PA, which in turn, stimulates PI(4)P 5-kinase and PLD itself. The increase in local $PI(4,5)P_2$ concentration can then define the site at which additional Arf and its GEF interact to facilitate the binding of Coatamer coat proteins. Again, this is not without a biochemical base. It is known that ATP is a required cofactor for $PI(4,5)P_2$ production. While ATP is not required for nucleotide exchange on Arf, ATP has been shown to stimulate the binding of

GTP loaded Arf to Golgi membranes *(11)* and amplify the assembly of coated vesicles *in vitro (5)*. This apparent ATP requirement may be the result of the need for ATP in Phosphoinositide Transfer Protein (PITP) dependant shuttling of PI between phospholipid membranes, known to be required for vesicle formation in Golgi complexes *(12)*.

## 5.3 Future Directions

### *5.3.1 Membrane Association*

One of the primary goals of this research has been to study the effect of membrane binding on Arf1 structure, and the implications of that structure relating to Arf1 activation. Current methods have employed truncation as a means of mimicking possible effects seen upon contacting the vesicle surface. While this is not without precedent, our findings remain far from conclusive. Structural studies on membrane associated Arf1 requires the production of N-terminally acylated samples for membrane anchoring. This poses multiple problems, both biochemically and structurally. First, E. *coli*, used for protein production, does not contain the protein machinery required for myristoylation, necessitating chemical addition post purification or the use of co-expression systems that provide the needed myristoyl-transferase. Efforts in co-expression have been unproductive, resulting in less than 10 percent myristoylation of Arf1. This approach is further marred by difficulties in separation of myristoylated from un-myristoylated samples. We have chosen, therefore, to explore chemical attachment of acyl chains post purification. Secondly, structural studies on membrane-

anchored systems would require the use of more solid-state NMR methods, as the line-shapes observed in membrane associated systems are anticipated to be very broad.

It would therefore be advantageous to design a method in which the high resolution spectra commonly observed in solution could be retained while extracting orientational data on membrane associated forms of Arf. Assuming an ability to separate residual dipolar couplings emanating from membrane bound and solution forms, one could, in principle, produce a structural model of membrane associated Arf. This is currently being attempted in our lab. The addition of fatty acyl chains of varying length (2- 10 carbons) in a site-specific manner (N-terminal) can be used as a means of varying the degree of alignment with the liquid crystal by hydrophobic insertion into an oriented bicelle media. In samples containing a low bicelle concentration (3-4%), ordering of Arf in solution would be small and residual dipolar couplings of membrane-associated fractions would dominate observations, as transient membrane anchoring would effectively increase the observed level of order for these fractions. By varying hydrophobic chain length, the fraction associated could be varied to reach a compromise between dominance of membrane-bound factions and degradation of resolution from immobilization at the membrane surface. Thus, orientational information obtained from this system would allow for the generation of membrane-associated models with minimal influence of freely soluble populations.

### *5.3.2 PI(4,5)$P_2$ Association*

While we have alluded to structural changes in Arf1 upon PI(4,5)$P_2$ interaction in previous chapters, a three-dimensional model for the complex remains to be derived. A model

for Arf can be generated using $^{1}H$-$^{15}N$ RDCs measured in the presence of saturating levels of PI(4,5)$P_2$, in much the same way we generated a structure for Δ17Arf1•GDP. However, a model for bound PI(4,5)$P_2$ presents a different set of problems. This ligand is not easy to label with NMR active isotopes, and it would be difficult to achieve concentrations of bound species needed for work at natural abundance. One solution is to use a ligand that exists in rapid exchange and operate with a 10-20 fold excess of ligand. This is typical of conditions used for transferred NOE studies where bound state contributions dominate due to the enhanced efficiency of NOE transfers when the ligand is bound to the protein *(13)*. Such studies could be carried out with PI(,4,5)$P_2$. However, it may be possible to add RDC data and CSA ($^{31}P$) data *(14)* as well, using a bound state enhancement strategy analogous to that described above. The use of ultra-high magnetic fields coupled with direct field alignment of the protein could provide the selective enhancement. It has been shown that proteins with metal centers spontaneously align in the presence of strong external magnetic fields *(15)*. Moreover, replacement of the more common diamagnetic metals with paramagnetic ions, having a higher magnetic susceptibility anisotropy, has showed promise *(16, 17)*. Substituting the required divalent metal ($Mg^{2+}$) coordinating GDP in Arf1 with a paramagnetic counterpart ($Co^{2+}$), could offer the needed magnetic susceptibility. The measured tensorial interactions within the ligand would belong solely to those in the bound state of PI(4,5)$P_2$ as alignment is through the protein directly. Thus a model for ligand orientation could be generated.

### *5.3.3 Method Development and Macromolecular Complexes*

The ability to refine existing structures using easily obtained experimental constraints in order to generate structures of closely related conformational forms, such as those that occur in Arf1 on nucleotide exchange, membrane association, and phosphoinositide binding, has far reaching utility. While the ability to determine independent structures of low to medium molecular weight isolates directly from complete NMR data sets remains fairly routine, extension to large molecules and macromolecular complexes using standard structural approaches is limited. Where closely related structures exist, incorporation of easily obtained data, such as RDC constraints, into energy minimization regimes, analogous to those described in chapter 3, can allow for an adequate description of biochemically important transformations in many systems. RDCs have been collected on quite large proteins and complexes *(18)*; hence applications to this wider range of systems may be possible. It is clear that residual dipolar coupling, although a new addition to the arsenal of tools available to the spectroscopist, has become central in the determination of three-dimensional models. The application of residual dipolar couplings is likely to expand in the years to come as research objectives extend to larger protein systems where traditional methods of structure determination prove inadequate.

**References**


1. Kahn, R.A., J.K. Yucel, and V. Malhotra. (1993) ARF signaling: a potential role for phospholipase D in membrane traffic. *Cell*, **75**(6): p. 1045-8.
2. Shome, K., C. Vasudevan, and G. Romero. (1997) ARF proteins mediate insulin-dependent activation of phospholipase D. *Curr Biol*, **7**(6): p. 387-96.
3. Godi, A., P. Pertile, R. Meyers, P. Marra, G. Di Tullio, C. Iurisci, A. Luini, D. Corda, and M.A. De Matteis. (1999) ARF mediates recruitment of PtdIns-4-OH kinase-beta and stimulates synthesis of PtdIns(4,5)P2 on the Golgi complex. *Nat Cell Biol*, **1**(5): p. 280-7.
4. Paris, S., S. Beraud-Dufor, S. Robineau, J. Bigay, B. Antonny, M. Chabre, and P. Chardin. (1997) Role of Protein-Phospholipid Interactions in the Activation of ARF1 by the Guanine Nucleotide Exchange Factor Arno. *J Biol Chem*, **272**(35): p. 22221-22226.
5. Roth, M.G. (1999) Lipid Regulators of Membrane Traffic through the Golgi Complex. *Trends in Cell Biology*, **v.9**(5): p. 174-179.
6. Barylko, B., S. Gerber, D. Binns, N. Grichine, M. Khvotchev, T. Studhof, and J. Albanesi. (2001) A Novel Family of Phosphatidylinositol 4-kinases Conserved from Yeast to Humans. *J. Biol. Chem.*, **276**(11): p. 7705-7708.
7. Ishihara, H., Y. Shibasaki, N. Kizuki, T. Wada, Y. Yazaki, T. Asano, and Y. Oka. (1998) Type I Phosphatidylinositol-4-phosphate 5-Kinases. *J. Biol. Chem.*, **273**(15): p. 8741-8748.
8. Jenkins, G.H., P.L. Fisette, and R.A. Anderson. (1994) Type I phsphatidylinositol 4-phosphate 5-kinase isoforms are specfically stimulated by phosphatidic acid. *J. Biol. Chem.*, **269**(15): p. 11547-11554.
9. Houle, M.G., R.A. Kahn, P.H. Naccache, and S. Bourgoin. (1995) ADP-ribosylation factor translocation correlates with potentiation of GTP gamma S-stimulated phospholipase D activity in membrane fractions of HL-60 cells. *J Biol Chem*, **270**(39): p. 22795-800.
10. Jones, D.H., J.B. Morris, C.P. Morgan, H. Kondo, R.F. Irvine, and S. Cockcroft. (2000) Type I Phosphatidylinositol 4-Phosphate 5-Kinase Directly Interacts with ADP-ribosylation Factor 1 and Is Responsible for Phosphatidylinositol 4,5-Bisphosphate Synthesis in the Golgi Compartment. *J. Biol. Chem.*, **275**(18): p. 13962-13966.
11. Tsai, S.C., R. Adamik, R.S. Huan, J. Moss, and M. Vaughan. (1993) Effects of Brefeldin A and Accessory proteins on Association of ADP-ribosylation Factors 1, 3, and 5 with Gogli membranes. *J. Biol. Chem.*, **268**(15): p. 10820-10825.
12. Sha, B. and M. Luo. (1999) PI transfer protein: the specific recognition of phospholipids and its functions. *Biochem. e. Biophys. Acta*, **1441**: p. 268-277.
13. Post, C.B. (2003) Exchange-trnasferred NOE spectroscopy and bound ligand structure determination. *Curr. Opin. St. Biol.*, **13**(5): p. 581-588.
14. O'Neil-Cabello, E., Z.R. Wu, D.L. Bryce, E.P. Nikonowicz, and A. Bax. (2004) Enhanced spectral resolution in RNA HCP spectra for meaurement of (3)J(C2/P) and (3)J(C4/P) couplngs and P-31 chemical shift changes upon weak alignment. *J. Bio. NMR*, **30**(1): p. 61-70.

15. Tolman, J.R., J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc Natl Acad Sci U S A*, **92**(20): p. 9279-83.
16. Barbieri, R., I. Bertini, G. Cavallaro, Y.M. Lee, C. Luchinat, and A. Rosato. (2002) Paramagnetically induced residual dipolar couplings for solution structure detrmination of lanthanide binding proteins. *J. Am.chem. Soc.*, **124**(19): p. 5581-5587.
17. Carrell, C.J., X.T. Wang, L.M. Jones, W.L. Jarrett, V.L. Davidson, and f.S. Mathews. (2004) Crystallographic and NMR investigation of cobalt-substituted amicyanin. *Biochemistry*, **43**(29): p. 9381-9389.
18. Tugarinov, V., P.M. Hwang, and L.E. Kay. (2004) Nuclear magnetic resonance spectrsocopy of high-molecular weight proteins. *Ann. Rev. Biochem.*, **73**: p. 107-146.

## Appendices

### Appendix A. Refinement Script for Torsion Angle Energy Minimization using RDCs

```
{* ======== Parameters, Coordinates, and Structure Files ======== *}

 {===>} structure @d17gdp_mg_water_sani.psf end       {*Read the structure file*}

 {===>} coor @d17refinegdp_mg_water_sani.pdb

parameter
   @param/parhcsdx.pro                        {*protein parameters*}
   @param/parameter.elements                        {*magnesium*}
   @param/param11.wat                                    {*water*}
   @param/par_axis_3.pro                          {*tensor axes*}

        {* Learn parameters for GDP from the cartesian coordinates.*}
        {* Learn equilibrium parameters from known cartesian coordinates.*}


   learn initiate sele=(resname GDP and known ) mode=nostatistics end
   learn accumulate end
   learn terminate end

{* Set energy constants. We are using generic values.*}

   BOND (resname GDP  ) (resname GDP ) 500. TOKEN
   ANGLE (resname GDP ) (resname GDP )
              (resname GDP ) 500. TOKEN
   IMPR (resname GDP )  (resname GDP )
          (resname GDP )  (resname GDP ) 500. TOKEN TOKEN

{* Set nonbonded parameters for the Lennard-Jones potential. *}

   NONBonded ( name C* and resname GDP ) 0.1  3.5     0.1  3.5
   NONBonded ( name H* and resname GDP ) 0.05 1.4     0.05 1.4
   NONBonded ( name O* and resname GDP ) 0.1  3.4     0.1  3.4
   NONBonded ( name N* and resname GDP ) 0.1  3.4     0.1  3.4
   NONBonded ( name P* and resname GDP ) 0.58 3.38    0.58 3.38
end

 {===>} coor @d17refinegdp_mg_water_sani.pdb    {* Read current coordinates.*}

 ! remove bogus coordinates for C terminal oxygens. Comment out if not needed.
```

```
        !delete select (resid 18 and name HT#) end
        delete select (resid 181 and name O#) end
        delete select (resid 181 and name OT#) end

{* ============== Protocol Definitions ================== *}


evaluate ($ini_sani1 = 0.001)
evaluate ($fin_sani1 = 0.5)             ! RDC1 starting and ending forces

evaluate ($ini_sani2 = 0.001)
evaluate ($fin_sani2 = 0.5)             ! RDC2 starting and ending forces

evaluate ($increment=500)               ! defines slope of RDC increase
            ! (i.e. how many points between starting and ending forces)

evaluate ($hold=1000)                           ! How many steps of minimization performed at
                                each RDC force constant
evaluate ($initial=1800)        ! How many steps of minimization performed   during initial
    refinement
                            !(RDC force constant fixed at ini_sani value defined above)
evaluate ($cartesian=500)               !short round of cartesian minimization to allow small
deviations from ideality
evaluate ($cart_final=1000)             ! steps in final round of cartesian minimization
evaluate ($file ="d17gdp_clore.pdb")! sets the output PDB file name used at the end of
minimization


                !---------- No Need to Modify---Ramping Definitions ----------
evaluate ($cycle=1.0)                   ! counter used to increment RDCs.  Do not touch.
evaluate ($sani1_fac = ($fin_sani1/$increment)*$cycle)          ! RDC1 ramping factor.
Do not Touch.
evaluate ($sani2_fac = ($fin_sani2/$increment)*$cycle)          ! RDC2 ramping factor.
Do not Touch.
evaluate ($ksani1 = $ini_sani1+$sani1_fac)          ! Actually sets RDC forces for set 1. Do
not Touch.
evaluate ($ksani2 = $ini_sani2+$sani2_fac)          ! Actually sets RDC forces for set 2. Do
not Touch.


{* ========= Read in RDC Constraint files and Set Coefficients ========= *}

sani                            {*Tells xplor what type of data is to be read in.  SANI or DIPO
used for RDCs*}
reset                           {*Clear existing buffer to allow reading new RDCs*}
```

```
nres=2000                    {*Number of RDC constraints used.  Set greater than or equal
to number of constraints*}


class=RDC1
force=$ksani1                  {*sets force constant.  Defined above*}
coefficient=0.0 -6.1 0.29      {*DFS,anisotropy <Da>, Rhombicity <Dr/Da>*}
potential=harmonic               {* set potential <harmonic or square>}
@d17gdp_etbic_rdc_2.tbl                  {*sani restraint table*}
end

sani
class=RDC2
force=$ksani2                            {*sets force constant*}
coefficient=0.0 -11.0 0.41     {*DFS,anisotropy <Da>, Rhombicity <Dr/Da>*}
potential=harmonic               {* set potential <harmonic or square>}
@d17gdp_peg_rdc_2.tbl                    {*sani restraint table*}
end

flags include  sani end


sani
 class RDC1 force $ksani1                    {*initializes forces for RDCs*}
end

sani
class RDC2 force $ksani2
end



{* == Initial Powell Minimization in Cartesian Space using initial force constants= *}

 {* = No Need to Modify Below this Point.  All Variables Defined Above= *}

constraints fix=(resid 600) end
constraints fix=(resid 700) end

sani class RDC1 force $ini_sani1 end
sani class RDC2 force $ini_sani2 end

                    minimize powell
                    nstep= $initial
                    nprint=25
        drop=0.001
```

```
        tolgradient=0.00001
        debug=false
end


  {* =Torsion Angle/Cartesian Minimization with linearly ramping force constants =*}


set message = normal end
set seed = 123456 end

vector do (vx = maxwell(.3)) ( all )
vector do (vy = maxwell(.3)) ( all )
vector do (vz = maxwell(.3)) ( all )

vector do (fbeta = 100.) ( all )
vector do (store1 = decode(resid)) (name ca)
vector show min (store1) (name ca)
evaluate ($first_residue = $result)
vector show max (store1) (name ca)
evaluate ($last_residue = $result)



while ($cycle < $increment) loop ramp
        evaluate ( $cycle = $cycle + 1 )

        evaluate ($ini_sani1 = $ini_sani1)        evaluate ($fin_sani1 = $fin_sani1)
        evaluate ($ini_sani2 = $ini_sani2)        evaluate ($fin_sani2 = $fin_sani2)


        evaluate ($increment=$increment)
        evaluate ($hold=$hold)
        evaluate ($cartesian = $cartesian)

        evaluate ($sani1_fac = ($fin_sani1/$increment)*$cycle)
        evaluate ($sani2_fac = ($fin_sani2/$increment)*$cycle)


        evaluate ($ksani1 = $ini_sani1+$sani1_fac)  ! RDC set 1
        evaluate ($ksani2 = $ini_sani2+$sani2_fac)  ! RDC set 2


                sani class RDC1 force $ksani1 end
                sani class RDC2 force $ksani2 end
```

```
        dyna inte
                group (resid 600 )
                hinge rotate (resid 600)

                group (resid 700 )
                hinge rotate (resid 700)


        itype=powell
        nstep = $hold
        nprint = 25

    set echo off message off end                !Turn off echo to log

  ! set rigid groups.  reads through coordinate file and sets groups
  !based on the following definitions.

        evaluate ( $res = $first_residue )
        while ($res le $last_residue) loop group
        group (resid $res and resname PHE and
        (name CG or name CD1 or name CD2 or name CE1 or name CE2 or name CZ))
        group (resid $res and resname PRO and
          (name N or name CA or name CB or name CG or name CD))
        group (resid $res and resname HIS and
          (name CG or name ND1 or name CD2 or name CE1 or name NE2))
        group (resid $res and resname TYR and
        (name CG or name CD1 or name CD2 or name CE1 or  name CE2 or name
         CZ))
        group (resid $res and resname TRP and
          (name CG or name CD1 or name CD2 or name NE1 or
                name CE2 or name CE3 or name CZ2 or name CZ3 or
                name CH2))


    evaluate ( $res = $res + 1 )
        end loop group

        set echo on message on end                  !Turn on echo to log

        auto torsion
        end

constraints fix=(resid 600) end
constraints fix=(resid 700) end
```

```
                        minimize powell
                        nstep= $cartesian
                        nprint=25
        drop=0.001
        tolgradient=0.00001
        debug=false

 end

end loop ramp


{* == Powell Minimization in Cartesian Space using high force constants == *}

constraints fix=(resid 600) end
constraints fix=(resid 700) end

evaluate ($cart_final = $cart_final)

evaluate ($final_sani1 = $fin_sani1*4)
evaluate ($final_sani2 = $fin_sani2*2)

sani class RDC1 force $final_sani1 end
sani class RDC2 force $final_sani2 end

                        minimize powell
                        nstep= $cart_final
                        nprint=25
                        drop=0.001
                        tolgradient=0.00001
                        debug=false

 end


{* Cartesian Space Minimization NO RDCs (return to ideal peptide plane geometry) *}

constraints fix=(resid 600) end
constraints fix=(resid 700) end

evaluate ($sani1_nordc = 0.0)
evaluate ($sani2_nordc = 0.0)

sani class RDC1 force $sani1_nordc end
sani class RDC2 force $sani2_nordc end
```

```
                    minimize powell
                    nstep=500
                    nprint=25
                drop=0.001
                tolgradient=0.00001
                debug=false

 end

 {* === Torsion Angle Minimization with High RDC force constants == *}


set message = normal end
set seed = 123456 end

vector do (vx = maxwell(.3)) ( all )
vector do (vy = maxwell(.3)) ( all )
vector do (vz = maxwell(.3)) ( all )

vector do (fbeta = 100.) ( all )
vector do (store1 = decode(resid)) (name ca)
vector show min (store1) (name ca)
evaluate ($first_residue = $result)
vector show max (store1) (name ca)
evaluate ($last_residue = $result)


        evaluate ($final_sani1 = $fin_sani1*4)
        evaluate ($final_sani2 = $fin_sani2*2)


                sani class RDC1 force $final_sani1 end
                sani class RDC2 force $final_sani2 end

        dyna inte
                group (resid 600 )
                hinge rotate (resid 600)

                group (resid 700 )
                hinge rotate (resid 700)


        itype=powell
        nstep = 750
        nprint = 25
```

```
    set echo off message off end                !Turn off echo to log

 ! set rigid groups.  reads through coordinate file and sets groups
  !based on the following definitions.

        evaluate ( $res = $first_residue )
        while ($res le $last_residue) loop group
        group (resid $res and resname PHE and
        (name CG or name CD1 or name CD2 or name CE1 or name CE2 or name CZ))
        group (resid $res and resname PRO and
          (name N or name CA or name CB or name CG or name CD))
        group (resid $res and resname HIS and
          (name CG or name ND1 or name CD2 or name CE1 or name NE2))
        group (resid $res and resname TYR and
          (name CG or name CD1 or name CD2 or name CE1 or
                name CE2 or name CZ))
        group (resid $res and resname TRP and
          (name CG or name CD1 or name CD2 or name NE1 or
                name CE2 or name CE3 or name CZ2 or name CZ3 or
                name CH2))


    evaluate ( $res = $res + 1 )
          end loop group

          set echo on message on end                !Turn on echo to log

          auto torsion

         end


 {* ========== Write Coordinates ========== *}

print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
  print thres=5. angles
  evaluate ($rms_angles=$result)
  print thres=5. impropers
  evaluate ($rms_impropers=$result)
  couplings print threshold 0.5 class phi end

  write coordinates output =$file end


stop
```

## Appendix B. Chemical Shift Assignments for Δ17Arf1•GDP

| AA Type | Residue # | HN | N | $C_\alpha$ | $C_\beta$ |
|---|---|---|---|---|---|
| MET | 18 | 8.77 | 123.372 | | |
| ARG | 19 | 8.789 | 122.915 | 58.3 | 30.07 |
| ISO | 20 | 8.73 | 125.659 | | |
| LEU | 21 | 8.17 | 127.442 | 52.77 | |
| MET | 22 | 9.067 | 124.681 | 56.69 | |
| VAL | 23 | | | | |
| GLY | 24 | 6.56 | 101.375 | 43.79 | |
| LEU | 25 | 10.196 | 123.462 | 55.76 | 41.17 |
| ASP | 26 | 8.778 | 118.224 | 55.6 | 40.47 |
| ALA | 27 | 10.702 | 121.961 | 53.86 | 17.18 |
| ALA | 28 | 7.534 | 121.959 | 55.51 | 20.06 |
| GLY | 29 | 8.6 | 101.505 | 45.52 | |
| LYS | 30 | 9.748 | 121.425 | 60.03 | |
| THR | 31 | 9.26 | 121.91 | | |
| THR | 32 | 8.85 | 118.169 | 68.13 | 70 |
| ISO | 33 | 7.852 | 116.486 | 58.04 | |
| LEU | 34 | 9.152 | 128.121 | 55.9 | 43.5 |
| TYR | 35 | 8.155 | 115.1 | 61.41 | 38.12 |
| LYS | 36 | 7.739 | 114.581 | 56 | |
| LEU | 37 | 7.678 | 115.456 | 58.75 | 42.14 |
| LYS | 38 | 7.584 | 119.624 | 53.31 | 32.51 |
| LEU | 39 | 7.916 | 114.397 | 56.65 | |
| GLY | 40 | 7.08 | 102.1 | 43.18 | |
| GLU | 41 | 8.489 | 120.839 | | |
| ISO | 42 | 8.259 | 123.097 | 65.01 | |
| VAL | 43 | | | | |
| THR | 44 | | | | |
| THR | 45 | 8.938 | 120.12 | 61.05 | |
| ISO | 46 | 8.344 | 120.754 | 58.06 | |
| PRO | 47 | | | | |
| THR | 48 | 7.049 | 109.48 | | |
| ISO | 49 | 7.861 | 124.209 | | |
| GLY | 50 | 8.784 | 114.41 | | |
| PHE | 51 | 8.265 | 123.274 | | |
| ASN | 52 | 7.463 | 116.146 | 57.31 | |
| VAL | 53 | 8.684 | 118.331 | 60.6 | 40.02 |
| GLU | 54 | 8.285 | 120.398 | 56.75 | |

| | | | | | |
|---|---|---|---|---|---|
| THR | 55 | 8.133 | 117.862 | 62.1 | |
| VAL | 56 | 8.476 | 122.214 | | |
| GLU | 57 | 8.463 | 124.733 | 54.95 | 32.63 |
| TYR | 58 | 9.279 | 124.844 | 53.26 | 38.45 |
| LYS | 59 | 8.016 | 115.238 | 57.56 | |
| ASN | 60 | 8.905 | 117.027 | 54.15 | 38.87 |
| ISO | 61 | 8.705 | 121.223 | 61.17 | 40.01 |
| SER | 62 | | | | |
| PHE | 63 | 9.261 | 125.385 | | |
| THR | 64 | 8.283 | 118.372 | 63.73 | 64.3 |
| VAL | 65 | 9.403 | 125.959 | 60.32 | 35.56 |
| TRP | 66 | 7.49 | 122.996 | | |
| ASP | 67 | 8.801 | 115.682 | 53.38 | 41.49 |
| VAL | 68 | 7.85 | 112.853 | 59.45 | |
| GLY | 69 | 8.357 | 106.493 | 44.86 | |
| GLY | 70 | 8.769 | 108.814 | 46.09 | |
| GLN | 71 | 8.563 | 118.063 | 56.13 | 29.11 |
| ASP | 72 | 8.402 | 119.373 | 59.14 | 40.67 |
| LYS | 73 | 7.16 | 116.484 | 57.84 | 27.12 |
| ISO | 74 | | | | |
| ARG | 75 | 9.598 | 115.24 | 58.85 | 30.14 |
| PRO | 76 | | | | |
| LEU | 77 | | | | |
| TRP | 78 | | | | |
| ARG | 79 | 6.552 | 120.899 | | |
| HIS | 80 | 9.01 | 115.938 | 55.71 | 27.69 |
| TYR | 81 | 7.522 | 114.705 | 59.3 | 38.79 |
| PHE | 82 | 7.576 | 113.988 | 56.44 | 38.22 |
| GLN | 83 | 7.545 | 117.227 | 57.66 | 28.85 |
| ASN | 84 | 8.04 | 115.813 | 53.47 | 38.87 |
| THR | 85 | 7.932 | 114.482 | | 71.98 |
| GLN | 86 | | | | |
| GLY | 87 | 7.687 | 103.432 | 44.76 | |
| LEU | 88 | 9.247 | 126.812 | | |
| ISO | 89 | 9.582 | 123.801 | 58.9 | 39.42 |
| PHE | 90 | 9.302 | 128.045 | 58.05 | 40.86 |
| VAL | 91 | 8.404 | 127.258 | 61.51 | 32.78 |
| VAL | 92 | 8.811 | 122.048 | 59.42 | 36.33 |
| ASP | 93 | 9.153 | 122.924 | 53.54 | 40.96 |
| SER | 94 | 8.469 | 122.734 | 61.33 | 63.11 |
| ASN | 95 | 9.149 | 116.549 | 53.13 | 42.2 |

| | | | | | |
|---|---|---|---|---|---|
| ASP | 96 | 6.956 | 118.116 | 51.81 | 37.15 |
| ARG | 97 | 8.081 | 123.037 | 58.99 | 29.94 |
| GLU | 98 | 8.49 | 117.383 | 63.78 | 36.48 |
| ARG | 99 | | | | |
| VAL | 100 | 7.074 | 120.358 | 65.56 | 31.42 |
| ASN | 101 | 8.386 | 118.455 | 56.4 | 40.83 |
| GLU | 102 | 7.717 | 121.532 | 58.95 | 28.74 |
| ALA | 103 | 8.09 | 120.77 | 56 | 17.75 |
| ARG | 104 | 8.159 | 115.58 | 60.07 | 29.63 |
| GLU | 105 | 8.116 | 116.59 | 59.52 | 29.12 |
| GLU | 106 | 8.338 | 116.083 | 58.49 | 29.2 |
| LEU | 107 | 8.441 | 121.859 | 58.04 | 32.88 |
| MET | 108 | 8.488 | 115.563 | 57.28 | |
| ARG | 109 | 7.991 | 118.187 | 59.54 | 30.53 |
| MET | 110 | 7.743 | 117.484 | 58.64 | 32.05 |
| LEU | 111 | 8.044 | 112.844 | 55.99 | 40.5 |
| ALA | 112 | 7.16 | 118.817 | 52.3 | 19.11 |
| GLU | 113 | 7.38 | 117.553 | 55.06 | 28.47 |
| ASP | 114 | | | | |
| GLU | 115 | 7.387 | 115.842 | 57.61 | 30.22 |
| LEU | 116 | | | | |
| ARG | 117 | 7.507 | 121.81 | 60.1 | 30.57 |
| ASP | 118 | 8.53 | 113.815 | 53.92 | 41.71 |
| ALA | 119 | 7.232 | 121.072 | 52.9 | 18.13 |
| VAL | 120 | 7.572 | 121.669 | 62.57 | 32.78 |
| LEU | 121 | 8.577 | 125.689 | 53.71 | 45.71 |
| LEU | 122 | 9.252 | 127.623 | 52.5 | 46.19 |
| VAL | 123 | | | | |
| PHE | 124 | 8.706 | 124.27 | 52.7 | 39.65 |
| ALA | 125 | 8.686 | 127.847 | 50.39 | 17.02 |
| ASN | 126 | 8.738 | 120.474 | 52.8 | 41.23 |
| LYS | 127 | 7.656 | 115.666 | 57.7 | 28.61 |
| GLN | 128 | 8.007 | 111.606 | 57.1 | 28.89 |
| ASP | 129 | 9.768 | 112.726 | 54.32 | 40.68 |
| LEU | 130 | 7.272 | 121.108 | 53.29 | 39.64 |
| PRO | 131 | | | | |
| ASN | 132 | 8.582 | 113.774 | 52.55 | 38.5 |
| ALA | 133 | 7.348 | 120.755 | 53.36 | 18.56 |
| MET | 134 | 8.231 | 121.419 | 56.54 | 35.94 |
| ASN | 135 | 8.698 | 117.612 | 51.62 | 38.83 |
| ALA | 136 | 8.915 | 118.121 | 55.88 | 18.1 |

| | | | | | |
|---|---|---|---|---|---|
| ALA | 137 | 8.421 | 122.777 | 55.7 | 17.71 |
| GLU | 138 | 8.398 | 121.218 | 58.48 | 29.49 |
| ISO | 139 | 8.48 | 118.179 | 63.8 | 36.46 |
| THR | 140 | 8.219 | 115.752 | 67.61 | |
| ASP | 141 | 7.419 | 119.608 | 57.48 | 41.35 |
| LYS | 142 | 8.597 | 118.095 | 58.66 | 33.15 |
| LEU | 143 | 8.327 | 113.469 | 55.49 | 41.15 |
| GLY | 144 | 7.349 | 102.484 | 46.37 | |
| LEU | 145 | 7.218 | 114.543 | 56.87 | 38.22 |
| HIS | 146 | 8.555 | 110.155 | 58.17 | 28.81 |
| SER | 147 | 7.7 | 111.978 | 59.01 | 63.72 |
| LEU | 148 | 7.127 | 120.481 | 55.12 | 41.05 |
| ARG | 149 | 8.467 | 121.087 | 55.1 | 32.02 |
| HIS | 150 | 7.82 | 120.838 | 52.67 | 29.57 |
| ARG | 151 | 7.914 | 116.915 | 53.79 | 32.67 |
| ASN | 152 | 9.321 | 127.003 | 52.83 | 38.53 |
| TRP | 153 | 7.91 | 120.171 | 54.022 | 34.98 |
| TYR | 154 | 8.149 | 120.315 | 56.92 | 40.2 |
| ISO | 155 | 7.112 | 122.05 | 57.74 | 41.21 |
| GLN | 156 | 8.865 | 125.145 | 52.62 | 33.14 |
| ALA | 157 | 9.046 | 132.056 | 52.54 | 19.29 |
| THR | 158 | 8.712 | 114.815 | 60.69 | 72.83 |
| CYS | 159 | 8.301 | 119.737 | 56.65 | 30.1 |
| ALA | 160 | 9.462 | 131.228 | 55.43 | 20.21 |
| THR | 161 | 7.11 | 102.097 | 63.54 | 67.02 |
| SER | 162 | 7.576 | 113.988 | 58.14 | 64.95 |
| GLY | 163 | 8.298 | 111.22 | 45.5 | |
| ASP | 164 | 7.877 | 122.67 | 56.2 | 41.31 |
| GLY | 165 | 8.626 | 112.175 | 47.48 | |
| LEU | 166 | | | | |
| TYR | 167 | 9.21 | 120.19 | 60.21 | |
| GLU | 168 | 9.532 | 120.837 | 90.89 | 28.5 |
| GLY | 169 | 8.184 | 105.417 | 48.19 | |
| LEU | 170 | 8.299 | 119.18 | 56.21 | 41.22 |
| ASP | 171 | 9.332 | 124.563 | 57.43 | |
| TRP | 172 | | | | |
| LEU | 173 | 8.088 | 118.581 | 58.33 | 42.8 |
| SER | 174 | 8.768 | 112.354 | 61.92 | 62.86 |
| ASN | 175 | 7.807 | 117.474 | 56 | 38.25 |
| GLN | 176 | 7.556 | 116.096 | 57.66 | 28.28 |
| LEU | 177 | 7.749 | 118.116 | | |

| | | | | | |
|---|---|---|---|---|---|
| ARG | 178 | 7.683 | 118.309 | | |
| ASN | 179 | 7.42 | 113.876 | 53.4 | 41.7 |
| GLN | 180 | 7.47 | 117.15 | 57.6 | |
| LYS | 181 | 7.998 | 126.448 | 57.72 | 33.54 |

## Appendix C. Residual Dipolar Couplings for Δ17Arf1•GDP

| AA Type | Residue # | Ether Bicelle Dipolar Coupling | $C_8E_5$ : Octanol Dipolar Coupling |
|---|---|---|---|
| MET | 18 | | |
| ARG | 19 | 1.019 | 8.223 |
| ISO | 20 | 2.374 | 8.248 |
| LEU | 21 | 1.308 | 14.069 |
| MET | 22 | 2.9244 | 3.541 |
| VAL | 23 | | |
| GLY | 24 | 0.989 | 6.61995 |
| LEU | 25 | 6.7447 | 4.3395 |
| ASP | 26 | -0.785 | |
| ALA | 27 | 0.3255 | -9.0887 |
| ALA | 28 | -0.617 | 1.064 |
| GLY | 29 | 0.208 | |
| LYS | 30 | | |
| THR | 31 | 0.9454 | |
| THR | 32 | | |
| ISO | 33 | -2.036 | |
| LEU | 34 | | |
| TYR | 35 | -0.837 | -10.379 |
| LYS | 36 | | |
| LEU | 37 | 0.153 | -39.11 |
| LYS | 38 | 0.736 | |
| LEU | 39 | 1.481 | |
| GLY | 40 | -1.428 | |
| GLU | 41 | -8.9788 | 7.165 |
| ISO | 42 | -0.15805 | |
| VAL | 43 | | |
| THR | 44 | | 11.237 |
| THR | 45 | 0.229 | -4.251 |
| ISO | 46 | 6.888 | |
| PRO | 47 | | |
| THR | 48 | 0.9454 | |
| ISO | 49 | 1.324 | |
| GLY | 50 | 4.006 | |
| PHE | 51 | -3.577 | 9.304 |

| | | | |
|---|---|---|---|
| ASN | 52 | | |
| VAL | 53 | 24.203 | |
| GLU | 54 | 16.822 | 12.991 |
| THR | 55 | 2.853 | 10.645 |
| VAL | 56 | -0.039 | -6.35 |
| GLU | 57 | 4.328 | 14.463 |
| TYR | 58 | 6.876 | |
| LYS | 59 | 1.778 | 2.142 |
| ASN | 60 | | |
| ISO | 61 | | |
| SER | 62 | | |
| PHE | 63 | 2.892 | |
| THR | 64 | 5.036 | |
| VAL | 65 | 2.872 | |
| TRP | 66 | -3.735 | 4.348 |
| ASP | 67 | | |
| VAL | 68 | 0.064 | |
| GLY | 69 | 2.658 | |
| GLY | 70 | 0.276 | -0.004 |
| GLN | 71 | 3.979 | 0.096 |
| ASP | 72 | 0.681 | |
| LYS | 73 | 4.686 | 2.678 |
| ISO | 74 | | |
| ARG | 75 | 0.432 | 1.414 |
| PRO | 76 | | |
| LEU | 77 | | |
| TRP | 78 | | |
| ARG | 79 | -4.117 | -1.423 |
| HIS | 80 | | |
| TYR | 81 | -4.958 | |
| PHE | 82 | | |
| GLN | 83 | 1.607 | 4.455 |
| ASN | 84 | 0.502 | |
| THR | 85 | -4.082 | |
| GLN | 86 | | |
| GLY | 87 | 3.696 | 10.846 |
| LEU | 88 | | 4.943 |
| ISO | 89 | 2.618 | 5.961 |
| PHE | 90 | 3.069 | 10.462 |
| VAL | 91 | -0.482 | 9.005 |
| VAL | 92 | -4.595 | 3.989 |

| | | | |
|---|---|---|---|
| ASP | 93 | -0.285 | 3.714 |
| SER | 94 | -1.491 | |
| ASN | 95 | 7.141 | 1.724 |
| ASP | 96 | 0.467 | -12.499 |
| ARG | 97 | 0.844 | -2.928 |
| GLU | 98 | 3.394 | 5.257 |
| ARG | 99 | | |
| VAL | 100 | -0.781 | -8.516 |
| ASN | 101 | -4.122 | |
| GLU | 102 | -2.059 | -5.64 |
| ALA | 103 | -1.006 | -13.383 |
| ARG | 104 | -2.105 | -8.019 |
| GLU | 105 | -5.078 | |
| GLU | 106 | -0.731 | |
| LEU | 107 | 4.481 | 3.425 |
| MET | 108 | 0.263 | -10.264 |
| ARG | 109 | -0.087 | 30.661 |
| MET | 110 | 0.547 | -3.883 |
| LEU | 111 | -4.375 | 4.405 |
| ALA | 112 | -1.27 | -6.833 |
| GLU | 113 | 4.723 | |
| ASP | 114 | | |
| GLU | 115 | | |
| LEU | 116 | | |
| ARG | 117 | | |
| ASP | 118 | 1.938 | |
| ALA | 119 | -1.887 | 0.404 |
| VAL | 120 | 0.835 | -15.189 |
| LEU | 121 | 1.746 | |
| LEU | 122 | 5.468 | 15.216 |
| VAL | 123 | | |
| PHE | 124 | 1.542 | 5.942 |
| ALA | 125 | -0.134 | 0.706 |
| ASN | 126 | 1.584 | 5.089 |
| LYS | 127 | -21.999 | |
| GLN | 128 | -0.085 | |
| ASP | 129 | -3.338 | -13.772 |
| LEU | 130 | -2.114 | -3.889 |
| PRO | 131 | | |
| ASN | 132 | 1.423 | |
| ALA | 133 | 0.995 | 2.887 |

| | | | |
|---|---|---|---|
| MET | 134 | 1.623 | |
| ASN | 135 | -2.831 | -3.116 |
| ALA | 136 | 0.64 | |
| ALA | 137 | 12.84885 | |
| GLU | 138 | 3.302 | |
| ISO | 139 | 2.172 | 3.527 |
| THR | 140 | 0.143 | 2.585 |
| ASP | 141 | 1.695 | 4.517 |
| LYS | 142 | 2.12 | |
| LEU | 143 | 1.682 | 5.152 |
| GLY | 144 | -1.432 | -5.595 |
| LEU | 145 | 0.897 | -1.849 |
| HIS | 146 | -4.592 | -12.334 |
| SER | 147 | | |
| LEU | 148 | 3.075 | -0.857 |
| ARG | 149 | | |
| HIS | 150 | | |
| ARG | 151 | | |
| ASN | 152 | -6.197 | 2.024 |
| TRP | 153 | -0.195 | 10.027 |
| TYR | 154 | 4.993 | -5.416 |
| ISO | 155 | 1.558 | 8.065 |
| GLN | 156 | 11.225 | 10.296 |
| ALA | 157 | 0.115 | 7.59255 |
| THR | 158 | -3.287 | -7.673 |
| CYS | 159 | | |
| ALA | 160 | -4.367 | -8.599 |
| THR | 161 | 0.16175 | -5.829 |
| SER | 162 | | |
| GLY | 163 | 0.423 | -7.584 |
| ASP | 164 | -1.694 | |
| GLY | 165 | -0.132 | 5.131 |
| LEU | 166 | | |
| TYR | 167 | | |
| GLU | 168 | -0.936 | -9.379 |
| GLY | 169 | 1.246 | -0.939 |
| LEU | 170 | | |
| ASP | 171 | | |
| TRP | 172 | | |
| LEU | 173 | 0.91 | -4.398 |
| SER | 174 | -1.034 | -1.807 |

| | | | |
|---|---|---|---|
| ASN | 175 | | |
| GLN | 176 | -0.732 | -9.073 |
| LEU | 177 | -0.204 | |
| ARG | 178 | -1.14 | |
| ASN | 179 | | |
| GLN | 180 | 4.471 | 2.569 |
| LYS | 181 | 2.994 | 17.347 |