Characterization of the Golli-interacting Protein and its interactions with BG21 and the 21.5-kDa isoform of Myelin Basic Protein

By
Sergio Jaramillo Tatis

A Thesis
Presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Master of Science
In
Molecular and Cellular Biology

Guelph, Ontario, Canada

© Sergio Jaramillo Tatis, 2013
ABSTRACT

Characterization of the Golli-Interacting Protein and its interactions with BG21 and the 21.5-kDa isoform of MBP

Sergio Jaramillo Tatis
University of Guelph

Advisor: Professor G. Harauz

The gene in oligodendrocyte lineage (golli) is responsible for a number of intrinsically disordered proteins including the understudied golli protein BG21 and the 21.5-kDa isoform of the myelin-basic-protein (MBP-21.5). The golli-interacting-protein (GIP) was discovered in a search for potential BG21 interacting-partners. It was shown to interact with BG21, although the inability to produce a full-length version of GIP led to inconclusive results. GIP is an acidic phosphatase belonging to the family of RNA-polymerase-2 C-terminal-phosphatases. Here, a novel protocol for the purification of a functional full-length GIP has been developed with a yield of 4-mg/L culture with 85% purity. An enzymological approach was used to study the GIP-BG21 and GIP-MBP-21.5 interactions. No GIP-MBP-21.5 interactions were found. The phosphorylation state of BG21 was important for the GIP-BG21 interactions. BG21 enhanced GIP ($K_a=30$ μM) whereas Protein-Kinase-C-phosphorylated BG21 inhibited GIP ($K_i=3.5$ μM, $K_i'=8.0$ μM), suggesting a potential role of BG21 and GIP in transcription regulation.
ACKNOWLEDGEMENTS

It has been an arduous and challenging two and a-bit years. This was one of the most difficult challenges I have endured. It required an immense amount of thought, stamina, and most of all, a lot of patience. I suppose this all means that I actually did work in science and research for over the past two years.

I have a slight, just a slight, feeling that none of this would have been possible without the help of so many, beginning with my supervisor, Dr. George Harauz. He provided me the opportunity to use his lab, and resources to go off in the lab and do my enzymology with GIP, while my lab-mates worked on MBP.

Speaking of my lab-mates, both former and current, they all deserve a big hug for their company, conversations, and advice throughout this whole chapter. Janine Voyer, Dr. Graham Smith, Dr. Luna Rahman, Lauren Seymour, Marie-Lise Jobin, Agata Zienowicz, Ehsan Fayaz, Caroline Velte, Erin Stephenson, Danielle Lanthier, Andrew Jenkins, Dr.2b Kirilo Bessonov, Dr.2b Ivana Komljenovic, Dr.2b Miguel De Avila, Dr. Kenrick (Ricky) Vassall, and Dr. Vladimir (Vova) Bamm.

V. Bamm in particular has been my mentor during my stay here in this lab. From my first day as an undergrad student, he taught me almost everything I know regarding protein chemistry and purification. As well, Dr. Vassall has been a great company in the office, and a resource of knowledge, helping to bounce ideas back and forth with this project.

To my advisory committee, Dr. Rod Merrill and Dr. Matt Kimber, many thanks are in place. Both provided a lot of helpful insight into my project. More importantly, each one played a significant role in my development as a biochemist. Dr. Merrill was my instructor in my undergrad and taught me my enzymology. Dr. Kimber gave me an opportunity to work in his lab for one summer, where I had the privilege of learning about protein purification and crystallography.

Of course my family and friends also deserve an important mention here. Although we did not see each other often, the conversations we had over the phone with my mother and father were calming and a much needed break from it all. My friends provided a lot of fun moments and good breaks.

Last, but most certainly not least, my family here in Guelph, my dear Pepper and Stephanie. Pepper, a silly little black Labrador, she came into my life almost two years ago. Her furry company has helped relieve a lot of stress and induced many smiles over this time. However, the person that deserves the highest praise is my dear Stephanie. She has sacrificed so much of her sanity over the last two-plus years to help me through this process. Ya casi corazon. The rock in my life, the girl I will be proud to call my wife on September 14 2013. Gracias.

Sergio Jaramillo Tatis
University of Guelph
August 2013
# Table of Contents

ABSTRACT ........................................................................................................................................... ii

ACKNOWLEDGEMENTS ....................................................................................................................... iii

LIST OF FIGURES AND TABLES .......................................................................................................... vii

ABBREVIATIONS ................................................................................................................................. viii

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

1.1 Multiple Sclerosis ......................................................................................................................... 2

1.2 Myelin ............................................................................................................................................ 3

1.3 The golli gene ............................................................................................................................... 5

1.3.1 Intrinsically-disordered proteins ......................................................................................... 7

1.3.2 Classic Myelin Basic Proteins ............................................................................................. 8

1.3.2.1 MBP-21.5 is different from other classic MBPs ............................................................ 10

1.3.3 The golli proteins ............................................................................................................... 11

1.4 The Golli Interacting Protein Discovery ..................................................................................... 13

1.4.1 NLI-IF/SCP1 orthology .................................................................................................... 15

1.4.1.1 CTD Phosphatases ................................................................................................... 18

1.4.2 GIP is a phosphatase ......................................................................................................... 21

1.5 Objectives and Hypothesis ........................................................................................................ 23

Chapter 2: Purification and Characterization of Full-Length GIP .................................................... 26

2.1 Overview ...................................................................................................................................... 27

2.2 Materials and Methods .............................................................................................................. 28

2.2.1 Purification strategy ........................................................................................................... 28

2.2.2 Enzymology ....................................................................................................................... 31

2.2.3 Full-length GIP expression and purification ...................................................................... 33

2.2.4 Protocol for determining GIP kinetics ............................................................................... 35

2.3 Results ........................................................................................................................................ 36

2.3.1 Expression of SUMO-GIP ................................................................................................. 36

2.3.2 Cell lysis and Solubility ..................................................................................................... 36

2.3.3 Purification of the SUMO-GIP Fusion ................................................................................. 41

2.3.4 Cleaving the SUMO-GIP Fusion and Purifying GIP .......................................................... 41

2.3.5 Enzymatic Characterization Using an Artificial Substrate .................................................. 43

2.4 Discussion .................................................................................................................................. 48
Appendix I

Chapter 4: Conclusions and Future Directions

4.1 General conclusions .................................................................................. 101
4.2 Future directions ....................................................................................... 104

Bibliography .................................................................................................. 107

Appendix I ..................................................................................................... 116
i. General mixed inhibition mechanism ................................................................. 116
ii. Dixon plot ........................................................................................................... 117
iii. Cornish-Bowden plot ...................................................................................... 118
Appendix II ............................................................................................................. 119
Figure A2.1 ............................................................................................................ 119
LIST OF FIGURES AND TABLES

Chapter 1

- Figure 1.1 - Myelin Sheath ................................................................. 12
- Figure 1.2 - golli gene ................................................................. 14
- Figure 1.3 - Multiple Sequence Alignment of GIP .......................... 24
- Figure 1.4 - Classifying SCP1 .................................................. 28
- Figure 1.5 - Reaction Snapshot .................................................. 30

Chapter 2

- Figure 2.1 - Expression of SUMO-GIP .................................... 45
- Figure 2.2 - Solubilization of SUMO-GIP .................................. 46
- Figure 2.3 - Effect of DTT on SUMO-GIP ............................ 48
- Figure 2.4 - Purification of SUMO-GIP .................................... 50
- Figure 2.5 - Cleavage of SUMO-GIP and GIP purification .... 52
- Figure 2.6 - Velocity curve for GIP with p-NPP ..................... 54
- Figure 2.7 - Effect of pH on GIP ............................................. 55
- Figure 2.8 - Secondary structure prediction of GIP ............ 58
- Table 2.1 - Summary of kinetic parameters ......................... 70

Chapter 3

- Figure 3.1 - Malachite green .................................................. 79
- Figure 3.2 - Effect of BG21 and MBP-21.5 on GIP ............... 82
- Figure 3.3 - Inhibition of GIP by BG21 .................................. 83
- Figure 3.4 - Linear analysis of BG21 inhibition ....................... 84
- Figure 3.5 - Natural substrates ............................................. 86
- Figure 3.6 - Velocity curve for GIP with CTD-peptide .......... 88
- Figure 3.7 - Effect of BG21 on GIP with CTD-peptide ......... 89
- Figure 3.8 - Effect of PO₄-BG21 with CTD-peptide .......... 91
- Figure 3.9A - Model of GIP with p-NPP and CTD-peptide .... 101
- Figure 3.9B - Model of BG21 interactions with GIP and p-NPP ... 102
- Figure 3.9C - Model of BG21 interactions with GIP and CTD-peptide ... 103
- Figure 3.9D - Model of PO₄-BG21 interactions with GIP and CTD-peptide ... 104

Appendix II

- Figure A2.1 - Preliminary CD analysis and thermal melting of ΔN-GIP .... 127
ABBREVIATIONS

CCVI - Chronic Cerebrospinal Venous Insufficiency
CD - Circular Dichroism
CNS - Central Nervous System
CTD - C-Terminal Domain
DTT - Dithiothreitol
E. coli - Escherichia coli
EAE - Experimental Allergic Encephalomyelitis
FCP1 - first CTD phosphatase 1
FPLC - Fast Performance Liquid Chromatography
GIP - Golli-Interacting Protein
golli - gene in the oligodendrocyte lineage
HAD - Halo Acid Dehalogenase
HPLC - High Pressure Liquid Chromatography
IDP - Intrinsically-disordered protein
IMAC - Immobilized Metal-Affinity Chromatography
IPTG - isopropyl-β-D-thiogalactopyranoside
Kapβ - Karyopherin-β
lacI - lac repressor gene
MAPK - Mitogen-Activated Protein Kinase
MBP - Myelin Basic Protein
MS - Multiple Sclerosis
MWCO - Molecular Weight Cut-Off
Ni^{2+}-NTA - Nickel nitrilo-triacetic acid
NLI-IF - Nuclear LIM-Interacting Interactor Factor
NLS/NES - Nuclear Localization/Export Signal
p-NP - para nitrophenol
p-NPP - para nitrophenol phosphate
PBS - Phosphate Buffered Saline
PCR - Polymerase Chain Reaction
PEG - Polyethylene Glycol
PKC - Protein Kinase C
PLP - Proteolipid Protein
PMSF - phenylmethylsulfonylfluoride
PNS - Peripheral Nervous System
PTM - Post-Translational Modification
RNAP2 - RNA Polymerase-2
SCP1 - Small CTD Phosphatase 1
SDS-PAGE - Sodium Dodecyl Sulfate - Poly-Acrylamide Gel Electrophoresis
SH3 - Src-homology-3
SUMO - Small Ubiquitin-Like Modifier
Chapter 1: Introduction
1.1 Multiple Sclerosis

Jean-Martin Charcot, a French neurologist, was the first person to characterize the disease multiple sclerosis (MS) in 1868. He identified demyelinated lesions with perivascular inflammation along the white matter of the central nervous system (Charcot et al., 1868). To this day, this definition of multiple sclerosis stands strong, indicative of a neurodegenerative and auto-immunological disease.

Since its discovery, various forms of multiple sclerosis have been identified: relapsing-remitting, primary progressive, secondary progressive, progressive relapsing, and the fulminating Marburg's syndrome (MSSC. 2011). The main differences between these types of MS lie in the way the disease progresses over time; either the patient experiences chronic relapses or they deteriorate linearly. Patients can experience various symptoms, including dizziness, depression, numbness, and tingling sensations, at different times of their life (MSSC. 2011). This high variance in symptoms makes MS difficult to establish. MS is most often diagnosed in people between 15 and 40 years of age, with women being three times more likely to be affected than men. Multiple sclerosis in Canada is of high concern, with an average of three new people being diagnosed every day, giving this country one of the highest onset rates in the world (MSSC. 2011).

There is no answer to the question of what causes multiple sclerosis. However, several theories have been postulated, most recently the highly controversial chronic cerebrospinal venous insufficiency (CCSVI) or "liberation" theory proposed by Zamboni et al., in 2008. Genetic screens on MS patients have been unable to localize MS susceptibility to a particular gene, suggesting that it is polygenic (Chataway et al. 1998). On the other hand, there is a vast
inequivalence in the geographical incidence of the disease, suggesting that the environment can play a role as well (Rosati 2001). The more traditional thoughts on the origin of multiple sclerosis involve the degeneration (or failure to regenerate) of the insulating layer around neuronal axons and the generation of auto-antigens (Kim et al. 2003; Husted 2006). For the purpose of this thesis, the latter theory is of relevance. Regardless of the specific theories, one thing appears to be certain: multiple sclerosis is a complex multi-factorial disease dependent on various genetic and environmental factors.

1.2 Myelin

The mammalian nervous system is made up of two major regions, the peripheral nervous system (PNS) and the central nervous system (CNS). The CNS includes the brain and spinal cord, whereas the PNS links the CNS to limbs and organs throughout the body. Both of these components are composed of two main cell types, neurons and glial cells. Neurons are cells in charge of transmitting information across the body through electrochemical impulses, whereas glial cells function to help nurture, maintain, and protect neurons (Baumann and Pham-Dinh 2001).

Oligodendrocytes are the main glial cells in the central nervous system. Their name is derived from the fact that each cell contains a large number of dendritic processes. These processes extend to the axon of a neuronal cell and encase it, forming a compact membranous structure known as myelin (Figure 1.1). A single oligodendrocyte can extend its membrane processes and myelinate multiple axons (Butt et al. 1995). In order to myelinate neurons, oligodendrocytes need to reach maturity through a number of developmental stages (Simons and Lyons 2013; Baumann and Pham-Dinh 2001).
**Figure 1.1.** Depiction of the myelin sheath extending from oligodendrocyte processes wrapping around neuronal axon. Diagram shows the intraperiod (minor dense) lines and major dense lines. Zoom-in box showing the 'classic' MBP-18.5 isoform joining the apposing cytoplasmic oligodendrocyte leaflets.

Figure reproduced from of N. Ishiyama (2000) M.Sc. Thesis with permission.
Myelin is a lipid-rich (70:30, wt:wt, lipid:protein) extension of the oligodendrocyte membrane that wraps the axons forming an insulating sheath, facilitating saltatory conduction and, in turn, the electrochemical impulses carried by neurons. The sheath contains alternating electron-light (minor dense line) and electron-dense (major dense line) regions (Figure 1.1). The major dense line is formed by the apposed cytoplasmic leaflets of the oligodendrocyte membrane, usually 3-4 nm in width (Sjöstrand 1949; Harauz, Ladizhansky, and Boggs 2009). The myelin lipid composition is particularly high in cholesterol (44%) and also high in glycosphingolipids (20%) (Raff, Mirsky, and Fields 1978; Zalc et al. 1981; Baumann and Pham-Dinh 2001). A number of proteins are present in myelin, with the two major ones in the mature sheath being the proteolipid protein (PLP), also known as lipophilin, and myelin basic protein (MBP), accounting for 80% of the total protein by weight (Baumann and Pham-Dinh 2001). These proteins play an important role in the formation and maintenance of the myelin sheath.

1.3 The golli gene

The gene in the oligodendrocyte lineage (golli) codes for several proteins involved in the regulation and compaction of the myelin sheath. At about 105 kb (in mice, 179 kb in humans) in length, golli can be thought of as a large transcriptional unit containing smaller genes (Figure 1.2) (Campagnoni et al. 1993). By utilizing extensive alternative splicing and three individually regulated transcription start sites, it is able to code for a number of different products at different stages of oligodendrocyte development (Campagnoni et al. 1993). Through different splicing patterns of its 11 exons (1, 2, 3, 4, 5a, 5b, 5c, 6, 7, 8, 9a, 9b, 10, 11), this gene is responsible for two families of proteins, the 'classic' myelin basic proteins (arising from the
Figure 1.2. The gene in oligodendrocyte lineage (golli) showing the alternative splicing patterns of the introns and the 11 exons. The golli-MBP proteins arise from the first transcription start site (tss1). The second and third transcription start sites (tss2 and tss3) are responsible for the 'classic' MBPs. Both the conserved 133-amino acid golli domain and MBP domains in BG21 and J37 are shown. TP8 contains only the first 47 amino acids of the conserved golli domain, the remaining 22 amino acids are a unique domain as a result of a frameshift with respect to the downstream classic MBP exons.

Figure adapted with permission from Campagnoni and Skoff, 2001.
third transcriptional start site) and the golli proteins (arising from the first transcriptional start site) (Figure 1.2) (Campagnoni et al. 1993).

1.3.1 Intrinsically-disordered proteins

Every protein's function is result of a well-defined protein structure. This was the dogma that presided over all protein science prior to the late 1990's. It was during this time that researchers started to question the generality of this dogma, as they found more and more proteins with missing electron density in their structures derived from X-ray crystallography. These regions were thought to have no defined structure. Since then, researchers have found many more proteins with these unstructured regions, as well as whole proteins with no defined three-dimensional structure. This rapidly expanding group of proteins is now collectively known as intrinsically-disordered proteins (IDPs) (Dunker et al. 2001).

Intrinsically-disordered proteins lack stable secondary or tertiary structure, though they can adopt one depending on their environment and or interacting partners. A high proportion of charged and proline residues, combined with a low proportion of hydrophobic residues, results in a high net charge and a low mean hydrophobicity, which contributes to their unstructured nature (Wright and Dyson 1999; Uversky, Gillespie, and Fink 2000; Harauz et al. 2004).

The functionality of IDPs is based on them being very flexible, allowing them to interact either simultaneously or independently with several different partners, resulting in various functions (Harauz et al. 2004; Wright and Dyson 1999). They frequently adopt structured elements along regions of the protein upon interactions with a partner (Uversky 2002; Receveur-Bréchot et al. 2006). As an interacting partner binds to the IDP, it increases the propensity of regions within the IDP to undergo a disorder-to-order transition (Tompa et al.
IDPs may have more than one binding region, giving them a much larger binding surface area and also the ability to bind several partners. Another advantage that IDPs have due to these properties is a high specificity for their binding partners combined with a low affinity, leading to faster rates of association and dissociation (Uversky 2009; Dunker et al. 2002). These properties make IDPs important for several tasks such as signal transduction, or acting as molecular scaffolds for other proteins (Wright and Dyson 1999; Harauz et al. 2004).

1.3.2 Classic Myelin Basic Proteins

The myelin basic protein family has been extensively studied for the last 35 years. Five major isoforms of classic MBP arise from a 36 kb single transcript (exons I-VII in classic gene nomenclature; 5b-11 in golli gene) through alternative splicing and are identified by their approximate molecular masses, 14, 17, 17.2, 18.5 and 21.5 kDa (Figure 1.2) (Campagnoni et al. 1993; Barbarese, Carson, and Braun 1978). All MBP proteins belong to the growing family of intrinsically-disordered proteins (IDPs) (Harauz et al. 2004; Harauz, Ladizhansky, and Boggs 2009).

The classic 18.5-kDa isoform of MBP has been the most heavily studied since it is the prevalent form found in the CNS (R. Smith 1992). MBP-18.5 is a peripheral membrane protein with a very high positive net charge of +19 at neutral pH (Kim et al. 2003; Moscarello et al. 1994). However, it is also subjected to numerous post-translational modifications (PTMs) such as phosphorylation, deimination, methylation, and glycosylation. These PTMs all reduce its net charge, resulting in many different charge components (previously called charge isomers) depending on the degree and types of PTMs. These variants can be separated through a cation exchange column and have been named C1 through C8, where C1 (net charge +19) is the least
modified form of MBP-18.5 and C8 is the most modified (net charge +13) (Kim et al. 2003; Moscarello et al. 1994). MBP-18.5 is mainly present in the major dense line of the myelin sheath where it can tightly adhere together the oligodendrocyte membrane leaflets via electrostatic interactions with the negatively-charged phosphate groups of the lipids (Figure 1.1) (Harauz and Boggs 2013; Wood and Moscarello 1997; Harauz, Ladizhansky, and Boggs 2009).

Solution nuclear magnetic resonance spectroscopy and hydrophobic moment studies have shown three segments of the intrinsically-disordered MBP adopting an α-helical structure along the N-terminus (α1, A22-K56), a proline rich central region (α2, S72-S107), and the C-terminus (α3, S133-S159) (Bamm et al. 2011). The central α2-segment contains an immunodominant T-cell recognition site and B-cell epitope as well as being a Src-homology-3 (SH3) domain ligand. In multiple sclerosis, lesions resulting from myelin degeneration contain T-cells and B-cells reactive against myelin antigens (Wucherpfennig et al. 1997), such as this α2 segment in MBP. The MBP-18.5 isoform has been shown to induce autoimmune encephalomyelitis (EAE, a widely accepted murine model for multiple sclerosis) in mice, further emphasising its antigenic properties (Harauz, Ladizhansky, and Boggs 2009).

In MS, phosphorylation of MBP correlates with a decrease in the severity of MS (Yon et al. 1996; Kim et al. 2003). As with other intrinsically-disordered proteins, MBP is multifunctional. Aside from the physical compaction of the myelin sheath, it also acts as a molecular scaffold in myelin. Interactions with SH3-domain proteins, calcium-dependent calmodulin, and with the cytoskeletal proteins actin and tubulin have been studied (Harauz, Ladizhansky, and Boggs 2009; Boggs and Rangaraj 2000; Bamm et al. 2011). It has also been
shown to inhibit calcium influx through voltage operated calcium channels (G. S. T. Smith et al. 2011), and potentially to interact with Fyn tyrosine kinase (G. S. T. Smith, De Avila, et al. 2012).

MBP-18.5, being a peripheral membrane protein, needs to be transported to the myelin membrane. However, this is not a simple task when dealing with such a highly positively-charged molecule. Fortunately, oligodendrocytes have developed a solution to this issue; they translocate MBP mRNA and translate it in regions where myelin compaction is taking place (Colman et al. 1982; Trapp et al. 1987). This process only occurs after oligodendrocyte process extension and myelination around axons (Campagnoni et al. 1993).

1.3.2.1 MBP-21.5 is different from other classic MBPs

Isoforms of MBP containing exon II (6 in golli gene numbering) (17 and 21.5-kDa) are expressed much earlier in the development of oligodendrocytes and myelin formation (Kamholz, Toffenetti, and Lazzarini 1988; Roth et al. 1987). Interestingly, immunofluorescence studies have shown that these isoforms are not myelin-associated peripheral membrane proteins, and instead they are localized to the nucleus (Pedraza et al. 1997). MBP-21.5 is the prevalent isoform in the nucleus. These isoforms have also been found in the cytoplasm of the radial component of myelin where actin and tubulin are also present (Karthigasan et al. 1994). Pedraza et al. (1997) showed this dual localization and also showed active transport of MBP-21.5 into the nucleus from the cytoplasm, demonstrating this process to be energy- and temperature-dependent. These workers also speculated that phosphorylation of serine in exon II may inhibit this process, since phosphorylation of other nuclear proteins has been observed to block their nuclear import (Miller et al. 1991; Moll et al. 1991).
The karyopherin-β (Kapβ) family of proteins are the major mediators in the nucleocytoplasmic transport of proteins. They operate by recognizing a nuclear localization (NLS) and/or export (NES) signal in proteins, then coupling with guanosine tri-phosphate (GTP) to shuttle their cargo molecules across the nuclear pores (Chook and Süel 2010). Exon II (classic gene numbering; exon 6 in golli nomenclature) codes for an additional 26 amino acids which contain a non-traditional PY nuclear localization signal (G. S. T. Smith, Seymour, et al. 2012).

The role of MBP-21.5 in the nucleus has yet to be determined; however, it is currently being studied. The intriguing timing of the expression of MBP-21.5 and its transport to the nucleus just before myelination and myelin compaction leads to the idea of it having some sort of regulatory function in this process. Furthermore, ongoing studies are showing that MBP-21.5 has an effect on oligodendrocyte proliferation, even in non-expressing cells in the same culture dish (G. S. T. Smith et al. 2013). This property suggests that it may induce or regulate the formation of a secreted factor that induces cell proliferation.

1.3.3 The golli proteins

In 1993, Campagnoni et al. found a series of genes upstream of the well known MBP genes in mice. They showed that MBP was only part of the larger golli gene described above. Three proteins (BG21, TP8, and J37 in the mouse) arise from the genes upstream of what they termed classic MBP, and are jointly known as the golli proteins. These proteins share exons with the downstream classic MBPs, and both BG21 and J37 are similar to MBP. Unlike the other two golli proteins, TP8 is not similar to MBP, due to a frameshift with respect to the MBP coding sequence (Figure 1.2) (Campagnoni et al. 1993). The three golli proteins share a common N-terminal golli domain made up of 133 amino acids (TP8 only has the first 47). This region is
translated from the first three exons of the *golli* gene (Campagnoni et al. 1993). The main differences between the golli proteins arise from the different exons shared with the MBP region downstream. The golli isoform J37 utilizes exons 5b, 7, 8, and 11 to make up a 117 (1-102 and 155-168) amino acid MBP domain, for a total of 250 residues (27 kDa) (Figure 1.2) (Campagnoni et al. 1993). On the other hand, the golli isoform BG21 uses exons 5b and 5c to code for 62 amino acids, including the first 56 residues of the classic MBPs, for a total of 195 amino acids (21 kDa) (Figure 1.2) (Campagnoni et al. 1993).

Campagnoni *et al.* (1993) also found that these proteins are expressed at the onset and peak of myelination in mice, before expression of the membrane-associated MBPs and myelin compaction. Both TP8 and BG21 are expressed early in embryonic life, whereas J37 is expressed neonatally before full oligodendrocyte differentiation. The golli proteins have been found to be expressed in tissues and cells other than the CNS and oligodendrocytes, such as neurons, thymus, spleen, macrophages, and both T and B cells (Pribyl et al. 1993). Interestingly, golli proteins have been localized in the nucleus and cytoplasm, and in some tissues this localization is developmentally dependent (Pribyl et al. 1996; Landry et al. 1996) just like exon II-containing MBPs. Although the precise role of golli proteins has yet to be elucidated, it is speculated that they are involved in cellular signalling due to their nucleocytoplasmic shuttling and interacting partners (Landry et al. 1996; Fernandes et al. 2004). Reyes and Campagnoni showed two functional domains within golli proteins. They showed that the MBP domain is responsible for the nuclear targeting of golli proteins. Meanwhile, the golli domain affects both oligodendroglial and neuronal cells differently by elaborating membrane sheets, and increasing the number and length of processes in each one, respectively (Reyes and Campagnoni 2002).
'Classic' MBP's ability to induce EAE has long implicated it as a putative autoantigen in MS, though exactly how it is presented to T-cells without crossing the blood-brain barrier is not known. The golli proteins J37 and BG21 are expressed outside the CNS, and also contain classic MBP domains as mentioned above (Campagnoni et al. 1993). In 1996, the Campagnoni group showed that these golli proteins were also recognised by different T-cell lines if presented to them (Pitchekian-Halabi et al. 1996). This observation may explain how the cells of the immune system are primed to recognize MBP and lead to an autoimmune response. Intriguingly, BG21 is the main golli product in T-cells. It has also been shown to be a substrate for Protein Kinase C (PKC), acting as a negative regulator of its pathway to regulate the T-cell receptor and its antigen recognition sensitivity (Feng et al. 2004).

Structural studies have been done on both BG21 and J37 (Kaur et al. 2003; Bamm et al. 2007; Ahmed et al. 2007). These investigations have shown that, like the rest of the MBPs, the golli proteins belong to the family of intrinsically-disordered proteins. Under membrane-mimetic conditions (trifluoroethanol) and in lipids, they both adopted secondary structural elements, such as the α-helices also demonstrated in classic MBPs. Their N-terminus also contains a calmodulin-binding site and both have been shown to bind calmodulin with a 1:1 stoichiometry in a calcium-dependent manner, albeit much more weakly than MBP-18.5.

1.4 The Golli Interacting Protein Discovery

The nucleocytoplasmic shuttling of golli proteins in oligodendrocytes and neurons suggests a role in cell signalling or transcriptional regulation. Both of these functions require multiple interacting partners as part of a pathway or a multi-molecular complex. The search for golli interacting partners began with a yeast two-hybrid approach by Fernandes et al., in 2004.
The yeast two-hybrid technique is a powerful tool to identify interacting partners. It exploits the DNA binding and activation domains of transcription factors in yeast, such as GAL4, which can be separated (Chien et al. 1991). The principle of the technique involves the creation of plasmids encoding fusion or hybrid proteins. The first plasmid codes for the DNA-binding domain of a transcription factor like GAL4 fused to the protein of interest, the 'bait' (X). A second plasmid codes for the hybrid protein between the activation-domain of GAL4 and a candidate interacting protein (Y). Both plasmids would then be transformed in yeast. If protein X does interact with protein Y, then the DNA binding domain and activation domain of GAL4 would come into proximity, leading to the transcription of a selective gene containing the GAL4 binding site. Cells that are able to grow in selective medium show that proteins X and Y interact.

What makes this technique so powerful is that unknown interacting partners can be found simply by creating a cDNA library and using it to create a library of hybrid plasmids. By transforming this plasmid library with the bait plasmid, one can screen for interacting partners.

In 2004, Fernandes et al. used the 133-amino acid specific golli domain as bait for this approach. They constructed a cDNA library from a high-transfection efficiency rat strain and fused it to the activation domain of GAL4. From this experiment, they found a positive clone consisting of a ~2.3 kb region in rat, which they then used to screen for the corresponding gene in murine oligodendrocytes. By this approach, they found a 2.8 kb gene containing an open reading frame corresponding to a 261 amino acid protein, which they dubbed 'golli interacting protein' or GIP. Bioinformatic studies of GIP showed that the N-terminus of GIP (residues 1-89) seems to be less conserved than the rest of the protein (residue 90-261). GIP contained some sequence identity (~20%) with a number of proteins, though one stood out with 99% identity
This protein was first identified as a human nuclear protein, the nuclear LIM-interactor interacting factor (NLI-IF) (Marquet et al. 2000). Later it was identified as a small RNA polymerase II (RNAP2) C-terminal domain (CTD) phosphatase (SCP1).

1.4.1 NLI-IF/SCP1 orthology

This protein is not usually described in the literature as the NLI-IF, and its function related to this is still widely unresolved. It was postulated that it interacts with the nuclear LIM interactor (NLI) (Marquet et al. 2000). This latter protein is known to interact with very high affinity with the LIM domain of several transcription factors, thus the speculation is that NLI-IF/SCP1 is involved in transcription, as it is also found to be expressed in 15 different tissues (lowest in the brain) (Marquet et al. 2000; Jurata, Pfaff, and Gill 1998). The LIM domain is a widely-conserved domain among transcription factors, and is composed of two Zn-fingers. Transcription factors with LIM domains cannot function on their own, and require other interacting partners such as NLI (Jurata, Pfaff, and Gill 1998). Curiously, NLI itself shows no transcriptional activity (Jurata, Pfaff, and Gill 1998). Marquet et al. proposed that NLI could act as an 'adaptor' protein between LIM domain-containing transcription factors and non-LIM interacting molecules, such as NLI-IF/SCP1.

Fernandes et al. (2004) showed through a series of immunoprecipitation experiments that GIP interacts with both NLI and BG21, specifically the 133-amino acid golli domain of the latter, \textit{in vitro}. Furthermore, these interactions did not seem specific to a region of GIP, since both the unique (1-89) and conserved (90-261) regions of GIP immunoprecipitated with BG21.
Figure 1.3. Multiple sequence alignment of the murine golli-interacting protein and its closest relatives; hSCP1 (human ortholog) and the phosphatase domain of yFCP1 (yeast homolog) (residues 180-330). SCP1 and GIP have 99% sequence identity whereas this domain of FCP1 has about 20% identity with both of these proteins. Comparing these proteins, A29 in SCP1 and V29 in GIP are very similar; however, there are fewer conservative substitutions between A53 (SCP1) and P53 (GIP) as well as between Q73 and H73.

Figure generated using ClustalW for the multiple sequence alignment and ESPript to produce the image. ESPript/ENDscript Gouet, P., Robert, X. & Courcelle E. (2003).
They also showed that GIP, NLI, and BG21 interacted as a triplex, with GIP acting as the bridge molecule since the golli protein does not interact with NLI. These observations were found to be consistent in the cell, by means of whole-cell lysate immunoprecipitation studies with N19 oligodendrocytes transfected with NLI. Expression of GIP was found in several tissue types, including every tissue in which golli proteins were expressed. Both GIP and BG21 expression also had very similar time courses in the developing brain. Immunocytochemistry studies were employed to confirm nuclear co-localization between BG21 and GIP (Fernandes et al. 2004).

As mentioned earlier, this protein as the NLI-IF has not been heavily cited, as it was later identified as the small CTD phosphatase 1 (SCP1) (261 amino acids). As SCP1, this protein has been much more thoroughly studied. SCP1 belongs to a small family of phosphotransferases that dephosphorylate the C-terminal domain of RNA polymerase II in eukaryotes in a Mg$^{2+}$-dependent manner (Chambers and Dahmus 1994; Lin, Dubois, and Dahmus 2002; Kamenski et al. 2004; Zheng et al. 2005; M. Zhang et al. 2010; Y. Zhang et al. 2006). The CTD is unique to RNA polymerase II in eukaryotes and is not found in prokaryotic polymerases (Corden 1990). It comprises several (up to 52 in mammals) conserved tandem heptad repeats Tyr-Ser-Pro-Thr-Ser-Pro-Ser ($Y_1S_2P_3T_4S_5P_6S_7$) (Egloff and Murphy 2008). These repeats are heavily subjected to reversible PTMs including glycosylation and phosphorylation. The PTMs on different residues, along with proline cis/trans isomerisation, of the tandem repeats appear to create a code that seemingly can be read to mediate processes such as the recruitment of transcription, and elongation factors, and mRNA maturation (Egloff and Murphy 2008). Phosphorylation of the serines in the heptad repeats seems to play a large role. Phosphorylation of $S_5$ recruits the 5' methylguanosine capping enzyme, whereas $S_2$ phosphorylation recruits the polyadenylation
capping enzyme (Schroeder et al. 2000). Oddly enough, RNAP2 cannot be recruited to initiate transcription if it is phosphorylated (Egloff and Murphy 2008). Therefore, it is thought that CTD phosphatases, such as SCP1, are responsible for the recycling of RNAP2 in the transcription cycle. Through a number of promoter-reporter constructs, SCP1 in particular was shown to actually inhibit transcription, whereas a mutant lacking phosphatase activity enhanced transcriptional activity (Yeo et al. 2003). Another study from Yeo et al., in 2005 showed that SCP1 was expressed in neuronal progenitor cells and also targeted neuronal genes. Using an SCP1 mutant lacking phosphatase activity, they blocked the dephosphorylation of the CTD and derepressed neuronal gene expression, which in turn led to cell differentiation (Yeo et al. 2005).

As mentioned earlier, these two proteins (GIP and SCP1) have 99% sequence identity and are essentially the same protein, orthologs from the mouse and human. For the purpose of this thesis, it is important to clarify the nomenclature between GIP and SCP1. The protein identified as GIP is the murine ortholog. The name of GIP arose from Fernandes et al. (2004) after it was involved with the golli protein, BG21. Although SCP1 has more presence in the literature than GIP, the name GIP will be used throughout this thesis when referring to the murine ortholog. This was done in order to emphasize the connection to the golli proteins. Meanwhile, the name SCP1 will be used when referring to the human ortholog. This was done in order to maintain consistency with the literature already found on this protein.

1.4.1.1 CTD Phosphatases

Phosphatases are mainly divided into two classes, tyrosine phosphatases and serine/threonine phosphatases. The Tyr-phosphatases utilize a conserved Cys residue in their active site for the dephosphorylation of tyrosines (B. Zhang et al. 2003). Serine/threonine
phosphatases utilize a metal-dependent catalysis of the dephosphorylation of serine and/or threonine residues (Figure 1.4A) (Cohen 1997; Barford, Das, and Egloff 1998). The crystal structure of SCP1, however, showed striking differences along the Mg$^{2+}$ binding and catalytic site compared to other members of this class of phosphatases (Figure 1.4B) (Y. Zhang et al. 2006). The CTD phosphatases are now classified as their own unique family of phosphotransferases belonging to the super-family of haloacid dehalogenases (HAD) (Gohla, Birkenfeld, and Bokoch 2004). This is an ancient family of enzymes with several different functions, though they were first named after enzymes responsible for the cleavage of carbon-halogen bonds (Kurihara et al. 2008). This superfamily also includes phosphotransferases that account for one fifth of the total human phosphatase catalytic subunits (Seilfred et al. 2013). Though they have very little primary sequence identity with the other phosphotransferases in this super-family, their active sites superimpose very well (Figure 1.4C) (M. Zhang et al. 2010).

In 2004, a truncated form of SCP1 lacking the first 78 N-terminal residues (with a double selenomethionine-substituted mutant L165M/L205M) was first crystallized by Kamenski et al. using the hanging drop method. The crystals had a plate-like morphology and diffracted well, from which the structure was refined at a resolution of 2.3 Å. This structure contained a 5-stranded β-sheet flanked by two α-helices, along with a short 2-strand β-sheet and a short 3_10-helix. There is a large depression along the end of β1 that leads to the active site of the enzyme, and thus it is thought to be a docking/entry site for the target phospho-protein (Kamenski et al. 2004).

What makes CTD phosphatases so unique is their signature motif ψψψDXDX(T/V)ψψ
Figure 1.4. Classifying the small CTD phosphatase SCP1 as a unique family of phosphatases belonging to the super family of the haloacid dehalogenases (HAD). (A) Catalytic site of the serine/threonine class of phosphatases including the two Mn$^{+2}$ ions bound and four essential aspartates. (B) Catalytic site of the SCP1 with a CTD heptapeptide, contains only one Mg$^{+2}$ ion; the topology of the aspartate residues does not match that of the Ser/Thr phosphatases. (C) Ribbon superposition of SCP1 and other phosphotransferases of the HAD super family despite very little primary sequence similarity. Scp1 (yellow, 2ght), PSP (green, 1l7m), β-PGM (cyan, 1lvh), PMM (magenta, 2fuc), P-Type ATPase (pink, 2zbd), and Eya2 (gray, 3geb). Red sphere is a Mg$^{+2}$ ion in the catalytic site of SCP1.

Figure adapted from Zhang et al., 2010.
(where ψ is a hydrophobic amino acid) and their phosphoryl-aspartate intermediate (Y. Zhang et al. 2006; Kamenski et al. 2004). The current proposed two-step reaction mechanism for this family of phosphatases involves the first aspartate residue acting as a nucleophile against the phosphate group yielding the intermediate species (Figure 1.5B) (Kamenski et al. 2004). This high energy intermediate is then hydrolyzed through the second aspartate (general base) to restore the enzyme resting state (Figure 1.5C) (Allen and Dunaway-Mariano 2004). Both Mg$^{2+}$ and water molecules play important roles in coordinating these atoms (Figure 1.4B) (Allen and Dunaway-Mariano 2004). In 2010, Zhang et al., managed to capture snapshots of this reaction through mutagenesis, X-ray crystallography, and use of the artificial substrate para-nitrophenylphosphate (p-NPP) (Figure 1.5).

1.4.2 GIP is a phosphatase

In 2008, Bamm et al. developed a method of over-expression and purification of recombinant murine GIP. The GIP gene was cloned into a pET vector utilizing a C-terminus LE-H$_6$ tag, provided by Drs. Anthony and Celia Campagnoni (UCLA), for purification with a nickel nitrilo-triacetic acid (Ni$^{2+}$-NTA) resin. An *Escherichia coli* Rosetta (DE3)-RIPL cell strain was used for expression with induction by 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Recombinant murine GIP was insoluble when expressed as the full-length protein, and had to be purified under denaturing conditions and then re-folded. This procedure resulted in a relatively low yield of 7.4 mg/L culture. Bamm et al. then proceeded to create a truncated form of GIP lacking the N-terminus (residues Q77-G260). This truncation resulted in perfectly soluble ΔN-GIP, yielding ~31 mg/L of culture and 97% purity based on a quantitative high-performance liquid chromatography (HPLC) analysis (Bamm and Harauz 2008).
Figure 1.5. Snapshot of the reaction of SCP1 with a C-terminal domain peptide as substrate.

(A) Mutation of D96N to capture structure of SCP1 with bound, phosphorylated CTD-peptide (in pale grey) inside the catalytic cleft. (B) Mutation of D206A to capture the formation of the phosphoryl-aspartate intermediate. (C) Mutation of D206A to capture the hydrolysis of the intermediate and release of the phosphate group.

Figure adapted from Zhang et al., 2010.
Circular dichroism (CD) analysis of both full-length and ΔN-GIP showed primarily β-strand secondary structure in both proteins, though ΔN-GIP showed a slightly higher α-helical content. Moreover, Mg$^{2+}$ did not significantly alter its secondary structure. Bamm et al. (2008) also showed that both full-length GIP and ΔN-GIP had Mg$^{2+}$-dependent phosphatase activity with relatively similar $K_m$ values (2.41 ± 0.29 mM and 6.58 ± 0.9 mM, respectively) but drastically different $k_{cat}$ values (0.05 ± 0.01 s$^{-1}$ and 4.82 ± 0.16 s$^{-1}$, respectively) using the artificial substrate, $p$-NPP. These numbers indicated a significant difference in kinetic specificity between rmGIP and ΔN-GIP. The Mg$^{2+}$ was strictly required for full-length GIP activity but not for ΔN-GIP ($pK_{Mg}$ ~2.8). Bamm et al. proposed improper refolding of GIP following denaturing conditions for this difference, inhibiting its ability to coordinate the Mg$^{2+}$ atom (Bamm and Harauz 2008).

In addition to these studies, V. Bamm investigated the interaction between the golli protein BG21 and GIP using a glutaraldehyde cross-linker and polyacrylamide gel electrophoresis, and also by observing its effect on ΔN-GIP phosphatase activity. The cross-linking experiment found a 1:1 stoichiometry between BG21 and ΔN-GIP; however, this interaction was very weak. The ΔN-GIP phosphatase activity was dually affected by BG21, in that at concentrations below 6 μM BG21 enhanced activity, but above 6 μM it inhibited activity (Bamm and Harauz 2008).

1.5 Objectives and Hypothesis

There are still many queries regarding GIP, its activity and its interactions with other proteins such as BG21 and MBP 21.5. There are also questions regarding the biological effects that GIP has with these potential interactions with various macromolecules. In the constraints of this thesis, only a few of these issues were investigated. However, they were the first topics that needed to be addressed in order to understand the whole story. The two requirements
were to (a) Develop an expression and purification protocol for full-length recombinant GIP under native conditions, and to (b) Characterize the phosphatase activity of full-length GIP with the artificial substrate p-NPP, and its dependence on pH. This objective and the knowledge from our lab's previous work led to the first hypothesis:

1) *Full-length recombinant murine GIP from native purification conditions has better phosphatase activity than refolded rmGIP from denaturing conditions.*

Once this characterization was done, the next step was to revisit the interactions with BG21 (and also to investigate potential interactions of MBP-21.5) with full-length GIP. The role of the non-conserved N-terminal domain in these interactions was unknown before this project was started. Comparing how full-length GIP interacts with BG21, in contrast to the truncated ΔN-GIP, should provide better understanding of this role. To understand further how GIP interacts with BG21 and/or MBP-21.5, several strategies and techniques were applied. Both BG21 and MBP-21.5 were phosphorylated with different kinases and subjected to treatment with GIP, to determine if these proteins were potential natural substrates. Enzyme kinetics in the presence of BG21 and/or MBP-21.5 was studied as well to see their effect on GIP activity. Lastly, an enzymological approach was also employed to study interactions between these proteins and help address the second hypothesis:

2) *Interactions between full-length GIP and BG21 are stronger than with ΔN-GIP.*

*Phosphorylated BG21 is also a suitable substrate for GIP. Furthermore, GIP interacts with MBP-21.5 in a similar fashion as with BG21.*
Disproving or proving these hypotheses will be the first step in paving the way for much more intense research in hopes of understanding the golli-interacting protein and its role in oligodendrocyte and myelin biology.
Chapter 2: Purification and Characterization of Full-Length GIP
2.1 Overview

Dr. V. Bamm in our group, worked first towards developing a protocol to over-express and purify the golli-interacting protein (GIP) in its entirety. His study yielded an insoluble protein and he was forced to purify this enzyme under harsh denaturing conditions and attempt to refold it. This method of purification is a last-resort method as it is very difficult to ensure proper refolding of the enzyme. Indeed, he showed that purifying GIP this way resulted not only in low yields, but also an inability to efficiently bind Mg$^{2+}$, likely due to improper folding. His next strategy for purification involved constructing a truncated form of GIP, lacking the first 76 residues, and denoted ΔN-GIP. This truncated ΔN-GIP proved to be very soluble and V. Bamm managed to purify it with better yields. Truncated ΔN-GIP showed very good activity and was used for V. Bamm's previous protein-protein interaction studies. These studies led to unexpected results that may be explained by the absence of the N-terminus.

In order to study GIP and its interactions properly, a new protocol for the purification of full-length and unmodified GIP was needed. Once a protocol was in place, GIP was then characterized through its de-phosphorylation activity. The artificial substrate p-NPP was used to characterize the kinetics of GIP and subsequently its dependence on pH. By studying the kinetics of full-length GIP, parallels were drawn with the truncated ΔN-GIP.
2.2 Materials and Methods

2.2.1 Purification strategy

The biggest predicament encountered in the purification of GIP was protein solubility. The pET-SUMO Champion™ expression system from Invitrogen was employed as a new strategy for the native purification of GIP. This system has been devised to help with protein expression, solubilisation, and purification of recombinant proteins in *E. coli* (Zuo et al. 2005; Lee et al. 2009; T. R. Butt et al. 2005; Marblestone et al. 2006; Malakhov et al. 2004). The SUMO-expression system exploits a small ubiquitin-like modifier (SUMO) to help improve expression and solubility of recombinant proteins. The SUMO family of proteins is highly stable and conserved (T. R. Butt et al. 2005). Butt et al. (2005) reasoned that these proteins act almost like a chaperone on an insoluble protein, as they have a very hydrophobic core, whereas the exterior is very hydrophilic. Overall, its structure is composed of an α-helix surrounded by several β-strands yielding a compact globular fold (T. R. Butt et al. 2005). An N-terminal fusion of SUMO (13.5 kDa) to a less soluble and stable protein can help to stabilize and increase the yield of recombinant expression of a given protein (T. R. Butt et al. 2005).

The SUMO-fusion purification strategy is very simple. It involves cloning the target protein into the pET-SUMO vector provided by Invitrogen (Carlsbad, CA, USA). The bacteriophage T7 promoter is commonly used for the over-expression of the target genes and is utilized by the pET-SUMO vector. The T7 RNA polymerase is required for the expression of genes linked to this promoter. This polymerase is provided by *Escherichia coli* BL21-CodonPlus-(DE3)-RP. This strain of *E. coli* contains the DE3 bacteriophage lysogen, which in turn expresses the T7 RNA polymerase. This polymerase is controlled by the *lac* promoter, which can be
induced by IPTG. The lysogen also carries the *lac* repressor gene (*lacI*) to prevent unwanted over-expression of the polymerase. An addition to this system by the pET-SUMO vector is placement of a *lac* operator downstream of the T7 promoter. The reason is to provide an additional binding site for the lac repressor (from the *lacI* gene) to further repress any basal expression of the T7 RNA polymerase in BL21-(DE3) cells until the inducer IPTG is added.

The pET-SUMO vector uses resistance to kanamycin as a selection marker. The cloning strategy of the target gene into the vector relies on 'TA' cloning. TA cloning involves a linearized vector with 5' T overhangs. While amplifying the target gene via the polymerase chain reaction (PCR), a *Taq* polymerase is used. This polymerase has a template-independent activity, responsible for the addition of a deoxyadenosine (A) to the 3' ends of the PCR products. This 3' A is now able to ligate to the 5' T ends of the vector. This property allows for quick and efficient cloning without the need for restriction endonucleases. As soon as the SUMO-target plasmid is built, it can be transformed into *Escherichia coli* BL21-CodonPlus-(DE3)-RP for recombinant protein expression, as mentioned above.

Once the SUMO-fusion protein is expressed, cells are harvested and lysed under native conditions. At this point, the fusion protein should be in the soluble matter. The SUMO tag contains a further poly-histidine tag attached to the N-terminus, allowing for purification of the complex via Ni\(^{2+}\)-affinity chromatography. Following this initial purification step, a SUMO-protease is used to cleave the SUMO tag from the fusion protein.

A common issue with fusion proteins in general is removing the fusion partner. Many proteases are not able to cleave the complex efficiently and accurately, and often leave additional amino acids at the N-terminus of the target protein, as a result of a sequence-specific
cleavage mechanism. The SUMO protease does not have any of these problems. The SUMO protease belongs to the superfamily of cysteine proteases (T. R. Butt et al. 2005), and is very robust, specific, and accurate. The SUMO protease has been shown to cleave a wide range (6-110 kDa) of SUMO-fusion proteins. Cleavage with SUMO protease can also be cost-effective, as it is very efficient even at ratios of 1:10,000 of enzyme to fusion proteins (Marblestone et al. 2006). Over 100 SUMO-fusions have been cleaved without any non-specific cleavage even when increasing the ratio of substrate to protease to 1:1 (T. R. Butt et al. 2005; Malakhov et al. 2004). Another advantage of this protease is how robust it is, being able to cleave under a variety of different conditions; it has been shown to cleave over a wide range of purification buffers, pH (5.5-10.5) and temperature (4-37°C), even under 2 M urea when used with proteins that are being refolded (T. R. Butt et al. 2005). One condition that it does require is reducing conditions, as it is a cysteine protease. However, the biggest advantage the SUMO protease provides is that it does not leave additional residues at the N-terminus of the target protein. Its mechanism for cleavage is not sequence-dependent but rather dependent on tertiary structure. The protease contains a hydrophobic tunnel through which the fusion protein, or substrate, must pass through (T. R. Butt et al. 2005). This structure aids in the selectivity of the protease and allows it to cleave many different fusion substrates. One drawback of this system is that the SUMO protease is unable to cleave fusion proteins where the target protein contains an N-terminal proline residue. The constricted proline residue is unable to fit through this tunnel, preventing cleavage (T. R. Butt et al. 2005).

Considering that both the SUMO protease and the now cleaved SUMO tag contain a poly-histidine tag, the cleavage reaction can be subjected to a second round of Ni$^{2+}$-affinity
chromatography. These two will bind to the column while the target protein will elute in the flow-through and/or wash steps of the chromatography. Given that the SUMO protease cleaves based on tertiary structure, the target protein is left unmodified with no linker residues after the SUMO tag is removed. This is another advantage of the SUMO purification strategy; the target protein is purified with no added residues and should be in its native state.

2.2.2 Enzymology

As for the majority of enzymes, the Michaelis-Menten model was applied to study the kinetic properties of GIP. GIP has only one substrate binding site and thus the single substrate model was appropriate,

\[
E + S \xrightleftharpoons{k_1}{k_{-1}} ES \xrightarrow{k_2} E + P
\]

where \(E\) is the free enzyme, \(ES\) is the enzyme-substrate complex, \(P\) is the product and \(k_1, k_{-1}, k_2\) are rate constants.

A few assumptions have to be taken into account with this model. It is important to measure the rate of reaction at the beginning of the reaction, when the consumption of substrate and the production of product are linear. This helps establish that the ES complex is in a steady state. It is also assumed that the product formation is directly correlated to the concentration of the ES complex. This assumption means that the rate of reaction observed is mainly the rate of the breakdown of substrate into product, allowing for the reverse reaction to be ignored. The reaction catalyzed by GIP is heavily favoured towards the removal (not the addition) of the phosphate group, another reason to ignore the reverse reaction. Finally, it is assumed that the concentration of substrate bound to enzyme is negligible compared to free
substrate concentration. This situation is easily achieved via the experimental set-up, by adding excess substrate with respect to enzyme.

The activity of GIP and its dependence on pH was also investigated here, by measuring the activity and obtaining kinetic parameters at various pH points. The distribution of the different ionization states of GIP, at different pH values, reflects the effect of pH on the activity of GIP (Takai and Mieske 1991; Z.-Y. Zhang et al. 1994; Doyle 1948; Delory and King 1943; Hengge, Edens, and Elsing 1994; Gahan, Dawson, and Fielding 1978).

For the enzymatic characterization of GIP, the artificial substrate p-nitrophenyl phosphate (p-NPP) was used. Phosphoric esters have long been used in the analysis of various phosphatases. Of these compounds, p-nitrophenyl phosphate has become the most widely used substrate for these enzymes. It is commonly used in undergraduate laboratories as a means to teach the concept of enzymology. In research, it is frequently employed to demonstrate phosphatase activity and furthermore to characterize this activity in vitro (Gahan, Dawson, and Fielding 1978; Delory and King 1943; Z.-Y. Zhang et al. 1994; Hengge, Edens, and Elsing 1994).

The chemical structure of p-NPP is a simple benzene ring with a nitro group in the para position with respect to a phosphate group, which is joined via an ester linkage. Phosphatases cleave off the phosphate group, releasing inorganic phosphate and p-nitrophenolate (p-NP) which quickly becomes protonated to form p-nitrophenol. In the enolate form, when the oxygen is un-protonated, the negative charge helps to stabilize the resonance of the benzene ring, changing its absorbance properties. The enolate moiety of p-NP absorbs light strongly at around 405 nm with an extinction coefficient of 18300 M⁻¹ cm⁻¹. However, p-NP has a pKₐ of
7.16, meaning that the pH of the reaction needs to be high to observe properly the amount of product. Alkaline phosphatases function at pH values above this threshold; therefore, a continuous assay to monitor the formation of product is possible. Unfortunately, this assay is not possible for acidic phosphatases such as GIP, since their enzymatic reactions will take place at pH values lower than 7.16. As a result, a discontinuous assay is necessary, meaning that after a designated amount of time, the reaction will be terminated with NaOH to raise the pH above the pKₐ of p-NP. This time will be short enough so the product formation is linear, and thus the proper kinetic parameters can be obtained.

2.2.3 Full-length GIP expression and purification

The pET-SUMO-CHAMPION expression system from Invitrogen was used for the recombinant expression of GIP. *Escherichia coli* BL21-CodonPlus(DE3)-RP cells transformed with the recombinant plasmid were grown at 37°C in 1 L of Luria Bertani medium containing kanamycin (kan)⁺- chloramphenicol (cam)⁺ (25 µg/mL and 33 µg/mL, respectively) until an A₆₀₀ of 0.8 was reached. The culture was then induced with 1 mM IPTG and grown for a further 4 h. Cells were harvested by centrifugation at 7,500 rpm for 12 min at 4°C. The cell pellet was immediately resuspended and homogenized by stirring for 1 h at 4°C in 30 mL of native lysis buffer (1X PBS, 10 mM imidazole, 300 mM NaCl) with 30 mg lysozyme, 1 mM phenylmethylsulfonylfluoride (PMSF), and 5 mM dithiothreitol (DTT). Homogenized cells were then lysed on ice using a French press system.

The crude lysate was then clarified by centrifugation at 15,000 rpm for 12 min at 4°C. Crude protein was filtered through a 0.45-µm membrane before chromatography. A 1.0-mL bed
volume of Ni\textsuperscript{+2}-NTA resin was used in a Fast Performance Liquid Chromatography (FPLC) system to purify the fusion protein. The system utilizes the QuadTech\textsuperscript{TM} (Bio-Rad Laboratories Inc., CA, USA) detection system monitoring the protocol at 214, 260, and 280 nm. The method used involved a calibrating step of 10 mL of 50:50 mixture of column buffer and elution buffer (1xPBS, 500 mM imidazole, pH 8.0) (1 mL/min), zero baseline of detector, followed by equilibration step of 20 mL column buffer (1xPBS, 150 mM NaCl, 10 mM imidazole, pH 8.0) (1 mL/min). Protein containing crude lysate was then loaded (1 mL/min), followed by a wash step using 20 mL column buffer (1 mL/min). Finally, bound protein was eluted with 10 mL of a 50:50 mixture of column buffer and elution buffer. Flow-through and wash fractions were collected in volumes of 7 mL, while elution was collected in 1-mL fractions. Fractions were analyzed using SDS-PAGE. All buffers used were filtered through a 0.45-µm membrane.

Selected fractions were pooled and dialyzed (12,000-14,000 MWCO tubing) against 1 L of 1X PBS, 10%, glycerol, pH 8.0 twice. DTT was added to the dialyzed protein to a final concentration of 0.5 mM. The total amount of SUMO-GIP was estimated by using the extinction coefficient (\( \varepsilon = 21,430 \text{ M}^{-1} \text{ cm}^{-1} \)) and the \( \text{A}_{280} \). SUMO protease was added (2 \( \mu \text{g} \) protease: 1 mg substrate) to the reaction mixture and incubated at 30°C for 2 h. SDS-PAGE was used to confirm full cleavage. The protease reaction mixture was then subjected to the Ni\textsuperscript{+2}-NTA column as described above. Flow-through and wash fractions were collected in volumes of 1 mL, while elution was collected in 5-mL fractions. Fractions were analyzed using SDS-PAGE. Selected fractions were pooled and quantified by measuring its \( \text{A}_{280} \) and using the extinction coefficient (\( \varepsilon = 20063 \text{ M}^{-1} \text{ cm}^{-1} \)). Final pure GIP was then aliquoted and flash frozen with \( \text{N}_2 (l) \) and stored at -80°C.
2.2.4 Protocol for determining GIP kinetics

A basic discontinuous colourimetric assay was used to study the kinetic activity of GIP, by measuring the absorbance of the reaction product p-NP at 405 nm. Reaction mixtures were assembled in 50 mM Tris-Ac pH 5.5, 10 mM MgCl₂, and increasing concentrations of the substrate p-NPP (0-15 mM) with a final volume of 100 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were started by the addition of 50 nM GIP. After 10 minutes, 1 M NaOH (final concentration) was added to terminate the reaction and the A₄₀₅ was measured. All reactions were performed in triplicate and repeated with several preparations of GIP.

Kinetic parameters were obtained by the non-linear fitting of the Michaelis-Menten equation to the data of A₄₀₅/t (mO.D./s) against increasing concentrations of substrate.

$$\frac{A_{405}}{t} = \frac{[S] \times V_{max}}{[S] + K_m}$$  

[Eq. 1]

Where [S] is the concentration of substrate, $V_{max}$ is the theoretical maximum velocity of the reaction and $K_m$ is the Michaelis-Menten constant, representing the concentration of substrate required to achieve $V_{max}$. Non-linear fitting was done using OriginPro 8.0 (OriginLab, MA, USA). In order to determine if the non-linear curve fitting was in good agreement with the data, the chi-squared ($\chi^2$) value of the fit was determined.
2.3 Results

2.3.1 Expression of SUMO-GIP

The SUMO expression system was utilized to develop a purification protocol for full-length golli interacting protein. The SUMO-GIP construct was successfully built, as mentioned above, by Dr. Graham Smith, as published for MBP-derived peptides (Bamm et al. 2011). After obtaining the construct, it was transformed into *Escherichia coli* BL21-CodonPlus-(DE3)-RP for expression.

*Escherichia coli* BL21-CodonPlus-(DE3)-RP cells were grown at 37°C to O.D of 0.8 before inducing expression of SUMO-GIP with 1 mM IPTG. Expression of the fusion protein was carried out under two different conditions: 1) Expressing cells were kept at 37°C for 4.5 h, and 2) expressing cells were transferred to a 15°C incubator and kept there over-night for 15 h. The gel in Figure 2.1 shows a pre-IPTG and post-IPTG sample for each set of expression conditions. SUMO-GIP is ~42.5 kDa and is clearly seen over-expressed under both conditions around the 42-kDa marker. This gel suggests that both of these expression strategies yielded similar levels of SUMO-GIP. There are other over-expressed bands observed in the gel, around 22 kDa and 38 kDa.

2.3.2 Cell lysis and Solubility

Once SUMO-GIP was expressed, the cells were harvested, and immediately re-suspended and homogenized in lysis buffer for 1 h at 4°C. Following this step, a French press was utilized to lyse the cells. After clarifying the lysate and testing the cell pellet and supernatant, SUMO-GIP was found in the insoluble matter in the pellet. As a result, three
Figure 2.1. A 14% SDS-PAGE gel of the expression of SUMO-GIP. The SUMO-GIP was transformed into *Escherichia coli* BL21-CodonPlus(DE3)-RP cells for expression. Cells were grown to an appropriate O.D. before inducing the expression of the fusion protein by adding 1 mM IPTG. Expression of protein was performed under two conditions, 37°C for 4.5 h and 15°C for 18 h. Lane-1) Pink Plus Pre-stained Protein Ladder from GeneDirex. Lane-2) Pre-IPTG sample for expression at 37°C for 4.5 h. Lane-3) Post-IPTG under same conditions. Lane-4) Pre-IPTG sample for expression at 15°C for 18 h. Lane-5) Post-IPTG under same conditions.
Figure 2.2. 14% SDS-PAGE gel of the solubilisation of SUMO-GIP from cell lysate. Upon harvesting, the cell pellet was immediately resuspended and homogenized by stirring for 1 h at 4°C in 30 mL of native lysis buffer (1X PBS, 10 mM imidazole, 300 mM NaCl) with 30 mg lysozyme, 1 mM phenylmethylsulfonylfluoride (PMSF) and various additives. Homogenized cells were then lysed on ice using a French-press system. (A) Additive to improve solubility was 2% Tween-20: Lane-1) Pink Plus Pre-stained Protein Ladder from GeneDirex. Lane-2) Whole cell lysate. Lane-3) Supernatant from clarified lysate. (B) Additive to improve solubility was 0.3% Sarkosyl: Lane-1) Whole cell lysate. Lane-2) Supernatant from clarified lysate. (C) Additive to improve solubility was 5 mM dithiothreitol DTT: Lane-1) Pink Plus Pre-stained Protein Ladder from GeneDirex. Lane-2) Whole cell lysate. Lane-3) Supernatant from clarified lysate.
different conditions were added to the lysis buffer in order to improve the solubility of SUMO-GIP. The first was 2% Tween-20, followed by 0.3% sarkosyl, and finally 5 mM DTT. Figure 2.2 shows that Tween-20 was able to partly solubilize SUMO-GIP, as it is observed in the supernatant. However, about 80% of the protein is still present in the insoluble fraction. Sarkosyl was much more efficient than Tween-20, as it solubilized ~60% of SUMO-GIP. DTT was better than Tween-20 but not as efficient as sarkosyl, managing to help solubilize ~50% of the fusion protein. From these results, DTT was chosen as the preferred method to help lyse cells for the finalized protocol. The other over-expressed bands observed in Figure 2.1 are observed with greater intensity here; it is likely that they are either truncated forms of SUMO-GIP or protease-cleaved products.

The effect of DTT on the solubility of SUMO-GIP is likely due to its reducing properties. To observe any potential aggregates, lysates solubilized with 5 mM DTT and also without DTT were run on a non-reducing SDS-PAGE (Figure 2.3). This gel revealed that SUMO-GIP aggregates with other impurities in the lysate in the absence of DTT, without any specificity. Under non-reducing conditions, the band corresponding to SUMO-GIP decreases in intensity, while new bands appear at higher molecular masses. SUMO-GIP does not appear to dimerize, since no bands around 84 kDa were observed. The main higher-order band observed is at around 56 kDa, suggesting that SUMO-GIP likely aggregated with the proteins observed at the bottom of the gel (14-18 kDa). Much larger complexes are seen at the very top of the gel at 175 kDa or bigger. These assemblies are at least 4 times the size of SUMO-GIP, and could potentially be tetramers or a combination of SUMO-GIP and other proteins in the lysate. One of these combinations is likely with the protein observed around 29 kDa, considering that there is a big
Figure 2.3. Non-reducing 14% SDS-PAGE gel of the cell lysate with expressed SUMO-GIP. Upon harvesting, the cell pellet was immediately resuspended and homogenized by stirring for 1 h at 4°C in 30 mL of native lysis buffer (1X PBS, 10 mM imidazole, 300 mM NaCl) with 30 mg lysozyme, 1 mM phenylmethylsulfonylfluoride (PMSF). Lane-1) Pink Plus Pre-stained Protein Ladder from GeneDirex. Lane-2) Whole-cell lysate homogenized with 5 mM DTT. Lane-3) Whole-cell lysate homogenized without reducing conditions.
decrease in intensity of this band under non-reducing conditions.

2.3.3 Purification of the SUMO-GIP Fusion

After the lysis procedure, the lysate was clarified via centrifugation, and then filtered through a 0.45-μm membrane before chromatography. The first step of the purification involved IMAC using Ni$^{2+}$ as the metal. The (His)$_6$ tag at the N-terminus of SUMO has a high affinity towards Ni$^{2+}$ and thus SUMO and SUMO-fusions will remain in the column until they are eluted. High concentration of imidazole will out-compete the (His)$_6$ tag for Ni$^{2+}$, displacing it from the column.

In order to visualize the results of the purification, SDS-PAGE was performed. The gel in Figure 2.4 represents a typical profile of this first chromatography step. The lane represents the crude protein before loading onto the column. The flow-through and wash steps were collected together and are shown here in the second lane. Subsequent lanes in the gel show individual fractions collected from the elution step of the chromatography. Here, SUMO-GIP is evident near the 42-kDa mark. There are also other bands between 14 and 22 kDa with intensity equal to SUMO-GIP. These are believed to be shorter fragments of SUMO-GIP, from either truncation during expression or proteolysis products. These fragments do not appear in the crude lysate with such intensity. Moreover, these peptides still contain the N-terminus of the fusion protein since they were able to bind to the column.

2.3.4 Cleaving the SUMO-GIP Fusion and Purifying GIP

Following the first chromatography step, the imidazole concentration in the pooled fractions collected with SUMO-GIP needed to be reduced from ~250 mM to ~10 mM. This was
Figure 2.4. A 14% SDS-PAGE gel of the Ni\(^{2+}\)-chromatography profile of SUMO-GIP. Clarified cell lysate was passed through a 4.5 mL bed volume of Ni\(^{2+}\)-NTA resin was using in a Fast Performance Liquid Chromatography system. Column buffer used was 1xPBS, 150 mM NaCl, 10 mM imidazole, pH 8.0. Bound protein was eluted with 15 mL of 50:50 mixture of column buffer and elution buffer (1xPBS, 500 mM imidazole, pH 8.0). Flow-through and wash fractions were collected in one large fraction, whereas elution was collected in 1-mL fractions. Lane-1) Pink Plus Prestained Protein Ladder from GeneDirex Lane-2) Crude protein Lane-3) FT/W Lanes-4-15) Elution fractions.
done via dialysis, and also by gel filtration chromatography. Dialysis led to significant protein precipitation and was avoided in the final protocol. Once the imidazole concentration was reduced, glycerol and DTT were added to final concentrations of 10% and 0.5 mM, respectively. The SUMO-GIP was then subjected to treatment with the SUMO protease to cleave the fusion protein.

Figure 2.5 shows the fusion protein sample just before and after 1.5 h of SUMO protease treatment in lanes 2 and 3, respectively. Here, it is evident that full cleavage occurred with no fusion protein remaining around 42 kDa. Instead, it is observed that a band appears at ~29 kDa, representing GIP alone. The cleaved SUMO appears at just above the 14-kDa marker. The SUMO band has a much higher intensity than GIP, which is explained by all the SUMO-GIP peptides that were also present in the mixture being cleaved. Their presence results in a much larger total amount of SUMO than full-length GIP. The remaining GIP peptides are not large enough to be seen on this gel, as they should be ~2-6 kDa in length. The following lanes (5-13) in the gel show pure full-length GIP in the flow-through and wash fractions. Pure SUMO appears in the last two lanes (14 and 15), which correspond to the elution step of this chromatography. The yield was determined to be ~4-5 mg of GIP per 1 L of cell culture by absorbance measurement, whereas the purity is estimated at 85% from the SDS-PAGE. HPLC analysis of purity was not possible as GIP consistently crashed (i.e., precipitated) in the column guard.

2.3.5 Enzymatic Characterization Using an Artificial Substrate

Frozen GIP aliquots were thawed and desalted into 50 mM Tris-Ac, pH 5.5, and 10 mM MgCl₂, for experimental use. The GIP concentration was obtained using the A₂₈₀ and the
Figure 2.5. A 14% SDS-PAGE gel of the cleavage reaction by the SUMO protease on SUMO-GIP and the second Ni\textsuperscript{2+} chromatography step. Conditions for the protease reaction included 10% glycerol and 0.5 mM DTT at 30°C for 1.5 h. A ratio of 2 µg of SUMO protease to 1 mg of SUMO-GIP was used for the reaction. A 1 mL bed volume of Ni\textsuperscript{2+}-NTA resin was used in a Fast Performance Liquid Chromatography system. Column buffer used was 1xPBS, 150 mM NaCl, 10 mM imidazole, pH 8.0. Bound protein was eluted with 15 mL of 50:50 mixture of column buffer and elution buffer (1xPBS, 500 mM imidazole, pH 8.0). Flow-through and wash fractions were collected in 1-mL fractions, whereas elution was collected in three 5-mL fractions. Lane-1) Pink Plus Prestained Protein Ladder from GeneDirex. Lane-2) SUMO-GIP prior to protease treatment. Lane-3) SUMO protease reaction after 1.5 h. Lane-4) Crude mixture before chromatography. Lanes-5-13) Flow-through and wash fractions. Lanes-14-15) Elution fractions.
extinction coefficient (20063 M$^{-1}$ cm$^{-1}$). To measure the kinetic parameters of GIP on the artificial substrate $p$-NPP, a discontinuous tube-based assay was used. Reaction mixtures containing increasing concentrations of $p$-NPP had a final volume of 100 μL, and were incubated at 37°C in a water bath. The reaction was initiated for every tube, every 15 s, by the addition of 50 nM GIP and each tube terminated after 10 min. Plotting the absorbance at 405 nm over time against $p$-NPP concentration yielded a classic hyperbolic function. The Michaelis-Menten equation was fitted to this curve via non-linear regression done in OriginPro 8.0 (OriginLab, Massachusetts, USA). This experiment was done in triplicate and repeated four different times using two separate GIP preparations. The fitting revealed an average $K_m$ of 4.59 ± 0.30 mM, a $k_{cat}$ of 4.01 ± 0.36 s$^{-1}$, and in turn a $k_{cat}/K_m$ number of 0.764 ± 0.08 mM$^{-1}$ s$^{-1}$. A typical velocity curve is shown in Figure 2.6. The most important aspect of the purification regarding the final activity of GIP is the overall time spent purifying. Either freezing the cell pellet for storage or spending more than 72 h purifying GIP resulted in much less active enzyme.

After obtaining the basic kinetic parameters of GIP with $p$-NPP, a pH analysis was performed. Initially, the substrate concentration was kept constant at 10 mM, while the pH of the buffer was changed. This was done in order to find out the optimal pH under which GIP operates. Panel A in Figure 2.7 revealed that GIP is most active at pH 5.5, whereas it is essentially ineffective at pH 4.5 and also pH 7. Following this, a kinetic profile of the pH dependence of GIP was done by performing the same kinetics assay at various pH points. Panels B, C, and D in Figure 2.7 show the effect of pH on $K_m$, $k_{cat}$, and $k_{cat}/K_m$ respectively.
Figure 2.6. A velocity curve for the de-phosphorylation of the artificial substrate $p$-NPP by GIP, obtained using a discontinuous assay. Reaction mixtures were done in 50 mM Tris-Ac pH 5.5, 10 mM MgCl$_2$, and increasing concentrations of $p$-NPP (0-15 mM) with a final volume of 100 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize temperature. Reactions were initiated by the addition of 50 nM GIP. After 10 minutes, 1 M NaOH was added to terminate the reaction and the $A_{405}$ was measured. This experiment was done in triplicate and repeated two different times using two separate GIP preparations; error bars represent the standard deviation of these individual experiments. Data were fitted to the Michaelis-Menten kinetics model using a non-linear fit with OriginPro 8 (OriginLab, MA, USA). The $\chi^2$ value of 4.01E-4 indicates an excellent fit of the model to the measured data.
Figure 2.7. Analysis of the effect of pH on the activity of GIP. Reaction mixtures were prepared in 50 mM Tris-Ac pH 5.5, 10 mM MgCl₂ with a final volume of 100 μL. (A) Profile of pH on GIP, with a constant concentration of p-NPP at 10 mM. For the kinetic analysis, increasing concentrations of p-NPP (0-15 mM) were used. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize temperature. Reactions were initiated by the addition of 50 nM GIP. After 10 minutes, 1 M NaOH was added to terminate the reaction and $A_{405}$ was measured. This experiment was done in triplicate and repeated two different times using two separate GIP preparations; data were fitted to the Michaelis-Menten kinetics model using a non-linear fit with OriginPro 8 (OriginLab, MA, USA) in order to obtain $K_m$ and $k_{cat}$. (B) Effect of pH on the $K_m$ of GIP. (C) Effect of pH on the $k_{cat}$ of GIP. (D) Effect of pH on the $k_{cat}/K_m$ of GIP.
2.4 Discussion

The golli-interacting protein is the murine ortholog of the human small CTD phosphatase 1. These two proteins have identical sequences, except for three amino acids near the N-terminus (Figure 1.3). Due to these similarities, SCP1 was taken as an excellent model to guide the research presented. SCP1 belongs to a specific group of phosphatases, the CTD phosphotransferases. The first protein discovered that belonged to this group was the first CTD phosphatase 1 (FCP1) in yeast (Mandal et al. 2002). SCP1 has high sequence identity to the phosphatase domain of FCP1, except that it has a unique N-terminal extension. As its name implies, SCP1 (small CTD-phosphatase 1) is much smaller in size, lacking a breast cancer domain present in FCP1 (Yeo et al. 2003). These phosphatases were not classified in their own unique group until the crystal structure of SCP1 was solved (Kamenski et al. 2004). This structure revealed differences in the active and Mg\(^{2+}\)-binding sites with respect to other Ser/Thr phosphatases (Y. Zhang et al. 2006; Kamenski et al. 2004).

Obtaining a structure of a protein via crystallography is a challenging task and requires a lot of protein at high concentrations. SCP1 was crystallized at a concentration of 20 mg/mL (Kamenski et al. 2004). This achievement was largely due to the high yield of protein from recombinant expression and purification. Conditions used to purify SCP1 included a lysis step with a Tris buffer with 300 mM NaCl, and glycerol (Kamenski et al. 2004). As a starting point here, these conditions were replicated to purify GIP, with no success. The construct used to determine the crystal structure of SCP1, had the N-terminal-most 77 residues omitted. Kamenski et al. decided to truncate SCP1 and remove the N-terminus since they struggled to purify full-length SCP1. They found that SCP1 underwent proteolysis between residues 58 and
65 during the expression of the recombinant protein in *E. coli* (Kamenski et al. 2004). It is likely that the N-terminus would not have been visible in the electron density map had it been crystallized, as it is predicted to be primarily random coil and very mobile. The server used for this prediction was the Porter server from University College, Dublin, and the prediction for GIP shows very good agreement with the known structure of SCP1 (Figure 2.8). Once the N-terminus was removed, expression and subsequent yield of SCP1 was dramatically increased (Kamenski et al. 2004).

Unfortunately, the literature is unclear about recombinant SCP1 used in most studies. It was not always specified when the SCP1 used is the truncated form or the full-length protein. When it is not specified, the protocol described to purify SCP1 is either vague or very similar to the protocol used for the truncated protein (Zhang et al. 2012). Conversely, when it is specified that the protein used is full-length, the purification protocol is also similar to that used for truncated SCP1 (Yeo et al. 2003). Furthermore, there is no mention about difficulties in the purification process and the authors claim a high yield for full-length protein (Yeo et al. 2003). This is a very different story from that later reported by Kamenski *et al.* and also that found by V. Bamm with GIP in 2008. It is no surprise that V. Bamm struggled to develop a protocol for the recombinant expression and purification of full-length GIP. He was forced to use denaturing conditions and attempted to refold GIP. This proved to produce a much less active form of GIP. Following this, he too removed the N-terminus of GIP and saw a striking improvement in yield and activity (Bamm and Harauz 2008).
Figure 2.8. The secondary structure prediction of full-length GIP from the Porter server from University College, Dublin. For comparison, the secondary structure footprint deduced from the crystal structure of truncated SCP1 (residues 77-261) is shown. The conserved active site motif is shown by the blue box, with the important aspartyl residues (D96 and D98) highlighted. Random coil is shown in green, β-strands are in yellow, and α-helices are in red.
2.4.1 Expression and Solubility

The first task of the research presented here was to develop a purification protocol for full-length GIP. The biggest challenge in the past was the insolubility of full-length GIP. The SUMO fusion tag strategy was chosen here to help with the solubility of the expressed enzyme. It has been shown to help with the expression and solubility of a multitude of other proteins (Malakhov et al. 2004; T. R. Butt et al. 2005; Zuo et al. 2005; Lee et al. 2009; Bamm et al. 2011; Rahman, Smith, et al. 2011; Rahman, Bamm, et al. 2011; Rahman et al. 2012). Here, it is shown that GIP with the SUMO tag is expressed equally well at 37°C for 5 h as compared to 15°C and 18 h. There was also no change in expression levels seen when expressing in the Rosetta strain compared to BL21 cells (not shown). This insensitivity of expression conditions by SUMO-GIP is not surprising, since it is known that the SUMO tag can help with protein expression (T. R. Butt et al. 2005; Zuo et al. 2005).

Initially, the same conditions used to purify truncated GIP were used to purify SUMO-GIP. This was done to assess the effect of the SUMO tag on the solubility of GIP. This experiment showed that SUMO by itself did not have a huge effect and that SUMO-GIP was still insoluble. Therefore, different conditions were tested to improve this solubility. The first strategy involved using a gentle non-ionic detergent, such as Tween-20. This detergent was able to help solubilize about 20% of SUMO-GIP. The next step was to use a harsher anionic detergent. Sarkosyl was used to help solubilize the fusion protein. This adaptation improved the amount of protein found in the soluble matter to about 60%. Although this result was expected, this was not a desired approach to obtain soluble protein. Sarkosyl, being anionic, can severely interfere with the IMAC purification step. The negatively-charged detergent will interact with
the positively-charged Ni\textsuperscript{2+} ions in the column. If sarkosyl was the only option, it would have to be removed. There are a few strategies to do this, such as trapping it inside Tween-20 micelles as well as precipitating it out of solution with Mg\textsuperscript{2+} ions. However, these options will also likely interfere with the enzyme, or dramatically reduce the yield. Therefore, new avenues to improve the solubility of SUMO-GIP were explored before relying on this method.

2.4.2 DTT on the Solubilization of SUMO-GIP

Looking at the crystal structure of the human ortholog, SCP1 (3pgl), it was seen that two protomers were present in the smallest repeat unit (Figure 1.4) (Kamenski et al. 2004). The crystal structure only showed three of the six cysteine residues of SCP1, since it was only the truncated form of SCP1 that was able to be purified and crystallized, as mentioned above. From the structure, it was not evident whether these three residues are capable of forming disulfide bridges since they are buried.

As previously stated, the missing N-terminus is predicted to be disordered and very mobile. This section of the protein also contains three neighbouring cysteines. If these three cysteines are very mobile, it is very likely that under oxidizing conditions they would freely form disulfide bridges between monomers. This would create the possibility of dimerization of GIP or even higher order oligomers. In order to check this possibility here, cell lysates treated with or without 5 mM DTT were run on a non-reducing SDS-PAGE to see if any higher order species of GIP were visible in the lysate (Figure 2.3).

This experiment showed that SUMO-GIP readily aggregated with other impurities in the lysate (Figure 2.3). Several species of higher molecular weight were observed in the lane
without reducing conditions. However, no dimerization between SUMO-GIP molecules was observed, likely due to the SUMO tag. This tag is relatively large at ~13.5 kDa and it is fused to the mobile N-terminus of GIP, likely reducing the flexibility and mobility of this part of GIP. Furthermore, the sheer size of SUMO can prevent the interactions between N-termini from different SUMO-GIP molecules. Regardless of the lack of dimerization, SUMO-GIP is still able to form aggregates of various sizes with other proteins, even forming very large complexes which were seen at the very top of the gel and which were unresolved. After seeing these results, it was decided to try using 5 mM DTT as a means of solubilizing SUMO-GIP. Figure 2.2 shows that, indeed, 5 mM DTT improved the solubility of SUMO-GIP by 50%. DTT, as mentioned earlier, is a strong reducing agent and would prevent cysteine residues from cross-linking. It is commonly used in purifications for the purposes of minimising oxidation damage. In this case, it helped solubilize SUMO-GIP. It was likely able to do this by preventing non-specific association of SUMO-GIP with other proteins. These proteins themselves could be membrane- or lipid-associated, and thus very insoluble. By forming a complex with SUMO-GIP, these interactions could further decrease the solubility of SUMO-GIP. Although the SUMO tag by itself was unable to solubilize GIP, it made it possible for a mild reagent such as DTT to achieve this goal.

2.4.3 Purification of the SUMO-GIP Fusion

Following the solubilization of SUMO-GIP, the next step in the purification was the first round of IMAC, using Ni\(^{2+}\) metal. This round of chromatography yielded fairly pure product, except for truncated or degraded SUMO-GIP seen between 14 and 22 kDa (Figure 2.4). These fragments were a very interesting phenomenon observed throughout the troubleshooting
phases of the protocol. They were constantly observed during this step of the purification with equal or greater intensity than the band corresponding to full SUMO-GIP. This observation shows that the overall quantities of these impurities are equal to or greater than full-length SUMO-GIP. Furthermore, they have the same elution pattern as SUMO-GIP, implying that these fragments contain a (His)_6-tag in order to bind the column with this affinity. Curiously, it is not possible to see these same bands when looking at the crude protein mixture before it is loaded into the column. However, they appear after eluting from the column. At first glance, it appears that SUMO-GIP is being cleaved in the column, which cannot be true. After seeing this phenomenon, both fresh resin and columns were used in the purification. This change yielded the same result. Another possibility is that there is a protease being expressed and purified along with SUMO-GIP. Once SUMO-GIP is immobilized and concentrated in the column, this protease can cleave the fusion protein. To prevent any degradation by rogue proteases, a cocktail of protease inhibitors (which included pepstatin A and PMSF) was used during the lysis step. These inhibitors do not protect against every protease, but do inhibit a wide range of proteases.

By analyzing the sizes of these fragments, and removing SUMO, it was thought that there was a rough section of the sequence where GIP was being cleaved. The ExPASy PeptideCutter (Swiss Institute of Bioinformatics) was used with the sequence of GIP to determine possible cleavage sites and proteases. This analysis revealed that the most likely proteases responsible for the cleavage occurring at this region of SUMO-GIP were aspartyl proteases. The protease inhibitor normally used for these proteases is pepstatin A. Pepstatin A was already included in the lysis step without any improvement in the degradation of SUMO-
GIP; however, the cleaved products were mainly observed after chromatography. As a result, pepstatin A was added to the column buffer when purifying SUMO-GIP to ensure the constant presence of the protease inhibitor during chromatography. Unfortunately, this modification to the protocol yielded no improvement in the problem.

Another explanation for this problem would be that, as SCP1, SUMO-GIP is cleaved during expression by an *E. coli* protease. Meanwhile, the fragments produced as a result of this cleavage, have a higher affinity to the column than the full protein, due to their smaller size. These fragments may be not visible in the gel at first; however, as they bind the column they become concentrated. By outcompeting SUMO-GIP in binding the column, they are able to be observed in the elution with great intensity. This explanation is supported by the SDS-PAGE of the chromatography samples. The lane representing the flow-through and wash steps shows a good amount of SUMO-GIP that did not bind the column, describing an overloaded column (*Figure 2.4*). The next logical step was to re-run this flow-through and wash through the column again, now without these cleaved products. Inexplicably, these fragments were again visible in the elution of this second chromatography with the same intensity ratio previously observed. These gels were not included since they looked more or less identical to the gel shown in *Figure 2.4*, albeit with less total protein.

2.4.4 Cleaving the SUMO-GIP fusion

Without a clear explanation for the presence of these lower molecular mass products, the troubleshooting of the protocol continued. The next step was to utilize the SUMO protease to remove the SUMO tag from GIP. This is a very straightforward reaction; as explained above,
this protease is very robust and its only caveat is the requirement of reducing conditions. In order to help with overall protein stability during the reaction, 10% glycerol was also used. Another important requirement before removing the SUMO tag is reducing the concentration of imidazole to <10 mM in preparation for a second IMAC purification. In order to achieve these global buffer changes, two strategies were used. The first strategy involved dialysis into the same buffer with 10% glycerol, while concurrently removing the imidazole, then adding DTT. This procedure worked well, sometimes. Unfortunately, SUMO-GIP often precipitated out of solution during this step, resulting in anywhere from 10-90% loss in final yield. The second strategy involved concentrating SUMO-GIP into less than 1 mL, then subjecting it through a gel filtration desalting column (5,000 MWC, 1 mL loading volume) to remove the imidazole. Following this chromatography step, DTT and glycerol were added to their respective final concentrations (0.5 mM and, 10%, respectively). This strategy resulted in an even greater loss of protein, though it was consistent, at about 50% loss. For the final protocol, the dialysis strategy is mentioned since it had the potential for higher yield.

The cleavage reaction by the SUMO protease was very reliable. After 1.5 h, all of the SUMO-GIP fusion protein was separated into its two proteins. The smaller impurities were also cleaved, confirming that they were SUMO-GIP fragments. The band near 14 kDa, corresponding to SUMO, appears with greater intensity than that of GIP (~29 kDa) (Figure 2.5). This observation is explained by all the impurities. These impurities also contained SUMO; thus, the total amount of SUMO is much greater than that of full-length GIP. Meanwhile, the GIP components of the impurities are not observed, although these are likely still present. Since these fragments would be <7 kDa in size, they are too small to observe by this SDS-PAGE.
2.4.5 Removing SUMO from Protein Mixture

Following the cleavage of the SUMO tag, it needed to be separated from GIP in solution. This was done by subjecting it through a second round of IMAC, again using Ni$^{2+}$ as the metal. This step in the purification never provided major difficulties. The only issue that was sometimes observed was SUMO leaking into the flow-through and wash fractions. Whenever this happened, it meant that there was still enough imidazole in the SUMO protease reaction mixture to out-compete and displace SUMO from the column. This is not an issue from the chromatography, but rather insufficient buffer exchange from the previous step, before the SUMO protease reaction. Whenever this issue was encountered, the buffer exchange was repeated, and so was the second IMAC step.

2.4.6 Full-length GIP Purification Assessment

The end result of this purification protocol was a relatively pure full-length GIP. From SDS-PAGE, the average purity of the final GIP solution was ~90%. However, knowing that those shorter GIP fragments are likely present, a safer estimate and likely more accurate purity of GIP is around 85%. The GIP fragments present correspond to the mobile N-terminus of GIP. This region does not contain many aromatic residues; thus, protein quantification via $A_{280}$ will be representative of full-length GIP. Furthermore, although some of these fragments may be long enough to include the active site of GIP ($\psi\psi\psi\psi\psi\psiDXDX(T/V)\psi\psi$), they are not large enough to fold properly and be catalytically active. As a result, for the enzymological approach of studying GIP, this is a sufficient level of purity. The maximum yield ever achieved with this protocol was ~10 mg/L of culture. Unfortunately, this was only achieved once and could not be replicated after
several trials. A more realistic yield from repeating this protocol multiple times is ~4-5 mg of GIP per 1 L of cell culture. Again, for the enzymological studies, this is a sufficient amount of GIP. For expanding this protocol to suit other techniques where the demand of protein is greater, expressing larger volumes of culture will be required. One thing that will help is that GIP can be concentrated using the Amicon (10,000 MWC) concentrators. The highest concentration attained was ~50 μM or ~1.5 mg/mL. Although this amount is not enough for some techniques, this was only achieved by concentrating the yield from one liter of culture. It is possible that by increasing the scale of the culture, a higher concentration could be achieved.

2.4.7 Enzymatic Characterization using an Artificial Substrate

The only way of assessing whether the purification protocol developed was a success, was to ensure that the protein purified is catalytically active. To measure the activity of GIP as a phosphatase, the artificial substrate p-NPP was used first. As mentioned earlier, this substrate is universally used for the kinetic analysis of various phosphatases. Unfortunately, this colourimetric assay needed to be a discontinuous assay for GIP; thus, a tube-based discontinuous method was utilized here. This involved initiating each sample point in a tube by adding GIP, then terminating the reaction in every tube after exactly 10 min with NaOH. This process involved a considerable amount of manual labour and pipetting, meaning more opportunity for error. Initially the assay was automated in the plate reader, utilizing its injection ports to initiate and then terminate the reaction. However, this automated method produced a significantly worse kinetic efficiency number when compared to the manual method (results not shown). It is suspected that the temperature control of the plate reader is not as consistent
as that of the water bath. Another obvious reason for this discrepancy is the possibility that the mixing in the 96-well plate by the shaking mechanism of the machine is not as efficient as the mixing via pipetting in the manual method.

Ultimately, the manual method of performing this basic assay was chosen to study the activity of GIP. By this method, the velocity graph of GIP toward $p$-NPP showed an excellent hyperbolic curve (Figure 2.6). This curve was fitted to the basic Michaelis-Menten equation, representing the classic single-substrate binding model. The $R^2$ (0.9895) and reduced $\chi^2$ (3.395E-5) values of the fitting indicate a very good agreement between the data and the model. Averaging the kinetic parameters obtained from repeating this experiment several times, including two separate protein purifications resulted in a $K_m$ value of 4.57 ± 0.30 mM and $k_{cat}$ value of 3.95 ± 0.37 s$^{-1}$, as well as the specificity constant ($k_{cat}/K_m$) of 0.76 ± 0.08 mM$^{-1}$ s$^{-1}$. All of these values are in very good agreement with those of SCP1 with the same substrate (Table 2.1). This concordance confirms that the purification protocol developed yields a fully-functional full-length GIP - success, in other words.

The kinetic parameters obtained by V. Bamm in 2008 of the truncated form show that truncated GIP is actually more active compared to full-length (Table 2.1). It is possible that this very mobile N-terminus may have a detrimental effect on the binding and catalysis of GIP on $p$-NPP. However, the N-terminus is likely to be important for other functions such as protein interactions and overall stability of GIP. The former possibility will be analyzed and discussed in Chapter 3, while the latter effect is currently being investigated via CD spectroscopy and will be further discussed in Chapter 4.
The active site of the CTD phosphatases is characterized by the signature motif ψψψDXDX(T/V)ψψ (where ψ is a hydrophobic residue). The two aspartates (D96 and D98) have been shown to be catalytically important (Figure 1.5) (Kamenski et al. 2004; Yeo et al. 2003). The first aspartate residue here acts as a nucleophile, helping to form a phospho-aspartate intermediate, which is then hydrolyzed by the second aspartate residue to release inorganic phosphate. Removing D96 via mutagenesis (D96N) abolished all activity, whereas a mutation of the second aspartate (D98N) managed to retain ~30% of activity of SCP1 (Yeo et al. 2003). This mechanism is functional under acidic conditions, placing the CTD phosphatases under the acidic type.

To help confirm this mechanism here with GIP, a pH profile of enzymatic activity was done. This process involved holding the substrate concentration constant while varying the pH. The pH profile for GIP with p-NPP is nearly a "carbon copy" of the profile from FCP1, the yeast homolog, with the same substrate (Figure 2.7 A) (Hausmann and Shuman 2002). SCP1 also has a very similar profile, though it was done with a synthetic peptide from the CTD as a substrate (Yeo et al. 2003). When looking at the effect that pH has on $K_m$ and $k_{cat}$, the $K_m$ is less affected by the pH. This result shows that GIP retained the ability to bind the substrate up to a pH of 7.5. Meanwhile, the effect of pH on the $k_{cat}$ resembles more the effect seen in the profile at a constant substrate concentration. This result indicates that pH is very important for the catalysis of the substrate, not the binding. This explanation makes sense when looking at the active site of GIP, and those two key aspartate residues, suggesting that the phospho-aspartate intermediate is likely being formed at D96, but that it cannot be hydrolyzed by D98. This conjecture could be difficult to prove as mutagenesis on D96 abolishes the activity of the
enzyme. Largely, this pH profile further demonstrates that the purification protocol designed is suitable for the purification of active, full-length GIP. Furthermore, it shows that GIP is catalytically very similar to SCP1 and also FCP1, providing evidence for GIP also being part of the CTD phosphatase family, as expected. This possibility is further investigated in Chapter 3, where the interactions between GIP and other proteins are investigated.
Table 2.1. Summary of kinetic parameters for full-length GIP, ΔN-GIP, and ΔN-SCP1.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat} / K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jaramillo-Tatis et al., 2013.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-length (native purification)</td>
<td>4.57 ± 0.30</td>
<td>3.95 ± 0.37</td>
<td>0.76 ± 0.08</td>
</tr>
<tr>
<td><strong>Bamm et al., 2008.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-length (denature/refold purification)</td>
<td>2.84 ± 0.29</td>
<td>0.05 ± 0.01</td>
<td>0.0183 ± 0.004</td>
</tr>
<tr>
<td><strong>Bamm et al., 2008.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔN-GIP</td>
<td>2.24 ± 0.02</td>
<td>4.82 ± 0.16</td>
<td>2.15 ± 0.07</td>
</tr>
<tr>
<td><strong>Zhang et al., 2006.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔN - SCP1</td>
<td>5.80 ± 0.57</td>
<td>4.20 ± 0.21</td>
<td>0.72 ± 0.061</td>
</tr>
<tr>
<td><strong>Zhang et al., 2010.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔN - SCP1</td>
<td>3.6 ± 0.8</td>
<td>2.5 ± 0.1</td>
<td>0.7 ± 0.2</td>
</tr>
</tbody>
</table>
2.4.8 Chapter Summary

In this chapter, a new protocol for the purification of full-length GIP was developed. The protocol involved using the SUMO-fusion protein system to improve the solubility of GIP, with the help of 5 mM DTT. Like SCP1, GIP undergoes proteolysis in *E. coli*, likely during expression since protease inhibitors could not prevent it. The impurities that resulted from this proteolysis were not major factors in the final purification of GIP. The purification step that had the biggest impact in the final yield of GIP was the buffer exchange, acclimatizing SUMO-GIP for the SUMO protease reaction and the following IMAC step. The removal of the SUMO tag via the SUMO protease was very robust and reliable. The last chromatographic step presented no issues; it always resulted in pure GIP being present in the flow-through and wash fractions, whereas SUMO came out in the elution. Ultimately, this protocol yielded ~4 mg of GIP per 1 L of culture, with a purity of ~85%. This enzyme was also fully active as shown by the comparable kinetic parameters (*K_m*, *k_cat*, and *k_cat/K_m*) and pH profile to its ortholog, SCP1, and homolog, FCP1. Moreover, there was no obvious protocol for the purification of full-length SCP1 for us to follow. Considering the similarities to GIP, our protocol should also serve as the first protocol for the over-expression and purification of full-length human SCP1.
Chapter 3: Protein Interactions
3.1 Overview

There is much evidence to suggest that GIP has a potential role in transcription regulation, starting from its discovery by Campagnoni et al. in 2004, showing that it localizes and interacts with the golli protein BG21. The nucleocytoplasmic shift of the golli proteins implies a role in transcriptional regulation or cell signalling. Their early developmental expression suggests that they could regulate the differentiation of neurons and elaboration of membrane sheets in oligodendrocytes (Reyes and Campagnoni 2002). Alternatively, MBP-21.5 is also localized to the nucleus of oligodendrocytes and is expressed at the same time as the golli proteins (G. S. T. Smith et al. 2013; G. S. T. Smith, Seymour, et al. 2012; Harauz and Boggs 2013). This classic isoform also undergoes a nucleocytoplasmic shift, giving rise to the same mechanism of gene regulation in oligodendrocytes, perhaps cell differentiation and expression of the membrane-associated MBPs involved in myelin compaction. Ongoing studies are showing that MBP-21.5 has an effect on cell proliferation due to some secreted factor. It is entirely possible that GIP interacts with MBP-21.5 in a similar manner as it does with BG21. Looking closely at the closest homologs of GIP, one can also see a connection with transcriptional regulation.

GIP’s human ortholog, SCP1, is a proven CTD phosphatase closely involved in the process of transcriptional regulation via its direct interactions with RNA polymerase II (Yeo et al. 2003; Yeo et al. 2005). SCP1 has been shown to inhibit transcription in human neurons through its phosphatase activity, as inactive mutants not only did not repress transcription but actually enhanced it (Yeo et al. 2003; Yeo et al. 2005). Studies involving GIP and the CTD have yet to be documented, though there is no evidence suggesting that these interactions would not exist.
Protein-protein interactions between GIP and BG21 have been studied thus far by two groups. However, the two studies show slight disagreement. First, the Campagnoni group used immunoprecipitation to demonstrate an interaction between these two proteins. Here, it is important to note that immunoprecipitation studies have come into question over recent years, requiring other techniques for confirmation (Mackay et al. 2007). Furthermore, immunoprecipitation studies (and for that matter yeast-two hybridization) cannot differentiate between weak and non-specific interactions, especially when dealing with intrinsically-disordered proteins such as BG21 which contain a very high net positive charge, leading to a lot of non-specific binding. Subsequently, Bamm et al. (2008), used a glutaraldehyde cross-linker to investigate this interaction, and thereby showed that it was specific but very weak. During this analysis, they used a truncated form of rmGIP lacking the first 76 residues of the N-terminus. Although Campagnoni demonstrated no domain specificity of the interactions between GIP and BG21, the possibility that this interaction does not require the presence of the N-terminal coil domain cannot be ruled out. Alternatively, as Bamm et al., suggested, other factors might be required to facilitate this interaction. Other factors might be another interacting partner(s) (perhaps NLI, or MBP-21.5), or post-translational modifications such as glycosylation or phosphorylation (analogous to the SCP1 interaction with RNAP2), or both. Perhaps GIP is capable of removing phosphate groups of phosphorylated BG21 or MBP-21.5. It is also conceivable that either MBP-21.5 or BG21 (two intrinsically-disordered proteins) can act as molecular scaffolds for transcription regulation of genes involved in oligodendrocyte proliferation or differentiation, with GIP helping to regulate the formation of this complex through its phosphatase activity.
As described in the previous chapter, a new protocol for the purification of a fully-active, full-length form of GIP has been successfully developed. This protocol can be used to produce GIP and properly study its interactions with BG21 and MBP-21.5 in vitro. An enzymological approach to study these interactions was used. In 2008, Bamm et al showed that BG21 had a dual effect on the activity of ΔN-GIP, showing activation below 6 μM BG21, and inhibition above this threshold (Bamm and Harauz 2008). A similar study was done with full-length GIP here, measuring the effect that both BG21 and MBP-21.5 have on the activity of GIP using the artificial substrate. Another way of studying these potential interactions was to determine whether these two proteins were substrates of GIP if they were phosphorylated. Both MAPK and PKC were used to phosphorylate BG21 and MBP-21.5. Following this experiment, it was also determined that GIP is a CTD phosphatase. This was done by utilizing a phosphorylated peptide with the heptad repeats of the CTD, followed by analyzing the effect of BG21 on GIP with its natural substrate.
3.2 Materials and Methods

3.2.1 Purification of BG21 and MBP-21.5

Glycerol stocks of *Escherichia coli* BL21-CodonPlus(DE3)-RP culture containing the pET-22b(+) plasmid with BG21 (or MBP-21.5) were grown at 37°C in 1 L of minimal 9 medium containing ampicillin (amp)* and chloramphenicol (cam)* (25 μg/mL and 33 μg/mL respectively) until $A_{600} \approx 0.8$. The culture was then induced with 1 mM of IPTG and grown for a further 5 h. Cells were harvested by centrifugation at 7500 rpm for 12 min at 4°C, then frozen and stored at -20°C until needed. The cell pellet was resuspended and homogenized by stirring for 1 h at room temperature in 40 mL of lysis buffer (8 M urea, 100 mM NaH$_2$PO$_4$, 500 mM NaCl, 10 mM Tris, 1% Tween-20, 1 mM PMSF, pH 8.0). The lysate was then sonicated for 3 minutes on high power with 30 second cooling periods between 30 second bursts. The lysate was clarified via centrifugation at 15,000 rpm for 15 min at 4°C and crude filtered (0.45-μm membrane) before chromatography.

A 1-mL bed volume of Ni$^{2+}$-NTA resin was used in a Fast Performance Liquid Chromatography (Bio-Rad Laboratories Inc., CA, USA) system to purify the fusion protein. The system utilizes the QuadTech™ (Bio-Rad Laboratories Inc., CA, USA) detection system monitoring the protocol at 214, 260, and 280 nm. The method used involved an equilibration step of 150 mL of column buffer (8 M urea, 100 mM NaH$_2$PO$_4$, 500 mM NaCl, 10 mM Tris, pH 8.0) (1 mL/min), followed by setting the zero baseline of the detector. Protein crude was then loaded (1 mL/min) followed by a wash step of 20 mL column buffer (1 mL/min). Finally, bound protein was eluted with 10 mL of elution buffer (100 mM NaH$_2$PO$_4$, 500 mM NaCl, pH 4.5).
Flow-through and wash fractions were collected in volumes of 7 mL, while elution was collected in 1-mL fractions. All buffers used were filtered through a 0.45-µm membrane before use.

Elution fractions were pooled and serially dialyzed (3,000-5,000 MWCO tubing) against 2 L of elution buffer and water for a total of six exchanges, the last three being water alone. Dialyzed pure protein was lyophilized, and then stored at -20°C until needed.

3.2.2 Measuring effect of BG21 and MBP-21.5 on GIP activity

Lyophilized BG21 and MBP-21.5 were stored at -20°C; enough protein was weighed and dissolved in the reaction buffer (50 mM Tris-Ac, 10 mM MgCl₂ pH 5.5). Reconstituted protein was then centrifuged at maximum speed for 10 min (benchtop centrifuge) in order to remove any major precipitatant. Protein concentration was determined by A₂₈₀ and the extinction coefficients for BG21 and MBP-21.5 (ε=6,970 M⁻¹ cm⁻¹ and ε= 19,940 M⁻¹ cm⁻¹, respectively).

The activity assay for GIP was performed the same way as described in the previous chapter, using the artificial substrate p-NPP. However, the first step in this analysis was done with increasing concentrations of BG21 or MBP-21.5 while the substrate concentration was held constant. This step determined the concentrations to be used when studying these effects.

To measure the kinetic effect of BG21 on GIP, the assays were performed as before, increasing the concentration of substrate with different concentrations of BG21, based on the experiment above. The velocity graphs were then fitted to the Michaelis-Menten model in order to obtain the apparent kinetic parameters. Similarly, other known methods of analyzing kinetic data, such as the reciprocal, Dixon, and Cornish-Bowden plots, were used to help identify different mechanisms of interaction and subsequent interaction constants. The latter two plots are described in Appendix I.
3.2.3 Malachite green assay for detection of inorganic phosphate

Malachite green is an organic molecule composed of three benzene rings attached via a central carbon. Two of these rings also have a nitro-dimethyl group at the para position. The resonance structures of malachite green involve the protonation of these nitrogens. This gives malachite green ionic properties allowing it to be coupled with other ions, forming a salt. It is commonly combined with ammonium molybdate in colourimetric assays designed to detect inorganic phosphate. These assays work on the principle of the orthophosphate molecule binding and forming a complex with ammonium molybdate under harsh acidic conditions. This composite then binds the malachite green ion, forming a new malachite green-phosphomolybdate complex. This new complex absorbs light in the range of 620-650 nm, appearing green. The absorbance at this range is proportional to the free inorganic phosphate concentration. When this assay is performed, the amount of inorganic phosphate is determined by comparing to a standard curve of NaH₂PO₄ salt. The structure of malachite green and a scheme of the chemistry is shown in Figure 3.1.

Modern versions of this assay use a single reagent solution which is composed of 3 parts malachite green reagent and 1 part ammonium molybdate (in acid), due to the stoichiometry of the reaction (Geladopoulos, Sotiroudis, and Evangelopoulos 1991). Unfortunately, this solution must be made fresh right before use, since the salt formed between these two molecules precipitates, eventually. The more concentrated the components, the quicker the precipitate forms, but also the sensitivity of the assay is reduced. The concentrations used in this study are the same used by Geladopoulos et al. (1991), at these concentrations, minimal precipitate was observed. However, in the presence of the orthophosphate, the malachite green-
Figure 3.1. A) The chemical structure of malachite green ion, the compound used for detecting inorganic phosphate. B) The chemical reaction scheme of the malachite green assay used.

Figure courtesy of http://bitesizebio.com.
phosphomolybdate precipitate will be observed after a few minutes. In order to minimize any interference on measurements by this precipitate, the reactions should be measured immediately after adding the detection reagent.

3.2.4 Protocol for malachite green assay

The malachite green assay was performed as a discontinuous colourimetric assay in order to study the kinetic activity of GIP on the CTD-peptide. Frozen GIP aliquots were thawed and desalted into 50 mM Tris-Ac, pH 5.5, and 10 mM MgCl₂, for experimental use. The GIP concentration was obtained using the $A_{280}$ measurement and the extinction coefficient 20,063 M⁻¹ cm⁻¹. Reaction mixtures were done in 50 mM Tris-Ac pH 5.5, 10 mM MgCl₂, and increasing concentrations of CTD-peptide (0-90 μM) with a final volume of 50 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were started by the addition of 12.5 nM GIP. After 5 minutes, the reactions were terminated by boiling, then frozen until developed.

Three parts of 0.045% malachite green in H₂O were combined with one part 4.2% ammonium molybdate in 4 M HCl to form the developing reagent. The reagent was then centrifuged at maximum speed for 10 min to remove any formed precipitate. The 50-μL reaction mixtures and standards were aliquoted into a 96-well plate, and a multi-channel pipette was used to add 100 μL of the developing reagent solution to all wells simultaneously. Absorbance at 620 nm was then immediately measured. A standard curve of NaH₂PO₄ (0-70 μM) was used to quantify the inorganic phosphate present. All reactions were done in triplicate and repeated with several preparations of GIP. Kinetic analysis was carried out as described in the previous chapter (section 2.2.4).
3.3 Results

3.3.1 Effects of BG21 and MBP-21.5 on GIP

An enzymological approach was used to study the interactions between GIP and both BG21 and MBP-21.5. The first step involved repeating V. Bamm's experiment in 2008, where he measured the effect that BG21 had on the activity of ΔN-GIP, but with now a functional full-length GIP. In this assay, the concentration of the artificial substrate p-NPP was held constant at 10 mM. Meanwhile, increasing concentrations of BG21 and also of MBP-21.5 were used. To analyze the effect on the activity of GIP, the velocity of the reaction recorded was normalized as a percentage, with no potential interacting protein being 100%.

From Figure 3.2 it is evident that BG21 had an inhibitory effect on GIP. The major drop in activity started with about 4 μM BG21, and saturation occurred at about 15 μM BG21. BG21 managed to reduce the activity observed down to 40%. Meanwhile, MBP-21.5 did not show any evident pattern affecting the activity of GIP. This assay showed that the activity went up as high as 130% but also down as low as 80%, while not changing much with as high as 30 μM MBP-21.5. From these results, the inhibitory effect of BG21 on GIP was further investigated, while MBP-21.5 was not used further.

To study the inhibitory effect of BG21 on GIP, the same assay used in Chapter 2 was used, at various concentrations of BG21. The concentrations chosen were 0, 4, 6, 8, and 12 μM BG21 based on the previous experiment and V. Bamm's experiment with ΔN-GIP (Bamm and Harauz 2008). Figure 3.3 shows the various velocity graphs at each BG21 concentration. The curves are progressively lower with increasing concentrations of BG21. All curves have minimal
Figure 3.2. Profile of the effect of BG21 and MBP-21.5 on the activity of GIP. Reaction mixtures were prepared in 50 mM Tris-Ac, pH 5.5, 10 mM MgCl₂ with 10 mM p-NPP in a final volume of 100 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were initiated by the addition of 50 nM GIP. After 10 minutes, 1 M NaOH was added to terminate the reaction, and the A₄₀₅ value was measured. (A) Effect of (0-30 μM) BG21 on the percent of GIP activity, with no BG21 being 100%. (B) Effect of (0-30 μM) MBP-21.5 on the percent of GIP activity, with no MBP-21.5 being 100%. These experiments were done twice, each time with a separate GIP preparation.
Figure 3.3. Inhibition of GIP by BG21. Reactions were run in 50 mM Tris-Ac, pH 5.5, 10 mM MgCl₂ with increasing concentrations of p-NPP (0-15 mM) in a final volume of 100 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were initiated by the addition of 50 nM GIP. After 10 minutes, 1 M NaOH was added to terminate the reaction and the A₄₀₅ value was measured. This experiment was repeated with different concentrations of BG21 (0-12 μM). Each of these experiments was done in triplicate, and repeated two different times using two separate GIP preparations; error bars represent the standard deviation of these individual experiments. Data were fitted to the Michaelis-Menten model using a non-linear fit with OriginPro 8 (OriginLab, MA, USA). The χ² values of each fitted curve indicate an excellent fit of the model to the measured data.
Figure 3.4. Cornish-Bowden (A) and Dixon (B) plots of the inhibition of GIP in the presence of various concentrations of BG21.
error and they each fit the Michaelis-Menten model very well, revealing the apparent kinetic parameters. Reciprocal plots of this inhibition showed no particular pattern, suggesting a mixed inhibition mechanism. The Dixon (1/V vs. [I]) and Cornish-Bowden ([S]/V vs. [I]) plots were used to determine $K_i$ and $K'_i$, respectively (Figure 3.4) (Dixon 1953; Cornish-Bowden 1974). These plots are described in more detail in Appendix I where it is shown that the point of intersection of the lines at each substrate concentration correspond to the negative $K_i$ (or $K'_i$) value. From these plots, a $K_i$ of $16.6 \pm 0.26 \mu M$ and $K'_i$ of $5.7 \pm 0.05 \mu M$ were found for BG21 on GIP with $p$-NPP.

3.3.2 BG21 and MBP-21.5 as substrates

The next step in studying the various protein-GIP interactions was to determine a potential natural substrate for GIP. In order to do this, recombinant BG21 and MBP-21.5 were phosphorylated with both MAPK and PKC. The phosphorylation reactions were carried out and analyzed via HPLC and later confirmed by mass spectrometry. The addition of phosphate groups caused a slight shift in the HPLC profile of each protein. Another potential natural substrate of GIP would be the CTD of RNAP2 which comprises a number of heptad repeats. A synthetic di-heptad peptide with the second serine at position 5 (of the repeat, 12 overall) phosphorylated was ordered from Biomatik Corp. (Cambridge, ON).

Following the preparation of these potential substrates, 90 μM of each were subjected to potential de-phosphorylation by GIP. The malachite green assay was performed to detect any potential inorganic phosphate produced by the removal of the phosphate group from these proteins. Figure 3.5 shows that the removal of phosphate from the CTD-peptide was readily catalyzed by GIP. Meanwhile, both BG21 and MBP-21.5 had very low levels of de-
**Figure 3.5.** Measurement of amount of inorganic phosphate release from potential natural substrates of GIP. Reaction mixtures were done in 50 mM Tris-Ac, pH 5.5, 10 mM MgCl$_2$ with constant concentration of each substrate (90 μM) in a final volume of 50 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were initiated by the addition of 50 nM GIP. After 10 minutes, the reactions were terminated by boiling and transferred to a 96-well plate where 100 μL of the fresh malachite green reagent was added and the absorbance at 620 nm was read immediately. Substrates shown are PKC-phosphorylated BG21 and MBP-21.5, and the synthetic 14-mer CTD-peptide consisting of two heptad repeats with the serine at position 5 (of the second repeat) phosphorylated. This experiment was done in triplicate; error bars represent the standard deviation of these individual experiments.
phosphorylation. From this result, only the catalysis of the CTD-peptide by GIP was characterized. Only proteins phosphorylated with PKC are shown, since no de-phosphorylation occurred with MAPK-phosphorylated BG21 and MBP-21.5.

3.3.3 GIP as a CTD phosphatase

To study the kinetics of GIP towards the CTD-peptide, the malachite green assay was used with increasing concentration of the substrate. Plotting the absorbance at 620 nm over time against CTD concentration yielded a classic hyperbolic function. The Michaelis-Menten equation was fitted to this curve via non-linear regression done in OriginPro 8.0 (OriginLab, Massachusetts, USA). This experiment was done in triplicate and repeated two different times using two separate GIP preparations. The fitting revealed a $K_m$ of $69.5 \pm 0.34 \, \mu M$, a $k_{cat}$ of $2.54 \pm 0.06 \, s^{-1}$, and in turn a $k_{cat}/K_m$ number of $36.5 \pm 0.54 \, mM^{-1} \, s^{-1}$. A typical velocity curve is shown in Figure 3.6.

3.3.4 Effect of BG21 on GIP with CTD-peptide

Having a 'natural' substrate for GIP to work with, the next logical step was to see whether BG21 had the same inhibitory effect on the activity of GIP with this substrate. The same experiment where measuring the percent activity of GIP with increasing concentrations of BG21 was done. This experiment revealed a very different and very interesting result. BG21 did not inhibit GIP with the CTD as substrate, but rather linearly enhanced the activity as high as 170% (Figure 3.7 A). The concentration of BG21 tested was as high as 100 $\mu M$, yet no saturation of the enhancement was observed. Subsequently, the kinetic experiments were set up with 0-25 $\mu M$ BG21. The reciprocal plot of the velocity curves reveals a mechanism of activation.
Figure 3.6. Velocity curve for kinetic analysis of GIP with the synthetic CTD-peptide as a 'natural' substrate. Reaction mixtures were done in 50 mM Tris-Ac, pH 5.5, 10 mM MgCl₂ with increasing concentration of substrate (0-90 μM) in a final volume of 50 μL. Reaction mixtures were incubated at 37°C for 5 minutes to stabilize the temperature. Reactions were initiated by the addition of 12.5 nM GIP. After 5 minutes, the reactions were terminated by boiling and transferred to a 96-well plate, where 100 μL of the fresh malachite green reagent was added and the absorbance at 620 was read immediately. This experiment was done in triplicate and repeated two different times using two separate GIP preparations; error bars represent the standard deviation of these individual experiments. Data were fitted to the Michaelis-Menten model using a non-linear fit with OriginPro 8 (OriginLab, MA, USA). A standard curve of NaH₂PO₄ was used to quantify the amount of inorganic phosphate released. The χ² value of 1.75E-5 indicates an excellent fit of the model to the measured data.
Figure 3.7. Studying the effect of BG21 on GIP with the synthetic CTD-peptide as a 'natural' substrate. The GIP and malachite green assay was performed as before. A standard curve of NaH$_2$PO$_4$ was used to quantify the amount of inorganic phosphate released. (A) Profile of the effect of increasing concentrations of BG21 (0-100 μM) on the percent activity of GIP, with 90 μM CTD. (This experiment was only performed once.) (B) Lineweaver-Burke plot of the kinetic data shows an activation mechanism that mimics a classic competitive inhibition pattern. (C) Dixon plot shows a mirror image of a classic competitive inhibition pattern as BG21 enhances GIP activity. (D) Cornish-Bowden plot shows a mirror image of a classic competitive inhibition pattern as BG21 enhances GIP activity. The kinetic experiment was done in triplicate, and repeated twice.
resembling competitive inhibition (Figure 3.7 B). The Dixon and Cornish-Bowden plots were again used for the analysis of this mechanism (Figure 3.7 C and D). This analysis revealed equal $K_a$ and $K'_a$ values of $30.1 \pm 0.11 \mu M$.

In order to see whether phosphorylation affected this effect, the same experiment was done with PKC phosphorylated BG21. The concentrations used in the profile experiment were, again, as high as 100 $\mu M$. Interestingly though, phosphorylated BG21 inhibited the activity of GIP with CTD. Activity was reduced as low as 70% after only 15 $\mu M$. Above this concentration, there was a slight and gradual increase in activity, saturating at 75%. In order to find the mechanism of inhibition, the concentrations of phosphorylated BG21 used for the kinetic analysis were 0-16 $\mu M$. The reciprocal plot reveals a mixed mechanism of inhibition. Again the Cornish-Bowden and Dixon plots were utilized to analyze these data. A $K_i$ of $3.5 \pm 0.09 \mu M$ and $K'_i$ of $8.03 \pm 0.24 \mu M$ were found for the inhibition of GIP by PO$_4$-BG21 when using CTD as substrate.
Figure 3.8. Studying the effect of BG21 on GIP with the synthetic CTD-peptide as a 'natural' substrate. The GIP and malachite green assay was performed as before. A standard curve of NaH$_2$PO$_4$ was used to quantify the amount of inorganic phosphate released. (A) Profile of the effect of increasing concentrations of PO$_4$-BG21 (0-100 μM) on the percent activity of GIP, with 90 μM CTD-peptide. (This experiment was only performed once.) (B) Dixon plot analysis of the mixed inhibition by PO$_4$-BG21 on GIP with CTD. (C) Cornish-Bowden plot analysis of the mixed inhibition by PO$_4$-BG21 on GIP with CTD-peptide. The kinetics experiment was done in triplicate, and repeated twice.
3.4 Discussion

Both BG21 and MBP-21.5 belong to the class of intrinsically-disordered proteins. These proteins do not have a native defined structure like other well known proteins (Dyson 1996; Wright and Dyson 1999). They do, however, have the ability to shape-shift and adopt certain structural elements upon exposure to other factors. These factors can range from small metal ions to large proteins (Uversky 2009). The classic MBP isoform of 18.5 kDa is a prime example of this phenomenon. It is able to form three separate α-helices upon interaction with certain partners such as membranes, Ca^{2+}-calmodulin, and actin and tubulin, among others (Bamm et al. 2011; Reyes and Campagnoni 2002). This isoform of MBP and its fluctuating structure has been the subject of much research lately. It has become known as a 'hub' for other proteins, providing unique functions as it interacts with each partner. As MBP-18.5 shows, IDP's provide an immense degree of multifunctionality, allowing for one single protein to modify itself and carry out more than just one task. Unlike MBP-18.5, not much is known about any interacting partners of BG21 or MBP-21.5, much less the roles that these proteins might have.

The work done by Fernandes and colleagues (Fernandes et al. 2004) revealed a potential interacting partner in GIP. Their immunoprecipitation studies suggested strong interactions between these two proteins. However, a few years later in 2008, Bamm et al. showed a slightly contradictory result. He used a gel-shift assay with glutaraldehyde to show very specific interactions between GIP and BG21, but this association was rather weak. Unfortunately this work was done with a truncated form of GIP, lacking the first 76 residues, due to the inability to produce recombinant full-length and fully active GIP. As shown in the previous chapter, a new protocol was successfully developed in this thesis work for this purpose.
3.4.1 Inhibition of Full-length GIP with p-NPP by BG21

As part of his work with ΔN-GIP and BG21, V. Bamm showed that BG21 had a dual effect on the activity of ΔN-GIP using p-NPP as the substrate. At concentrations below 6 μM BG21 enhanced the activity, while above 6 μM it inhibited ΔN-GIP. In this thesis, the same effect was measured on full-length GIP. As expected, BG21 inhibited the ability of full-length GIP to dephosphorylate p-NPP. However, only inhibition was significantly observed. The difference between the effect of BG21 on ΔN-GIP compared to full-length GIP lies within that N-terminus. Perhaps this section of GIP is responsible for helping either recognition or binding of other proteins, and thus protein-protein interactions. Since MBP-21.5 has a lot in common with BG21, it was of interest to see if it would interact with GIP as well. The same experiment was done with MBP-21.5; however it yielded no significant effect and no obvious evidence of interactions between it and GIP.

The main difference between BG21 and MBP-21.5 is the net charge. BG21 is neutral whereas MBP-21.5 is highly positively-charged at neutral pH, as are the other classic MBP isoforms. GIP is an acidic phosphatase with an overall negative charge, at neutral pH, and a pI around 5.7 GIP (Bamm and Harauz 2008) did not seem to interact with the positively-charged MBP-21.5, but it did interact with the neutral BG21. This observation implies that the interactions between BG21 and GIP are not likely to depend upon global electrostatics in an obvious way.

Since only inhibition (no dual effect as seen with ΔN-GIP) full-length GIP was observed with BG21, a full kinetic study was possible. This study revealed a mixed inhibition mechanism. In 1973, Cornish-Bowden developed a new linear plot for analyzing inhibition kinetics that
complemented the traditional Dixon plot (Cornish-Bowden 1974). Using this plot allowed for the easy determination of the inhibition constant between the inhibitor and the enzyme-substrate complex, whereas the traditional Dixon plot revealed the inhibition constant between the inhibitor and the free enzyme. There is a relatively large level of mixed inhibition as evident from the large difference between the $K_i$ and $K'_i$ (~16 μM and 5.7 μM, respectively) values. This observation means that BG21 is able to bind the GIP-p-NPP complex with greater affinity (~3-fold) than GIP alone. Regardless whether p-NPP was present or not, these data suggest that the interaction between GIP and BG21 is strong enough to be observed via immunoprecipitation, as Fernandes et al. showed in 2004. This agreement with the immunoprecipitation studies explains that the findings from V. Bamm et al. were likely due to the missing N-terminus. The lack of this domain of GIP had a detrimental effect on the strength of the interaction between these two proteins, but not its specificity (Bamm and Harauz 2008).

3.4.2 Phosphorylated BG21 and MBP-21.5 as Substrates

Another avenue explored when studying the protein interactions of GIP was to determine if phosphorylated BG21 and MBP-21.5 were potential natural substrates. It was revealed that neither BG21 nor MBP-21.5 was significantly dephosphorylated by GIP. The MBP-21.5 isoform was phosphorylated by MAPK and also by PKC. MAPK is specific to the threonine residues in the SH3-ligand region of MBP. No de-phosphorylation was observed from MAPK-phosphorylated MBP-21.5. In contrast, PKC-phosphorylated protein did have a small degree of phosphate removed (Figure 3.5). This result was not surprising, since the CTD phosphatases are serine phosphatases, and although serine phosphatases are usually grouped with threonine
phosphatases (Lirui W. 2010; Cohen 1997), these families are not always interchangeable.

As mentioned, PKC-phosphorylated BG21 also had similar levels of de-phosphorylation. These levels were very low, and were only observed at very high concentrations of phosphorylated proteins, indicating that this was more likely an artificial phenomenon. At this point in the study, it was decided to continue with BG21 alone, since MBP-21.5 did not show any significant interactions with GIP as was thought possible before the work was done. It is evident that MBP-21.5 does not function in the same manner as BG21 with GIP, despite the many similarities that it shares with BG21. Further research needs to be carried out with MBP-21.5 in order to help elucidate its role in oligodendrocytes and myelin, but as shown here, a direct interaction with GIP is not evident \textit{in vitro}.

3.4.3 GIP as a CTD-phosphatase

In order to demonstrate that GIP is a CTD phosphatase, a small single phosphorylated peptide consisting of two heptad repeats of the CTD domain of RNAP2 was synthesized. In 2006, Zhang \textit{et al.} wanted to determine which phosphate of the CTD repeat SCP1 preferentially de-phosphorylated. To do this, they utilized several synthesized CTD-peptides with different serines phosphorylated. They concluded that SCP1 had greater affinity and activity towards the de-phosphorylation of the serine at position 5 in the heptad repeat. They also showed that this activity increased as the peptides included more repeats with S$_5$ phosphorylated, though enough activity was observed with even a 7-residue peptide with only one repeat (Y. Zhang \textit{et al.} 2006). From this previous work and for economical reasons, it was decided here to order a 14-mer containing two repeats with only the serine at position 5 of the second repeat
phosphorylated for our study. The reason that the serine of the second repeat was phosphorylated, and not the first, was simply because the study with SCP1 had used this exact peptide with this particular serine phosphorylated. They did not use a peptide where it was the first repeat that was phosphorylated; it was also not discussed why they chose the second repeat to have the phosphate group. As was done earlier with BG21 and MBP-21.5, this peptide was subjected to treatment with GIP at 90 μM to determine if the phosphate group was removed. In contrast to the other two proteins, the CTD-peptide was readily de-phosphorylated by GIP.

A kinetic study of this reaction revealed a $K_m$ of 69.5 ± 0.34 μM, a $k_{cat}$ of 2.54 ± 0.06 s⁻¹, and in turn a $k_{cat}/K_m$ number of 36.5 ± 0.54 mM⁻¹ s⁻¹. These numbers suggest that the CTD-peptide is likely its natural substrate, as the catalytic efficiency number for GIP with this peptide ($\sim 36$ mM⁻¹ s⁻¹) is $\sim$50-fold stronger than with the artificial substrate ($\sim 0.76$ mM⁻¹ s⁻¹). More importantly, these values are in line with what was reported with SCP1 with this substrate. However, there is still a significant difference between GIP and the reported $K_m$ for SCP1 (36.53 mM⁻¹ s⁻¹ and 11.4 mM⁻¹ s⁻¹, respectively). The difference lies in a $\sim$3-fold higher $K_m$ value (214 ± 0.51 μM), which in turn decreases the specificity number of SCP1 by 3-fold. It is important to note that this study used the truncated form of SCP1 which is identical to truncated GIP, since the differences between these two orthologs lie in the N-terminus. This suggests that the N-terminus helps binding of the CTD substrate. This is another piece of evidence to suggest the N-terminus of GIP is important for the mediation of protein interactions with GIP, though not essential.
3.4.4 Effect of BG21 and PO₄-BG21 on GIP with CTD-peptide

Continuing with the investigation of the interactions between BG21 and GIP, the effect of BG21 on GIP with this new 'natural' substrate was investigated. This revealed a surprising result, as BG21 enhanced the activity of GIP in a linear fashion. This linearity would suggest essential activation; however, this is really a case of non-essential activation, since GIP is able to efficiently catalyze the CTD-peptide on its own. Saturation of this enhancement was simply not observed at these concentrations of BG21 even as high as 100 μM. The kinetic analysis of this enhancement was carried out to determine the mechanism of interaction between BG21, and GIP with the CTD. The reciprocal plot showed a mechanism that mimics competitive inhibition. The Dixon and Cornish-Bowden plots in combination confirmed this conclusion. These plots are mirror images of a true competitive mechanism.

A competitive mechanism implies that the inhibitor and the substrate are mutually exclusive when binding the enzyme. For an activator, this is impossible, since the activator must be interacting with the enzyme and the substrate in order to enhance the reaction between enzyme and substrate. A mechanism that describes this phenomenon is known as competitive activation. The term "competitive" is only used because of the resemblance to competitive inhibition, even though no actual competition is occurring. Here, the overall effect (opposite from competitive inhibition) is that, as the activator binds the enzyme, it increases the ability or affinity of the enzyme to bind the substrate, without affecting the ability of the enzyme to catalyze the substrate. This means the GIP-BG21-CTD complex is equally capable of producing product as the GIP-CTD complex. As a result, as the BG21 concentration is increased, the apparent $K_m$ is decreased, whereas the rate or $V_{max}$ of catalysis is unchanged. This mechanism
gives the impression that there is more GIP available for catalysis, as it is able to bind more substrate.

After seeing this unexpected effect of BG21 on GIP with the CTD-peptide (Figure 3.7), the same experiment was done with BG21 that had been phosphorylated with PKC (Figure 3.8). This yielded interesting results, showing that PO₄-BG21 inhibited GIP with the CTD as substrate. This profile showed that PO₄-BG21 inhibited GIP until a saturation point at about 20 μM of inhibitor. Past this concentration, a slight recovery of ~5% of activity was observed. It is of note that at high concentrations of PO₄-BG21, GIP was able to de-phosphorylate BG21 (Figure 3.5). What seems like a recovery of activity is more likely the release of phosphate resulting from the small level of dephosphorylation of BG21 at these concentrations. With this possibility in mind, the kinetic analysis was carried out with concentrations of BG21 that fit within the linear portion of this inhibition pattern. The results revealed a mixed mechanism of inhibition. Under this mechanism, the values of $K_i$ and $K'_i$ differ by a factor of 0.44. This implies that BG21 is able to bind the free enzyme as well as the GIP-CTD complex, although it prefers the free enzyme. Upon PO₄-BG21 binding, a minimal effect on the ability of GIP to bind the CTD-peptide is seen. This results in an insignificant change in the $K_m$ of GIP towards the peptide. Moreover, once this BG21-GIP-CTD complex is formed, it is still able to de-phosphorylate the CTD substrate, albeit with much less efficiency, as is reflected by the dramatic change in apparent rate of catalysis.

3.4.5 Working Model for BG21-GIP Interactions

To summarize, the interactions of BG21 and GIP were studied using an enzymological approach. Utilizing the artificial substrate $p$-NPP, BG21 inhibited GIP through a mixed inhibition
mechanism. The $K_i$ and $K'_i$ values imply that BG21 interacts with GIP about 3-fold stronger when it is associated with the substrate than with the free enzyme. When using the CTD-peptide as substrate, BG21 has a very different effect. BG21 is unable to bind GIP with as high affinity as when using the artificial substrate. This is shown by its $K_a$ of ~30 µM compared to the $K_i$ of ~16 µM when using $p$-NPP. Interestingly, it is able to bind the free GIP equally as well as the GIP-CTD complex. However, upon binding, it improves the affinity of GIP towards the substrate, dramatically enhancing the activity of GIP without affecting the hydrolysis of the phosphoprotein intermediate. When BG21 is phosphorylated by PKC, it binds the free enzyme with greater affinity than the GIP-CTD complex. This is the opposite from that found with the artificial substrate. Moreover, there is a large difference when comparing the $K_i$ values from BG21 with the artificial substrate (~16 µM) to that of PO$_4$-BG21 with the CTD-peptide (~3.6 µM). Phosphorylated BG21 shows a 5-fold increase in affinity towards the free enzyme over BG21. However, when looking at the $K'_i$ values, the difference is not very significant (0.7-fold), meaning that the phosphorylation of BG21 does not seem to have a great effect on the binding of GIP to its substrate. This point is further supported by the mixed mechanism by which PO$_4$-BG21 inhibits GIP, where the $K_m$ of the reaction is not significantly affected. To understand how BG21 interacts with GIP, a model was developed and is shown in Figure 3.9.

When looking at this model, it is important to remember that, as was shown earlier, GIP is much better at catalyzing dephosphorylation of the more 'natural' CTD-peptide than the artificial substrate $p$-NPP (Figure 3.9 A). Here, conformational changes in GIP upon interactions between the various partners are depicted in the shape change. These conformational changes are designed to help explain and visualize how BG21, along with the substrates, influence how
GIP is able to carry out its reaction and explain all the phenomena observed throughout the kinetic analysis. In 2004, Kamenski et al. analyzed the crystal structure of SCP1, and showed that SCP1 contained two main structurally-conserved domains among some members of the HAD superfamily. The first domain was identified as the 'core' domain where the active site is present. The other domain is known as the 'insertion' domain where the substrates are believed to be inserted through and into the active site. This domain is made up by one of the two β-sheets (three strands, spanning residues 102-134) in SCP1 (Kamenski et al. 2004). Since this region of SCP1 is likely responsible for substrate binding, this is also very likely the region where BG21 interacts with GIP when the binding to the substrate is affected. Similarly, whenever the substrate binds, this region likely undergoes a conformational change that may affect how BG21 or PO₄BG21 binds GIP. In this model, this insertion domain is highlighted by the lighter colour and it is shown how it could undergo conformational changes upon interactions of GIP with its partners. What could not be explained by this model is the difference in interaction constants with the free enzyme (Ki) between the analysis of BG21 with p-NPP and BG21 with the CTD-peptide. Looking at these mechanisms, the constants (Ki and Ka) describe only the interaction between BG21 and free GIP. However the inhibition constant of BG21 was ~16 μM, whereas the activation constant with BG21 was ~30 μM. This 2-fold difference could not be explained by this model and perhaps is not significant.
**Figure 3.9.A** - Model representing how GIP is able to bind and catalyze the artificial substrate ($p$-NPP, blue triangle) compared to the CTD-peptide (blue shape). GIP undergoes a slight conformational change within the insertion domain (light orange) upon binding of the substrate.
Figure 3.9.B - A model representing how BG21 and p-NPP bind to GIP, resulting in a mixed inhibition mechanism where $K_i > K_i'$, $K_m < K_m'$, and $k_2 < k_2'$. As BG21 (green) binds GIP, it causes a conformational change within the insertion domain. This change, results in a reduced ability of GIP to bind and catalyze the cleavage of the artificial substrate. Meanwhile, the conformational change along the insertion domain upon the binding of p-NPP allows BG21 to bind the GIP-substrate complex with higher affinity than the free enzyme. Overall, the BG21-GIP-p-NPP complex cannot catalyze the cleavage of the substrate as efficiently. This effect, combined with BG21 reducing the ability of GIP binding the substrate, results in the inhibition pattern observed.
Figure 3.9.C - A model representing how BG21 and the CTD-peptide bind to GIP, resulting in a 'competitive' activation mechanism where $K_a = K_a'$, $K_m > K_m'$, and $k_2 = k_2'$. As BG21 binds GIP, it causes the same conformational change within the insertion domain as before. This time, it greatly enhances the ability of GIP to bind the CTD-peptide. This change however does not change the ability of GIP to catalyze the reaction. Meanwhile, the conformational change along the insertion domain upon the binding of the CTD has no effect on the binding between GIP and BG21. Overall, the BG21-GIP-CTD complex is equally apt at efficiently catalyzing the substrate as free GIP, but its increased affinity towards the substrate results in the enhancement of activity observed.
Figure 3.9.D - A model representing how PO₄-BG21 and the CTD-peptide bind to GIP, resulting in a mixed inhibition mechanism where $K_i < K_i'$, $K_m = K_{m'}$, and $k_2 > k_{2'}$. PO₄-BG21 is able to bind the free enzyme with greater affinity than non-phosphorylated BG21. This interaction is also stronger than that with the GIP-CTD complex. As it does this, it does not cause a major conformational change on the insertion domain of GIP, thus not greatly affecting the binding of GIP to the CTD. However the small change does result in GIP not able to efficiently catalyze the CTD substrate. Meanwhile, the conformational change GIP the insertion domain undergoes upon the substrate binding does prevent BG21 binding as efficiently to GIP. Overall, this leads to the inhibition pattern observed.
Overall, as this model shows, BG21 interacts strongly with GIP ($K_i=16\ \mu M$, $K_a= 30\ \mu M$). However, these interactions are greatly enhanced when a new partner is involved. The only partner that did not enhance the affinity of BG21 towards GIP was the CTD-peptide. The artificial substrate p-NPP greatly improved this affinity, from a $K_i$ of $\sim16\ \mu M$ to a $K_i'$ of $\sim5.4\ \mu M$. However, what had the largest effect on this affinity was the phosphorylation of BG21. Phosphorylated BG21 was able to interact with the free enzyme much better, reducing the $K_i$ from $\sim16\ \mu M$ to $\sim3.5\ \mu M$. It also reduced the $K_i'$ of BG21 with the GIP-CTD complex from $\sim30\ \mu M$ to $\sim8\ \mu M$.

3.4.6 Chapter Summary

This study shows the CTD would be a natural substrate for GIP. Considering this conclusion, BG21 is able to bind GIP and enhance its ability to bind the CTD of RNAP2 resulting in enhanced CTD de-phosphorylation. This binding is unaffected if BG21 is phosphorylated by PKC, but instead the ability of GIP to catalyze its reaction is reduced, leading to less de-phosphorylation of CTD. Phosphorylated BG21 also has greater affinity towards GIP, which increases its efficiency at inhibiting GIP. This means that the phosphorylation of BG21 by PKC acts as a quick-brake mechanism of reducing the de-phosphorylation of the CTD.

There have been numerous studies on the CTD code and what the phosphorylation states involve in the biology of the eukaryotic cell (Egloff and Murphy 2008). Phosphorylation of the CTD at Ser5 recruits the 5' methylguanosine capping enzyme, whereas phosphorylation at S$_2$ recruits the polyadenylation capping enzyme (Schroeder et al. 2000). More importantly, RNAP2 cannot be recruited to the transcription initiation complex when it is phosphorylated.
This makes the CTD phosphatases an essential part of transcription as they recycle the RNAP2 after each transcriptional process (Egloff and Murphy 2008).

What this present study shows is a link between the golli protein BG21 and regulation of transcription. Prior to this, it was only speculated that BG21 would have a role in this process. As an activator of GIP, BG21 would help increase the turnover of RNAP2. This would lead to an increase in transcriptional activity. However, if PKC phosphorylates BG21, it would instead reduce the activity of GIP, resulting in less RNAP2 being de-phosphorylated and recycled, and in turn reducing transcriptional activity. It is important to note that the research presented here does not provide evidence for these effects. It merely illustrates a possible mechanism by which BG21 and GIP (and PKC) can affect the cellular process of gene transcription. This is a carefully balanced process, and the work by Yeo et al. in 2003 helps put this in perspective by showing the effects of over-expression levels of SCP1. This group discovered that over-expression of wild type SCP1 significantly inhibited the transcription of several promoter-reporter gene constructs studied. Meanwhile, over-expression of an inactive mutant form of SCP1 greatly enhanced the transcription of these genes. They reasoned that the recruitment of inactive SCP1 to the transcription complex helped increased levels of S5 phosphorylation. This would in turn result in an increase of cap formation and promoter clearance. Consequently, over-expressed active SCP1 could lead to excess de-phosphorylation of RNAP2 and, in turn, lessening the efficacy of the capping of the mRNA transcript.

Another possibility for these interactions lies outside of the CNS. As mentioned in Chapter 1 (Section 1.3), BG21 is the main golli protein present in T-cells. Here, it has been shown to be a substrate for PKC, acting as a negative regulator of its pathway to regulate the T-
cell receptor and its antigen recognition sensitivity (Feng et al. 2004). The work presented here with GIP supports this importance of BG21 phosphorylation in cell-mediated events.

Overall, the work presented in this chapter shows how protein-substrate and protein-protein interactions can be mediated, whether it is the presence of the N-terminus in GIP that helps the enzyme bind to its substrate and BG21, or how the phosphorylation state of BG21 can have such opposite effects on the activity of GIP. When full-length GIP interacts with BG21, the strength of interaction is much greater than when this domain is absent. However, phosphorylation of BG21 further increases this affinity of BG21 to free GIP about 10-fold, leading to inhibition of the phosphatase activity. Similarly, when not phosphorylated, BG21 enhances the activity of GIP towards the CTD by helping this association of GIP and the substrate. This goes to show how sensitive protein-protein interactions are, as they can be affected by a modification as simple as the addition of phosphate groups, and what potential effect these interactions can have in cell processes such as gene transcription.
Chapter 4: Conclusions and Future Directions
4.1 General conclusions

The initial goal of this project was to develop a successful protocol for the purification of enzymatically-active full-length murine-GIP. After many months of trials and purifications this goal was successfully achieved, as a protocol was developed that yielded ~4 mg of GIP per 1 L of culture, with a purity of ~85%. Considering that GIP and SCP1 are orthologs, this protocol should serve as a means for the purification of full-length SCP1 as well, for which no specific protocol has hitherto been described. Characterization of the GIP enzyme revealed it was fully active as it had specificity numbers comparable to truncated SCP1, its human ortholog. A pH profile on the activity of GIP revealed that its optimum pH for activity was 5.5, and that it was fully inactive at pH <4.5 and pH >7.5. This profile resembled a carbon copy of that of the yeast homolog FCP1. Moreover, this pH analysis revealed that the pH is not essential for binding the substrate, but it is critical for the catalysis of the substrate.

Revisiting the first hypothesis of this study:

1) Full-length recombinant murine GIP from native purification conditions has better phosphatase activity than refolded rmGIP from denaturing conditions.

This hypothesis was proven to be true as the full-length GIP purified under this new protocol had a specificity number 38-fold higher than that of GIP from denaturing conditions presented in V. Bamm's work in 2008.

The next goal of this project involved using this active full-length GIP to study its interactions with BG21, as well as investigate potential interactions with MBP-21.5. Prior to this
work, the role of the N-terminus of GIP was completely unknown. Drawing a comparison between this study and that done with ΔN-GIP in 2008 revealed that the N-terminus of GIP is very important for its interactions with other partners such as the CTD of RNAP2, and BG21.

Next, it was of interest to identify a natural substrate for GIP. To do this experiment, phosphorylated variants of BG21 and MBP-21.5 were used as substrates, as well as a phosphorylated CTD-peptide containing two tandem repeats. Revisiting the second hypothesis of this study:

2) Interactions between full-length GIP and BG21 are stronger than with ΔN-GIP.
Phosphorylated BG21 is also a suitable substrate for GIP. Furthermore, GIP interacts with MBP-21.5 in a similar fashion as with BG21.

This hypothesis was multifaceted, and the first experiments examined how BG21 and MBP-21.5 affected the way GIP catalyzed the hydrolysis of the artificial substrate p-NPP. BG21 inhibited GIP via a mixed inhibition mechanism, whereas MBP-21.5 did not have any noticeable effect on this reaction. Comparing this result to the study done with ΔN-GIP in 2008, it was shown here that the interactions between full-length GIP and BG21 are stronger than that with ΔN-GIP. Meanwhile, no evidence was observed that GIP interacts with MBP-21.5, despite its similarities to BG21.

Following this part of the study, it was shown that phosphorylated BG21 and MBP-21.5 were not suitable substrates of GIP. However, the CTD-peptide was readily de-phosphorylated, providing evidence that GIP is a CTD phosphatase.
The next natural step was to investigate how BG21 affected GIP with this CTD-peptide as a substrate. BG21 proved to have a dual effect on GIP activity based on its phosphorylation state. When un-modified BG21 interacted with GIP, it enhanced the ability of GIP to bind the CTD-peptide, leading to a significant increase in overall activity. However, upon serine phosphorylation of BG21 by PKC, it was able to bind GIP more tightly, reducing the ability of GIP to catalyze the de-phosphorylation of the CTD-peptide. Overall, the phosphorylation state of BG21 influences the binding between BG21 and GIP, but not that between GIP and its substrate. This phosphorylation also acts a quick-brake mechanism to reduce the ability of GIP to de-phosphorylate the CTD of RNAP2. This result shows one manner in which BG21 and GIP could be involved in the balanced process of regulation of gene transcription, perhaps with genes involved in the proliferation or differentiation of oligodendrocytes and or neurons.

This study was the first step in paving the way for much more intense research in hopes of understanding the golli-interacting protein, BG21, and their role in oligodendrocyte and myelin biology.
4.2 Future directions

An interesting finding from the research presented here concerns the N-terminus of GIP. As initially mentioned, this region of GIP is predicted to be mainly random coil and to be very mobile. In V. Bamm’s study in 2008, it was found that when this region of GIP was removed, the interactions between GIP and BG21 were very weak but specific. Here, it is shown that the interactions between full-length GIP and BG21 were strong. At the same time, in the study by Zhang mentioned above, the SCP1 used there was also lacking the N-terminus, and this truncated SCP1 showed a 3-fold lower level of activity when compared to full-length GIP with the same substrate. All of this data implies that the N-terminus is helping the interactions between GIP and its partners. Interestingly enough, the activity of the truncated form of GIP compared to full-length was 3.5-fold better. It is possible that the N-terminus indeed helps decrease the activity of GIP, but in turn it helps with the protein interactions. Another possibility is that it also helps with the overall stability of the enzyme.

One way to determine this role is by using circular dichroism (CD) spectroscopy. This technique is based on the unique property of the backbone of proteins being sensitive to left-handed and right-handed circularly-polarized light. When the backbone is bent in the shape of a particular secondary structure element, such as an α-helix or β-strand or random coil, it absorbs the two polarized forms of this light differently. As such, this technique is helpful for studying the nature of the secondary structure of proteins. In turn, when a protein unfolds due to an external factor, such as temperature, these secondary structure elements will melt. This phenomenon can then be observed with this technique to obtain information about the overall stability of the protein. A melting temperature of the protein, and even the thermodynamics of
this event, can be measured. Comparing the thermal stability of truncated GIP to full-length GIP will provide insight into the potential role in the stability of GIP this N-terminus may provide. An experiment with truncated GIP has been performed and is presented in Appendix II. The same experiment with full-length GIP has yet to be done but it is the next step. This information will add another piece of evidence to determine the role that this N-terminus plays in the function of GIP.

The findings from this research about GIP, BG21, and how they interact with a peptide from the CTD of RNAP2, have opened the doors to more avenues of inquiry. Due to logistical issues, this study was only done with this peptide. In order to prove fully that GIP is a CTD phosphatase as well as to demonstrate any interactions that BG21 may have with these two parties, the entire phosphorylated RNAP2 CTD domain should be used. However, taking into consideration what was shown in 2003 by Zhang et al. with SCP1, using just the peptide was a great first step and it was an apt substrate for these *in vitro* studies. The only difference that they found by using this peptide was a lower activity towards the CTD. They showed that SCP1 had better activity towards CTD-peptides that contained more repeats with S₅ phosphorylated.

The major direction this project needs to take is the live cellular imaging route. An *in vitro* link between, PKC, the golli protein BG21 and transcription regulation via GIP has been established. However, it is of importance that these interactions are investigated *in cellula*. It would be interesting to test if over-stimulation of BG21 and GIP can lead to a decrease in transcription levels, as was shown with the wild-type SCP1 by Yeo et al. in 2003. Consequently, will over-expressing PKC in this system lead to an opposite effect as expected from these results? There are many questions like this that can be answered in an *in cellula* system that *in
vitro studies cannot provide. Biochemical and biophysical techniques will be required again to add more information to this on-going puzzle.
Bibliography


and is expressed in cells in the oligodendrocyte lineage in the brain.” *Journal of biological chemistry* 268 (7): 4930.


Hausmann, Stéphane, and Stewart Shuman. 2002. “Characterization of the CTD phosphatase Fcp1 from fission yeast preferential dephosphorylation of serine 2 versus serine 5.” *Journal*


Lin, P. S., M. F. Dubois, and M. E. Dahmus. 2002. “TFIIF-associating carboxyl-terminal domain phosphatase dephosphorylates phosphoserines 2 and 5 of RNA polymerase II.” Journal of


MSSC. Multiple Sclerosis Society of Canada. 2011. mssociety.ca.


through cDNA cloning.” *Journal of neuroscience research* 17 (4): 321–328.


Appendix I

i. General mixed inhibition mechanism

A general scheme (Courtesy of http://depts.washington.edu) for a mixed-inhibition system where E, S, ES, P are the free enzyme, substrate, enzyme-substrate complex and product, respectively. Similarly, I, EI, and EIS are the inhibitor, enzyme-inhibitor and enzyme-inhibitor-substrate complex, respectively. $K_m$ is the Michaelis constant, $K_i$ is the inhibitor constant, and $k_2$ is the rate constant. This mechanism is described by equation 2.

$$v = \frac{[S] \times V_{\text{max}}}{K_m \times \left(1 + \frac{[I]}{K_i}\right) + [S] \left(1 + \frac{[I]}{K_i'}\right)}$$  \hspace{1cm} \text{[Eq. 2]}
ii. Dixon plot

In 1953, Dixon et al. described a linear method that is now commonly used for the determination of the $K_i$ value of an inhibitor. The method is based on the reciprocal of Eq. 2. A plot of $1/V$ against $[I]$ at a particular substrate concentration yields a straight line defined by Eq. 3.

$$\frac{1}{V} = \left(\frac{K_m}{V_{\text{max}} \times [S] \times K_i}\right) \times [I] + \frac{1}{V_{\text{max}}} \left(1 + \frac{K_m}{[S]}\right)$$  \hspace{1cm} \text{[Eq. 3]}

This linear plot produces intersecting lines at each substrate concentration upon which the abscissa coordinate from the point of intersection corresponds to the $-K_i$, as shown below.

When two lines intersect, $1/V$ and $[I]$ will be equal as well as $V_{\text{max}}$. Therefore,

$$\frac{K_m}{[S]_1} + 1 + \frac{K_m \times [I]}{[S]_1 \times K_i} = \frac{K_m}{[S]_2} + 1 + \frac{K_m \times [I]}{[S]_2 \times K_i}$$  \hspace{1cm} \text{[Eq. 4]}

or

$$\frac{1}{[S]_1} \left(1 + \frac{[I]}{K_i}\right) = \frac{1}{[S]_2} \left(1 + \frac{[I]}{K_i}\right)$$  \hspace{1cm} \text{[Eq. 5]}

This situation is only true if $[S]_1=[S]_2$ or $[I]=K_i$; therefore, the intersection point of the lines corresponding to two different substrate concentrations reveals the $K_i$ value.
iii. Cornish-Bowden plot

Dixon considered a competitive mechanism and developed his linear plot to describe binding of the inhibitor to the free enzyme, but this method is not useful for distinguishing between competitive and mixed inhibition where the inhibitor and substrate are not mutually exclusive. In 1974, Cornish-Bowden described a new linear plot that described a simple way of measuring the $K_{i}'$ value of an inhibitor which represents the binding of the inhibitor to the enzyme-substrate complex. Multiplying the reciprocal of Eq. 2 by [S] yields the following.

$$\frac{[S]}{v} = \frac{Km}{Vmax} \left(1 + \frac{[I]}{Ki} \right) + \frac{[S]}{Vmax} \left(1 + \frac{[I]}{Ki'} \right)$$  \hspace{1cm} [Eq. 6]

A plot of $[S]/v$ vs. [I] yields a straight line for each substrate concentration. The point of intersection of the lines is given by

$$\frac{[S]}{v} = \frac{Km \times \left(1 - \frac{Ki'}{Ki} \right)}{Vmax} \hspace{1cm} \text{and} \hspace{1cm} [I] = -Ki'$$  \hspace{1cm} [Eq. 7 and 8]

Using the combination of the Dixon and the Cornish-Bowden plots, it is simple to measure the corresponding $Ki$ and $Ki'$ values for a mixed inhibition mechanism. In the case where $Ki < Ki'$, the intersection point in the Dixon plot is above the abscissa, whereas it is below it in the Cornish-Bowden plot. The opposite is true if $Ki > Ki'$. In the case of a true non-competitive inhibition system, $Ki = Ki'$. A true competitive mechanism of inhibition is characterized by these two plots if $Ki >> Ki'$ and an uncompetitive mechanism is the opposite, $Ki << Ki'$. 
Appendix II

Figure A2.1. Circular dichroism spectroscopy analysis for 5 μM ΔN-GIP in 10 mM phosphate buffer, 10 mM MgSO₄, pH 5.5. (A) Wavelength scans at 23°C. (B) Temperature ramp 10°C-82°C at 222 nm. (C) Temperature ramp 85°C-10°C at 222 nm. Experiment only performed once.

Panel A reveals mainly an α-helix structure, consistent with the CD results by Bamm et al. in 2008. The thermal melt of ΔN-GIP (Panel B) show the enzyme starting to unfold around 42°C. This process seems to be reversible as observed by the temperature ramp down in Panel C. From these results, a melting temperature of the enzyme can be estimated to be around 60°C, indicating a relatively stable protein.