ABSTRACT

Title of Dissertation: STRUCTURAL STUDIES OF THE ATYPICAL RIO2
KINASE AND N-ACETYLGLUTAMATE
SYNTHASE

Vatsala Sagar, Doctor of Philosophy, 2011

Dissertation Directed By: Assistant Professor Nicole LaRonde-LeBlanc
Department of Chemistry and Biochemistry

Rio2 is required for small subunit ribosomal RNA (rRNA) maturation. It is specifically involved in site D cleavage of the 20S rRNA to produce the mature 18S rRNA. Loss of Rio2p is lethal and a decrease of Rio2p activity results in 20S accumulation in the cytoplasm. One of the goals of this thesis was to crystallize Rio2 from a eukaryotic organism in order to determine the structural differences between eukaryotic and archaeal versions. Another goal was to define the importance of individual domains. Yeast Rio2 was the first eukaryotic protein purified, but it formed only microcrystals. Human Rio2 was purified next, but the solubility was too low to set up crystal trays. Finally, Rio2 from *Chaeotomium thermophilum* was purified and crystals were obtained. The structure revealed a possibly inhibitory alpha helix blocking the active site. The role of the N-terminal winged helix domain of Rio2 in yeast was investigated and found not to be necessary for binding of Rio2p to the ribosome.

The crystal structure of the first N-acetylglutamate synthase (NAGS) was also determined. The crystal structure of NAGS complexed with acetyl-CoA and with CoA plus N-acetylglutamate was determined at 2.5 and 2.6-Å resolution. Each NAGS consists of an N-terminal amino acid kinase domain (AAK) domain and a C-

terminal N-acetlyltransferase (NAT) domain connected by three amino acids. The monomers form a six membered ring with a trimer of dimers symmetry. The AAK domains form two dimeric contacts with other AAK domains. Each AAK domain interacts with the NAGS domain of another monomer at the polar ends of the ring. The NAGS domain contains the active site. The AAK domain is believed to bind arginine and also helps to bind acetyl-CoA. Structural insights suggest a one step mechanism in which both substrates bind and the acetyl group is directly passed from acetyl-CoA to the alpha amino group of gluatamate.

In addition, collaborative work on the structural characterization of the MphR(A) protein is reported.

STRUCTURAL STUDIES OF THE ATYPICAL RIO2 KINASE AND
N-ACETYLGLUTAMATE SYNTHASE

By

Vatsala Sagar

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Professor Nicole LaRonde-LeBlanc, Chair
Professor Douglass Julin
Professor David Fushman
Professor Paul Paukstelis
Professor Anne Simon, Dean's Representative

Acknowledgements

I am grateful to my advisor, Dr. Nicole LaRonde-LeBlanc, for helping me in this long journey that spanned approximately five years. She provided insights to solve multiple tough issues. She dedicated time to discuss research issues and provided insights. Along with the successes, she was always there supporting and guiding me to solve research issues that were originally difficult to solve. She encouraged and supported me in the use of novel approaches to resolve research problems that came along the way.

I would also like to thank all members of the lab, past and present who shared their knowledge with me and provided help.

I would like to thank all the members of this Thesis committee for their continued guidance and support. Dr. Fushman taught me a lot through his class on NMR and crystallography. I am grateful to Dr. Julin for helping me write my independent proposal, and to Dr. Paukstelis and Dr. Simon for dedicating their time to be on the review committee.

Thanks go to Dr. Allewell, Dr. Tuchman, Dr. Shi, Dr. Caldovic, and Dr. Morizono for their guidance in the research on N-Acetyl-L-glutamate Synthase cloning, purification, and crystallization. It was an interesting project and was completed within a year.

Thanks go to Dr. Zheng and Dr. Cropp for participating and collaborating with me on the crystal structure of MphR(A). I am still excited about the research and the ensuing publication.

Table of Contents

Chapter 3: Purification of Human Rio2

Chapter 4: Purification and Crystallization of Rio2 from Chaetomium thermophilum

Chapter 5: Structural Analysis of CtRio2

List of Tables

List of Figures

Abbreviations

A	Alanine
Å	Angstrom
AAK	Amino Acid Kinase
AcCoA	Acetyl Coenzyme A
afRio2	*Archaeoglobus fulgidus* Rio2
ADP	Adenosine-5'-diphosphate
AID	Autoinhibitory Domain
AMPK	AMP-activated Protein Kinase
APS	Advanced Photon Source
ATP	Adenosine-5'-triphosphate
BME	β-mercaptoethanol
C	Cysteine
cAMP	Cyclic Adenosine-3'-5'-monophophate
CBP	Calmodulin Binding Protein
CCP4	Collaborative Computational Project No. 4
CDK	Cyclin dependent kinase
CK2	Casein Kinase 2
C-hRio2	C-terminal 6x His Tag Human Rio2
CoA	Coenzyme A
COOT	Crystallographic Object Oriented Toolkit
C-Rio2p	C-terminally His Tagged Rio2p
cftRio2	*Chaetomium thermophilum* Rio2
D	Aspartic Acid
Da	Dalton
°C	Degrees Celsius
Δ72-C-hRio2	C-terminal 6x His Tag Human Rio2 with N-terminal Deletion of 72 Amino Acids
DNA	Deoxyribonucleic Acid
E	Glutamic Acid
EDTA	Ethylenediaminetetraacetic Acid
ePKs	Eukaryotic Protein Kinases
F	phenylalanine
FL K105A	Rio2p with K105A Mutation
FL K282A	Rio2p with K282A Mutation
FL R271A	Rio2p with R271A Mutation
G	Glycine
GST	Glutathione S-Transferase
H	Histidine
hRio2	Human Rio2
ITS	Internal Transcribed Spacer
I	Isoleucine
IPTG	Isopropyl-β-D-thiogalactopyranoside

K	Lysine
L	Leucine
LB	Luria Broth
LiAc	Lithium Acetate
M	Molar
M	Methionine
MAD	Multiple Wavelength Anomalous Diffraction
Mg	Milligram
$MgCl_2$	Magnesium Chloride
Mg^{2+}	Magnesium
mL	Milliliter
mM	Millimolar
Mn^{2+}	Manganese ion
N	Asparagine
NAG	N-acetlyglutamate
NAGK	N-acetylglutamate Kinase
NAGS	N-acetylgluatamate Synthase
NAT	N-acetyltranferase
NECAT	Northeastern Collaborative Access Team
NES	Nuclear Export Signal
N-hRio2	N-terminal 6x His Tag Human Rio2
OAT	Ornithine Acetyl Transferase
P	Proline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PHENIX	Python-based Hierarchical Environment for Integrated Xtallography
Q	Glutamine
R	Arginine
Rb	Retinoblastoma
Rbs	Ribosomal Binding Site
Rio2p FL	Full Length Rio2p
Rio2p WH	Wing Helix Domain of Rio2p
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
RPM	Revolutions Per Minute
rRNA	Ribosomal RNA
S	Serine
S	Svedburg Unit
SAD	Single Wavelength Anomalous Diffraction
SeMet	Selenomethionine
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
Short K105A	Rio2p Missing the Wing Helix Domain with K105A Mutation
Short K282A	Rio2p Missing the Wing Helix Domain with K282A Mutation
Short R271A	Rio2p Missing the Wing Helix Domain with R271A Mutation
Short Rio2p	Rio2p Missing the Wing Helix Domain
SSU	Small Ribosomal Subunit

T	Threonine
TAP	Tandem Affinity Purification
TEV	Tobacco Etch Virus
Tris	Tris(hydroxymethyl)aminomethane
UBF1	Upstream Binding Factor 1
UCE	Upstream Control Element
UV	Ultraviolet
V	Valine
W	Tryptophan
wHTH	Winged Helix Domain
Y	Tyrosine
μL	Microliter
μM	Micromolar

Chapter 1: Introduction

1.1: Importance of Ribosome Biogenesis in Tumor Growth and Cancer Progression

Ribosome biogenesis is the process of making ribosomes. It involves the coordinated function of over 200 proteins in the synthesis and processing of the four rRNAs, as well as assembly of those rRNAs with the ribosomal proteins. Ribosome biogenesis is critical for cell cycle progression and cellular proliferation in both normal cells and cancer cells (*1-4*). Ribosomal RNA (rRNA) processing occurs in the nucleoli (located within the nucleus) where ribosome production occurs (*5*). Nucleoli undergo enlargement in two types of cells: rapidly proliferating normal cells and cancer cells (*6*). In cancer cells, the average size of nucleoli is larger and occurs with a decreased cell doubling time (*7*). The size of the nucleoli is also directly correlated with the amount of rRNA transcription. Data show that the expression of nucleolar proteins involved in rRNA transcription control and rRNA processing is directly correlated to the nuclear size and the cellular proliferation rate (7). Thus, the size of a nucleus is an indicator of the rate of cellular proliferation. The enlarged nucleoli in cancer cells are associated with an enhanced ribosome production. For a proliferating cell to produce new daughter cells, an increase in the concentration of ribosomes is required. The enlarged nucleoli and the increase in rRNA production is a marker for conversion to a metastatic state in cancer cells (*8, 9*). Although ribosome biogenesis is an important part of cancer progression, the molecular details of how it occurs are not known.

1.2: Ribosome Biogenesis is Highly Controlled in Normal Cells but Becomes Deregulated in Tumorigenesis

Most of the energy of a proliferating cell is devoted to ribosome production (*10*). Ribosome production is a highly regulated process that is controlled by nutrient level changes and growth signals. The transcription of ribosomal genes is carried out by RNA Polymerase I (*5, 11*). An increase in production of rRNA occurs when the upstream binding factor 1 (UBF1) binds to the upstream control element (UCE) and the core promoter of the 35S rRNA gene in response to external stimuli (*3*). Binding of UBF1 to DNA is regulated by phosphorylation of the C-terminus (*12-15*). In nonproliferating cells, UBF1 is not phosphorylated and does not activate rRNA gene transcription (*15*). When UBF1 becomes phosphorylated by a kinase such as CDK4-cyclin D or CDK2-cyclin E, UBF1 binds to the DNA and directly interacts with Pol1, resulting in gene transcription (*13, 14*). Therefore, in order to control ribosome biogenesis, UBF1 has to be carefully regulated. Caesin Kinase II, ERK1/2 kinases and CDK-cyclin complexes have all been shown to phosphorylate UBF1 at the C terminus (*14*). These kinases are upregulated in many cancers. Retinoblastoma (Rb) is a tumor suppressor and interacts with UBF1 to prevent it from binding to the DNA promoter (*16-18*). Rb thereby inhibits the production of rRNA to regulate cell growth and proliferation. A point mutation in Rb in small cell lung carcinoma prevents it from interacting with UBF1 (*19, 20*) . The regulation of ribosome biogenesis must be understood in order to learn how to control this process in cancer cells.

1.3: RIO Kinases and Their Involvement in Ribosome Biogenesis

Rio2 protein kinases are atypical serine kinases that are required for the cleavage of 20S rRNA to 18S rRNA (*21-23*). Since, the yeast model of ribosome biogenesis is less complex and amenable to genetic manipulation, it has been used to study ribosome

assembly. rRNA synthesis begins with transcription of a 35S rRNA by RNA polymerase I. This 35S rRNA transcript contains 5.8S, 18S, and 25S rRNA, separated by internal transcribed spacers ITS1 and ITS2, and external transcribed regions 5' and 3' to the rRNA. (The transcription of 5S rRNA by RNA polymerase III occurs on a separate transcriptional unit).

The extra ribosomal proteins that assemble onto the 35S pre-rRNA cause it to sediment at 90S on a sucrose gradient. Proteins are assembled onto 35S rRNA for ribosomal protein loading, rRNA folding modifications, and mainly to direct cleavage of the rRNA to produce 20S rRNA. Cleavage first occurs at sites A0 and A1 on the 5' external transcribed spacer of the 35S rRNA. This is followed by cleavage at A2 in ITS1 to release the 20S rRNA particle needed for the 40S small ribosomal subunit. Proteins such as Enp1p, Dim1p, Nob1p, Hrr25p, Rrp12p, Rio2p, Trs1p and Yor145p associate with the pre40S ribosomal subunit (*37*).   Further processing of the 20S rRNA in yeast is performed in the cytoplasm by cutting at site D in ITS1 to release the 18S rRNA (Figure 1.1). Once the cleavages are completed, these ribosomal assembly factors dissociate. 40S rRNA processing is abrogated by mutations in 60S rRNA processing. However, the reverse is not true (*24-26*).

1.4: Rio2 is an Atypical Protein Kinase

Protein kinases phosphorylate serine and threonine residues or tyrosine residues (*27, 28*). All protein kinases contain a conserved kinase domain of 250-300 amino acids in length containing conserved secondary structures (*27-29*). RIO proteins are serine kinases that contain very few of the conserved residues of typical protein kinases. Thus they are labeled as atypical protein kinases (*21*). The RIO kinases consist of four

subfamilies of proteins, Rio1, Rio2, Rio3, and RioB. Rio1 and Rio2 are found in species from archaeal to human, while Rio3 is found only in multicellular eukaryotes (*21*). RioB is found only in eubacteria (*30, 31*). The RIO family of proteins consists of a RIO kinase domain with C and N terminal extensions specific to each subfamily of proteins. As seen from the structure of Rio1 and Rio2, the RIO domain consists of an N-terminal lobe containing a twisted β sheet and a long α helix that serves as the back of the ATP binding pocket, a hinge region, and a C-terminal lobe which houses the metal binding loop and the catalytic loop (*32-34*) (Figures 1.2 and 1.3). Rio2 contains a winged helix domain at the N-terminus, archaeal Rio1 contains three extra α helices, and the structure of the N-terminal extension of Rio3 is not yet known as no crystal structure has been determined (*32-34*). RioB does not contain an N-terminal extension to the RIO domain or a C-terminal extension.

Each RIO subfamily member contains unique sequence variations in the RIO domain. Distinct from eukaryotic protein kinases (ePKs), RIO proteins contain an insertion of 18-27 amino acids between αC and β3 called the flexible loop because in Rio2 this region is disordered (*32-34*). In addition, RIO kinases only contain three of the α helices of canonical ePKs in the C-lobe. In contrast to the structure of cAMP-dependent protein kinase (PKA), RIO proteins lack an activation loop (subdomain VIII) and lack subdomains X and XI (*32-34*). These subdomains are usually used by ePKs to bind to the peptide substrate. As such, the mode of binding to the peptide substrate of RIO kinases is unknown.

1.5: Rio2p Associates with the 40S Preribosome and is Involved in Cleavage of the 20S rRNA

Studies in which Rio2 was tagged on the C-terminus by a Tandem Affinity

Purification (TAP) tag allowed identification of proteins associated with Rio2 in the cell.

Mass spectrometry identified several proteins required for late 40S maturation such as

Trs1p, Rrp12p, and Dim1p (*35-37*). As such Rio2p was found to be stably associated

with the 40S preribosome and tagged Rio2p could be used to specifically pull down the

40S preribosome *(35-37).*

Rio2p can shuttle between the nucleus and cytoplasm and thus most likely

attaches in the nucleus and accompanies the 43S sub-complex to the cytoplasm (*21, 22,

40*). After the loss of Rio2p activity, 20S rRNA accumulates in the cytoplasm with a

reduction in 18S rRNA (*38*).

As Rio2p has serine kinase activity in vitro, its function might be strictly

enzymatic. The low concentration of Rio2p required to sustain growth in glucose from a

leaky Gal promoter of the wild type Rio2p, supports the conclusion that Rio2p might be

used catalytically, rather than stoichiometrically, to effect cleavage at site D (*37*).

Another RIO protein, Rio1p, is also important for cleavage of site D, but its

activity cannot replace that of Rio2p (*39*). Tandem affinity purification (TAP) of Rio2p

does not pull down Rio1p. There is no evidence for direct attachment of Rio1p to the

43S preribosomal particle.

1.6: The Winged Helix Domain is Structurally Similar to DNA Binding Proteins

The structure of Rio2 from archaeobacteria, *Archaeoglobus fulgidus*, (afRio2) has

been solved in complex with ATP or ADP and Mn^{2+}, the former at 1.84Å and the latter at

1.75Å (*32, 34*). The N-terminal domain of afRio2 contains a winged helix motif. This

motif consists of four α helices, followed by two β strands which are followed by a fifth α

helix. The connecting strand between the second and third α helix is a small β strand and forms a β sheet with the other two β strands. The loop between the second and third β strand is called a wing—leading to the term winged helix turn helix (*32, 34*) (Figure 1.4).

The winged helix motif is usually found in transcription factors that bind DNA such as MarR, SlyA, and the histone linker protein GH5. Although the root mean square difference (rmsd) between these structures is 1.2-1.3Å, they only share 7% to 14% identity (*34*). These proteins bind to DNA in two ways. The first way is through the amino acids in helix 3 interacting with the major groove of the DNA. This mode of DNA binding was first found in the crystal structure of HNF-3 with DNA (*41*). The second way of binding to DNA is through the amino acids in W1, β2, and β3 interacting with the major groove of DNA. This mode of DNA binding was first found in the crystal structure of hRFX1 and DNA (*42*). MarR and SlyA probably bind the major groove of DNA via helix 3 and histone H5 probably binds the major groove of DNA via W1, β2, and β3 (*43, 44*). Rio2 is closer to the structure of SlyA than with histone H5. Similar to SlyA, Rio2 has a fourth helix in the place of the second wing. Also the electrostatic surface of the winged helix domain is most closely identical between Sly A and afRio2. Furthermore, supporting a role for afRio2 in nucleic acid binding, the amino acids in helix 3 are positively charged (*34*).

1.7: Structure of Rio2p RIO Domain and Insights into Catalysis

The RIO domain consists of two lobes, the N-terminal lobe and the C-terminal lobe separated by a hinge region. The N-lobe of afRio2 is used to bind and position ATP for attack on the γ phosphate by a serine hydroxyl group. Like traditional kinases, afRio2 contains a 'DFG' loop to bind metal ions. It also contains a 'P-loop', which binds the

phosphate groups of ATP and contains the conserved sequence GXGXXS (GVGKES). The C-terminal lobe houses the catalytic loop, which contains the conserved N223 and D218 residues required for catalysis and metal binding (Figure 1.5). However, unlike traditional ePKs, afRio2 lacks an activation loop also known as sub-domain-VIII. In Erk2, upon phosphorylation of Y185, the loop changes conformation to allow the substrate to bind and to position the catalytic residues. Sub-domain IX, and X in ePKs are used for peptide substrate binding. As part of its classification as an atypical kinase, afRio2 lacks these sub-domains. Thus its mode of peptide substrate binding is still unknown (*28, 29, 32, 34*).

Upon co-crystallization of afRio2 with ATP and Mg^{2+}, a different conformation of the γ phosphate is seen than is seen in canonical protein kinases. When ATP binds between the N and C lobes, it causes a large movement of the two lobes with regards to each other causing the γ phosphate of ATP to be buried in a deep pocket. ATP binds deep underneath the nucleotide binding loop, and in this position, the γ phosphate is shifted away from the center of the active site. The distance between the catalytic base D218 and the phosphorus atom is 5.8Å in afRio2. In support of its role in catalysis, the mutation of D218 produces a largely inactive afRio2. In order for catalysis to take place, the γ phosphate may have to be unburied by a large conformational change of the enzyme (*32, 34*).

Further conformational changes to the enzyme might occur either upon binding of the peptide or upon allosteric changes when an unknown serine/threonine residue is phosphorylated by some other kinase present in yeast. In the case of Rio1p, kinase assays demonstrate that phosphorylation by Casein Kinase 2 (CK2) to the c-terminal

domain of Rio1p, increases its activity about twofold using a heterologous substrate (*45*).

As previously mentioned for ePKs, phosphorylation of the activation loop tends to

activate the kinase activity (*29*). AfRio2 is autophosphorylated at S128, but its effect on

catalysis had not been investigated (*32*).

1.8: Specific Aims

The aim of my research towards this thesis was to expand the knowledge of Rio2

catalysis and to analyze how Rio2 from archaea differs from Rio2 from eukaryotes. The

structure of the eukaryotic Rio2 needs to be elucidated as it contains features in addition

to those of the archaeobacteria Rio2.

The winged helix domain (wHTH) appears to be an important feature in Rio2.

The wHTH might bind to the RNA of the ribosome. The structures closest to wHTH

found with the DALI server are the transcription factors MarR and SlyA and the histone

linker protein GH5, all of which are known to interact directly with DNA.

The specific aims included:

1. To purify, crystallize and determine the x-ray crystal structure of a eukaryotic
   version of Rio2.

2. To determine the role of the winged helix domain in yeast Rio2.

In addition to the specific aims listed above that defined my thesis project, I have

included the work on two additional crystallography projects completed during my tenure

here that are considered part of my graduate work. They are listed below:

1. Determination of the x-ray crystal structure of MphR(A), a repressor protein of
   the MphA operon, in collaboration with Dr. Ashton Cropp. This is presented in
   Chapter 7.

2. Determination of the first x-ray crystal structure of a N-acetyl glutamate

synthase (NAGS) protein in the laboratory of Dr. Norma Allewell.

Purification and crystallization of Rio2, the main topic of this dissertation, has been

discussed in Chapters 2-6. The x-ray crystal structure of Mphr(A) has been described in

Chapter 7 and the x-ray crystal structure of NAGS has been presented in Chapter 8.

Discussion about future studies relating to Rio2 have been included in Chapter 9.
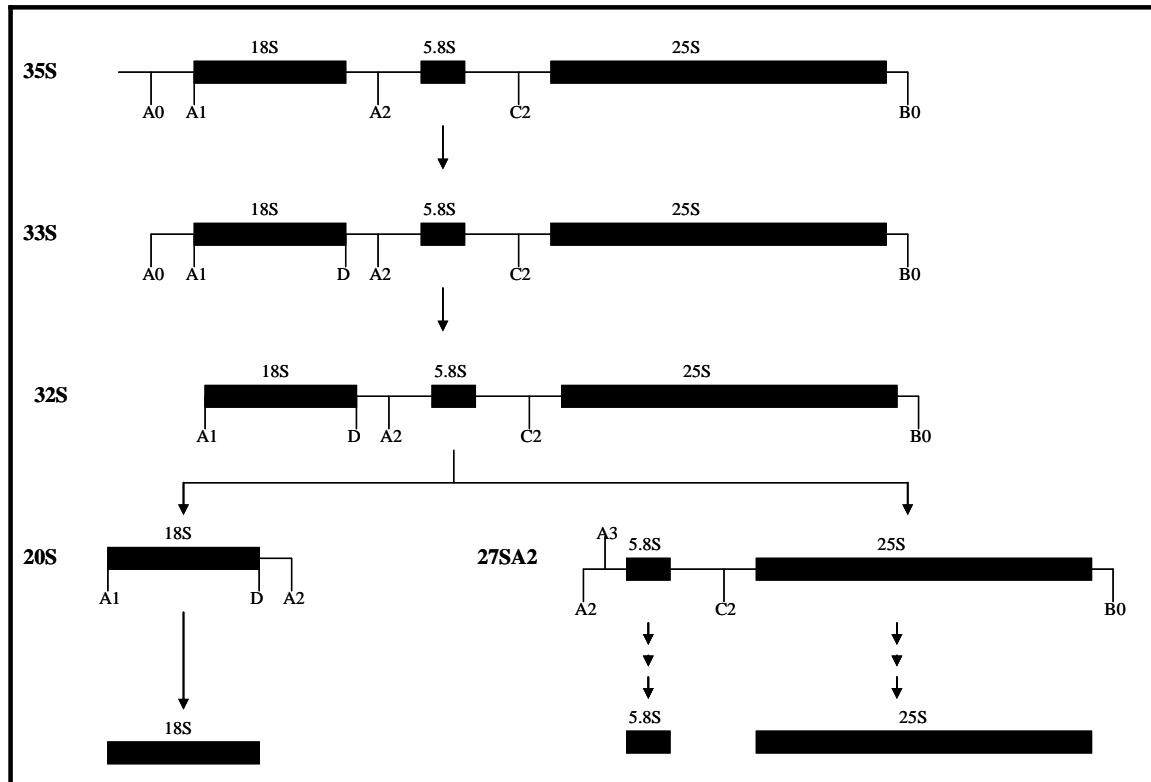
Figure 1.1: Ribosomal RNA Processing in Eukaryotes.
 Cleavages at A0, A1 and A2 release the precursor of the small subunit ribosomal rRNA, the 20S rRNA. In the cytoplasm, the 20S rRNA is processed by cleavage at site D to produce the 18S small subunit rRNA. Cleavages at A3, B0, and C2 release the precursor rRNAs of the large ribosomal subunit which undergo further processing to produce the mature 5.8S and 25S rRNAs.
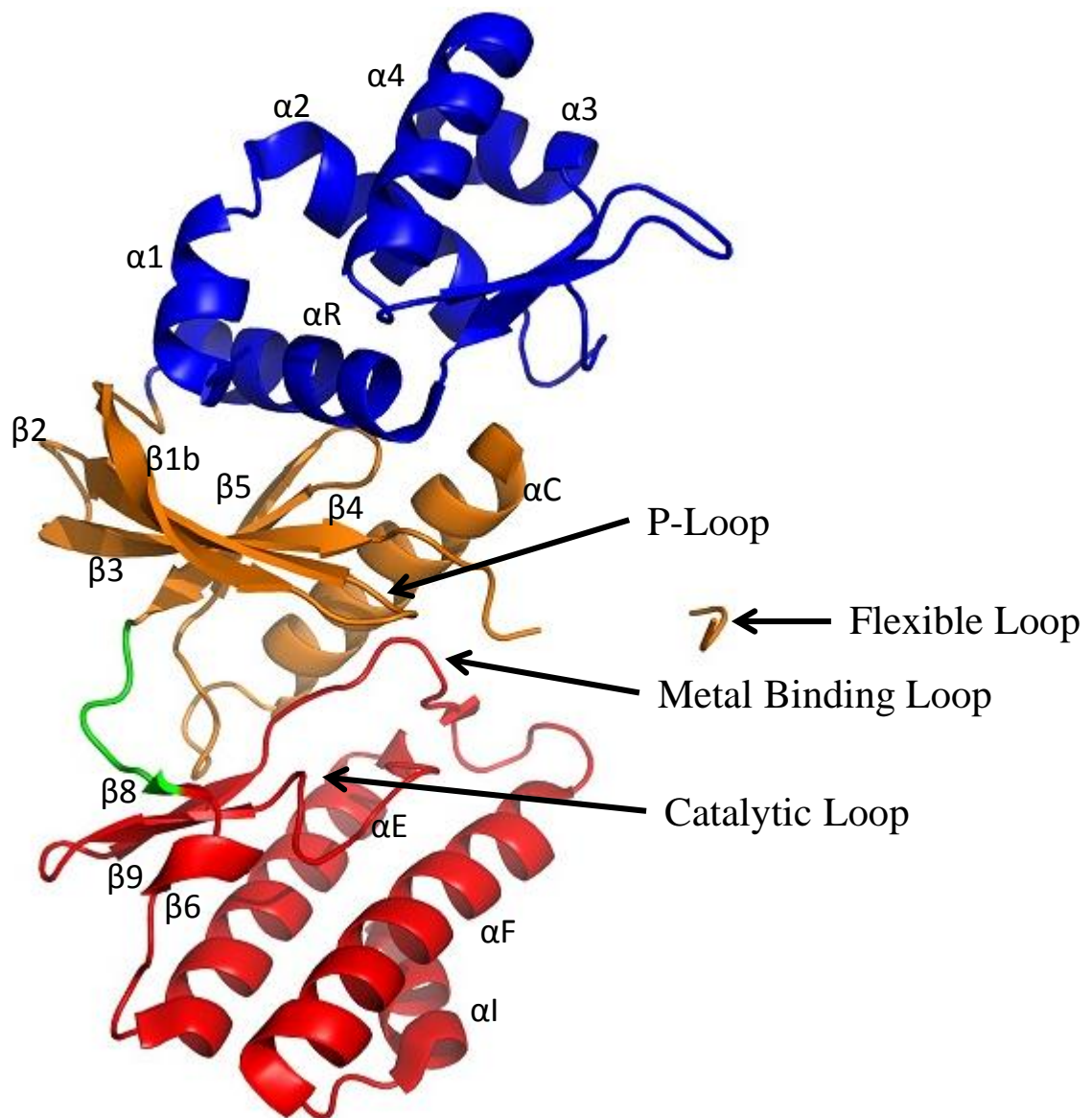
Figure 1.2: Crystal Structure of AfRio2

The crystal structure of afRio2 shows a bilobal kinase domain (brown and red) connected to a N-terminal winged helix domain (blue). The green "hinge" region connects the N- and C- terminal lobes of the kinase domain. The ATP binding loop "P-loop", metal binding loop used to bind Mg, and catalytic loop which contains the catalytic D218 are indicated by arrows. Electron density was not observed for the flexible loop. The cartoon representation was made by PYMOL using PDB structure1ZAO.
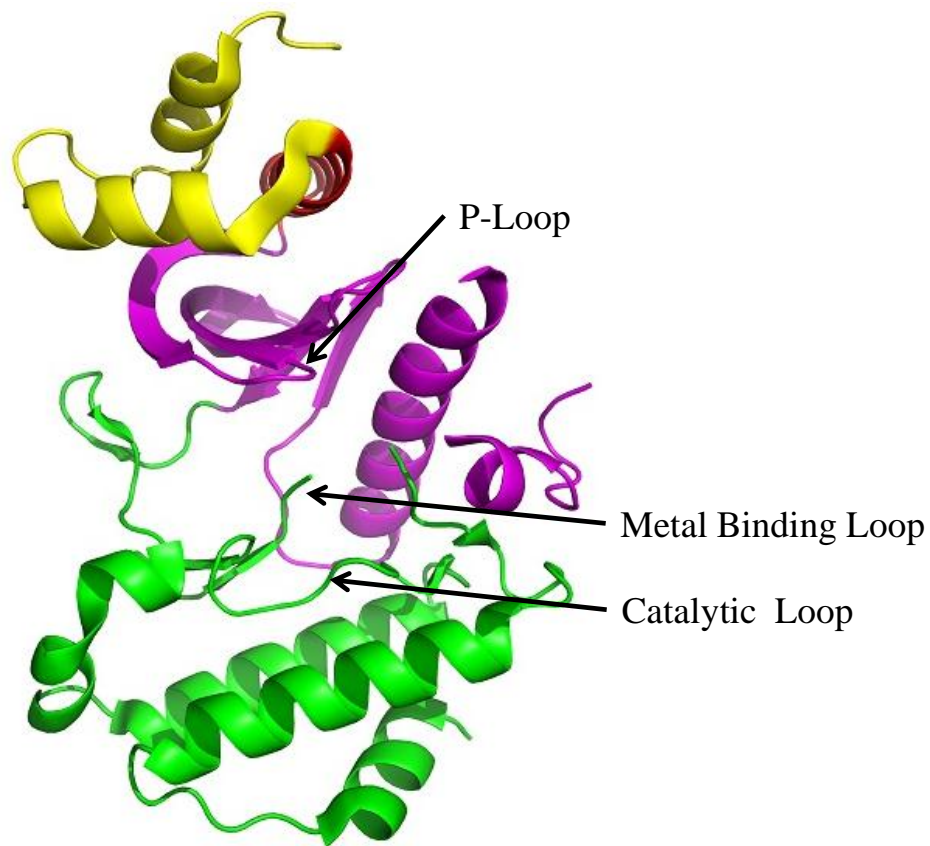
Figure 1.3: Crystal Structure of AfRio1

The crystal structure of afRio1 shows a bilobal kinase domain (magenta and green) connected to three N-terminal alpha helices (yellow and red). The ATP binding loop "P-loop", metal binding loop used to bind Mg, and catalytic loop which contains the catalytic D218 are indicated by arrows. The cartoon representation was made by PYMOL using PDB structure 1ZP9.
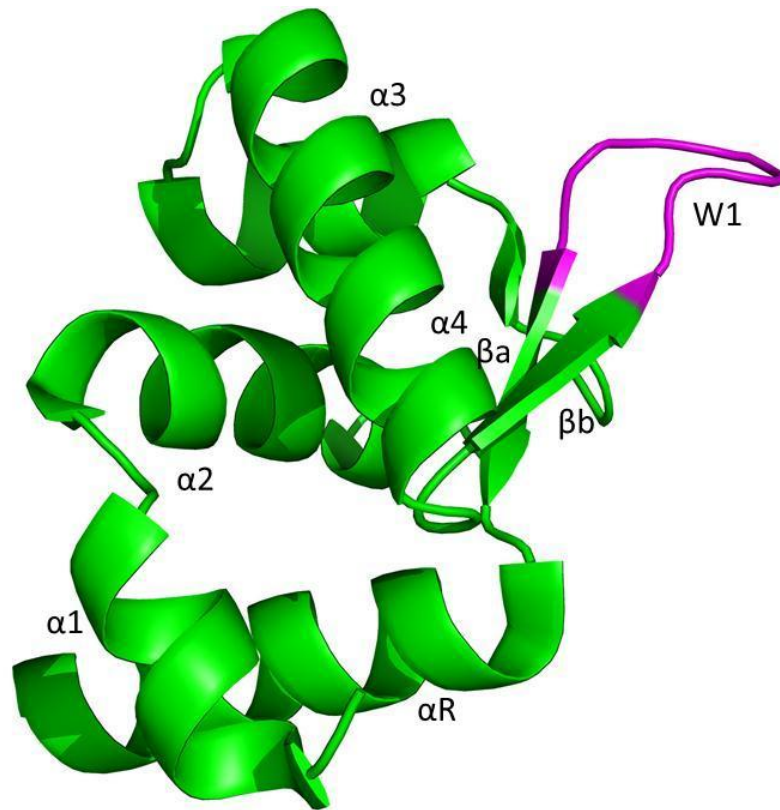
Figure 1.4: Structure of the Winged Helix Domain of AfRio2
The winged helix domain consists of four alpha helices followed by two beta sheets. A fifth alpha helix (αR) is part of the RIO domain. The wing is labeled W1 and is in magenta. This cartoon representation was constructed by PYMOL using PDB structure 1ZAO.

```
                                                                      P-loop
                                                                      xxxxxx
AFRio2   48 RNILKYLSDLRVVQN-RQKDYEGSTFTFIGLSLYSLHRLVRSGKVDAIGKLMGEGKESAVF---NCYSEK-FGEC
AFRio1    5 KKIESYLDKLRIKEK-DGEERKIYAEVLDGRTLKTLYKLSAKGYITAMGGVISTGKEANVFYADGVFDGK-PVAM
ScRio2   50 NRAISDLAKLSLISKMRNVKYDGYRLTYNGIDYLALKTMLNRDTVYSVGNTIGVGKESDIY---KVSDKN-GNPR
PKA      15                   VKEFLAKAKEDFLKKWESPAQNTAHLDQFERIKTLGTGSFGRVM---LVKHKETGNHY
CK1       6                      NVVGV----HYKVGRRIGEGSFGVIF---EGTNLLNNQQV


AFRio2  118 VVKFHKVGHTSFKKVKEKRDYGD-------LHFSVLAIRSARNEFRALQKLQ--GL-AVPKVYAWEG----NAVL
AFRio1   78 AVKIYRIE-TSEFDKMDEYLYGDERFDMRRISPKEKVFIWTEKEFRNLERAKEAGV-SVPQPYTYMK----NVLL
ScRio2  121 VMKIHRLGRTSFHSVRNNRDYLKKSNQG--ANWMHLSRLAANKEYQFMSMLYSKGF-KVPEPFDNSR----HIVV
PKA      70 AMKILDKQKVVKLKQIEH--------------------TLNEKRILQAV---NFPFLVKLEFSFKDNSNLYMV
CK1      40 AIKFEPRRSDAPQ------------------------LRDEYRTYKLLA--GCTGIPNVYYFGQEGLHNVLV
                                                                      Catalytic
                                                                      xxxxxx
AFRio2  179 MELIDAKEL---------YRVRV------ENPDEVLDMILEEVAKFYHRG-IVHGDLSQYNVLVSE-------EG
AFRio1  147 MEFIGEDELPAPTLVELGRELKE------LDVEGIFNDVVENVKRLYQEAELVHADLSEYNIMYID-------K-
ScRio2  189 MELIEGYPM---------RRLRK-----HKNIPKLYSDLMCFIVDLANSG-LIHCDFNEFNIMIKDKLEDENDCG
PKA     120 MEYVPGGEMFSH-----LRRI--G-RFSEPHARFYAAQIVLTFEYLHSLD-LIYRDLKPENLLIDQ------QGY
CK1      85 IDLL-GPSLEDL-----LDLC--GRKFSVKTVAMAAKQMLARVQSIHEKS-LVYRDIKPDNFLIGRPNSKN-ANM
            Metal
            xxxxx
AFRio2  231 IWIIDFPQSVEVGEEG--------------------------WREILERDVRNIITYFSRTYRT----------
AFRio1  208 VYFIDMGQAVTLRHPM--------------------------AESYLERDVRNIIRFFSK-YGV----------
ScRio2  248 FVVIDFPQCISIQHQD--------------------------ADYYFQRDVDCIRRFFKKKLKY----------
PKA     180 IQVTDFGFAKRVKGRTWXLCG----------TPEYLAPEIILSKGYNKAVDWWALGVLIYEMAAGYPPFFA---D
CK1     150 IYVVDFGMVKFYRDPVTKQHIPYREKKNLSGTARYMSINTHLGREQSRRDDLEALGHVFMYFLRGSLPWQGLKAA


AFRio2  269 EKDINSAIDRILQE
AFRio1  245 KADFEEMLKEVKGE
ScRio2  287 EPKPDSSMLDTEGFGDGYKYAYPDFKRDVKRTDNLDELVQASGFSKKH
PKA     242 QPIQIYEKIVSGKVR-----FPSHFSSDLKDLLRNLLQVDLTKRFGNLKNGVNDIKNHKWFATTDWIAIYQRKV
CK1     225 TNKQKYERIGEKKQSTPLRELCAGFPEEFYKYMHYARNLA----FDATP--DYDYLQGLFSKVLERLNTTE---
```

Figure 1.5: Structural Alignment of Rio2 with Eukaryotic Protein Kinases

Structure-based sequence alignment of the kinase domains of Rio1 and Rio2 from *A.fulgidus* (AFRio2 and AFRio1), Rio2 from *S.cerevisiae* (ScRio2), eukaryotic protein kinases, Protein Kinase A (PKA) and casein kinase (CK1). Residues highlighted in green are identical in all kinases. Residues highlighted in yellow are identical in Rio2 and Rio1 but not all kinases. The sequences of the ATP binding loop "P-loop", metal binding loop used to bind Mg, and catalytic loop which contains the catalytic D218 are indicated by x's.

Chapter 2: Purification and Crystallography of Rio2p

2.1: Introduction to the Purification of Yeast Rio2 (Rio2p)

The main goal of this project is to better understand the role of Rio2 in eukaryotic rRNA processing. The details of the the molecular function of Rio2p is not known. For instance, it is not known how Rio2p recognizes its peptide substrate. Traditional ePKs use subdomains IX and X to recognize and bind their peptide substrate. Rio2p lacks these domains (*28, 29, 32, 34*). The eukaryotic version of Rio2p differs from the archaeal version in that it contains an extended C-terminus (Figure 2.1). It can be hypothesized that the extended C-terminus of eukaryotic Rio2p proteins might be used to recognize the peptide substrate for specificity and binding. Alternatively, the extended C-terminus might be used for some allosteric or inhibitory role. Several attempts were made to obtain a structure of a eukaryotic version of Rio2p to determine how eukaryotic Rio2 functions and whether it functions differently from the archaeal version. Towards this goal, Rio2 from yeast (*Sacchoromyces cerevisae*), human, and *Chaetomium thermophilum* (ctRio2) were purified and screened for crystallization. Table 2.1 briefly outlines the problems and successes encountered with each protein.

In this chapter, the purification of eukaryotic Rio2p is described. During this process, I encountered the following difficulties: (a) Rio2p is rapidly degraded by proteases in the lysate;
(b) Rio2p binds to RNA nonspecifically; and (c) Rio2p exhibited poor solubility in the absence of solubilizing agents. To circumvent these problems, I purified two different constructs containing His tags either on the N-terminus or the C-terminus of the protein; added ATP to buffer solutions to solubilize Rio2; and added RNases to the lysate to degrade RNA that would otherwise bind to Rio2p.

2.2: Design

Two constructs were designed to determine which gave the best results in purification from degradation products (Figure 2.2). The first construct contained a N-terminal 6x His tag followed by a 2kDa linker and a TEV cleavage site (N-Rio2p). This construct was cloned into pDEST 527 and transformed into Rosetta (pLysS cells). The second construct contained a C-terminal 6x His tag directly attached to Rio2p (C-Rio2p).

2.3: Methods

2.3.1: Subcloning of C-Rio2p

A Rio2p construct with the C-terminal 6x His tag (C-Rio2p) (Figure 2.2B) yielded the most purifiable product. To accomplish this, a full length Rio2p gene was cloned in an *E. coli* expression plasmid that contained IPTG inducible promoter. C-Rio2p was amplified from the pDEST527 plasmid, which already contained an insert for the Rio2p gene which was obtained from Dr. Nicole LaRonde LeBlanc from her work at NCI. Two PCR reactions were performed to make the new construct. The first PCR used a forward primer to add a necessary ribosomal binding site (RBS) next to the 5' initiation codon of the Rio2p gene, because the vector does not contain a RBS. The first PCR also used a reverse primer to add the 6x His tag and two stop codons to the 3'end of the Rio2p gene. The second PCR used a forward primer to add the attB1 site to the 5' end of PCR1 and used a reverse primer to add the attB2 site to the 3' end of PCR1. The forward and reverse primers for the two PCR reactions are diagrammed in Figure 2.3. The PCRs were performed using KOD polymerase (Novagen). PCR products from both reactions were run on a 1% agarose gel shown in Figure 2.4.

The Gateway^{TM} system was used to clone C-Rio2p into pDEST 14, which is a bacterial T7 expression vector with no tags. The Gateway^{TM} system relies on recombination between the vector and insert. The process requires two cloning steps. First a cloning vector contained attP1 and attP2 sites known as an entry vector is recombined with the PCR product containing attB1 and attB2 sites inserted by PCR with primers containing the sequence (in the PCR product described above). This recombination produces a vector containing attL1 and attL2 sites on either side of the insert. Another recombination reaction was used to transfer the insert from the entry vector to the destination vector, pDEST 14, which contains attR1 and attR2 sites. The resulting vector was then used for protein expression.

To test the resulting clones for expression, Rossetta^{TM} cells were transformed with the pDEST 14 plasmid containing the C-Rio2p insert. Four individual colonies were grown in 5 mL cultures until the colonies reached an O.D.600 of 0.8 and were then induced overnight at 18°C with 1mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were pelleted by centrifugation and lysed with 5x SDS loading dye. Lysates of the four cell cultures, before and after induction, were run on a SDS-PAGE gel. All four colonies revealed expression of Rio2p. Cells were also lysed with 0.3x Bugbuster^{TM}, centrifuged, and the supernatants run on the SDS-PAGE gel. Solubility was not evident in supernatant without ATP (Figure 2.5).

2.3.2 Purification of C-Rio2p

The construct design for C-Rio2p is shown in Figure 2.9B. The optimized procedures for C-Rio2p expression in *E. coli* and the purification steps are described below. The pDEST 14-C-Rio2p vector, transformed into *E. coli* Rosetta^{TM} DE3 (pLysS)

cells, was used to inoculate an overnight culture in 100 mL of LB with 100 ug/ml

ampicillin and 34 ug/ml chloroamphenicol. The culture was grown overnight at 37°C in

a shaking incubator. 4 L of LB in four 4 L flasks with ampicillin and chloromphenicol

were inoculated with the overnight cultures of Rosetta^{TM} cells in 1/100 dilution and

grown at 37°C in a shaking incubator at 250 rpm. At O.D.600 of 0.6 the flasks were

transferred to a 20°C shaking incubator to cool. At O.D.600 of 0.8, the cultures were

induced with 1 mM IPTG and incubated overnight. The next day, the cultures were

centrifuged at 5000 rpm for 30 minutes. The cell pellet was collected and frozen at -

80°C. The frozen pellet was resuspended in 10 mL of buffer A (10% glycerol, 2.5 mM

$MgCl_2$, 50 mM Tris, pH 8.0, 300 mM NaCl, 1 mM ATP, 0.05% β-mercaptoethanol

(BME)) per gram of frozen pellet and 5 mg DNAase, 1 mg RNAase A, 20 μl RNAase

T1 (100,00 μ/mL) and 1 tablet EDTA free protease inhibitor tablet (Roche). Then 0.3X

Bugbuster^{TM} was added along with 20 mg lysozyme. The cell suspension was kept on ice

and stirred for thirty minutes to allow lysis. The lysate was centrifuged at 16000 rpm for

thirty minutes. The supernatant containing the soluble protein was passed over a 5 mL

GE Healthcare His Trap^{TM} column. Buffer B (10% glycerol, 2.5 mM $MgCl_2$, 50 mM

Tris, pH 8.0, 300 mM NaCl, 1 mM ATP, 1 M imidazole, 0.05% BME) was the elution

buffer. The column was first washed with 5% Buffer B and then eluted with 10% Buffer

B. The results of the purification are shown in Figure 2.6.

The fractions of the 10% Buffer B elution were pooled together and the solution

containing C-Rio2p was then diluted so that the final concentrations of solution

components were 200 mM NaCl, 50 mM HEPES, pH 7.0, 10% glycerol, 2.5 mM $MgCl_2$,

1 mM ATP. The resulting partially purified C-Rio2p was passed over a 5 mL GE

Healthcare HiTrap S^{TM} column. The loading buffer, Buffer C, for the cation exchange column was 10% glycerol, 200 mM NaCl, 2.5 mM MgCl$_2$, 50 mM HEPES, pH 7.0, 1 mM ATP, 0.01% BME. Buffer D, the high salt buffer, was 10% glycerol, 2.5 mM MgCl$_2$, 1 M NaCl, 50 mM HEPES, pH 7.0, 1 mM ATP. A gradient of the two buffer was run from 200 mM NaCl to 1 M NaCl over 100 mL. The results of the purification are shown in Figure 2.7.

C-Rio2p eluted around 524 mM NaCl. The two most pure and concentrated fractions eluting between 504 mM NaCl and 544 mM NaCl were pooled together and dialyzed against 20 mM Tris, pH 8.0, 300 mM NaCl, 2.5 mM MgCl$_2$, 10% glycerol, 1 mM ATP. Rio2p was then purified on a Superdex 200 Prep Grade gel filtration column equilibrated in dialysis buffer. The results of the purification are shown in Figure 2.8. Fractions 19, 20, 21 were concentrated to 20 mg/ml as determined by the Bradford Assay.

2.3.3: Crystallization of Rio2p

Crystal trays were set up with 10 mg/mL and 20 mg/mL C-Rio2p and 20 mM ATP substrate. The screens tested were Qiagen's PEG Suite, Hampton's Index Screen, Emerald Biosystems Wizard I, II, and III screens, and Qiagen's Cryos Suite. Seven days later crystals formed in the PEG Suite, condition number 68, 2:1 precipitant to protein. The formulation was 0.2 M magnesium formate, 20% PEG 3350. The hanging drop method was used to optimize crystal formation. In this method, the protein and precipitant solution were mixed in the ratio from the screen in a 4-µl drop. The drop was put on a slide that was then affixed over a well containing 1 mL of the precipitant solution. By vapor diffusion, some of the water gradually left the drop to join the

precipitant solution until equilibrium was achieved, thereby slowly concentrating the protein in the drop. The concentration of magnesium formate and the PEG concentrations were varied along 24 different wells in order to optimize crystal formation.

2.4: Results

2.4.1: Difficulties Encountered with Purification of Rio2p

In order for optimal crystallization of proteins, several criteria have to be met. Rio2p has to be purified to greater than 95% purity. Rio2p has to be fairly soluble. At least 10mg of protein is needed to screen the protein through all the desired different test crystallization conditions. The His tag used in Rio2p purification can also be an impediment to crystallization and should be removed.

It was difficult to purify Rio2p to 95% purity because Rio2p is easily degraded by proteases. The SDS-PAGE gel for the initial affinity purification of Rio2p using the His Trap^{TM} column showed numerous low molecular weight bands that appeared in a ladder. The reason that most of the contaminating bands were attributed to degradation is that most of these contaminating bands disappeared after addition of protease inhibitor tablets to the lysate. By adding three protease inhibitor tablets per 50 mL of lysate solution, the number of degradation products showing up on the SDS-PAGE gel after His Trap^{TM} slowly decreased as seen in Figure 2.6 compared to Figure 2.9. In order to purify Rio2p to 95% purity for crystallization screening, it had to be separated from its numerous degradation products.

Rio2p was not a very soluble or abundantly isolated protein. First attempts at purification resulted in 5mg of protein isolated per 2 L of culture. The solubility of Rio2p was initially less than 5 mg/mL.

2.4.2: Purification of the N-terminal His Tag Construct and a C-terminal His Tag Construct

Of the two constructs C-Rio2p and N-Rio2p, success came with the C-Rio2p construct. N-Rio2p was purified first on a His Trap^{TM} column, followed by a Q^{TM} column, and finally by gel filtration. N-Rio2p could not be purified of all its degradation products but was co-purified with a slightly lower molecular weight degradation product in gel filtration (Figure 2.10).

C-Rio2p could be made 95% pure by a purification method that involved use of the His Trap^{TM} column as the first step of purification followed by a cation exchange column. Although the pI of Rio2p is 5.13, it still binds to the cation exchange column— possibly because the N-terminal winged helix domain has an electrostatically positive surface potential despite the overall negative charge of the protein. Rio2p bound tightly to the S^{TM} column and eluted at 500 mM salt, in support of this hypothesis. The purification could occur because binding of C-Rio2p to the cation exchange column was via the N-terminal winged helix domain and the degradation products containing only C-terminal fragments would not bind. The first step of purification of C-Rio2p on the His Trap^{TM} column, via the C-terminal His tag, eliminated the degradation products containing the N-terminal domain.

The final polishing step to purify C-Rio2p involved passing it over a gel filtration column. Traces of small molecular weight degradation products or contaminating protein were then removed. Also traces of C-Rio2p that had unfolded during purification could

be removed from folded protein. Unfolded protein would have a larger Stokes' radius and elute in a smaller volume on the gel filtration column. Following this method, I was able to efficiently purify C-Rio2p (Figure 2.8).

2.4.3: Use of Chemical Lysis and Induction Overnight at 20°C to Combat Degradation

As discussed previously, it appeared as if Rio2p was being cleaved by proteases after lysis. To purify this kinase, many changes were made to stop degradation of Rio2p. The first of these changes was that initially the cultures of Rio2p were grown in 2 L culture volumes in 4 L flasks. The flasks were shaken at 220 rpm. We found that this method of growing cultures was inefficient at aerating the cultures, which led to an increase in degradation products. Thereafter, all cultures were grown in 1 L of media in a 4 L flask with shaking at 250 rpm.

As another way to combat the problem of degradation, the method of lysis was changed from a French Press to chemical lysis in an ice bath at 4°C with BugBuster^{TM}. The problem with a French Press is that it heats the lysate during its use. This leads to decreased protein stability, especially if the protein heats past its melting temperature. Use of the press also leads to an increase in proteolytic activity as it nears the temperature at which proteases are optimally active, i.e. 37°C.

Another change made to combat the problem of degradation, was to change the induction from 30°C for four hours to 20°C overnight. Since Rio2p is a yeast protein, and the optimal temperature for growing yeast is at 30°C, Rio2p expression can be induced at 30°C. However, yeast proteins with low melting temperatures, such as Rio2p, are more stable at lower temperatures. Inducing at 20°C overnight helps maintain low melting point proteins in a more folded state.   It is possible that lowering the temperature to 20°

could slow translation and help with proper folding, which could lead to a lower incident of protein degradation.

2.4.4: Approach to Increase the Solubility of Rio2p by Adding ATP or ADP and MgCl$_2$

ATP and MgCl$_2$ or ADP and MgCl were added to all buffers to increase the solubility of Rio2p. Since Rio2p is a serine kinase and archaeal Rio2p has been shown to autophosphorylate, ATP is a substrate of Rio2p. Including ATP or an analogue of ATP, such as AMP or ADP, in the solution changes the conformation of Rio2p into a more folded and stable structure that is more soluble. The increased stability of Rio2p with ATP or ADP was demonstrated by the fact that the melting temperature of C-Rio2p increased in the presence of ATP from 43°C (bound to AMP) to 47°C and increased even more in the presence of ADP to 51°C (Sagar,V. (2011). Personnel research.). C-Rio2p was purified two ways: by including 1 mM ATP in all buffer solutions; alternatively by including 1 mM ADP in all buffer solutions.

Mg^{2+} helps to stabilize the ATP and ADP bound structure in the active site. Thus 2.5 mM MgCl$_2$ was included in the lysis solution and all buffer solutions along with ATP or ADP.

The results of adding ATP and MgCl$_2$ or ADP and MgCl$_2$ were the same. The solubility of Rio2p was increased and the net yield of protein increased from 5 mg per 2 L culture to 20 mg/ 2 L culture. Addition of ATP or ADP and MgCl$_2$ combined with the change in the construct used helped increase the maximum level at which Rio2p can be concentrated from 5 mg/mL (N-Rio2p w/o ATP and MgCl$_2$ or ADP and MgCl) to 40 mg/mL (C-Rio2p w/ ATP and MgCl$_2$ or ADP and MgCl), which allowed for

crystallization screening of the protein. The SDS-PAGE gel in Figure 2.5 demonstrates that C-Rio2p has limited solubility in the supernatant in the absence of ATP or ADP.

2.4.5: Addition of RNases to Lysate Prevent Rio2 from Nonspecifically Binding to RNA

Rio2 proteins have been shown to nonspecifically bind RNA. The chromatograms of Rio2p, ctRio2 and afRio2 on a Q^{TM} column shows a large peak eluting at high salt concentrations. This peak does not show up on the protein gel and has a high 260/280 absorbance, suggesting it is RNA or DNA contamination that co-eluted with Rio2 from the first step of purification, the His Trap^{TM}. With RNA nonspecifically bound to Rio2p, the protein sample is not homogenous, and is less likely to crystallize. Thus RNases A and RNase T1 were included in the lysate to minimize binding of Rio2p to RNA. At the finish of purification, the A60/A280 ratio was 0.8 for fractions of Rio2p, characteristic of a protein unbound to RNA or DNA.

2.4.6: Results of Crystal Screens of C-Rio2p

The crystals from the sitting drop screen were reproducible in hanging drop after a period of three to four weeks, but did not grow larger in size than those found in the initial screen, remaining microcrystals (Figure 2.11).

Crystal screens were set up with 20 mM ADP instead of 20 mM ATP, since ADP (51°C) stabilizes C-Rio2p more than does ATP (47°C), as witnessed by the melting temperature. C-Rio2p was also screened for crystallization without glycerol. Including glycerol in solution increased the solubility of C-Rio2p but adversely increased the precipitant concentration needed for crystallization. Since the crystallization screens were set at one precipitant concentration, getting initial crystal hits with glycerol might be

24

a problem. In many of the wells with both ATP and ADP, with and without glycerol, quasi crystals formed that indicated the protein was becoming supersaturated in pocketed areas but not nucleating to form crystals.

2.5: Summary

During lysis, Rio2p degrades to smaller fragments which are difficult to remove during purification. Two constructs were made with a N and C terminal His tag to determine which produced the most purifiable protein product. C-Rio2p could be purified via His Trap^{TM} column and S^{TM} column to remove C-terminal (His Trap^{TM} column) and N-terminal (S^{TM} column) degradation products. Though the overall charge of the protein is basic it could still bind the S^{TM} column via the positively charged N-terminal wing helix domain.

ATP and Mg^{2+} was added to the solubilization buffers and increased the solubility of Rio2p from 5 mg/mL to 20 mg/mL. ATP and Mg^{2+} stabililized Rio2p in a more folded form as witnessed by the increase in melting temperature of Rio2p in the presence of these additives.

Rio2p binds to RNA nonspecifically and to remove RNA from the protein solution, RNases were included in the lysate. RNA contamination of Rio2p would impede it's crystallization. With these modifications, I could successfully purify a large quantity of Rio2p.

Rio2p did not yield diffraction quality crystals. The key to this problem might be found in the initial observations that Rio2p degrades easily. Rio2p might contain large unstructured regions that prevent the protein from forming a stable and rigid crystal lattice. Alignment of the sequence of Rio2 from *A. fulgidus* bacteria and Rio2p reveals

that they align well with small insertions in Rio2p (Figure 2.1). However, Rio2p contains

an extended C-terminus. Since the N-terminus forms a fairly well structured winged

helix domain, and the RIO domain is also fairly well structured, degradation from this

extended C-terminus might lead to the degradation of Rio2p seen on the SDS-PAGE gel

after His Trap^{TM}. This supposedly floppy C-terminal region might be the reason why a

C-terminal His tag without any linker still binds to the His Trap^{TM} as the C-terminal

sequence provides the flexible linker to the His tag. Redesigning the construct to include

a truncated C-terminus might lead to fewer degradation products. It may also eliminate

the unstructured regions of Rio2p and lead to successful crystallization of Rio2p. The

problem is knowing where to truncate Rio2p. Sequence alignment with archaeal Rio2p

and secondary structure prediction might provide a clue as to where to terminate the

sequence of Rio2p, but would not ensure that the protein will fold properly. Several

constructs might have to be tried before a solution to the problem of crystallization is

found. My approach to this problem was to move on from Rio2p and attempt to purify

(and crystallize) another eukaryotic version of Rio2—human Rio2, which will be

discussed in Chapter 3.

```
                    10        20        30        40        50        60
                    |         |         |         |         |         |
AFRio2  MNIAELYG-KMGKHSWRIMDAIFKNLWDYEYVPLQLISSHARIGEEKARN-ILKYLSDLR
Rio2p   MKLDTSHMRYLTTDDFRVLQAVEQGSRSHEVVPTPLIHQISGMRSQSGTNRAISDLAKLS
        *::     :      : ...:*:::*: :.  .:* **   ** . :  .:.. *   :. *:.*
                    70        80        90        100       110       120
                    |         |         |         |         |         |
AFRio2  VVQNRQ-KDYEGSTFTFIGLSLYSLHRLVRSGKVDAIGKLMGEGKESAVFNCYSEKFGEC
Rio2p   LISKMRNVKYDGYRLTYNGIDYLALKTMLNRDTVYSVGNTIGVGKESDIYKVSDKNGNPR
        ::.: :   .*:* :  *: *:. :*: ::. ..* ::*: :* **** ::: .:: .
                    130       140       150       160       170       180
                    |         |         |         |         |         |
AFRio2  VVKFHKVGHTSFKKVKEKRDYGD---------LHFSVLAIRSARNEFRALQKLQGLAVPKV
Rio2p   VMKIHRLGRTSFHSVRNNRDYLKKSNQGANWMHLSRLAAN-KEYQFMSMLYSKGFKVPEP
        *:*:*::*:***:.*:::*** .         .*:* **  .   . :* ::   :*: **:
                    190       200       210       220       230       240
                    |         |         |         |         |         |
AFRio2  YAWEGNAVLMELIDAKELYRVR-VENPDEVLDMILEEVAKFYHRGIVHGDLSQYNVLVSE
Rio2p   FDNSRHIVVMELIEGYPMRRLRKHKNIPKLYSDLMCFIVDLANSGLIHCDFNEFNIMIKD
        :  .   :  *:*****:.   : *:*   :*  ::  . :..:  : *::* *:.::*:::.:
                    250       260       270       280       290       300
                    |         |         |         |         |         |
AFRio2  E-------GIWIIDFPQSVEVGEEGWREILERDVRNIITYFSRTYRTEKDINS-------
Rio2p   KLEDENDCGFVVIDFPQCISIQHQDADYYFQRDVDCIRRFFKKKLKYEPKPDSSMLDTEG
        :        *: :*****.:.: .:.     ::*** *  .:*.:. : * . :*
                    310       320       330       340       350       360
                    |         |         |         |         |         |
AFRio2  --------------------AIDRILQE-------------------------------
Rio2p   FGDGYKYAYPDFKRDVKRTDNLDELVQASGFSKKHPGDRGLETAVESMRNAVYNSDDDMS
                            :*.::*
                    370       380       390       400       410       420
                    |         |         |         |         |         |
AFRio2  ------------------------------------------------------------
Rio2p   NDEAEEENGEGDYSEEDEYYDSELDNESSEDDSEDAQEEENERIIEALSSGVENLKMDKL


AFRio2  ------
Rio2p   GNYILE
```

Figure 2.1: Alignment between AfRio2 and Rio2p.

Alignment was performed with the program CLUSTALW using sequences from *A. fuligidus* (AFRio2) and Rio2p (from yeast). Identical, highly similar, and weakly similar residues are indicated by red, green, and blue letters respectively. The wing helix domain of afRio2 is highlighted in yellow, the N-terminal lobe in green and the C-terminal lobe in blue. The C-terminal extension of Rio2p is in the boxed region encompassing residues 293-496.

A.

N'	6x His Tag	TEV Site	Rio2p	C'

B.
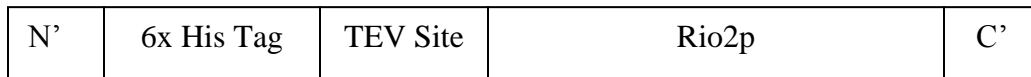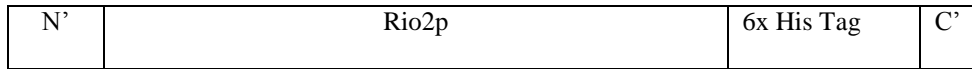
N'	Rio2p	6x His Tag	C'

Figure 2.2: Two Constructs Designed for Rio2p Purification
- A. N-Rio2p construct was designed with an N-terminal 6x His tag followed by a TEV cleavage site.
- B. C-Rio2p construct was designed with a C-terminal 6x His tag.

A.

TGTTTAACTTTAAGAAGGAGATATACATATG	AAATTGGATACTTCTCATATGAGATATTTG
RBS Site	Coding Sequence Rio2p

B.

CTACTA	GTGATGATGGTGGTGATG	CTCTAGTATATAGTTACCTAGTTTATCCCAT
2 Stop Codons	6x His Tag	Coding Sequence Rio2p

C.

GGGGACAAGTTTGTACAAAAAAGCAGGCTCG	TGTTTAACTTTAAGAAGGAGATATACATATG
Attb1 Site	Rbs Site

D.

GGGGACCATTTGTACAAGAAAGCTGGTC	CTACTA	GTGATGATGGTGGTGATG
Attb2 Site	Stop Codons	6x His Tag

Figure 2.3: PCR of C-Rio2p
In PCR1, Rio2p was amplified from the pDEST527 plasmid containing an insert for the Rio2p gene. A.
The forward primer for PCR1 was used to add an RBS site to the 5' end of the Rio2p gene. B. The reverse
primer for PCR1 was used to add a 6x His tag and two stop codons to the 3' end of the Rio2p gene. C. and
D The forward and reverse primers for PCR2. C. was used to add an attb1 site to the 5' end of PCR1
product. D. was used to add an attb2 site to the 3' end of PCR1 product.
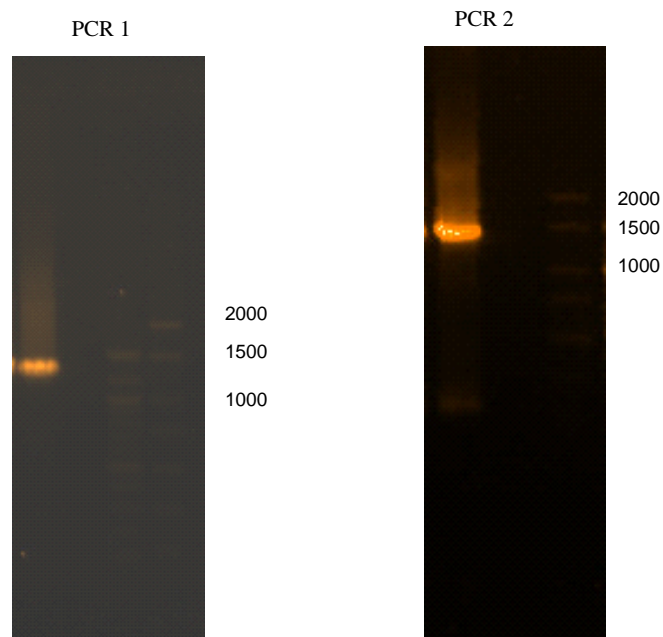
Figure 2.4: PCR Products
Both PCR1 and PCR2 products were run on a 1% agarose gel stained with ethidium bromide. In PCR1, Rio2p was amplified from the pDEST527 plasmid containing an insert for the Rio2p gene. Using PCR primers, A RBS site was added to the 5' end of the Rio2p gene and a 6x His tag and two stop codons were added to the 3' end of the Rio2p gene. In, PCR2, the gene product from PCR1 was amplified. An attb1 site was added to the 5' end of PCR1 product and an attb2 site was added to the 3' end of PCR1 product. The size of the expected product is 1275 nucleotides.
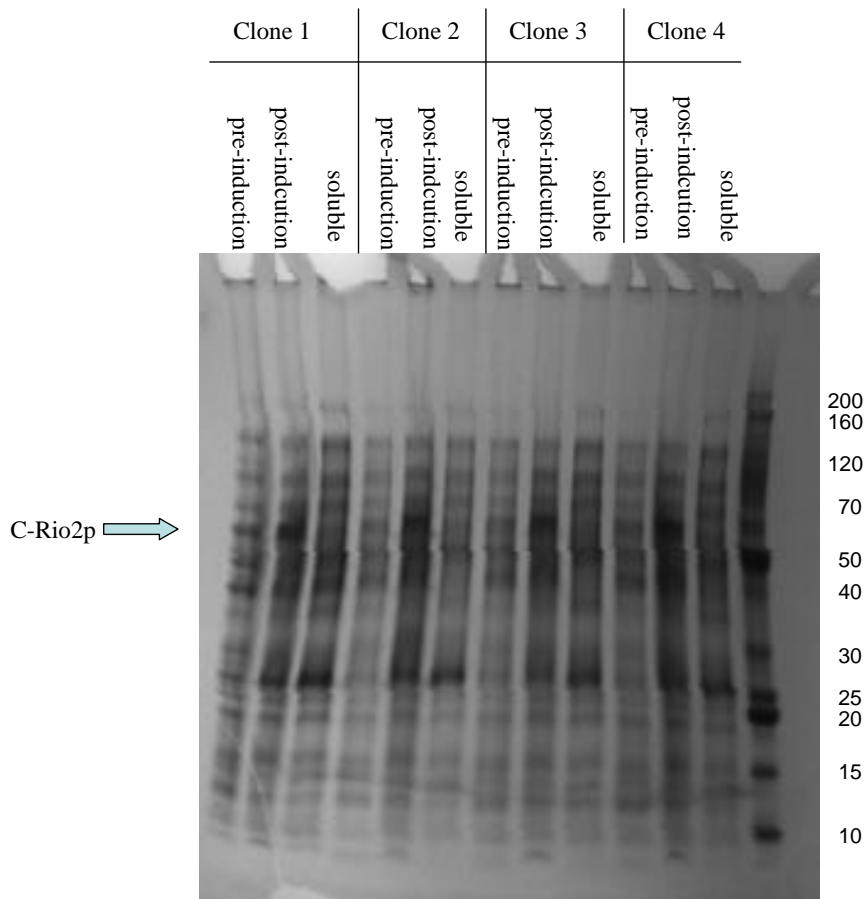
Figure 2.5: Induction of C-Rio2p
Rossetta[TM] cells were transformed with pDEST 14-C-Rio2p. Four individual colonies were grown in small cultures and induced overnight at 18°C with 1mM IPTG. An SDS-PAGE gel was run of samples of the four cultures before and after induction lysed with 5x SDS loading dye and stained with Coomassie Blue. Samples of the four cell cultures were also lysed with 0.3x Bugbuster ,centrifuged, and the four supernatants were run on the gel. Rio2p is expected to run at 50 kDa.
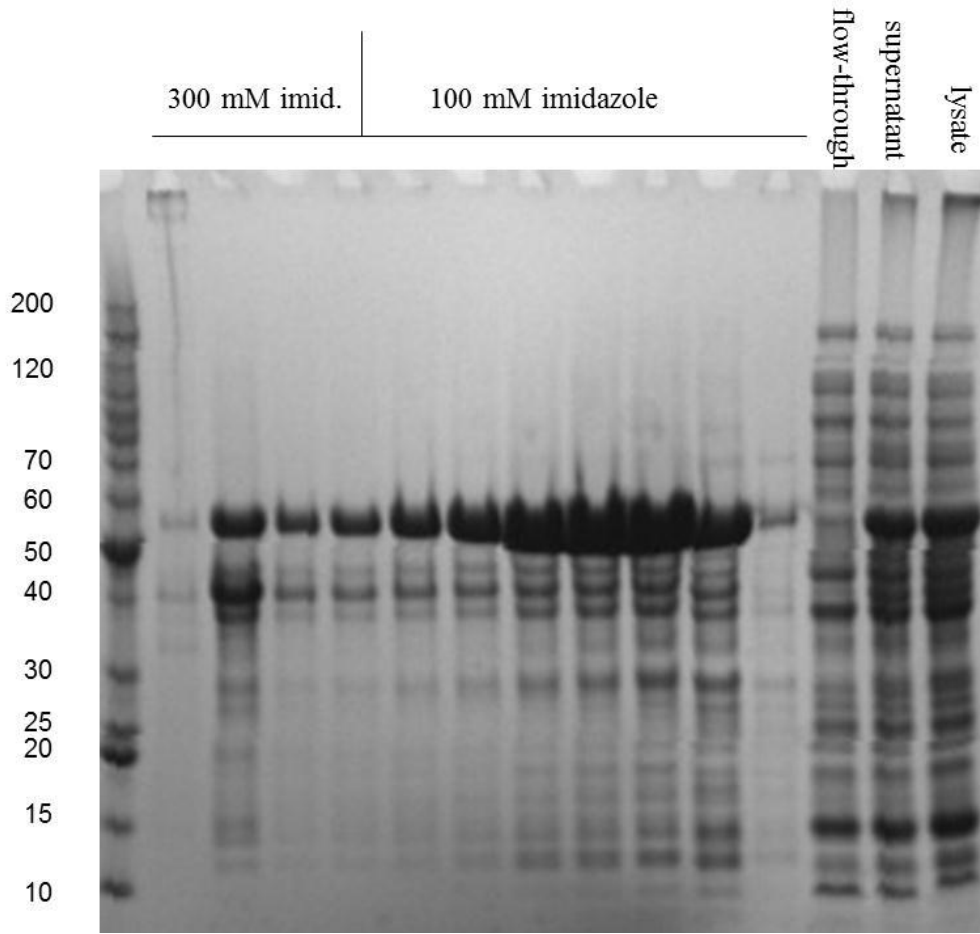
Figure 2.6: C-Rio2p Purification Step 1
Supernatant with overexpressed C-Rio2p was passed over a His Trap[TM] column. The column was washed with 50 mM imidazole. C-Rio2p was eluted from the column in a step gradient at 100 mM imidazole and 300 mM imidazole steps. The SDS-PAGE gel is stained with Coomassie Blue. Rio2p is expected to run at 50 kDa and is the most abundant band.
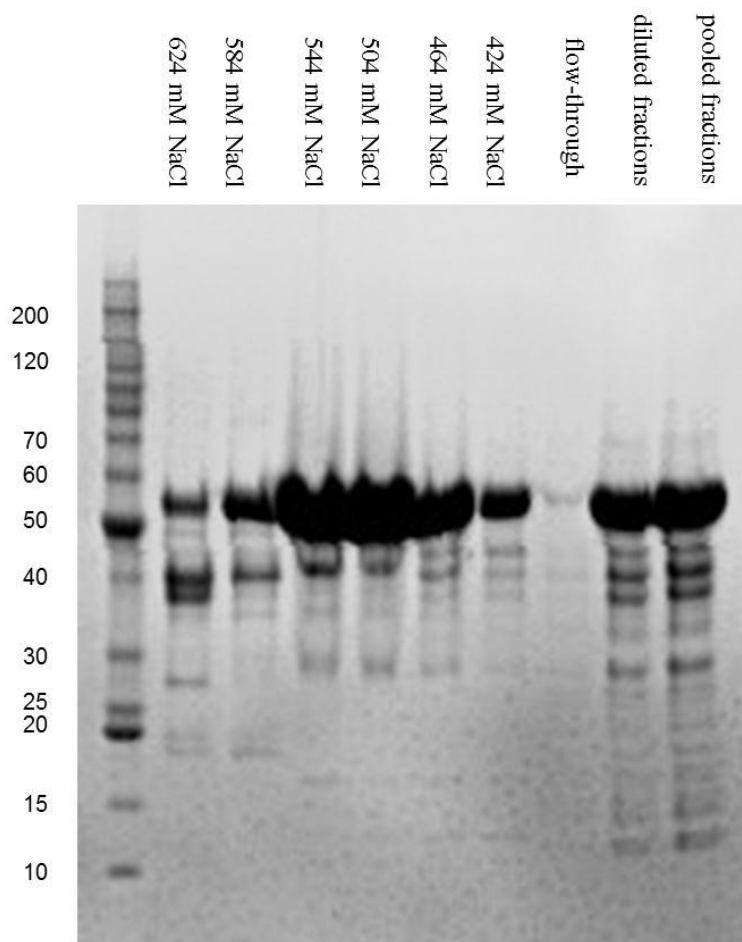
Figure 2.7: C-Rio2p Purification Step 2
After His Trap^TM, C-Rio2p was run on a cation exchange column S^TM column. The gel shows the pooled C-Rio2p fractions from the His Trap^TM column. The fractions were diluted to 200 mM NaCl (labeled on the gel as "diluted fractions") and loaded onto the S^TM column. C-Rio2p was eluted over a gradient of 200 mM to 1 M NaCl over 100 mL solvent. The SDS-PAGE gel, stained with Coomassie Blue, shows the flow-through from the S^TM column as well as individual fractions labeled with the approximate salt concentration at which they eluted.

A.
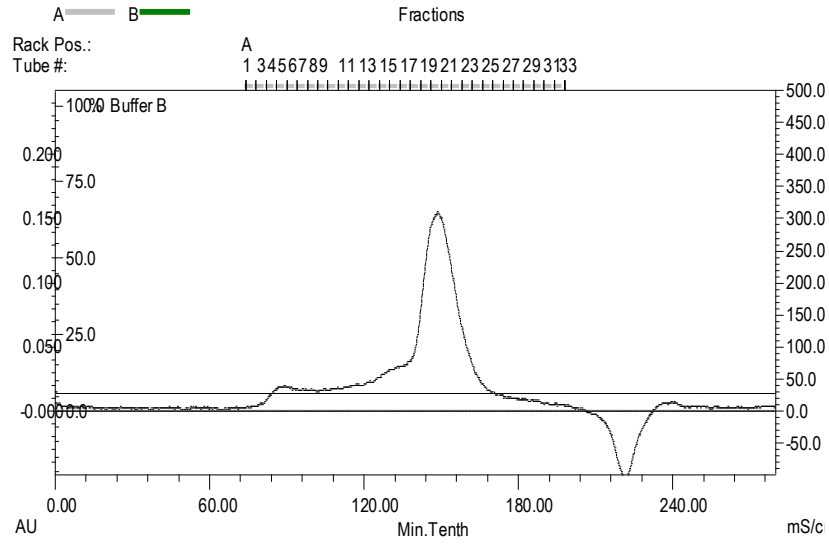
B.
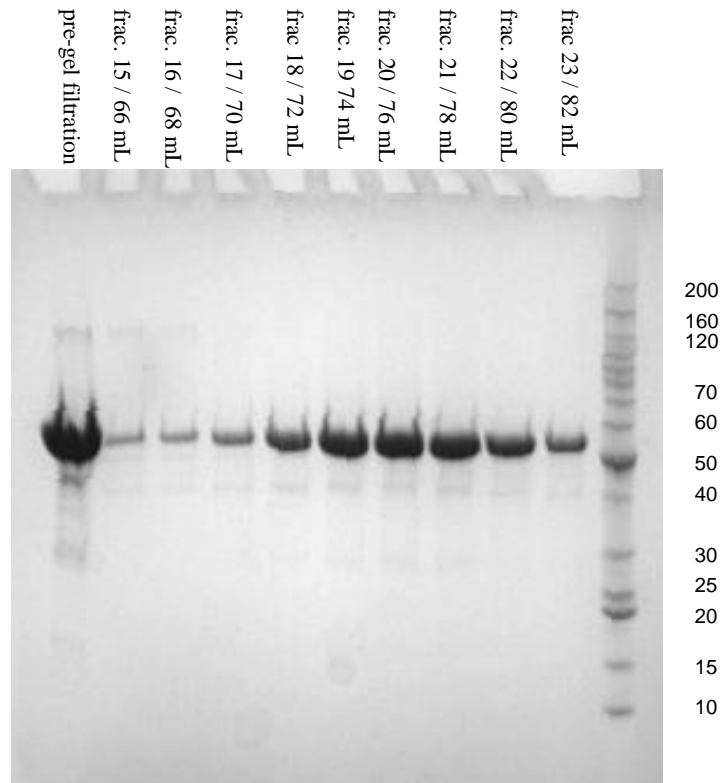
Figure 2.8: C-Rio2p Purification Step 3
C-Rio2p previously purified by His Trap™ column and by S™ column was loaded on a Superdex 200 Prep Grade gel filtration column.   A. The chromatogram of C-Rio2p shows the elution profile from the size exclusion column. B. The SDS-PAGE gel (stained with Coomassie Blue) is labeled with the fraction number and the elution volume of C-Rio2p elution from the size exclusion column.
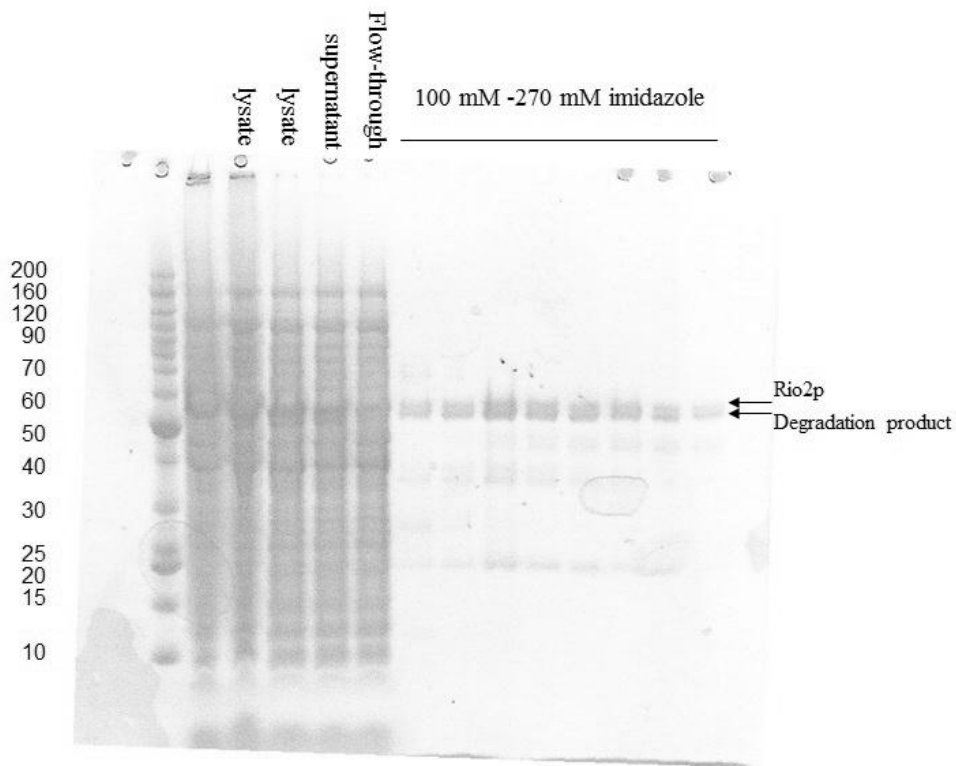
Figure 2.9: Rio2p Degradation Observed with N-Rio2p
This gel is a SDS-PAGE gel stained with Coomassie. Rosetta™ cells containing the plasmid for N-Rio2p were grown and lysed with 0.3x Bugbuster in the presence of three protease inhibitor tablets. The supernatant was passed over a His Trap™ column and N-Rio2p was eluted from the column over 100 mM-270 mM imidazole. The lysate, supernatant, flow-through, and fractions from the elution were run on a gel. The slightly lower molecular weight protein contamination that co-elutes with N-Rio2p on the His Trap™ column is labeled on the gel.
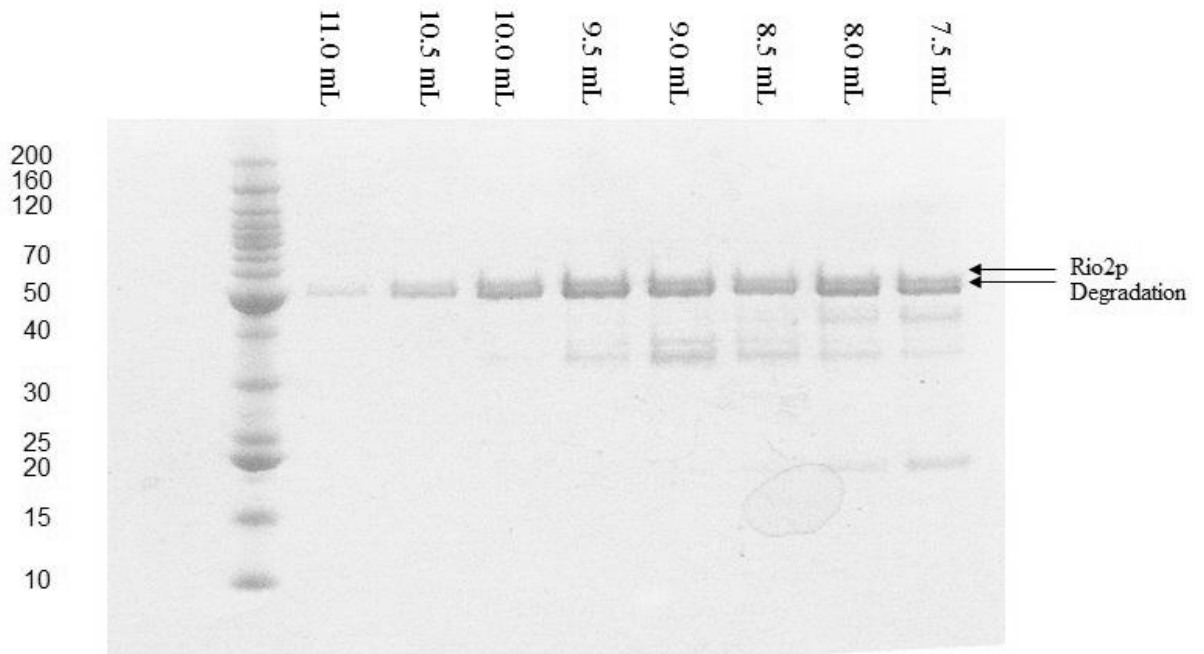
Figure 2.10: N-Rio2p Co-Purifies with Degradation Product
After purification by His Trap^{TM} and Q^{TM} column, N-Rio2p was loaded on a gel filtration column. The lanes are labeled with the elution volumes from the column. A close molecular weight degradation product that co-elutes with N-Rio2p is labeled. The SDS-PAGE gel is stained with Coomassie Blue.

Figure 2.11: Crystallization of C-Rio2p
Sitting drop screens of C-Rio2p were performed. Microcrystals of C-Rio2p formed in 0.2 M magnesium formate, 20% PEG 3350.

Protein	Construct	Crystal	Purification Results
Rio2p	x N Terminal His Tag ✓ **C Terminal His Tag**	• Crystals were not diffraction quality.	• PI=5.6 but purify by cation exchange column. • Add RNAases to lysate. • Add ATP to buffer to solubilize.
Human Rio2	x N terminal His Tag ✓ **C Terminal His Tag** x Deletion of Wing Helix Domain	• Poor Solubility. • Could not be concentrated. • Precipitated overnight.	• ATP, ADP, AMP, NP40 solubilized protein to a low extent.
Chaetomium Thermophilum Rio2 (CtRio2)	✓ **N Terminal His Tag**	• Diffraction quality crystals.	• Overload Ni column to remove contaminating proteins. • Anion exchange column to remove bound RNA.

Table 2.1: Eukaryotic Proteins Purified and Attempts Made to Crystallize
The table summarizes the results of Chapters 2-4. The first column delineates the protein purified. The second column delineates the constructs made with bold letters indicating the construct that gave purifiable protein. The third column delineates problems or successes made with the protein. The fourth column delineates the techniques used to overcome problems encountered during purification.

Chapter 3: Purification of Human Rio2

3.1: Introduction

Human Rio2 (hRio2) may contain structural differences from the archaeal version that may be critical to its kinase function. Most importantly hRio2 contains an extended C-terminus compared to AFRio2 (Figure 3.1). To elucidate the hRio2 function, I purified the protein using *E. coli* expression.

Initial studies showed that in E. coli hRio2 had a low level of expression, low solubility, and the recombinant protein co-purified with many degradation products. To overcome these problems I made three different versions of hRio2. In addition, I added different additives to increase the prospect of a soluble, purifiable protein product.

3.2: Design

Three recombinant constructs were designed for hRio2. The N-hRio2 construct was made by subcloning of the human Rio2 gene with an N terminal TEV cleavage site into Gateway^{TM} vector pDEST 527 that codes for a N-terminal 6x His tag. A schematic of the construct is shown in Figure 3.2A. The C-hRio2 construct was made by subcloning of the human Rio2 gene engineered with a C-terminal His tag into Gateway^{TM} vector pDEST 14. A schematic of the construct is shown in Figure 3.2B. In the third Δ71-C-hRio2 construct, the first 71 amino acids were deleted from the N-terminus by site directed mutagenesis and a 6x His tag was engineered at the C-terminus. Δ71-C-hRio2 was subcloned into Gateway^{TM} vector pDEST 14. A schematic of the construct is shown in Figure 3.2C.

These three constructs were designed to determine which construct would have the best induction, solubility, and purification. For yeast Rio2, it was found that the C-

terminal 6x His tag was best in all domains (Chapter 2). Constructs with both C-terminal

and N-terminal 6x His tags were investigated for hRio2 to determine, in this case, which

was better in terms of induction, solubility, and purification. Another construct, Δ71-C-

hRio2 with truncation of the winged helix domain, was also investigated to determine if it

would help with the low solubility of hRio2. Since it was found that removing the N-

terminus of Rio1and Rio3 made the proteins more soluble (Kiburu, I. and Thomas, S.

(2011). private communication, unpublished raw data.), it was hypothesized that the same

would occur by removing the winged helix domain from hRio2. Moreover, the construct

of Rio2p minus the winged helix domain was soluble in yeast cells (Chapter 6).

3.3: Methods

3.3.1: General Methods

All three contructs were transformed into Rosetta DE3 (pLysS)^{TM} cells. They

were grown in 2L cultures of LB and induced with 1 mM IPTG. Cells were chemically

lysed with Bugbuster and then passed over a His Trap^{TM} column. If protein was

purifiable, the protein was then passed over a S^{TM} column.

3.3.2: Subcloning of C-hRio2

HRio2 with the C-terminal 6x His tag (C-hRio2) was the construct that produced

the most purifiable hRio2. The PCR for formation of this construct follows. In order to

express this protein in *E. coli* the full length protein was subcloned into an inducible

bacterial expression vector. As described in previous chapters, The Gateway^{TM} system

was used to clone C-hRio2 into pDEST 14.

C-hRio2 was amplified from the pDEST 527 plasmid containing an insert for the

hRio2 gene obtained from Dr. Nicole LaRonde LeBlanc from her work at NCI. PCR was

performed in two steps as described in Chapter Two. The forward and reverse primers

for the two PCR reactions are diagramed in Figure 3.3. Thereafter an agarose gel was

run with the PCR products obtained from the screen of two positive clones (Figure 3.4).

3.3.4: Expression and Purification of hRio2

The C-hRio2 construct produced the most purifiable hRio2 (Figure 3.2B). The

procedure used for hRio2 expression and purification is as follows. C-hRio2 was cloned

into a pDEST14 vector. The vector was transformed into *E. coli* Rosetta (pLysS)[TM] cells.

An 80 mL starter culture of Rosetta[TM] cells with 100 μg/mL ampicillin and 32 μg/mL

chloroamphenicol was grown overnight at 37°C. Next day, the 80 mL starter culture was

used to inoculate 4 L of LB media in four 4 L flasks with ampicillin and

chloroamphenicol. When the O.D. 600 reached 0.6, the culture was transferred to an

18°C shaking incubator to cool the flasks. When the O.D. 600 reached 0.8, the culture

was induced with 1 mM IPTG and transferred to an 18°C incubator overnight. The

O.D.600 had reached 2.0 by the next morning and the cells were spun down and

transferred to a conical vial. The vial held approximately 3 g of cell pellet per L. The

vial was frozen and maintained in a −80°C freezer.

The frozen pellet was resuspended in 10 mL of Buffer A (0.05% BME, 500 mM

NaCl, 50 mM Tris, pH 8.0, 10% glycerol, 2.5 mM $MgCl_2$, 1 mM ATP) per gram of frozen

pellet. 0.5 mM PMSF, 2 mM Benzamidine, 1 EDTA free protease inhibitor tablet, 5 mg

DNAse/L cell culture, 1 mg RNase A/L cell culture, and 20 μL RNase T1 (100,00U/mL)

was added to the solution. The cell pellet was dissolved by vigorous vortexing in this

buffer. To the dissolved cell pellet was added 0.3x Bugbuster[TM] and 0.5 mg/ml

lysozyme. Lysis was performed on ice for an hour. Lysed cells were centrifuged at

16,000 rpm for 30 minutes. The pellet was discarded. The supernatant containing the soluble protein was filtered and loaded onto a GE Healthcare His Trap^{TM} 5 mL column. Elution Buffer, Buffer B, was 1M imidazole, 0.05% BME, 500 mM NaCl, 50 mM Tris, pH 8.0, 10% glycerol, 2.5 mM MgCl$_2$ and 1 mM ATP. The His Trap column was washed with 5% Buffer B and then 10% Buffer B and 5 mL fractions were collected. A gradient was run from 10% Buffer B to100% Buffer B in 100 mL and 5mL fractions were collected.

The fractions with hRio2 protein were pooled and the solution was diluted so that the final concentration of solution components was 200 mM NaCl, 50 mM HEPES, pH 7.0, 2.5 mM MgCl$_2$, 0.05% BME, 1 mM ATP and 10% glycerol. Buffer C, the loading buffer for the cation exchange column, was 10% glycerol, 2.5 mM MgCl$_2$, 50 mM HEPES pH 8.0, 200 mM NaCl, 0.05% BME, 1 mM ATP and Buffer D, the high salt buffer was 1M NaCl, 10% glycerol, 2.5 mM MgCl$_2$, 50 mM HEPES pH 8.0, 0.05% BME, 1 mM ATP . Protein was then loaded onto a S^{TM} column and a gradient (of Buffer C and D) from 200 mM to 1 M salt was run over 100 mL. 5 mL fractions were collected.

3.4: Results
3.4.1: Purification of N-hRio2

Induction and solubility of N-hRio2 were found to be adequate in some of the conditions tested. Induction of Rosetta^{TM} cells containing the pDEST 527-N-hRio2 was attempted both at 37°C for four hours and overnight at 18°C with 1 mM IPTG. N-hRio2 had the best induction at 37°C, as seen in Figure 3.5. Overnight induction at 18°C with 1mM IPTG was found to be too low to observe, as seen in Figure 3.6. N-hRio2 was solubilized with both 250 mM NaCl and 500 mM NaCl. The best solubility of N-hRio2

was in 500 mM NaCl. At this concentration of salt, N-hRio2 bound tightly to the His

Trap^{TM} column and eluted at 125 mM imidazole over several fractions. At 250 mM

NaCl, N-hRio2 bound to the His Trap^{TM} column weakly and eluted in the 40 mM wash in

only one fraction.

     While induction and solubility of N-hRio2 were sufficient, the amount of

degradation products that co-purified with N-hRio2 at 125 mM imidazole was extensive,

as can be seen from the SDS-PAGE gel of the His Trap^{TM} elution in Figure 3.7. For this

reason further attempts at purification of N-hRio2 were abandoned.

3.4.2: Purification of Δ71-C-hRio2.

     Rosetta^{TM} cells harboring the pDEST 14-Δ71-C-hRio2 plasmid were induced at

18°C overnight with 1 mM IPTG. Though the yeast construct was soluble without the

winged helix domain (Chapter 6), the human construct was not. Little to no protein

appeared in the supernatant after lysis of the *E. coli* cells. The supernatant containing 1

mM ATP was purified on the His Trap^{TM} column to see if a small amount of protein was

solubilized. No protein eluted in the gradient and no fractions were available from the

elution to run on the SDS-PAGE gel (Figure 3.8). The Δ71-C- hRio2 construct could not

be used to increase the solubility of hRio2.

3.4.3: Purification of C-hRio2.

     The induction level of C-hRio2 was adequate at 18°C overnight with 1 mM IPTG.

However, as discussed later, the solubility was low. I used the purification method of C-

Rio2p (Chapter 2) to purify C-hRio2. C-hRio2 was passed over a His Trap^{TM} column to

remove N-terminal degradation products and thereafter passed over a S^{TM} column to

remove C-terminal degradation products. Though the pI of C-hRio2 was 5.66, it was

found to still bind to the cation exchange column despite the overall negative charge of the molecule, possibly because the N-terminal winged helix domain has an electrostatically positive surface potential. C-hRio2 did bind tightly to the S^{TM} column and eluted at 530 mM salt (Figure 3.10). I purified C-hRio2 using the above two-step method. For this reason, N-hRio2 and $\Delta71$-C-hRio2 constructs were not pursued further.

The problem with the C-hRio2 construct was that while C-hRio2 was induced it had poor solubility. Though the protein could be isolated from the His Trap^{TM} in relatively pure form, the C-hRio2 precipitated 24-hours later. In fact, the C-hRio2 could not remain soluble for more than one day unless C-hRio2 was diluted to 0.3 mg/ mL and stored at that concentration. Furthermore, C-hRio2 once diluted, could not again be concentrated because the protein bound to the membrane filter of the ultraconcentrator.

3.4.4: Additives Tested for Their Ability to Solubilize C-hRio2

Several additives were added to the buffers and tested for enhanced solubility of C-hRio2. The additives tested were 0.1% NP40 with 10% glycerol, 1 mM AMP, 1 mM ATP, and 1 mM ADP. If hRio2 contained hydrophobic regions, adding detergent and glycerol in the lysate would solubilize these domains. 0.1% NP40 was the detergent used to solubilize Rio2p and the ribosome during TAP Tag pull downs (Chapter 6). Therefore, 0.1% NP40 was chosen to solubilize C-hRio2.

AMP should stabilize C-hRio2 in a manner similar to that in which ADP and ATP stabilized Rio2p. Though AMP is neither a substrate nor a product of autophoshorylation, it maintains the adenosine moiety of both compounds and contains a phosphate group. AMP was chosen in addition to ADP and ATP because it is a more stable molecule than either ADP or ATP and can remain in solution longer without

deteriorating. ATP can deteriorate to ADP; similarly ADP can deteriorate to AMP. Thus

Rio2 protein in solutions of ATP and ADP will, over time, become heterogeneous as

different portions of the protein will bind to ATP, ADP and AMP. By using AMP this

heterogeneity is avoided.

ATP was tested as an additive in the buffer solutions, because it was expected to

stabilize C-hRio2 to a greater degree than AMP. The melting temperature of Rio2p

bound to ATP is greater (47°C) than the melting temperature of Rio2p bound to AMP

(43°C) (Sagar, V. (2011)).

ADP was tested as an additive in the buffer solutions, because it was expected to

stabilize C-hRio2 to a greater degree than ATP. The melting temperature of Rio2p bound

to ADP (51°C) is greater than the melting temperature of Rio2p bound to ATP (47°C)

(Sagar, V. (2011)). ADP is also a more stable molecule than ATP.

3.4.4.1: Results of Adding NP40 to Buffer Solutions.

The additive 0.1% NP40 along with 10% glycerol was used to try to solubilize C-

hRio2. As seen from the SDS-PAGE gel in Figure 3.9, induction of C-hRio2 showed a

thick band.   A slightly smaller band was seen in the lane with the supernatant indicating

the protein was not fully soluble with this additive. 0.1% NP40 slightly enhanced the

solubility of C-hRio2. The elution of C-hRio2 from the His Trap^{TM} revealed only a small

to moderate amount of protein eluting from the column and the protein bands on the

SDS-PAGE gel were narrow. In Figure 3.10, evidence can be seen of the small amount

of soluble protein obtained after lysis through the SDS-PAGE gel of the elution of C-

hRio2 from the S^{TM} column. The protein band width shown on this gel is very thin.

Another problem with purifying the protein with 0.1% NP40, was that with NP40 C-hRio2 bound poorly to the column and results of subsequent purifications of C-hRio2 were found to be variable as to whether or not any protein bound to the His Trap™ column. Usually such ambiguous results might be indicative of RNA contamination of the protein that interfered with protein binding to the His Trap™ column.

3.4.4.2: Results of Adding AMP as an Additive to Buffer Solutions.

AMP was added to the buffer solutions to try to solubilize C-hRio2. However, the addition of AMP to the buffers had a more negative effect on the solubility of C-hRio2 than did NP40, as determined from the gels obtained from His Trap™ and S™ column purification. The size and appearance of the protein bands on the SDS-PAGE gels in Figures 3.11 and 3.12 are small and faint. The soluble concentration of protein in the supernatant obtained after adding AMP to the buffers was low, leading to dilute fractions eluting from the His Trap™ and S™ columns.

3.4.4.3: ATP as an Additive to Buffer Solutions.

ATP was added to the buffer solutions in order to solubilize C-hRio2. The solubility of C-hRio2 increased with ATP in the buffers more than with NP40 or AMP. After passing the protein over the His Trap™, most of the protein had precipitated by the next day when the SDS-PAGE gel was loaded, so the true extent of the effect of ATP on the solubility of C-hRio2 cannot be seen from the gel. The size and appearance of the protein bands on the SDS-PAGE gel are still moderate (Figure 3.13). The fact that C-hRio2 precipitated by the next day, shows that the protein is not stable even in the presence of 1 mM ATP.

3.4.4.4: ADP as an Additive to Buffer Solutions.

ADP was added to the buffer solutions in order to solubilize C-hRio2. The solubility of C-hRio2 with ADP in the buffers could not be determined due to purification issues. During elution of C-hRio2 from the His Trap^{TM} column, with ADP present in the buffers the pressure in the His Trap^{TM} column became too high. This could have happened because ADP stabilized a conformation of C-hRio2 that was prone to precipitation. To counteract this problem, a batch process was used for elution. The protein was loaded on the His Trap^{TM} column, then the resin was removed from the column, and the elution buffer was mixed with the resin. The resin was allowed to settle and the elution buffer was removed from the top of the column. The C-hRio2, then in the elution buffer, was diluted to 0.3 mg/mL to stabilize the protein before loading on a SDS-PAGE gel.

The next day, C-hRio2 was loaded onto an S^{TM} column (Fig 3.15). In this case also during elution from the S^{TM} column, the pressure became too high. Only two fractions eluted from the column during this time. Those two fractions were run on an SDS-PAGE gel and the fraction at 300 mM NaCl contained pure protein. However, only one protein fraction could be rescued from the column. The remaining C-hRio2 probably precipitated and was lost. Thus with ADP present in the buffers, the protein could not be purified in large amounts. Rather the presence of ADP in the buffers made C-hRio2 more prone to precipitation during purification.

3.5: Summary

Several different constructs were designed to determine which recombinant construct results in the most purifiable protein. Summary information on the three

constructs purified is available in Table 3.1 at the end of this chapter. Δ71-C-hRio2 was not soluble in the supernatant and could not be purified. Whereas N-hRio2 gave several degradation bands after His Trap^{TM}, C-hRio2 could be purified to greater than 95% purity using the His Trap^{TM} and S^{TM} column. The main issue with C-hRio2 was its low solubility in buffer solution. While the use of solubilizing agents such as NP40, ADP, and AMP resulted in lower solubility of C-hRio2 in the supernatant, the addition of ATP to buffers allowed a yield of large amounts of highly soluble C-hRio2. However, C-hRio2 could only be stored overnight without precipitating at a low concentration of 0.3 mg/mL. This is not sufficient for crystallization screening. A new construct for hRio2 needs to be developed to increase its stability.

Since the N-terminus is critical for even minimal solubility of human Rio2, it must remain intact. Rather the C-terminus could be truncated. Figure 3.1, shows alignment between hRio2 and *A. fulgidus* Rio2. The two proteins align with few gaps, but hRio2 contains an extended C-terminus. This C-terminus might be the reason that hRio2 is minimally soluble and stable at 0.3 mg/mL while *A. fulgidus* Rio2 is stable at 5 mg/ml. Instead of developing a new construct of hRio2, my approach to crystallizing Rio2 was to purify and crystallize a new protein from a different eukaryotic organism. That organism was *Chaetomium thermophilum* and results from this purification and crystallization were successful as discussed in the Chapter 4.

```
               10        20        30        40        50        60
                |         |         |         |         |         |
hRio2    MGKVNVAKLR-YMSRDDFRVLTAVEMGMKNHEIVPGSLIASIASLKHGGCNKVLRELVKH
AFRio2   ---MNIAELYGKMGKHSWRIMDAIFKNLWDYEYVPLQLISSHARIGEEKARNILKYLSDL
              :*:*:*    *.:..:*::  *:   .: ::* **  .**:*  * :  ..::*:  *  .
               70        80        90       100       110       120
                |         |         |         |         |         |
hRio2    KLIAWERTKTVQGYRLTNAGYDYLALKTLSSRQVVESVGNQMGVGKESDIYIVANEEGQQ
AFRio2   RVVQ-NRQKDYEGSTFTFIGLSLYSLHRLVRSGKVDAIGKLMGEGKESAVFNCYSEKFGE
              :::   :* *  :*   .*  * .  .:*: *    *::*: ** **** ::    .*: .
              130       140       150       160       170       180
                |         |         |         |         |         |
hRio2    FALKLHRLGRTSFRNLKNKRDYHKHRHNVSWLYLSRLSAMKEFAYMKALYERKFPVPKPI
AFRio2   CVVKFHKVGHTSFKKVKEKRDYG----DLHFSVLAIRSARNEFRALQKLQG--LAVPKVY
              .:*:*::*:*:***::*:****      :: :   *:  ** .**   :: *    .:***
              190       200       210       220       230       240
                |         |         |         |         |         |
hRio2    DYNRHAVVMELINGYPLCQIHHVEDPASVYDEAMELIVKLANHGLIHGDFNEFNLILDES
AFRio2   AWEGNAVLMELIDAKELYRVR-VENPDEVLDMILEEVAKFYHRGIVHGDLSQYNVLVSE-
              :: :**:****:.    *  ::: **:*  .* *  .:* ..*: ::*::***::.:*:::.*
              250       260       270       280       290       300
                |         |         |         |         |         |
hRio2    DHITMIDFPQMVSTSHPNAEWYFDRDVKCIKDFFMKRFSYESELFPTFKDIRRE⌐DTLDVE
AFRio2   EGIWIIDFPQSVEVGEEGWREILERDVRNIITYFSRTYRTEKDINSAIDRILQE------
              :  *  :***** *.... .   ::***:  *   :* : :  *.:: .::.  * :*
              310       320       330       340       350       360
                |         |         |         |         |         |
hRio2    VSASGYTKEMQADDELLHPLGPDDKNIETKEGSEFSFSDGEVAEKAEVYGSENESERNCL⌐
AFRio2   ------------------------------------------------------------
              370       380       390       400       410       420
                |         |         |         |         |         |
hRio2    ⌐EESEGCYCRSSGDPEQIKEDSLSEESADARSFEMTEFNQALEEIKGQVVENNSVTEFSEE⌐
AFRio2   ------------------------------------------------------------
              430       440       450       460       470       480
                |         |         |         |         |         |
hRio2    ⌐KNRTENYNRQDGQRVQGGVPAGSDEYEDECPHLIALSSLNREFRPFRDEENVGAMNQYRT⌐
AFRio2   ------------------------------------------------------------
              490       500       510       520       530       540
                |         |         |         |         |         |
hRio2    ⌐RTLSITSSGSAVSCSTIPPELVKQKVKRQLTKQQKSAVRRRLQKGEANIFTKQRRENMQN⌐
AFRio2   ------------------------------------------------------------
              550
                |
hRio2    ⌐IKSSLEAASFWGE⌐
AFRio2   -------------
```

49

Figure 3.1: Alignment between AFRio2 and hRio2

 Alignment was performed with the program CLUSTALW using sequences from *A. fuligidus* (AFRio2) and hRio2 (from yeast). Identical, highly similar, and weakly similar residues are indicated by red, green, and blue letters respectively. The wing helix domain of AFRio2 is highlighted in yellow, the N-terminal lobe in green and the C-terminal lobe in blue. The C-terminal extension of hRio2 is in the boxed region encompassing residues 295-553.

A.

N'	6x His Tag	TEV Site	Human Rio2	C'

B.

N'	Human Rio2	6x His Tag	C'

C.
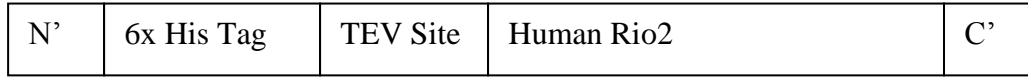
N'	(72-552) Human Rio2	6x His Tag	C'
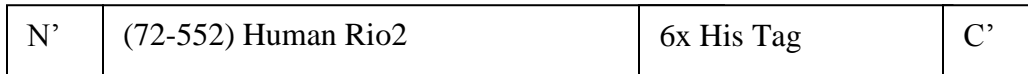
Figure 3.2: Schematic of Three hRio2 Constructs

    A.   Schematic of N-hRio2 with a 6x N-terminal His tag followed by a TEV cleavage site.

    B.   Schematic of C-Rio2 with a 6x C-terminal His tag.

    C.   Schematic of Δ71-C-hRio2 with deletion of the first N-terminal 71 residues of hRio2 and engineered with a C-terminal 6x His tag.

A.

TGTTTAACTTTAAGAAGGAGATATACATATG	GGGAAAGTGAATGTGGCCAAG
RBS Site	HRio2 Sequence

B.

TTATTA	GTGGTGGTGGTGGTGGTG	TTCTCCCCAAAAGCTAGCTGCTTCC
2 Stop Codons	6x His Tag	HRio2 Sequence

C.

GGGGACAAGTTTGTACAAAAAAGCAGGCTCG	TGTTTAACTTTAAGAAGGAGATATACATATG
Attb1 Site	RBS Site

D.

GGGGACCATTTGTACAAGAAAGCTGGG	TTATTA	GTGGTGGTGGTGGTGGTG	TTCTC
Attb2 Site	Stop Codons	6x His Tag	Human Rio2 Sequence

Figure 3.3: PCR of C-hRio2

In PCR1, Rio2p was amplified from the pDEST527 plasmid containing an insert for the hRio2 gene. A. The forward primer for PCR1 was used to add an RBS site to the 5' end of the hRio2 gene. B. The reverse primer for PCR1 was used to add a 6x His tag and two stop codons to the 3' end of the hRio2 gene. C. and D The forward and reverse primers for PCR2. C. was used to add an attb1 site to the 5' end of PCR1 product. D. was used to add an attb2 site to the 3' end of PCR1 product.
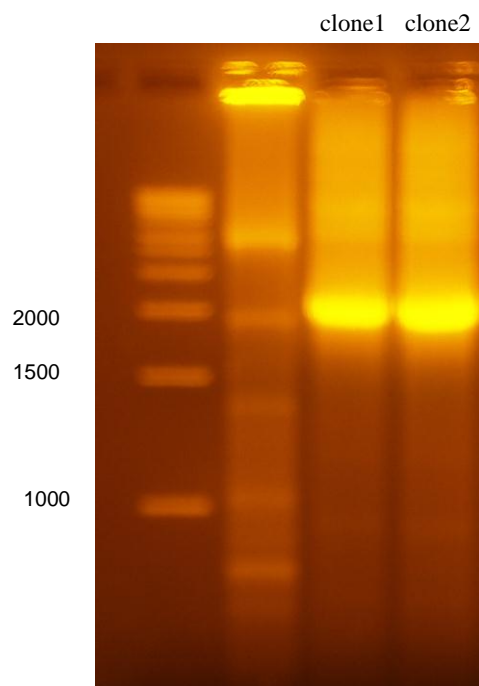
Figure 3.4: Subcloning of C-hRio2
pDEST 14-C-hRio2 plasmid was transformed into DH5α *E. coli* cells. The plasmids were isolated from DH5α cells and screened by PCR for the insert using the forward and reverse primers of PCR1 (Figure 3.2 A and B). The agarose gel, stained with ethidium bromide, shows the PCR product of two positive clones. The size of the expected product is 1656 base pairs.
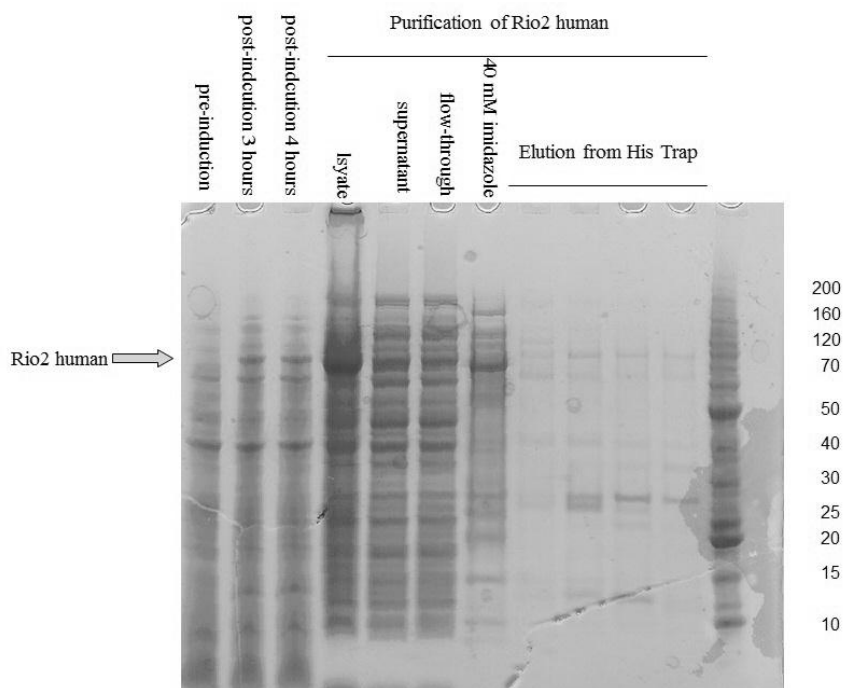
Figure 3.5: Induction of Rosetta Cells at 37°C and Purification of N-hRio2 by His Trap™

Rosetta cells expressing the plasmid for N-hRio2 were induced at 37°C for four hours with 1mM IPTG. A band for overexpressed N-hRio2 in the lysate after induction is indicated by an arrow. N-hRio2 was purified by His Trap™ Column. Fractions from the flow-through, 40mM imidazole wash, and elution were run on the SDS-PAGE gel, stained with Coomassie Blue.
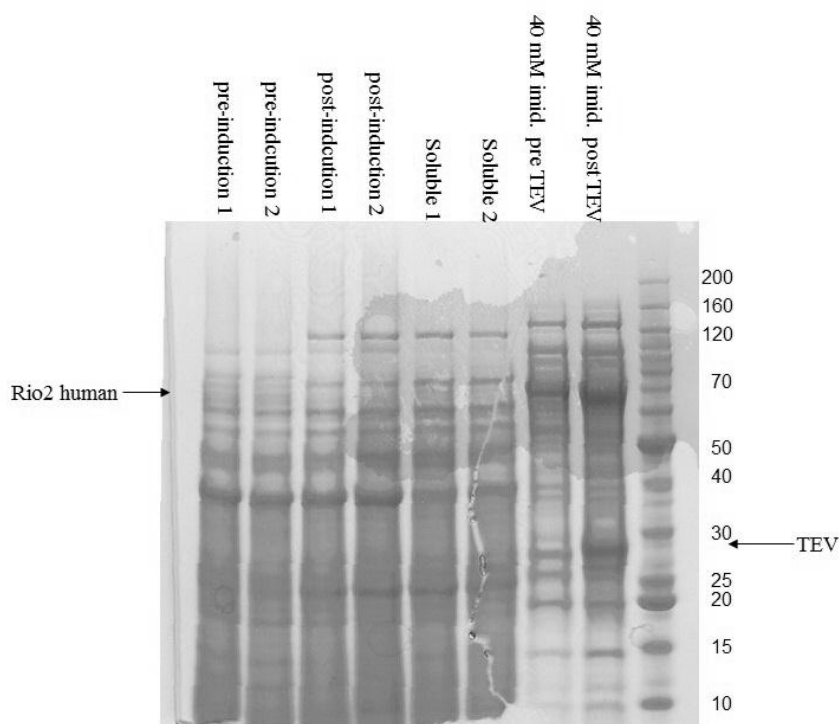
Figure 3.6: Induction of Rossetta Cells at 18°C

Rosetta cells expressing pDEST 527-N-hRio2 plasmid were induced at 18°C overnight with 1 mM IPTG. Two clones labeled 1 and 2 were tested for expression and solubility. Cells were lysed using 5x SDS loading dye before and after induction and samples were run on SDS-PAGE gel stained with Coomassie Blue. Cells were also lysed with 0.3x Bugbuster, and the supernatant was run on SDS-PAGE gel ("soluble"). The expected molecular weight of N-hRio2 is 66kDa. This SDS-PAGE gel also shows the 40 mM imidazole wash from Figure 3.5 purification of N-hRio2 on His Trap. This 40 mM imidazole fraction was treated with TEV protease.
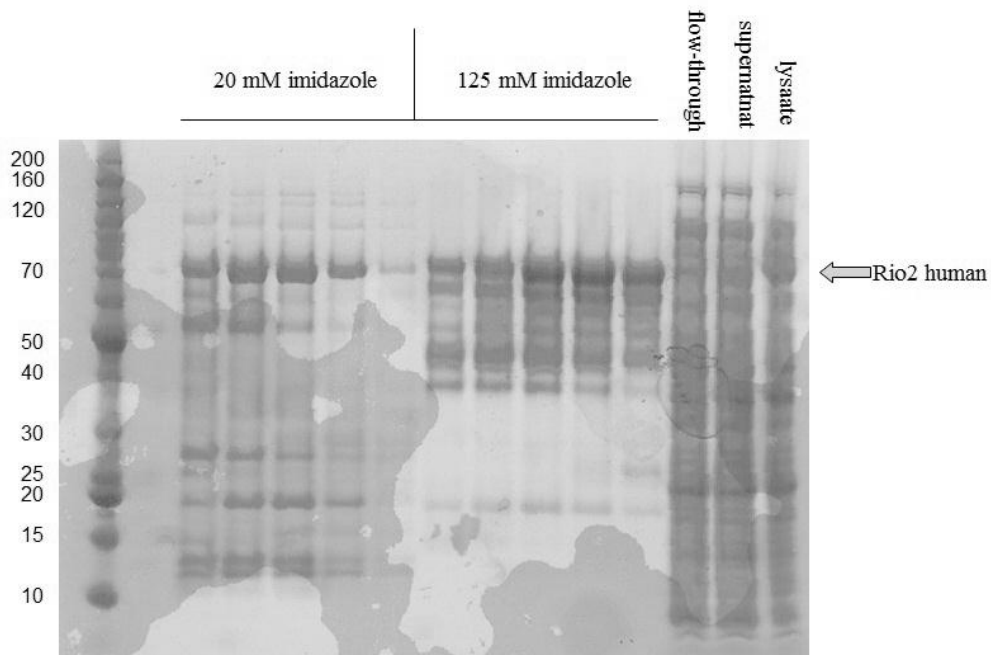
Figure 3.7: N-hRio2 Purified with 500 mM NaCl on His Trap^{TM}
N-hRio2 was solubilized with 500 mM NaCl in the lysis solution and then the supernatant was run on His Trap^{TM}. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the imidazole concentration at which the fraction eluted. Protein eluted both in the 20 mM wash and at 125 mM imidazole elution. N-hRio2 is expected to run at 66kDa and is seen as the most abundant band.
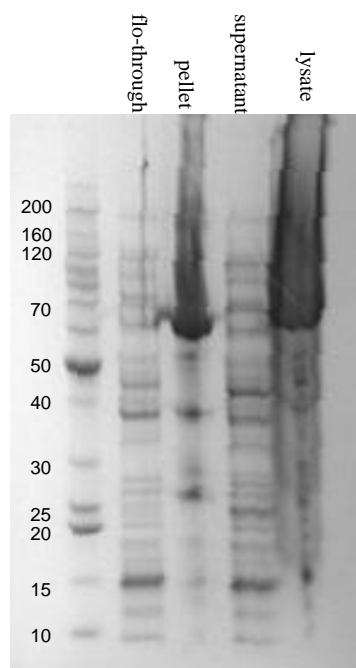
Figure 3.8: Δ71-C-hRio2 Purification

Rosetta cells containing the pDEST 14 plasmid for expression of Δ71-C-hRio2 were induced with 1 mM IPTG overnight at 18°C. The lysate and supernatant were run on the SDS-PAGE gel, stained with Coomassie Blue. The supernatant was loaded on a His Trap[TM] column. No protein eluted from the His Trap[TM] column (not run on SDS-PAGE gel). Δ71-C-hRio2 is expected to run at 55 kDa.
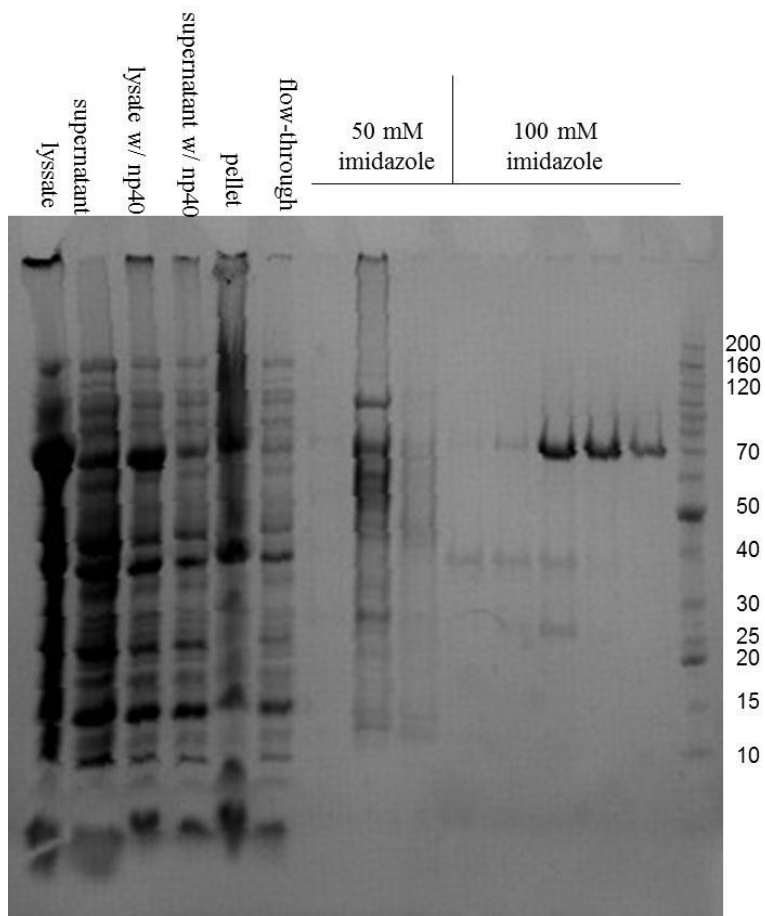
Figure 3.9: C-hRio2 Purified on His Trap™ with NP40 in the Buffers
The SDS-PAGE gel is stained with Coomassie Blue. Rosetta cells contained pDEST14 plasmid for expression of C-hRio2. Lysis of 2 L cultures of Rossetta cells was performed with and without 0.1% NP40 and 10% glycerol. The supernatant containing 0.1% NP40 was then loaded onto a His Trap™ column. The His Trap™ column was washed with 50 mM imidazole and C-Rio2p was eluted with 100 mM imidazole. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the concentration of imidazole each fraction eluted.
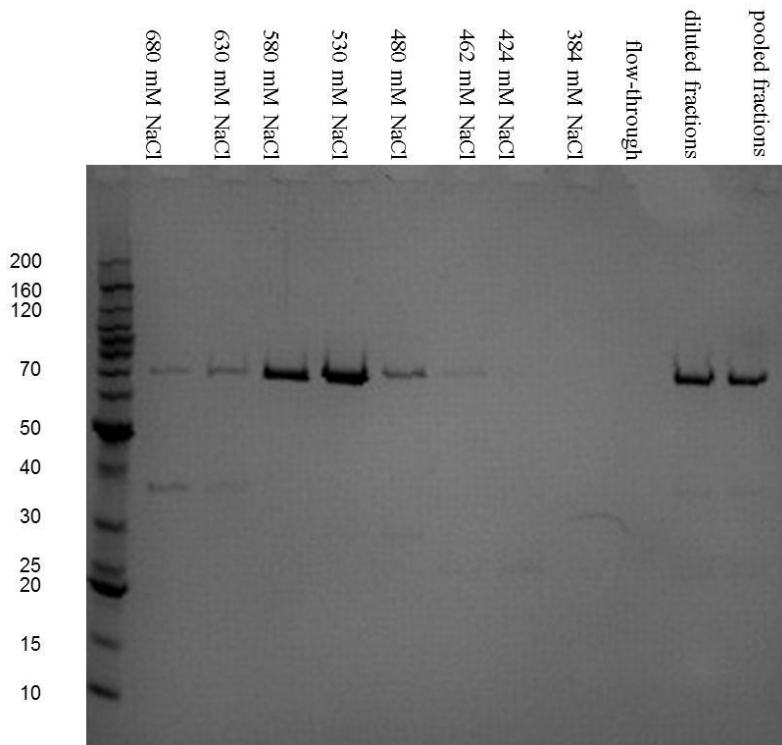
Figure 3.10: C-hRio2 Purified on S^{TM} Column
C-hRio2 was purified first on the His Trap^{TM} column. All buffers contained 0.1% NP40 and 10% glyerol. The fractions from the His Trap^{TM} column containing C-hRio2 were pooled together ("pooled fractions"), diluted to 200 mM NaCl ("diluted fractions") and run on a S^{TM} column. While C-Rio2p was bound to the S^{TM} column, the 0.1% NP40 was washed from the protein. C-hRio2 was then eluted. The SDS-PAGE gel, stained with Coomassie Blue, is marked with the salt concentration at which C-hRio2 eluted from the S^{TM} column.
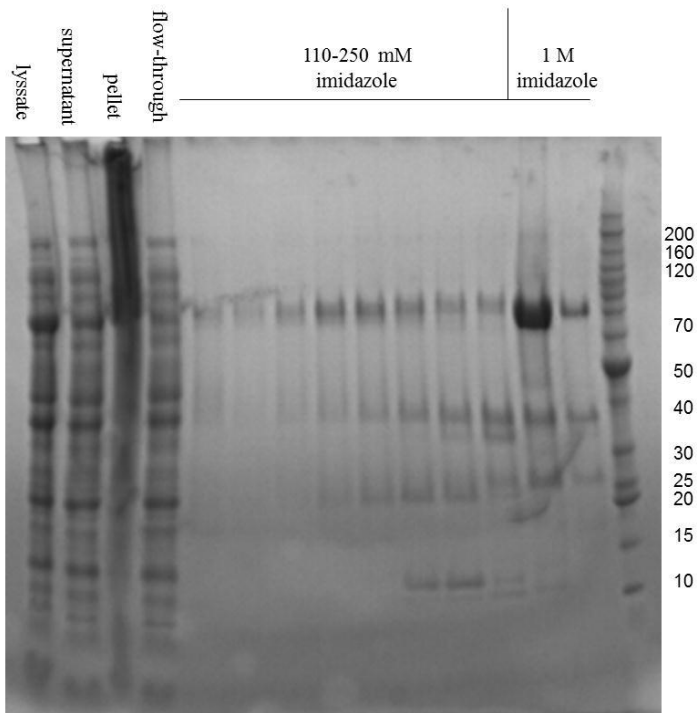
Figure 3.11: C-hRio2 Purified on His Trap^{TM} w 1 mM AMP in Buffers

1mM AMP was added to all buffers during the purification of C-hRio2. Lysate containing overexpressed C-hRio2 was passed over a His Trap^{TM} column. C-hRio2 elutes in dilute fractions in the gradient between 110-250 mM imidazole. The SDS-PAGE gel is stained with Coomassie Blue. C-hRio2 is expected to run at 63kDa.
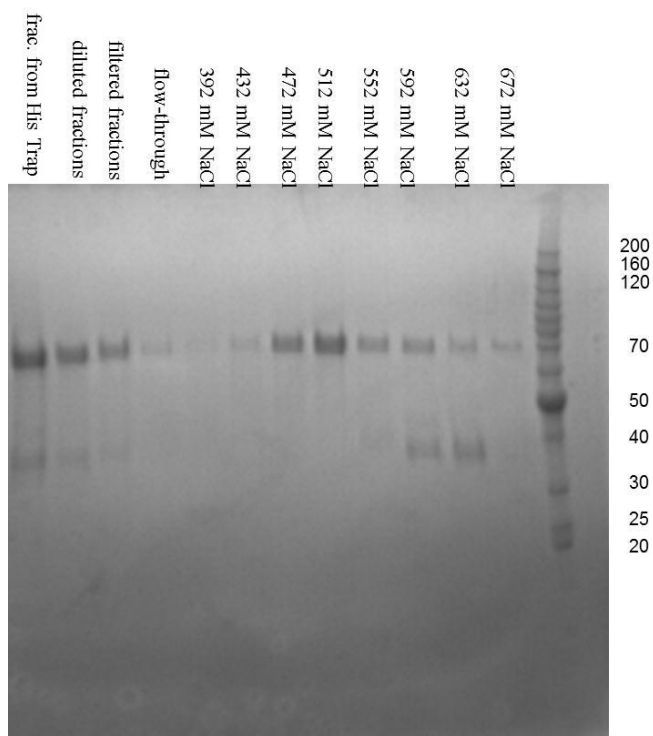
Figure 3.12: C-hRio2 Purified on S^{TM} Column with 1 mM AMP in Buffers
After purification on His Trap^{TM} column, C-hRio2 was loaded on a S^{TM} column and eluted over a gradient of 200 mM NaCl to 100 mM NaCl over 100 mL. 1 mM AMP was maintained in all buffers. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the salt concentrations the different fractions of C-hRio2 eluted at. C-hRio2 is expected to run at 63kDa.
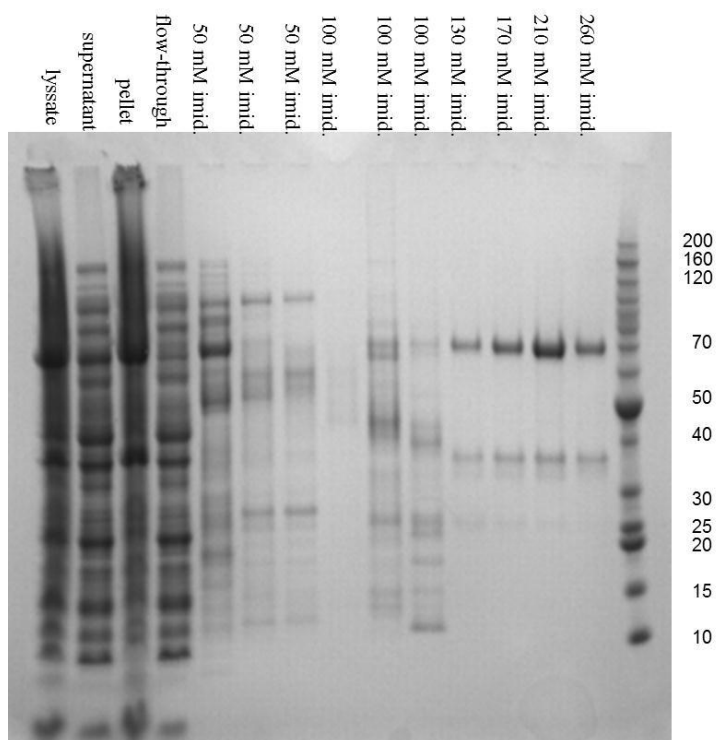
Figure 3.13: C-hRio2 Purified with on His Trap^{TM} with 1 mM ATP in Buffers
Lysate containing overexpressed C-hRio2 was loaded onto a His Trap^{TM} column. 1 mM ATP was included
in all buffers. The column was washed with 50 mM imidazole and 100 mM imidazole. C-hRio2 was
eluted at 130-260 mM imidazole. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the
imidazole concentration at which each fraction eluted from the His Trap^{TM} column. C-hRio2 is expected to
run at 63kDa.
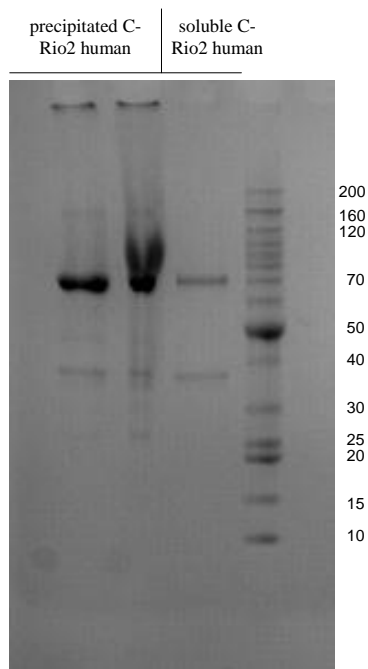
Figure 3.14: C-hRio2 Purified with 1 mM ATP on His Trap^{TM} Precipitates Overnight Supernatant containing overexpressed C-hRio2 was purified on His Trap^{TM}. 1mM ATP was present in all the buffers. The next day, the protein in the fractions from the His Trap^{TM} had precipitated. The precipitate was run on the SDS-PAGE gel stained with Coomassie Blue. The protein was then filtered and the soluble protein remaining was run on the SDS-PAGE gel.
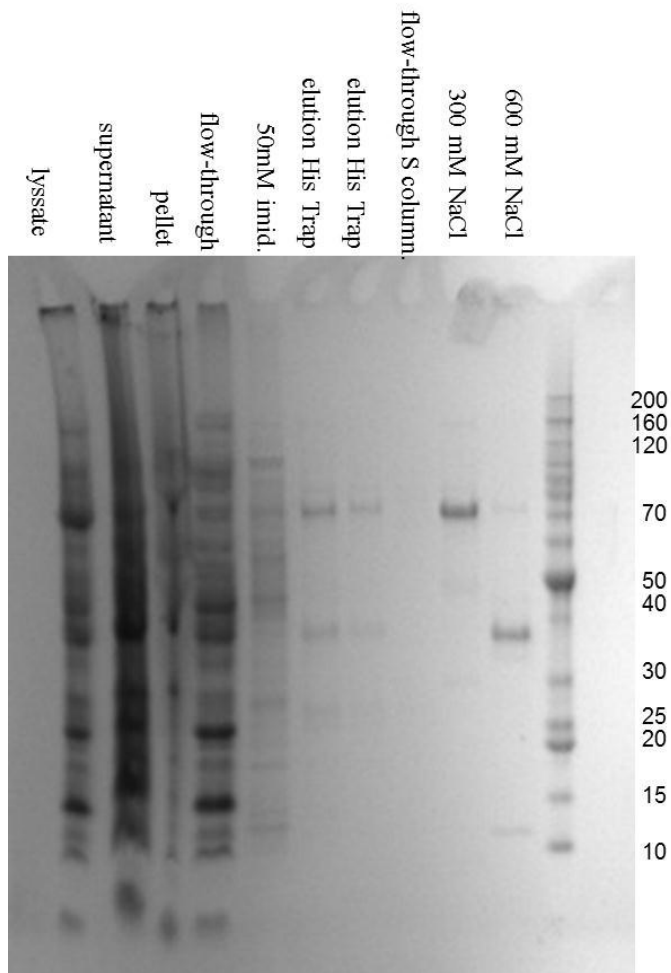
Figure 3.15: C-hRio2 Purified on His Trap^TM and S^TM Column with 1 mM ADP in Buffers

Supernatant containing overexpressed C-hRio2 was loaded onto a His Trap^TM column. The resin was removed from the column, and mixed with 500 mM imidazole elution buffer. The resin was allowed to settle and the elution buffer was removed from the top of the resin ("elution His Trap^TM). C-hRio2 was diluted to 0.3 mg/mL and then loaded onto a S^TM column. The pressure on the column rose and only one fraction of C-hRio2 eluting at 300 mM NaCl was able to be obtained from the column. The SDS-PAGE gel is stained with Coomassie Blue.

Construct	Residues Expressed	Expression	Induction Temp.	Pure?	MW	PI
N-hRio2	1-552	Yes	37°C	No	66,283	5.66
C-hRio2	1-552	Yes	18°C	Yes	63,283	5.66
Δ71-C-hRio2	72-552	No	18°C	N/A	55,297	5.18

Table 3.1: Summary Information of the Three Constructs Purified

Chapter 4: Purification and Crystallization of Rio2 from *Chaetomium thermophilum*

4.1: Introduction

Chapters 2 and 3 described the purification of Rio2p and hRio2 and the efforts to

obtain the structure of a eukaryotic Rio2. In the case of Rio2p the crystals were small,

and in the case of hRio2 poor solubility prevented crystal screening. In this chapter, I

discuss the purification and crystallization of Rio2 from a thermophilic fungus,

Chaetomium thermophilum (ctRio2). The construct shown in Figure 4.1, was obtained

from the laboratory of Dr. Ed Hurt in the Biochemistry Center, Heidelberg University,

Germany.   CtRio2, like Rio2p and hRio2, has an extended C-terminus (Figure 4.2).

There were issues in the purification and crystallization of ctRio2: it co-purified

with a close molecular weight degradation product; it nonspecifically bound to RNA in

the lysate; and crystals of ctRio2 were small and did not diffract well. These problems

were corrected by overloading the His Trap^{TM} column, which caused ctRio2 to elute

differently from the close molecular weight contaminant, using a Q^{TM} column to separate

RNA from ctRio2, and seeding with microcrystals to produce larger, better diffracting

crystals.

4.2: Methods

4.2.1: Purification of CtRio2

The optimized procedure for purification of ctRio2 (Figure 4.1) is outlined below.

The p7-ctRio2 vector (from the Hurt lab), transformed into *E. coli* Rosetta^{TM} DE3

(pLysS) cells, was used to inoculate an overnight 100 mL culture containing both 100

µg/ml ampicillin and 34 µg/ml chloroamphenicol at 37ºC.   The overnight culture was

diluted 1/100 into four 4 L baffle flasks containing 1 L of LB with ampicillin and

chloroamphenicol. The four flasks were then incubated at 37°C with shaking at 250 rpm.

At O.D.600 of 0.5, the flasks were moved to an 18°C shaking incubator to cool the

cultures to 18°C. At 0.6-0.8 O.D.600 the culture was induced with 1 mM IPTG and

incubated overnight.

The cells were centrifuged at 5,000 rpm for thirty minutes. The pellet was about

4-5 grams per L of cell culture. The cell pellets were frozen at -80°C. The frozen pellet

was resuspended in 10 mL of Buffer A (200 mM NaCl , 10% glycerol, 50 mM Tris, pH

8.0, 2.5 mM $MgCl_2$, 0.05% BME) per gram of cell pellet. 5 mg of DNAse, 1 mg of

RNase A, and 20 µl of RNase T1 (100,000 U/mL) per liter of cell pellet were added to

the solution. One Complete, EDTA free, protease inhibitor cocktail tablet (Roche) was

added to the cell suspension. 0.3X Bugbuster^{TM} and 0.5 mg of lysozyme/mL cell

suspension were added to lyse the cells. The cells were stirred at 4°C for thirty minutes.

The lysate was centrifuged at 16,000 rpm for one hour to pellet the cell debris. The

supernatant containing the soluble protein was filtered through a 0.25 micron filter.

The supernatant was passed over a GE Healthcare 5mL His Trap^{TM} column that

was pre-equilibrated with Buffer A. Buffer B (200 mM NaCl, 10% glycerol, 50 mM

Tris, pH 8.0, 2.5 mM $MgCl_2$, 1 M imidazole, 0.05% BME) was used to elute the protein.

3% Buffer B was used to elute CtRio2 from the column and 5 mL fractions were

collected. The column was washed with 10% Buffer B and then 100% Buffer B.

A SDS-PAGE gel was run of the fractions collected during the 3% Buffer B

elution. The concentrated and pure fractions were pooled together for TEV cleavage.

The protein concentration of the combined fractions was determined by A 280. A 280 of

one was assumed to be a 1 mg/mL concentration of ctRio2. CtRio2 was concentrated to

5 mg/mL on a Millipore Centrifugal Filter Unit at 4°C. Three mg of TEV protease was added to the ctRio2, which was then dialyzed overnight in Buffer A at 4°C.

The next day, the His Trap^{TM} column was equilibrated with Buffer A. CtRio2 was passed over the column and the flow-through was collected and assumed to be cleaved protein. The column was also eluted with 10% Buffer B to retrieve the uncleaved protein. The column was washed with 100% Buffer B.

The solution containing ctRio2 was diluted, the final concentrations of solution components were 200 mM NaCl, 50 mM Tris pH 8.0, 2.5 mM $MgCl_2$, 10% glycerol, 0.05% BME. Two GE Healthcare Q^{TM} columns were connected and were pre-equilibrated with the loading buffer, Buffer C, (200 mM NaCl, 50 mM Tris pH 8.0, 2.5 mM $MgCl_2$, 10% glycerol, 0.05% BME). Buffer D, the high salt buffer, was 1 M NaCl, 50 mM Tris pH 8.0, 2.5 mM $MgCl_2$, 10% glycerol, 0.05% BME. The ctRio2 was loaded onto the Q^{TM} column. A gradient of the two buffers was run from 150 mM NaCl to 1 M NaCl over 100 mL. CtRio2 eluted at 300 mM NaCl. The purest and highest concentration fractions were combined and dialyzed against 1 L of Buffer E (10 mM Tris pH 8.0, 150 mM NaCl, 2.5 mM $MgCl_2$, and 0.5% BME). The ctRio2 was concentrated to 10 mg/mL as determined by A280.

Selenomethionine substituted protein was grown using M9 SeMET High-Yield Growth Media Kit (Shanghai Medicilon Inc.) following the instructions provided. Selenomethionine substituted ctRio2 was purified same as described above.

4.2.2: Crystallization of CtRio2

Crystal screens used were Qiagen's PEG Suite, Hampton's Index Screen, Emerald Biosystems Wizard I, II, and III Screens, Hampton Research's Natrix Screen and

Qiagen's Cryos Suite. In these screens sitting drops were set up in 0.6 µl drops of well-to-protein ratios of 1:2, 1:1, and 2:1. A total of 384 conditions were tested in three different protein-to-well solution ratios. The screens were incubated at 20°C.

4.3: Results

4.3.1: Overloading the His Trap^{TM} Column Can Eliminate Degradation Products

The construct obtained from Dr. Hurt contained an N-terminal 6x His tag and a TEV cleavage site to remove the His tag from the final product (Figure 4.1). After transformation in Rosetta^{TM} cells, a robust expression of recombinant ctRio2 was obtained. Two purification strategies were used to purify ctRio2 and eliminate degradation products, and one of these methods was shown to have superior performance over the other.

In the first method, two His Trap^{TM} columns were used in tandem to avoid exceeding the binding capacity of the columns. While ctRio2 bound well to the column and eluted at 100 mM imidazole concentration, it copurified with many contaminants or degradation products (Figure 4.3). The His tag was cleaved with TEV protease and then ctRio2 was passed over the His Trap^{TM} column a second time to remove protein contaminants not cleaved by TEV. However, the first method (two columns) could not remove all contaminants because ctRio2 copurified with a degradation product slightly smaller than the full-length ctRio2 (Figure 4.4).

In the second method, one His Trap^{TM} column was used in order to overload the column and it was found to successfully purify ctRio2. Overloading the column changed the elution profile of ctRio2 so that it eluted much earlier (at 30 mM imidazole) and in more pure and concentrated fractions (Figure 4.5). Also the low molecular weight

degradation product that persisted with ctRio2 in method one continued to elute at 100 mM imidazole and was separated from full length ctRio2 which eluted at 30 mM imidazole. After TEV cleavage, ctRio2 was passed over the single His Trap™ column. The ctRio2 eluted in the flow-through while the contaminants remained on the column (Figure 4.6).

In conclusion, protein overloading allowed the elimination of protein contaminants including the degradation product. Since ctRio2 has higher affinity for the nickel column and binding sites on the nickel column are limited, ctRio2 will compete with protein contaminants for binding to the column. The result is that major protein contaminants will not bind to the column. In addition, since the ratio of ctRio2 to nickel atoms increased in method two, the binding affinity for ctRio2 decreased. As a result, the elution profile of ctRio2 changed in method two such that ctRio2 eluted earlier (at 30 mM imidazole) and was separated from the tighter binding close molecular weight degradation product that eluted at 100 mM imidazole.

4.3.2: Improving TEV Cleavage Increased Yield of CtRio2

TEV cleavage was improved in the purification procedure. Initially 1mg of TEV was used to cleave the large quantity of ctRio2 eluting from the His Trap™. The concentration of ctRio2 was 1 mg/ml during cleavage and cleavage was performed overnight at 4°C. Under these conditions, the TEV cleavage was not complete, as half the ctRio2 remained uncleaved and did not elute in the flow-through. The conclusion was that, the His tag was difficult to cleave from the N-terminus. The concentration of ctRio2 was then increased to 5 mg/ml and TEV protease was increased to 3 mg. With these changes, most of ctRio2 was found to be cleaved and eluted in the flow-through.

4.3.3: CtRio2 Binds to RNA Nonspecifically

An observation that helped to purify ctRio2 is that ctRio2 binds to RNA despite RNases being present in the lysate. The last step of purification of ctRio2 involved passing the protein over an anion exchange Q^{TM} column (Figure 4.7). In the chromatogram of the elution of the protein from the Q^{TM} column two peaks appear, one for ctRio2 at a lower concentration of 300 mM salt, and one for RNA at a higher concentration of 500 mM salt (Figure 4.7A). The second peak is most likely RNA because it has a high affinity for the Q^{TM} column—its A260/A280 absorbance ratio is over 2 and it shows no protein bands on the gel.

The Q^{TM} column is an important step for purification. If RNA continues to remain bound to ctRio2, it would prevent ctRio2 from forming crystals. The RNA is bound nonspecifically and consists of a variety of different RNAs. Any heterogeneity in the protein sample precludes the protein from crystallization.

4.3.4: Crystal Screens

In order to maximize the number of hits ctRio2 was screened at 10 mg/mL and at 20 mg/mL concentrations. CtRio2 was also screened with and without 20 mM ADP. The crystals began to appear after one month and they did not grow large enough for diffraction. However, the number of positive hits was large in both the PEG Screen and the Index Screen (Figures 4.8 and 4.9). Many factors could account for the delayed crystallization. Nucleation of the crystal is usually the rate limiting step, after which the crystal usually grows quite rapidly. The protein might undergo degradation to a more stable fragment before it crystallizes.

I attempted to reproduce positive hits by the hanging drop method. In the hanging drop method, the protein was mixed with the well solution in the ratios from the crystal screen in a 4 µl drop on a cover-slip. That coverslip was then affixed over a well containing 1 mL of the precipitant solution. Vapor from the drop on the coverslip diffused into the well solution where the concentration of the precipitant was greater. This allowed the protein on the cover-slip to become slowly concentrated, resulting in crystal formation.

Optimization screens were performed in 6 by 4, 24-well format. The concentration of precipitant was varied across the six wells and ranged from smaller to larger concentrations of precipitant than the screen solution. The pH was varied along each of the rows so that four different pHs were tested.

After months of incubation time, the hanging drop was ineffectual at reproducing the crystals formed with or without ADP and led me to try microseeding. Crystals were broken into tiny seeds and added to a new 4 µl protein/precipitant droplet. The microcrystals obtained in the Index Screen solution number 45 (0.1 M Tris pH 8.5, 25% PEG 3350) were used for the seed solution.

The 24-well hanging drop screen solutions were set up as before, as were the 4 µl protein/precipitant droplets. The protein droplets were allowed to incubate with the well solution for two days. Then 2 µl of the protein droplet was mixed with the microcrystals obtained from the sitting drop screen. The microcrystals in this solution were pipetted up and down to break the microcrystals into smaller seed crystals. 0.2 µl of that seed solution was used to seed 8 of the wells from the hanging drop optimization screen. The screen was then incubated at 20°C.

After one month with seeding, larger crystals appeared in the wells (Figure 4.10). The precipitant concentration in these wells was less than that from the original crystal screen. Diffraction data were collected with a few of the crystals at the University of MD source x-ray beam. The crystals were soaked in different amounts of glycerol and ethylene glycol to freeze them. 25% ethylene glycol was added to prevent the formation of ice rings in the diffraction pattern. The crystals diffracted to 3.6 Å. For better resolution data, the crystals were frozen and shipped to the synchrotron at Argonne, Illinois. Data were collected at the NE-CAT beamline on several crystals that diffracted at an average of 2.5 Å.

Later crystals were grown with 2 mM ATP. 10 mg/mL ctRio2 was incubated with 2 mM ATP (30x dilution) and 2.5 mM $MgCl_2$ (included in ctRio2 buffer solution) for one hour. The crystals were grown in a sitting drop screen. 0.4 µL of protein solution were mixed with 0.2 µL of precipitant solution (0.2 M Magnesium chloride hexahydrate, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350) and placed in a sitting drop next to a well containing 200 µL of precipitant solution. Crystals grew large enough for X-ray diffraction studies after 4-5 days at 20°C. Data were collected at the NE-CAT beamline on several crystals that diffracted at an average of 2.2 Å.

4.4: Summary

I was able to achieve greater than 95% purity of ctRio2 on His Trap^{TM} and Q^{TM} columns. Success was achieved by overloading the His Trap^{TM} column to remove major protein contaminants and degradation products. TEV cleavage was optimized to cleave most of the ctRio2 in solution. Passing of ctRio2 over the His Trap^{TM} a second time after TEV cleavage, removed the remaining contaminants. CtRio2 was next purified on a Q^{TM}

After one month with seeding, larger crystals appeared in the wells (Figure 4.10). The precipitant concentration in these wells was less than that from the original crystal screen. Diffraction data were collected with a few of the crystals at the University of MD source x-ray beam. The crystals were soaked in different amounts of glycerol and ethylene glycol to freeze them. 25% ethylene glycol was added to prevent the formation of ice rings in the diffraction pattern. The crystals diffracted to 3.6 Å. For better resolution data, the crystals were frozen and shipped to the synchrotron at Argonne, Illinois. Data were collected at the NE-CAT beamline on several crystals that diffracted at an average of 2.5 Å.

Later crystals were grown with 2 mM ATP. 10 mg/mL ctRio2 was incubated with 2 mM ATP (30x dilution) and 2.5 mM $MgCl_2$ (included in ctRio2 buffer solution) for one hour. The crystals were grown in a sitting drop screen. 0.4 µL of protein solution were mixed with 0.2 µL of precipitant solution (0.2 M Magnesium chloride hexahydrate, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350) and placed in a sitting drop next to a well containing 200 µL of precipitant solution. Crystals grew large enough for X-ray diffraction studies after 4-5 days at 20°C. Data were collected at the NE-CAT beamline on several crystals that diffracted at an average of 2.2 Å.

4.4: Summary

I was able to achieve greater than 95% purity of ctRio2 on His Trap™ and Q™ columns. Success was achieved by overloading the His Trap™ column to remove major protein contaminants and degradation products. TEV cleavage was optimized to cleave most of the ctRio2 in solution. Passing of ctRio2 over the His Trap™ a second time after TEV cleavage, removed the remaining contaminants. CtRio2 was next purified on a Q™

column to remove non-specifically bound RNA that could impede its crystallization.

Following this modified procedure, crystals of ctRio2 were obtained with ADP and with

ATP that diffracted at 2.5Å and 2.2Å respectively at the synchrotron, and finally a

structure solution was obtained. It will be discussed in Chapter 5.

N'	6x His Tag	TEV Site	ctRio2	C'
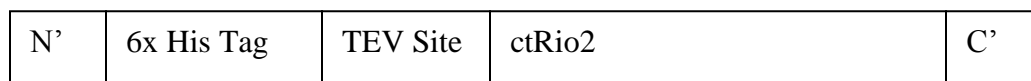
Figure 4.1: Schematic of CtRio2 Construct
CtRio2 construct was designed with an N-terminal 6x His tag followed by a TEV cleavage site.

```
                 10        20        30        40        50        60
                  |         |         |         |         |         |
AFrio2   MNIAELYG-KMGKHSWRIMDAIFKNLWDYEYVPLQLISSHARIGEEK---ARNILKYLSD
ctrio2   MKLDTRAMRHLTAEDWRVLTAVEMGSKNHEIVPTPLIEKIARLRGGSSGVHKSIATLAKA
         *::          ::   ..**::  *:    ::* **    **.. **:   .    ::.* .
                 70        80        90       100   P-loop       120
                  |         |         |         |  xxxxxx |         |
AFrio2   LRVVQNRQKDYEGSTFTFIGLSLYSLHRLVRSGKVDAIGKLMGEGKESAVFNCYSEKFGE
ctrio2   GLIARMKEAKYDGYRLTYGGLDYLALHTHAARKDVYSVGSRIGVGKESDIMIVADEKGKQ
         *:.: ::  .*:*   :*: **.  :**   .    *  ::*. :* **** ::    .**  :
                130       140       150       160       170       180
                  |         |         |         |         |         |
AFrio2   CVVKFHKVGHTSFKKVKEKRDYG----DLHFSVLAIRSARNEFRALQKL--QGLAVPKVY
ctrio2   KVLKIHRLGRISFRTVKANRDYLRNRSTGSWMYLSRLAAIKEFAFMKALYEEGFPVPEPI
          *:*:*::*: **:.** :***      :    *:   :* :** :: *   :*:.**:
                190       200       210       220   Catalytic  240
                  |         |         |         |      xxxxxx  |
AFrio2   AWEGNAVLMELIDAKELYRVR-VENPDEVLDMILEEVAKFYHRGIVHGDLSQYNVLVSEE
ctrio2   AQSRHTIVMSLVDALPMRQVSSVPDPASLYADLIALILRLAKHGLIHGDFNEFNILIREE
         *  . :::::*.*:** : :*   * :* .:    ::   : ::  :: ::*::***.::*:*: **
                250   Metal         270       280       290       300
                  |  xxxxx                        |         |         |
AFrio2   G-----------IWIIDFPQSVEVGEEGWREILERDVRNIITYFSR--------------
ctrio2   KDAEDPSSITLTPIIIDFPQMVSMDHPNAEMYFDRDVQCIKRFFERRFHFVSTTPGPFYK
                    ****** *.:.. .   ::*** : *   :*.*
                310       320       330       340       350       360
                  |         |         |         |         |         |
AFrio2   ----------------------TYRTEKDINSAIDRILQE--------------------
ctrio2   DAKKTVGKDGAKRLDAALEASGFTKKMAKDLEAAIREQQESRNDEEDSDDYEEDSDKEKA
                               * :  **::.** .  :.
                370       380       390
                  |         |         |
AFrio2   ----------------------------------
ctrio2   SNEDTDVDDGRLQVIGDHVQLGTEGDLKKLTINDGT
```

Figure 4.2: Alignment between AfRio2 and CtRio2
Alignment was performed with the program CLUSTALW using sequences from *A. fuligidus* (AFRio2) and
ctRio2. Identical, highly similar, and weakly similar residues are indicated by red, green, and blue letters
respectively. The wing helix domain of AfRio2 is highlighted in yellow, the N-terminal lobe in green and
the C-terminal lobe in blue. The C-terminal extension of ctRio2 is in the boxed region encompassing
residues 307-396.

A.

B.

ctRio2

degradation

Figure 4.3: Purification of CtRio2 on His Trap^TM Column – Method One
Supernatant containing overexpressed ctRio2 was passed over two connected His Trap^TM columns. The SDS-PAGE gels, stained with Coomassie Blue, are labeled with the imidazole concentration at which the fraction eluted. A. CtRio2 begins to elute from the His Trap^TM columns at 80 mM imidazole. B. A close molecular weight degradation product co-elutes with ctRio2.
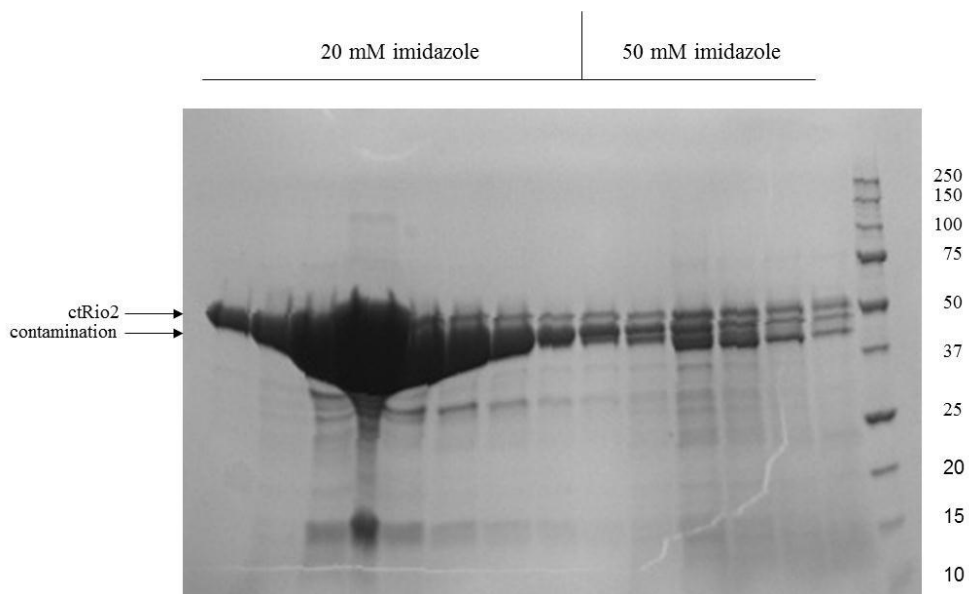
Figure 4.4: TEV Cleavage of CtRio2 – Method One

CtRio2 was cleaved with TEV protease and then passed over two connected His Trap™ columns a second time. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the imidazole concentration at which the fraction eluted. CtRio2 elutes at 20 mM imidazole and co-purifies with a degradation product of slightly lower molecular weight.

A.
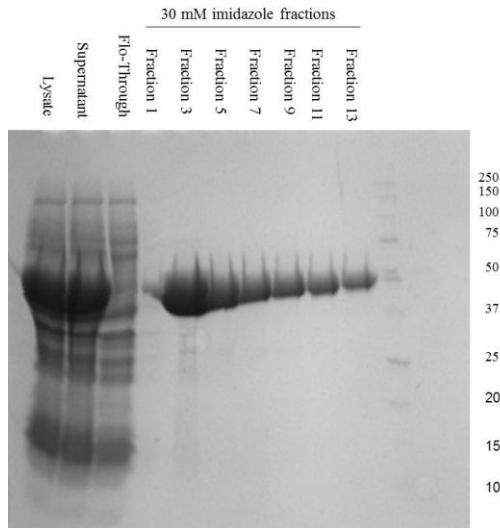
B.
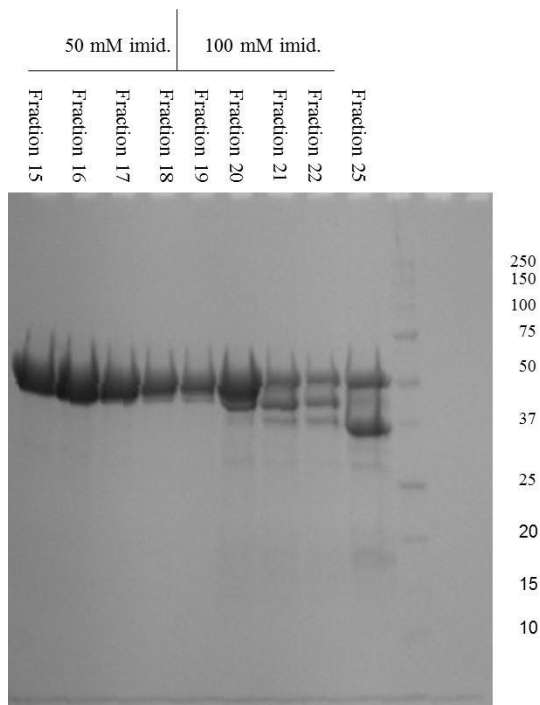
Figure 4.5: Purification of CtRio2 on His Trap^{TM} Column – Method Two
Supernatant containing overexpressed ctRio2 was passed over one His Trap^{TM} column. The SDS-PAGE gels, stained with Coomassie Blue, are labeled with the imidazole concentration and fraction number at which the fraction eluted. A. CtRio2 begins to elute from the His Trap^{TM} column at 30 mM imidazole. B. A close molecular weight degradation product elutes at 100mM imidazole.
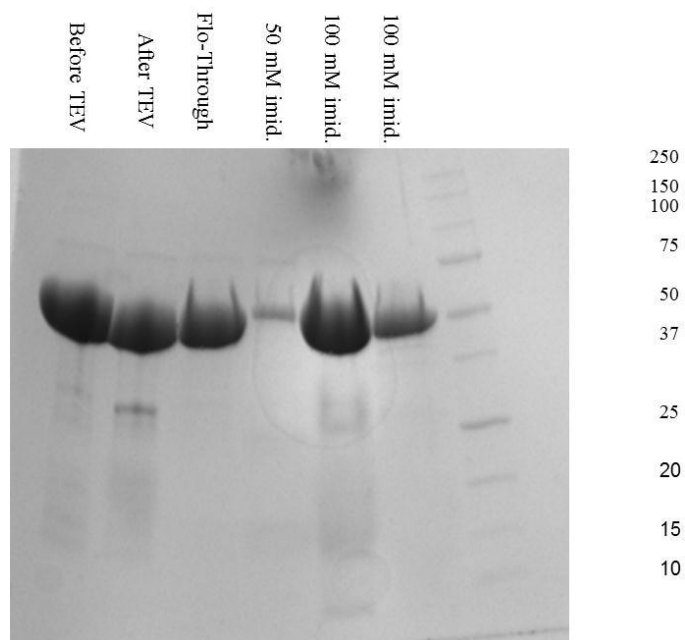
Figure 4.6: TEV Cleavage of CtRio2 – Method Two
CtRio2 was cleaved with TEV protease and then passed over a His Trap™ column a second time. The SDS-
PAGE gel, stained with Coomassie Blue, is loaded with a sample of ctRio2 before TEV cleavage ("Before TEV") and a sample of ctRio2 after TEV cleavage ("After TEV") before loading on the column. The flow-through from the His Trap™ column contains the cleaved ctRio2 and the fractions at 100 mM imidazole contain uncleaved ctRio2.

A.
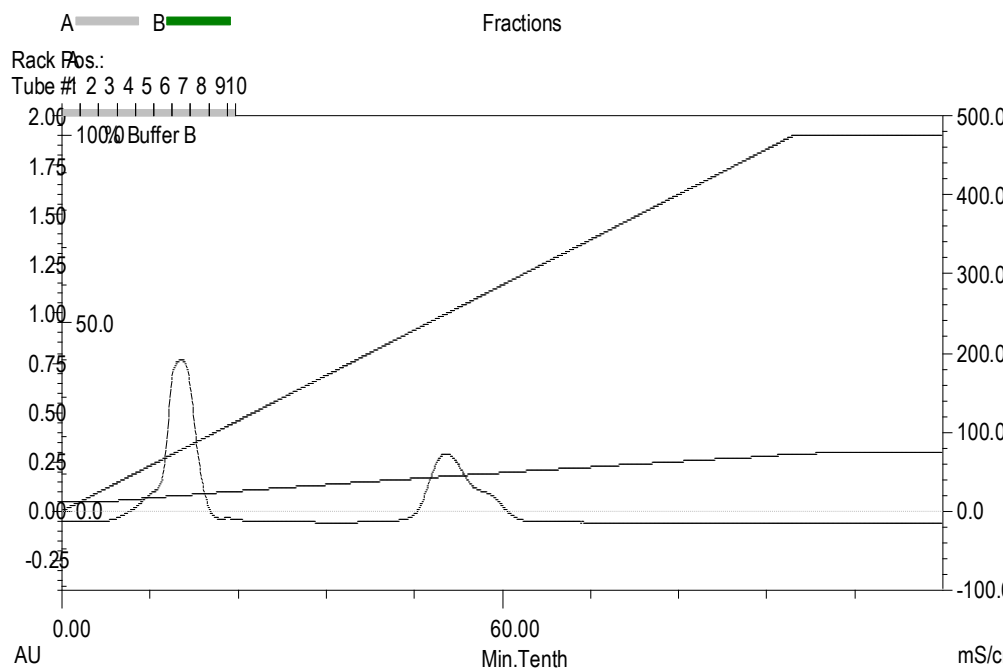
B.
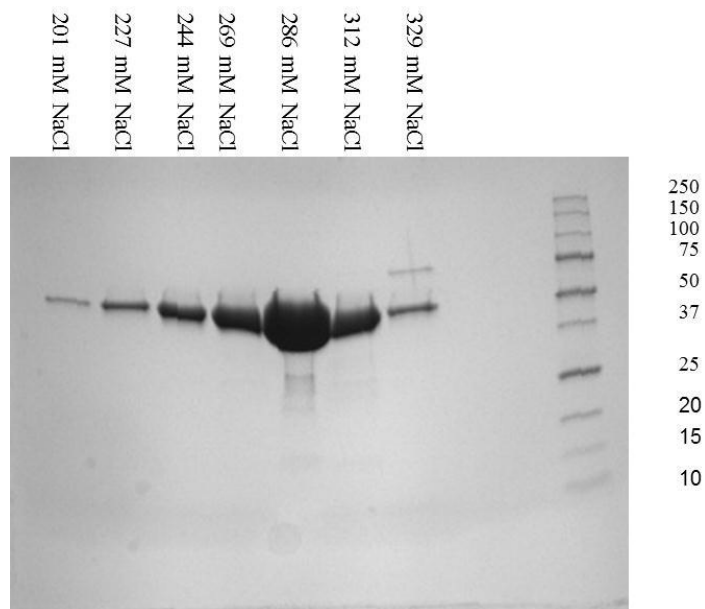
Figure 4.7: CtRio2 Purified on Q^{TM} Column.
CtRio2 was passed over a Q^{TM} column and then eluted over a 150 mM to 1 M NaCl gradient. A. The chromatogram of elution from the Q^{TM} column. Peak one is ctRio2. B. The SDS-PAGE gel, stained with Coomassie Blue, is labeled with the salt concentration at which the fractions eluted.

A.
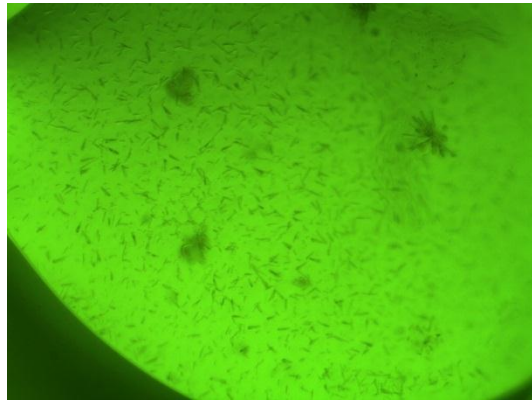
B.
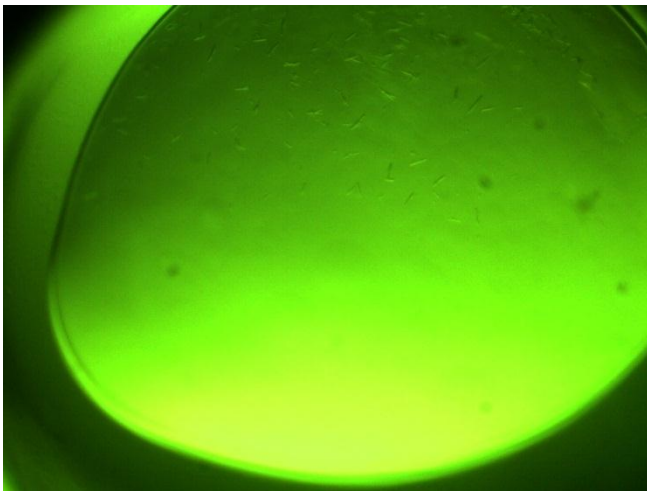
C.
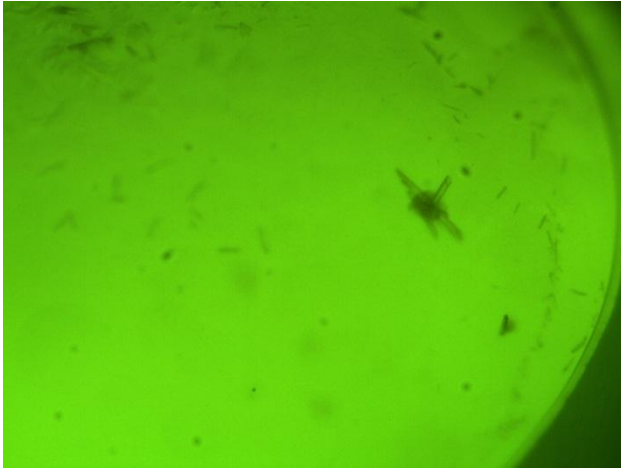
Figure 4.8: Apo CtRio2 Crystals
The crystals are formed with apo ctRio2 at a concentration of 10 mg/mL. Crystals are obtained from A. PEG Screen condition number 38 (0.1 M Sodium HEPES pH 7.5, 25 %(w/v) PEG 4000). B. PEG Screen condition number 35(0.1 M MES pH 6.5, 20 %(w/v) PEG 10000). C. Index Screen condition number 35 (1 M Ammonium sulfate, 0.1 M Hepes pH 7.0, 0.5 %.(w/v) PEG 8,000).

A.
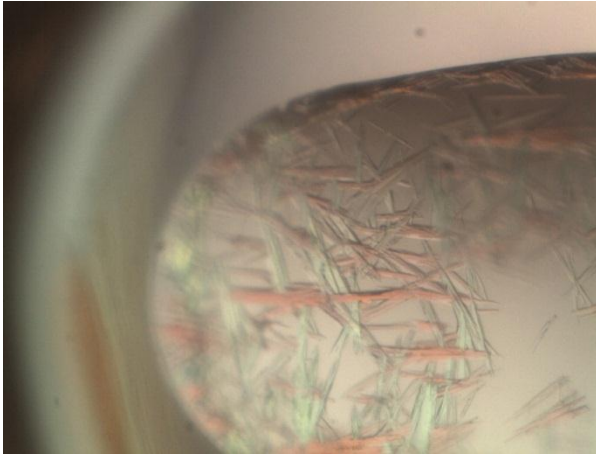
B.
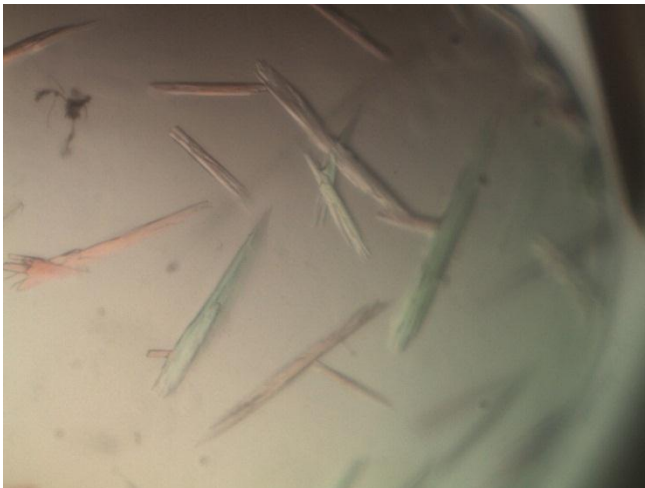
Figure 4.9: CtRio2 Crystals Grown with ADP
These crystals were produced with 20mM ADP and 20 mg/mL ctRio2. A. PEG Screen condition number 35, 2:1 protein to well solution (0.1 M MES pH 6.5, 20 %(w/v) PEG 10000). B. PEG Screen condition number 36, 2:1 protein to well solution (0.1 M MES pH 6.5, 15 %(w/v) PEG 20000).

A.

.

B.

Figure 4.10: CtRio2 Crystals Formed in Hanging Drop
The crystals were produced with 20 mM ADP and 10 mg/mL protein in hanging drop by microseeding from crystallization screen. A.Condition was 2:1 protein to well solution in 0.1 M Tris, pH 8.5, 16% PEG 3350. B. Condition was 2:1 protein to well solution in 0.1 M Tris, pH 8.5, 13% PEG 3350.

Chapter 5: Structure of the Eukaryotic Rio2 Kinase from CtRio2

5:1: Introduction

In this last chapter, I discuss acquiring x-ray diffraction data and analysis of the structural features of ctRio2. This is the first structure of a eukaryotic Rio2 kinase, which is significantly extended compared to the archaeal version for which we currently have structural information. Therefore, I performed structural analysis to determine the similarities and differences between ctRio2 and afRio2, and decipher the additional functionality provided by the extended sequences in eukaryotic Rio2.

5.2: Procedures

5.2.1: Obtaining the Crystal Structures of CtRio2

Both apo ctRio2 crystals and ctRio2 with ATP crystals were flash frozen in mother liquor containing 25% ethylene glycol. Diffraction data were collected at 100K with a MAR300 CCD detector at the NE-CAT beamline, 22-ID located at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA). The data were integrated and merged using HKL2000 (*46*). Tables 5.1 and 5.2 contain the relevant statistics of these two data sets.

The structure of apo ctRio2 was solved by molecular replacement using the known structure of afRio2 with ADP and Mn^{2+} bound. An all alanine model of afRio2 was used for molecular replacement. The program used for molecular replacement was AutoMR from the Phaser crystallography software suite (*47*). Refinement was done with phenix.refine in Phaser (*48-50*). R_{free} was monitored by using 8% of the reflections as a test set. The crystallographic refinement statistics are shown in Table 5.1. The figures that depict the structures of ctRio2 were created using PYMOL (*51*) .

The structure of ctRio2 with ATP was solved by molecular replacement using the partially refined structure of apo structure of Rio2 as a search model, also using AutoMR within Phaser (47). Autobuild (also within Phaser) was used to perform automatic model building using a map produced using the phases obtained from molecular replacement. The ligand was placed in the model and the structures were finalized by multiple rounds of rebuilding using COOT and refinement with phenix.refine (48-50). R_{free} was monitored by using 5% of the reflections as a test set. The refinement statistics are shown in Table 5.2.

5.2.2: Testing the Kinase Activity of CtRio2

A solution was prepared of 5 mM ATP with 10 mM $MgCl_2$ and 50 mM Tris pH 8.0. 1 in 10 serial dilutions of this stock solution were made to give 0.5 mM ATP, 0.05 mM ATP, and 0.005 M ATP. 1 µl or 5 µl of 10 mg/ml ctRio2 was mixed with 10 µl 2x Buffer (50 mM Tris, pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$). The total volume was brought to 17 µl with water. 1 µl of a 1/10 dilution of the γ-P^{32}-ATP (10 µCi per µl) was added to 2 µl of the varying concentrations of ATP. Then the 3 µl of each ATP solution was added to each reaction to give final ATP concentrations of 0.5 mM ATP, 0.05 mM ATP, 0.005 mM ATP, and 0.0005 mM ATP. The reaction was incubated at 37°C for 1 hour and then quenched with 8 µl of 5x loading dye. The samples were boiled for five minutes and then loaded on a SDS-PAGE gel. The gels were dried and then exposed overnight on a phosphor imaging plate.

5.3: Results

5.3.1: Overview of the Structure of CtRio2

Crystals of the apo structure and the ATP complex were isomorphous and belonged to the space group H3, with one molecule per asymmetric unit. Diffraction of the ATP complex and apo structure, extending to resolutions 2.2 Å and 2.5 Å respectively, was measured on a synchrotron source (NE-CAT, Argonne National Laboratories Advanced Photon Source).

Apo ctRio2 was crystallized in the presence of 20 mM ADP and 2.5 mM MgCl$_2$. However, no ADP was observed in the active site of the enzyme. CtRio2 was also crystallized with 2 mM ATP and 2.5 mM MgCl$_2$. Details of the crystallization methods were described in Chapter 4.

The core domain of ctRio2 mimicked the structure of afRio2 except for an extended C-terminal domain. Although some of the C-terminus was observed, electron density was not observed for most of it. There is no electron density for fifty-five of the extended C-terminal eighty-nine residues. It might be that the C-terminal domain is degraded or that it is highly flexible. Another hypothesis could be, since the apo structure was solved by molecular replacement with afRio2, it lacks phase information of the extended C-termini not seen in the afRio2 structure. As such the weak electron density of a more flexible C-terminus might not show up in the difference Fourier map. Phase information for these fifty-five residues might have to be obtained by heavy metal soaks of the native crystals or by obtaining selenomethionine substituted crystals. The problem with obtaining this information was that large native crystals needed for heavy metal soaks were not reproduced even by seeding. Also it was found that selenomethionine-substituted protein does not crystallize as efficiently as native protein and it has many fewer hits on the crystal screens. During crystal optimization in the

hanging drop format, the selenomethionine-substituted protein produced only small crystals with seeding. The protein that was visible in the crystal structure is analyzed below.

The N-terminus of ctRio2 contains a winged helix domain. It is comprised of four α helices followed by a β-strand wing β-strand motif, and another α helix (Figure 5.1). The RIO domain in ctRio2 (as in afRio2) consists of an N-terminal lobe and a C-terminal lobe with a flexible hinge region between them (Figures 5.2). ATP binds in the large cleft between the two lobes. The N-terminal lobe consists of a five-stranded antiparallel β sheet and a long α helix. The C-terminal lobe of the RIO domain consists of two antiparallel β strands and two α helices. The third α helix seen in afRio2 is missing in the ctRio2 structure. However, instead there is another C-terminal α helix that is not seen in the afRio2 structure partially obstructing the opening to the active site. This α helix, labeled αJ, is situated in the cleft between the C and N terminal lobes of the RIO kinase domain (Figures 5.2).

The active site of ctRio2 contains three loops—the P-loop, the "DFG" loop, and the catalytic loop. The P-loop is involved in binding ATP and orienting the γ phosphate in a position to be attacked by the nucleophilic serine in a phosphoryl transfer reaction. The "DFG" loop binds and positions the metal ions that coordinate the phosphate oxygens of ATP and help orient the nucleotide. The catalytic loop contains the conserved N and D residues necessary for catalysis and metal binding. In typical protein kinases the sequence of the P-loop is GXGXXG. In afRio2 the sequence is GEGKES, and in ctRio2 the sequence is GVGKES. The sequence of the DFG loop in both afRio2 and ctRio2 is IDFPQ. In afRio2, the catalytic loop contains the catalytic D218 and the

metal binding residue N223. It has the sequence DLSQYN. For ctRio2, the catalytic

residue is D232 and the metal binding residue is N237. The sequence is DFNEFN.

Figure 5.3 shows the labeled sequences of ctRio2.

5.3.2: Binding Role of the Extended C-terminus in CtRio2

CtRio2 contains an extended C-terminus. The C-termini of eukaryotic Rio2

proteins differ widely. Higher eukaryotes have an elongated C-terminus whereas lower

eukaryotes have a shorter C-terminus. The sequence alignment of the C-termini between

lower and higher eukaryotes is poor. Figure 5.4 shows the alignment of the most

conserved segment of the C-terminus between Rio2 from different eukaryotic organisms.

There is conservation in the region of αJ between these organisms. Most of ribosomal

processing is deduced from studies in yeast, thus Rio2p has been well studied. Its C-

terminus contains a weak nuclear export signal (NES) within residues 318-425, which

can supplement the endogenous NES of the Nmd3p adaptor protein (*37*). Another study

has assigned the NES signal in yeast specifically to residues 407-425 (*52*). Rio2p

temperature sensitive mutants show a 40S preribosomal subunit nuclear export

impairment at the restrictive temperature. Also leptomycin inhibits the export of NES

containing cargo. Upon leptomycin treatment of yeast cells, Rio2p accumulates in the

nucleus (*37*).

Human hRio2 protein also contains an NES signal in the C-terminus. After

depletion of hRio2, the 40S preribosome is stalled in the nucleus. The 40S preribosomes

get exported upon overexpression of hRio2 (but not hRio2) with a mutated or deleted

NES signal (*52*). Thus, by correlation, the extended C-termini of ctRio2 may simply be

involved in export of the 40S preribosome from the nucleus.

The C-terminus of ctRio2 also contains an extra α helix that binds between the two lobes of the kinase domain, indicating that the extension of the C-terminal domain might have an inhibitory effect on the kinase activity. This α helix is reminiscent of kinases with autoinhibitory domains. The α subunit of the AMP-activated protein kinase (AMPK) from *Schizosaccharomyces pombe* (Figure 5.5) contains a conventional serine/threonine kinase domain followed by an autoinhibitory domain (AID). The AID consists of three α helices and interacts with both the N- and C- lobes of the kinase domain, and buries 1500 $Å^2$ exposed surface area mainly through hydrophobic interactions. Without the AID, kinase activity is tenfold higher (*53*). CtRio2 has one α helix in the cleft between the two lobes of the kinase domain while the AID of AMPK binds to the back hinge region of the protein. The interactions of the two proteins with their potential AIDs are very different.

AMPK also contains an activation domain consisting of a single alpha helix similar to the one seen in ctRio2 that occupies the active site between the two lobes of the kinase domain. The C-terminal alpha helix in ctRio2 could be an activation domain that has an inhibitory effect on kinase activity. This kinase activity may be activated by phosphorylation of the helix or regions surrounding it.

Another example of a kinase with an AID is twitchin kinase (Figure 5.6). Twitchin kinase contains an autoinhibitory C-terminus with 58 residues (*54*). However, the AID of twitching kinase binds quite differently than the AID of AMPK and the α helix of ctRio2. It binds extensively to the C-terminal lobe of the kinase domain and uses a long region to bind to the cleft between the two lobes. This domain consists of a long α helix that occupies the peptide substrate binding site in the C-lobe of the kinase domain.

It also contains a long alpha helix followed by a short turn single α helix and a short β-strand that binds between the two lobes in the ATP binding site. In the case of twitchin kinase, structural changes occur upon binding to accommodate the AID. Binding of the AID causes the two lobes of the kinase domain to rotate 30° (*54*). It does not seem that the helix seen in the structure of ctRio2 causes structural changes to the kinase domain, as the positioning of lobes and helices is consistent with that of afRio2.

5.3.3: C-terminal αJ Interactions with the Kinase Domain

αJ from the extended C-terminus and the RIO domain are involved in many interactions in ctRio2. The sequence of αJ in ctRio2 is K227, KMAKDLEAAIREAQ, Q341. D232, L333, E334, A336, from ctRio2 are conserved in yeast (Figure 5.4). The side chain of D232 hydrogen bonds with the side chain of Y156 from αC. The side chain of L333 makes a hydrophobic interaction with the side chain of A161 from αC. The side chain of E334 makes a hydrogen bond with the side chain of Y275 from αF and the side chain of H269 from the loop region close to αF. The side chain of A336 makes a hydrophobic contact with the side chain of L157 from αC. The side chain of I337 (V344 in Rio2p) makes a hydrophobic contact with the side chains of P262 and N263 from the metal binding loop. The side chain of E339 makes a salt bridge with the side chain of R149 form αC. The side chain of Q340 hydrogen bonds to the backbone amide nitrogen of G132 from the flexible loop. Thus αJ binds tightly to the RIO domain of ctRio2 and is not expected to be flexible. Figure 5.8 illustrates the interactions of αJ with the kinase domain.

5.3.4: Binding of ATP to Rio2

In the structure of ctRio2 in a complex with ATP, Mg^{2+} is present in the active

site. As shown in Figure 5.9, the metal ion is coordinated by two phosphate oxygen

atoms from the α and β phosphate groups, a conserved catalytic residue N237, and a

conserved residue D260 in the metal binding loop. Water molecules complete the

coordination sphere of the metal ion. Figure 5.10 shows that the γ phosphate is held in

place by interaction with a conserved residue Q263 in the metal binding loop, and by a

conserved residue E110 in the P-loop (bond lengths 3.24 Å and 2.46 Å respectively).

The α phosphate is coordinated via a 2.08 Å bond to the metal ion (Figure 5.10).

K127 is the conserved lysine present in all protein kinases and forms a 2.64 Å hydrogen

bond with the α phosphate. The 3' hydroxyl of the ribose ring forms a 2.98 Å hydrogen

bond with F236 of the catalytic loop. The 6' amino group of adenine forms a 2.85 Å

hydrogen bond with S193.

The β phosphate is coordinated via a 2.14 Å bond to the metal ion (Figure 5.11).

The conserved S111 from the P-loop forms a 2.46 Å bond with the β phosphate and a

3.17 Å bond with the bridging oxygen between the α and β phosphate. K127 also forms a

3.24 Å bond with the bridging oxygen between the β and γ phosphate.

5.3.5: Position of γ Phosphate of ATP in AfRio2 and CtRio2

CtRio2 binds ATP in a different conformation than afRio2. The position of the γ

phosphate is significantly shifted relative to the position of the metal ion and the catalytic

residues (Figure 5.12). The structural alignment of afRio2 and ctRio2 reveals that the P-

loop of ctRio2 is positioned lower than in afRio2 (Figure 5.13). As a result the γ

phosphate of ATP in ctRio2 is positioned lower than in afRio2 (Figure 5.13). Residues

S111 and E110 of the P-loop coordinate the β and γ phosphates respectively of ctRio2.

The positioning of the γ-phosphate in afRio2 places the kinase catalytic base D 218 too far away to be able to participate directly in phosphoryl transfer (Figure 5.14). Whereas, the distance between the carboxyl oxygen of D218 and the phosphorus atom is ~4.48 Å in PKA, this distance is nearly 7.09 Å in afRio2. In ctRio2 the positioning of the γ-phosphate relative to the catalytic base D232 is similar to the positioning seen in PKA and the distance between the γ phosphate and D232 is 4.66 Å. Since the position of the γ phosphate in ctRio2 is like that of other ePKs such as PKA, it may represent an active state of the kinase while the position of the γ phosphate in afRio2, may represent an inactive state.

Structural alignment showing the conserved active site residues, ATP, and metal ions of ctRio2 and afRio2 (Figure 5.13) reveal that while the conserved residues do not change positions dramatically, the second Mn ion of afRio2 occupies the position of the phosphorous atom of the γ phosphate of ctRio2. Binding of the second Mn to afRio2 may preclude the γ phosphate of afRio2 from assuming a catalytically active state. afRio2 may become active when only one metal ion is bound as in ctRio2. Kinetic studies on afRio2 and ctRio2 have not been performed to determine whether excess Mg ion is inhibitory on kinase activity but based on this information, this is an obvious next step.

5.3.6: Structural Alignments

Structural alignment of ctRio2 and afRio2 (rmsd 2.204 Å for 217 atoms) reveals that the winged helix domain is shifted between the two proteins. This shift could be due to elongation of αC in ctRio2 (Figure 5.13). Alignment was performed of the C-terminal lobe of the kinase domain between afRio2 and ctRio2 using Swiss PDB viewer. The

alignment reveals a larger shift of the wing helix domain between the two proteins. The alignment also reveals β8 and β9 are extended in ctRio2 as compared to afRio2 (Figure 5.16).

Alignment of the best-fit fragment between the two proteins in the Swiss PDB viewer shows that the most structurally conserved feature between the two proteins is the active site as expected. The N-terminal lobe of the kinase domain and the active site portion of the C-terminal lobe of the kinase domain align, while the rest of the C-terminal lobe of the kinase domain and the N-terminal wing helix domain are shifted (Figure 5.15).

5.3.7: Kinase Activity of CtRio2

LC/MS data show that ctRio2 is phosphorylated once (Figure 5.17). The kinase activity was tested by incubating ctRio2 with γP^{32}-labeled ATP (Figure 5.18). Kinase activity was indicated by the autophosphorylation of ctRio2. It produced a radioactive band at the molecular weight of the protein on SDS-PAGE gel. Kinase activity was evident at 0.5 μM ATP with both 10 μg and 50 μg of ctRio2.

5.4: Summary

The crystal structure of ctRio2 was determined and the structures of ctRio2 and afRio2 were compared. The structure of ctRio2 in a complex with ATP aligns well with afRio2. However, ctRio2 contains an extended C-terminus containing αJ. αJ interacts with the kinase domain through both hydrophobic and hydrophilic interactions with αC, αF, the flexible loop and the metal binding loop. Since αJ occupies the cleft between the two lobes of the kinase domain, it may have an inhibitory effect on kinase activity by

preventing peptide substrates from binding. Examples of protein kinases containing inhibitory C-termini are AMP activated protein kinase (53) and twitchin kinase (54).

AfRio2 has two Mn ions in the active site. These metal ions coordinate all three phosphates of ATP. The structure of ctRio2 shows only one Mg ion. This Mg ion coordinates the α and β phosphates of ATP. The γ phosphate in the structure of ctRio2 is still well bound as it makes polar interactions with two amino acids.

The structural alignment reveals that the winged helix domain is shifted between afRio2 and ctRio2. My analysis revealed the position of the γ phosphate in ctRio2 is closer to the catalytic aspartate than in afRio2. The binding of the second Mn ion in afRio2 as well as the higher positioning of the P-loop may prevent the γ phosphate of ATP from binding in a productive manner. The positioning of the γ phosphate in ctRio2 closely mimics the position of the γ phosphate found in eukaryotic serine threonine protein kinases such as PKA. Thus ctRio2 might share the same mechanism of catalysis as traditional eukaryotic protein kinases.

Figure 6.6 shows the alignment between Rio2 from different eukaryotic species. The conserved residues are mapped on a surface representation of ctRio2 (Figure 5.19). The conserved surface patches might indicate important rRNA or protein interactions made with the ribosome. The crystal structure of ctRio2 allows us to postulate and study which residues might interact with the ribosome or ribosomal assembly factors.

.

Table 5.1 Data Collection and current refinement statistics for apo ctRio2. Crystal data: space group H3

a (Å)	117.40
b (Å)	117.40
c (Å)	71.29
$\alpha=\beta$ (°)	90.00
γ (°)	120.00
Resolution (Å)	29.3-2.48
R_{sym} (last shell)	0.074(0.516)
Reflections (last shell)	12315(1255)
Redundancy	2.9(2.9)
Completeness (%)	97.9(99.1)
Mean B-factor (Å)2	60.5
R/R_{free} (%)	21.78/28.51
(Last shell)	29.13/38.71
RMS deviations	
Lengths (Å)	0.010
Angles (°)	1.38

Table 5.2 Data Collection and current refinement statistics for ATP bound ctRio2. Crystal data: space group H3

a (Å)	119.62
b (Å)	119.62
c (Å)	71.78
$\alpha=\beta$ (°)	90.00
γ (°)	120.00
Resolution (Å)	34.5-2.20
R_{sym} (last shell)	0.067(0.539)
Reflections (last shell)	19451(1942)
Redundancy	3.2(3.2)
Completeness (%)	100(100)
Mean B-factor (Å)2	55.3
R/R_{free} (%)	19.61/24.73
(Last shell)	20.9/27.9
RMS deviations	
Lengths (Å)	0.009
Angles (°)	1.35

Table 5.3: Apo CtRio2 Scalepack Logfile

```
Shell          Average Redundancy Per Shell
  Lower Upper
  limit limit
  50.00  5.38                  3.0
   5.38  4.27                  2.9
   4.27  3.73                  2.9
   3.73  3.39                  2.9
   3.39  3.15                  2.9
   3.15  2.96                  2.9
   2.96  2.82                  2.9
   2.82  2.69                  2.9
   2.69  2.59                  2.9
   2.59  2.50                  2.9
  All hkl                      2.9
```

```
Shell              I/Sigma in resolution shells:
  Lower Upper       % of of reflections with I / Sigma less than
  limit limit    0     1     2     3     5    10    20    >20   total
  50.00  5.38   0.0   0.1   0.4   0.9   1.9   5.8  54.0  41.4   95.4
   5.38  4.27   0.0   0.4   0.7   1.5   3.4  17.8  95.1   1.3   96.4
   4.27  3.73   0.2   1.0   1.9   2.9   5.6  21.6  97.5   0.3   97.8
   3.73  3.39   0.7   2.5   4.8   8.5  14.9  40.7  98.1   0.1   98.2
   3.39  3.15   2.3   6.9  13.2  18.7  29.8  56.0  95.9   1.7   97.6
   3.15  2.96   5.5  11.1  19.3  27.5  42.0  73.4  98.6   0.2   98.8
   2.96  2.82   7.2  17.2  30.6  41.2  59.4  82.1  96.6   1.6   98.2
   2.82  2.69   8.9  23.8  38.5  52.7  69.7  91.2  98.2   0.6   98.8
   2.69  2.59  12.5  31.0  49.2  64.6  81.4  95.0  98.8   0.1   98.9
   2.59  2.50  14.4  36.1  59.5  74.5  90.9  98.0  99.1   0.0   99.1
  All hkl      5.2  13.0  21.8  29.3  39.9  58.1  93.1   4.8   97.9
```

```
Shell Lower Upper Average      Average    Norm. Linear Square
  limit     Angstrom      I   error  stat. Chi**2  R-fac  R-fac
       50.00   5.38  235.5   12.4    4.2  1.122  0.047  0.058
        5.38   4.27  138.1   10.1    3.7  1.094  0.068  0.084
        4.27   3.73   99.0    7.4    2.8  1.139  0.072  0.088
        3.73   3.39   48.3    4.2    2.3  1.118  0.084  0.092
        3.39   3.15   23.9    2.4    1.8  1.124  0.107  0.108
        3.15   2.96   14.2    2.0    1.7  1.346  0.163  0.154
        2.96   2.82    8.3    1.7    1.7  1.055  0.219  0.202
        2.82   2.69    6.0    1.7    1.7  0.906  0.280  0.260
        2.69   2.59    4.3    1.8    1.7  0.826  0.389  0.341
        2.59   2.50    3.3    1.8    1.8  0.787  0.518  0.473
  All reflections      57.4    4.5    2.3  1.052  0.074  0.067
```

Table 5.4: ATP-CtRio2 Scalepack Logfile

```
Shell          Average Redundancy Per Shell
  Lower Upper
  limit limit
  40.00  4.74                  3.2
   4.74  3.76                  3.2
   3.76  3.29                  3.2
   3.29  2.99                  3.2
   2.99  2.77                  3.2
   2.77  2.61                  3.2
   2.61  2.48                  3.2
   2.48  2.37                  3.2
   2.37  2.28                  3.2
   2.28  2.20                  3.2
  All hkl                      3.2
```

```
      Shell              I/Sigma in resolution shells:
  Lower Upper      % of of reflections with I / Sigma less than
  limit limit    0     1     2     3     5    10    20   >20   total
  40.00  4.74  0.0   0.1   0.3   0.4   1.1   5.7  90.5   9.2   99.7
   4.74  3.76  0.0   0.2   0.6   1.3   2.3   8.1  97.5   2.5  100.0
   3.76  3.29  0.5   1.1   2.2   3.5   6.4  19.6  96.7   3.3  100.0
   3.29  2.99  1.5   3.3   6.4   9.2  14.9  34.0  89.5  10.5  100.0
   2.99  2.77  2.4   6.2  11.2  17.0  28.4  52.9  95.3   4.7  100.0
   2.77  2.61  3.7   9.7  16.8  25.2  40.4  69.5  97.8   2.2  100.0
   2.61  2.48  6.0  16.5  29.4  39.8  57.4  84.3  98.6   1.4  100.0
   2.48  2.37  8.4  22.2  36.9  48.6  67.3  90.8  99.1   0.9  100.0
   2.37  2.28  9.4  24.6  44.3  58.2  77.9  94.8  99.5   0.5  100.0
   2.28  2.20 14.3  34.6  54.1  69.6  85.9  97.7  99.9   0.1  100.0
  All hkl     4.6  11.9  20.2  27.3  38.2  55.7  96.4   3.5  100.0
```

```
Shell Lower Upper Average     Average    Norm. Linear Square
  limit    Angstrom       I   error  stat. Chi**2 R-fac  R-fac
        40.00   4.74  357.8   22.1    5.9  0.949  0.041  0.048
         4.74   3.76  295.7   19.9    5.1  0.973  0.052  0.060
         3.76   3.29  147.9    9.7    3.9  1.017  0.061  0.067
         3.29   2.99   68.6    4.6    3.0  0.972  0.071  0.073
         2.99   2.77   34.2    3.1    2.6  0.901  0.103  0.101
         2.77   2.61   22.3    2.8    2.6  0.794  0.144  0.141
         2.61   2.48   14.4    2.7    2.6  0.837  0.224  0.215
         2.48   2.37   11.3    2.8    2.8  0.903  0.296  0.277
         2.37   2.28    9.0    3.0    2.9  0.870  0.385  0.370
         2.28   2.20    7.0    3.2    3.1  0.917  0.539  0.504
  All reflections      96.8    7.4    3.4  0.913  0.067  0.057
```
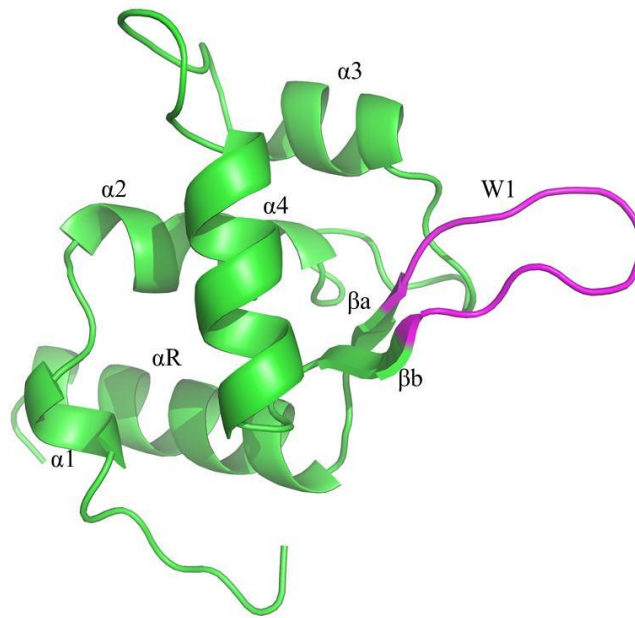
Figure 5.1: Structure of the Winged Helix Domain of CtRio2
The winged helix domain consists of four alpha helices followed by two beta sheets. A fifth alpha helix (αR) is part of the RIO domain. The wing is labeled W1and is in magenta. This cartoon representation was constructed by PYMOL.
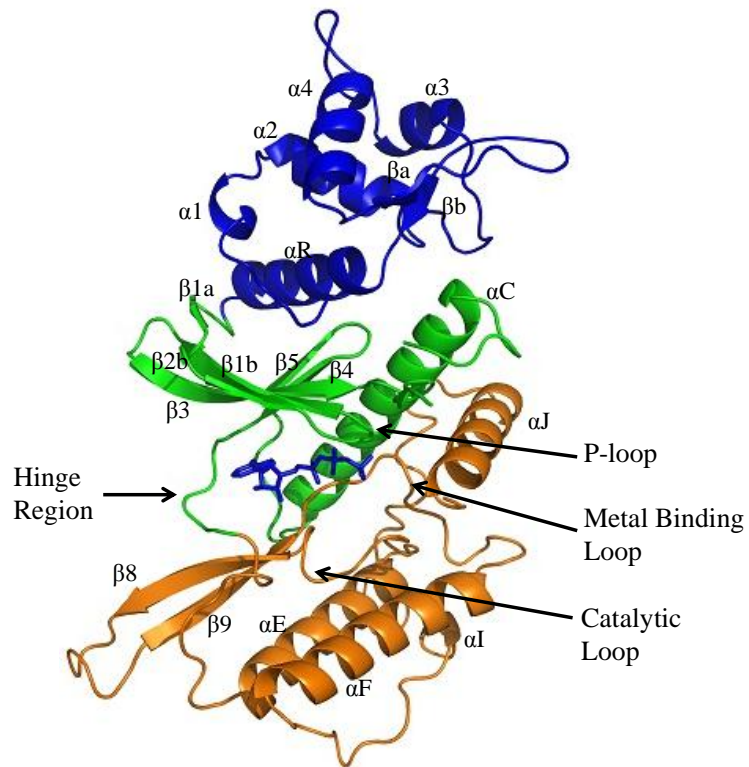
Figure 5.2: Crystal Structure of CtRio2

The crystal structure of CtRio2 shows a bilobal kinase domain (green and orange) connected to a N-terminal winged helix domain (blue). The "hinge" region connects the N- and C- terminal lobes of the kinase domain. The ATP binding loop "P-loop", metal binding loop used to bind Mg, and catalytic loop which contains the catalytic D232 are indicated by arrows. The cartoon representation was made by PYMOL.

```
                 1          11         21         31         41                   51
afRio2   MNIAELYGKM----GKHSWRIMDAIFKNLWDYEYVPLQLISSHARIGEEKARN------ILK-YLSDLRV
                  *                 *         *      *    *              *     *      *
                  *                 *         *      *    *              *     *      *
ctRio2   MKLDTRAMRHLTAEDWRVLTAVEMGSKNH-----EIVPTPLIE-KIARLRGGSSGVHKSIATLAKA-GLI
             4          14         24              34         44         54         64


                                                                      P-loop
                                                                      XXXXXX
                   61         71         81         91        101        111
afRio2   -------------VQNRQKDYEGSTFTFIGLSLYSLHRLVRSGKVDAIGKLM-GEGKESAVFNCYSEKFG
                      *                   *        * *  * **  **********
                      *                   *        * *  * **  **********
ctRio2   ARMKEAKYDGYRLTYGGL--------------DYLALHTHAARKDVYSVGSRIGVGKESDIMIVADEKGK
              74         84                         94        104        114


                   121        134              152        162        172        18
afRio2   ECVVKFHKVGHTKVKE-----------FSVLAIRSARNEFRALQKLQ--GLAVPKVYAWEGNAVLMELID
          * * * * * *              * * * * * * * * * **  * * * * *  * * * * * *
          * * * * * *              * * * * * * * * * **  * * * * *  * * * * * *
ctRio2   QKVLKIHRLGRISFRTVDYLRNRSTGSWMYLSRLAAIKEFAFMKALYEEGFPVPEPIAQSRHTIVMSLVD
             124        134        148        158        168        178        188


                                                      Catalytic           Metal
                                                       XXXXXX             XXXXXX
          2     192        202        212        222                      232
afRio2   AKELYRVR-VENPDEVLDMILEEVAKFYHRGIVHGDLSQYNVLVSEE-----------GIWIIDFPQSVE
          * * * * * * *  * * * * * * * * * * * * * * * * * * * * * *  *      * * * * * * *
          * * * * * * *  * * * * * * * * * * * * * * * * * * * * * *  *      * * * * * * *
ctRio2   ALPMRQVSSVPDPASLYADLIALILRLAKHGLIHGDFNEFNILIREEKDAEDESSITLTPIIIDFPQMVS
              198        208        218        228        238        248        258


              242        252        262        272        282
afRio2   VGEEGWREILERDVRNIITYFSRTYRTEKDINSAIDRILQE
          **   **  * * * * * * * * * * *     *     *
          **   **  * * * * * * * * * * *     *     *
ctRio2   MDHPNAEMYFDRDVQCIKRFFERRFHFVSTTPG-PFYKDAKKTVGKDGARLDAALEASGFTKKMAKDLEA
              268        278        288        298        308        319        329


ctRio2   AIREQQES
              339
```

Figure 5.3: Alignment between AfRio2 and CtRio2.

Structural alignment was performed using sequences from *A. fuligidus* (afRio2) and ctRio2. Alignment positions marked with two asterisks are defined as "matching residues", i.e. corresponding C-alpha-atoms lie within 1.8Å . The wing helix domain of afRio2 is highlighted in yellow, the N-terminal lobe in green and the C-terminal lobe in blue. The C-terminal extension of Rio2p is in the boxed region encompassing residues 307-346. The sequences of the ATP binding loop "P-loop", metal binding loop used to bind Mg, and catalytic loop which contains the catalytic D232 are indicated by x's.

```
ctrio2  251 LTPIIIDFPQMVSMDHPNAEMYFDRDVQCIKRFFERRFHFVSTT------PGPFYKDAKKTVG
zmrio2  240 IT--VIDFPQMVSVSHRNAQMFFDRDIECIYKFFNKRFNLRSVKNE--EQAGSENDDEGNSRP
hsrio2  242 IT--MIDFPQMVSTSHPNAEWYFDRDVKCIKDFFMKRFSYESELF-------PTFKDIRREDT
xlrio2  242 IT--MIDFPQMVSTSHANAEWYFDRDVKCIRDFFLKRFNYESELY-------PTFKDIQRECS
scrio2  247 CGFVVIDFPQCISIQHQDADYYFQRDVDCIRRFFKKKLKYEPKPDSSMLDTEGFGDGYKYAYP
cario2  245 FDFVVIDFPQCVSIEHPDAKEYFDRDVQGIRDFFSKKFRYNPQHDPTMFDTDGYGDGYKYAYP
                                                    αJ
                                        xxxxxx    xxxxxxxx
ctrio2  308 -----KDGAKRLDAALEASGFTKKMAKD-----LEAAIREQQ-ESR--ND----EEDSDDY-E
zmrio2  299 SFLSVEKGAGSLDKELAASGFTRKEKVD-----MDKYIEEV-VEG--H—DSS------SD-D-
hsrio2  296 -----------LDVEVSASGYTKEMQAD-----DEL-LHPL-GPDD-KNIETKEGSEFSFS-D
xlrio2  296 -----------LDVEIAASGFTKEMQED-----DEL-LQPS-GPSD-DEVEDDDDNDDDDN-D
scrio2  310 DFKRDVKRTDNLDELVQASGFSKKHPGD---RGLETAVESM-RNAV-YNSDDDM--SNDEA-E
cario2  308 NFKRDVIREKSLDVEVQASGYAKKKTGNREDKDLEKAVLGMRIKV-DEDDDLSEIDDDED-V-
```

Figure 5.4: Alignment Between CtRio2 and Eukaryotic Rio2

Alignment was performed using sequences from Rio2p (scRio2), *zea mays* Rio2 (zmRio2), human rio2 (hsrio2), *xenopus laevis* Rio2(xlRio2), *candida albicans* Rio2 (caRio2) and ctRio2. Identical, highly similar, and weakly similar residues are indicated by red, green, and blue letters respectively. The position of αJ is indicated by x's.

N-Terminal Kinase Domain ———→

Inhibitory Domain

Activation Domain ———→

C-Terminal Kinase Domain
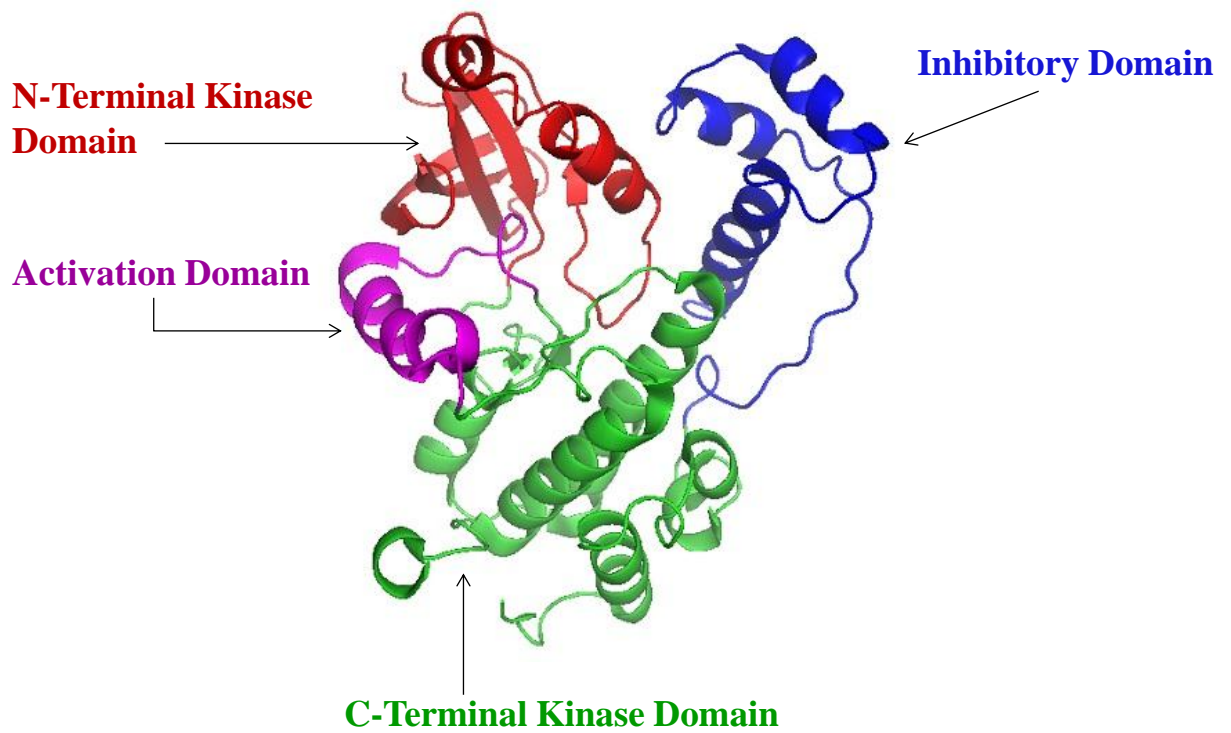
Figure 5.5 : Crystal Structure of AMP Activated Protein Kinase

The cartoon structure was generated by PYMOL using PDB structure 3H4J The N-terminal kinase domain, C-terminal kinase domain, activation domain, and inhibitory domain of AMPK are labeled in red, green, magenta, and blue respectively.
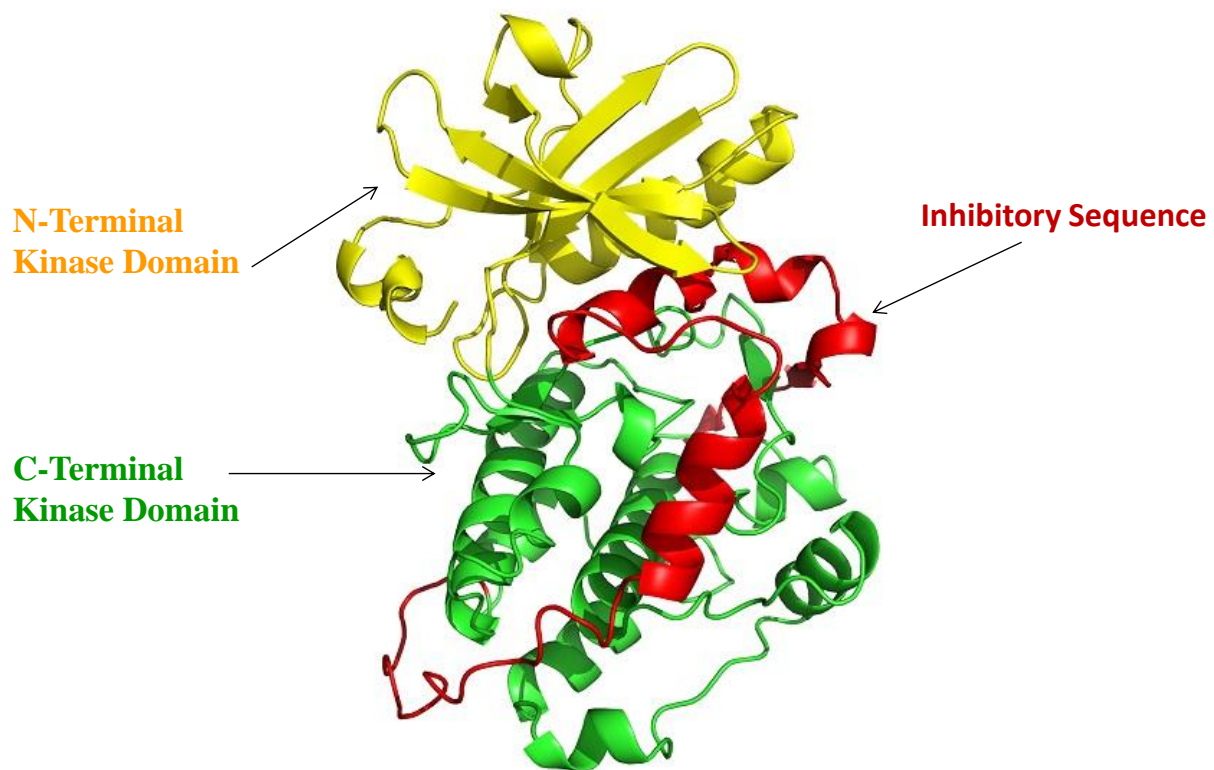
Figure 5.6: Crystal Structure of Twitchin Kinase

The cartoon structure was generated by PYMOL using PDB structure 1KOA. The N-terminal kinase domain, C-terminal kinase domain, and inhibitory domain of twitchin kinase are labeled in yellow, green, and red respectively.
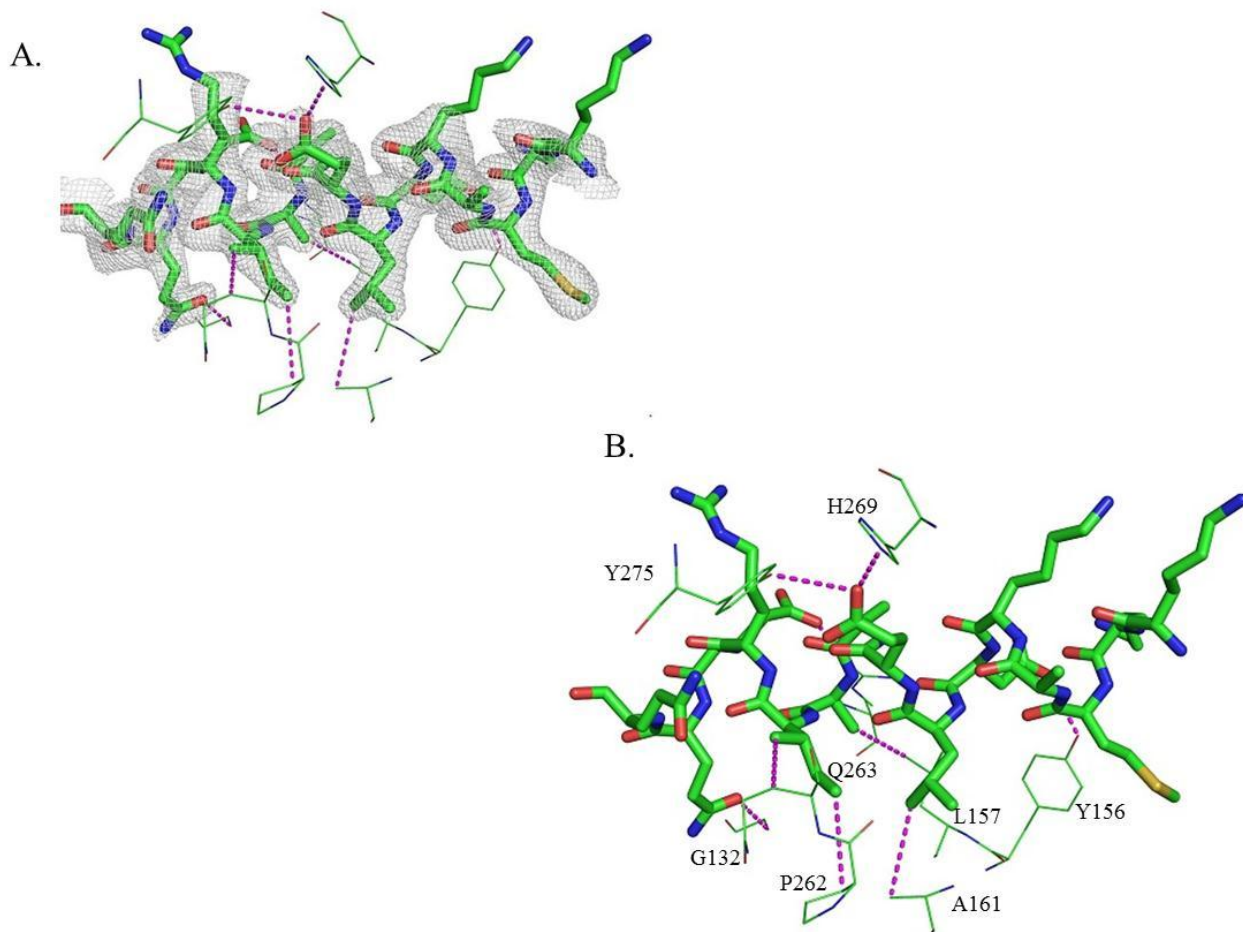
Figure 5.7: The Interface Between αJ and the Kinase Domain

The interaction between αJ (in sticks) and the kinase domain of ctRio2 (in lines) are shown in A and B. The figure was made using PYMOL. A. The 2Fo-Fc map of αJ is contoured at 1σ.
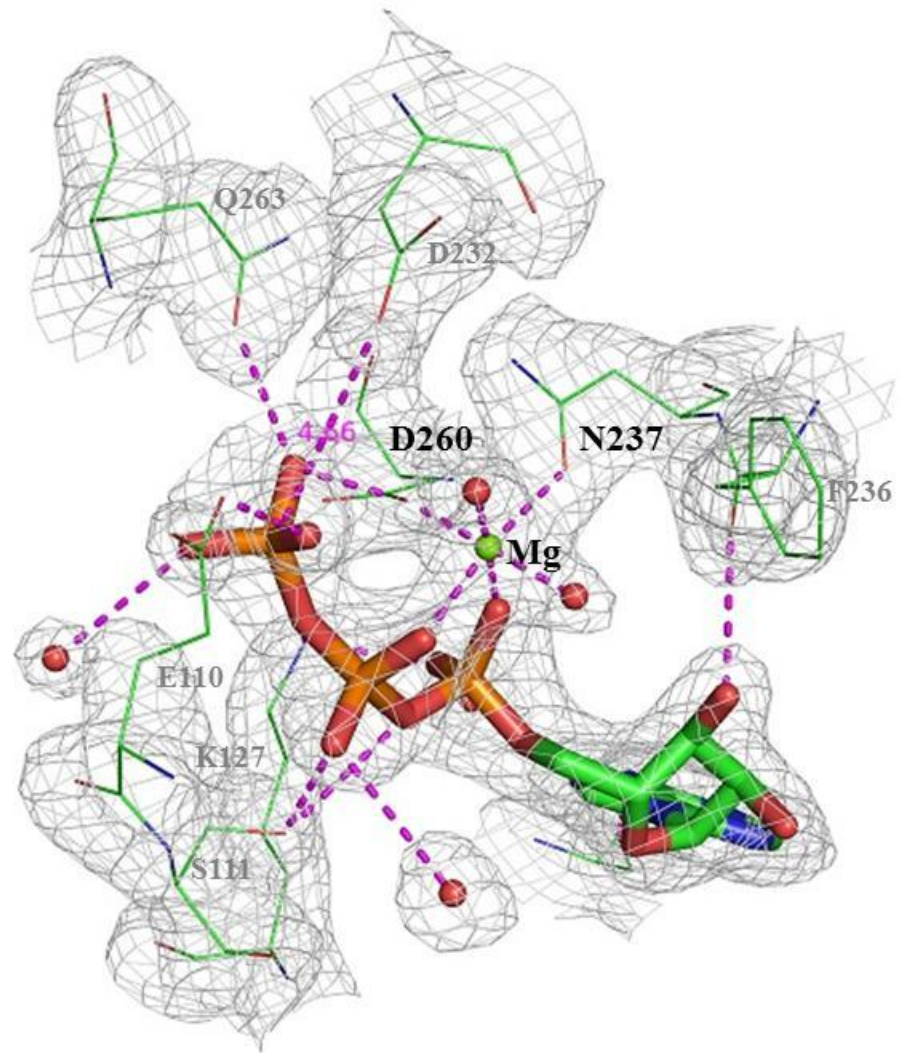
Figure 5.8: Coordination Sphere of Mg

Crystal structure of ctRio2 with 2Fo-Fc map contoured at 1σ. Mg is coordinated by D260, N237, α and β phosphates of ATP, and two water molecules.
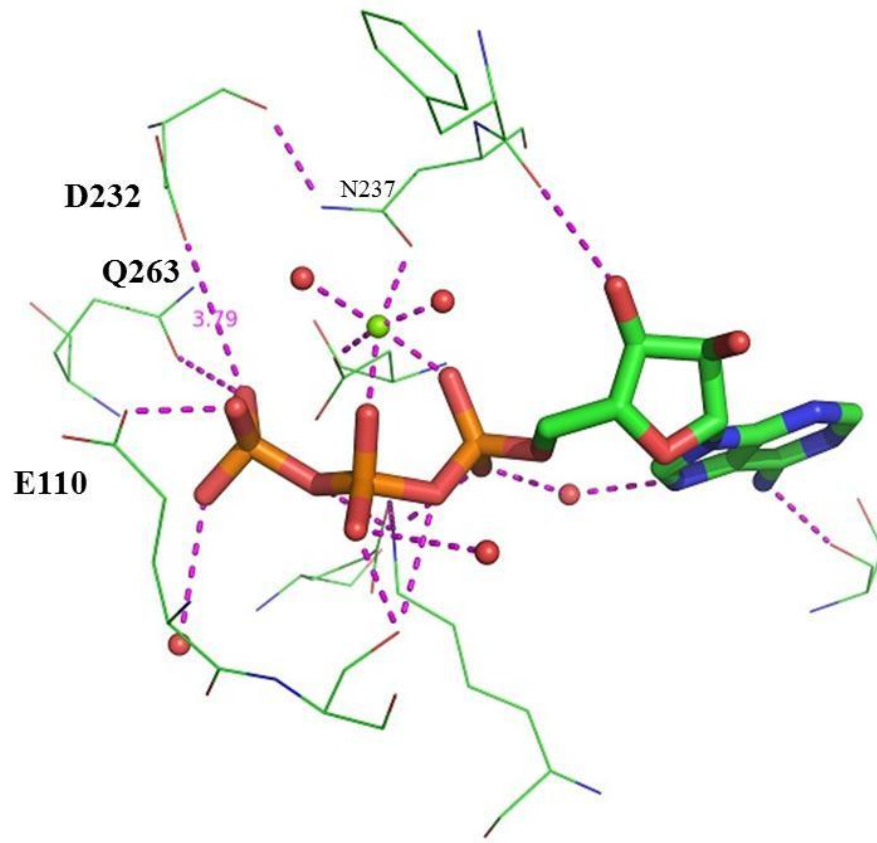
Figure 5.9: Polar Contacts with γ Phosphate of ATP

Crystal structure of ctRio2 with 2Fo-Fc map contoured at 1σ. The γ phosphate of ATP makes polar contacts with Q263 and E110. The catalytic D232 is 4.66 Å from the phosphorous atom of the γ phosphate.
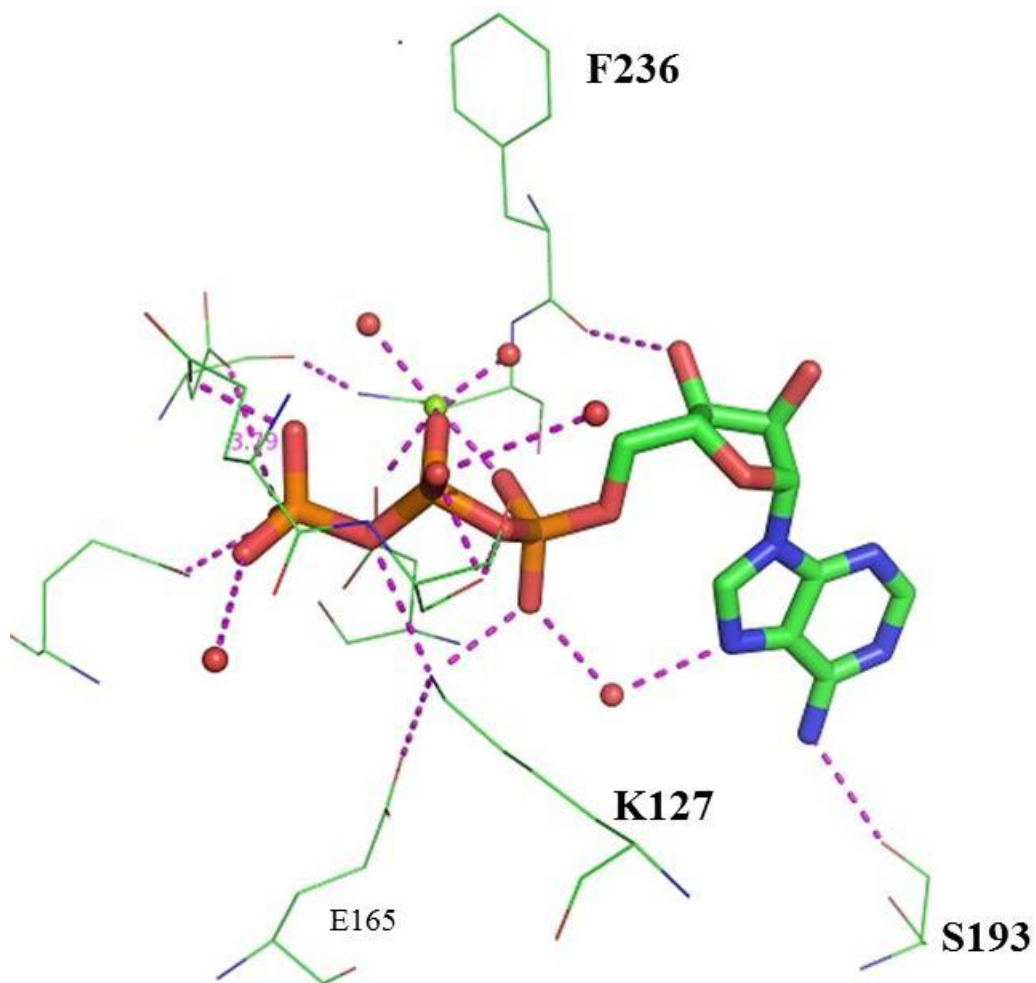
Figure 5.10: Polar Contacts with α Phosphate of ATP

Crystal structure of ctRio2 with 2Fo-Fc map contoured at 1σ. The α phosphate of ATP makes polar contacts with K127 and is coordinated by the Mg ion. The 6'amino group of adenine makes a polar contact with S193 and the ribose 3'hydroxyl group makes polar contacts with F236.
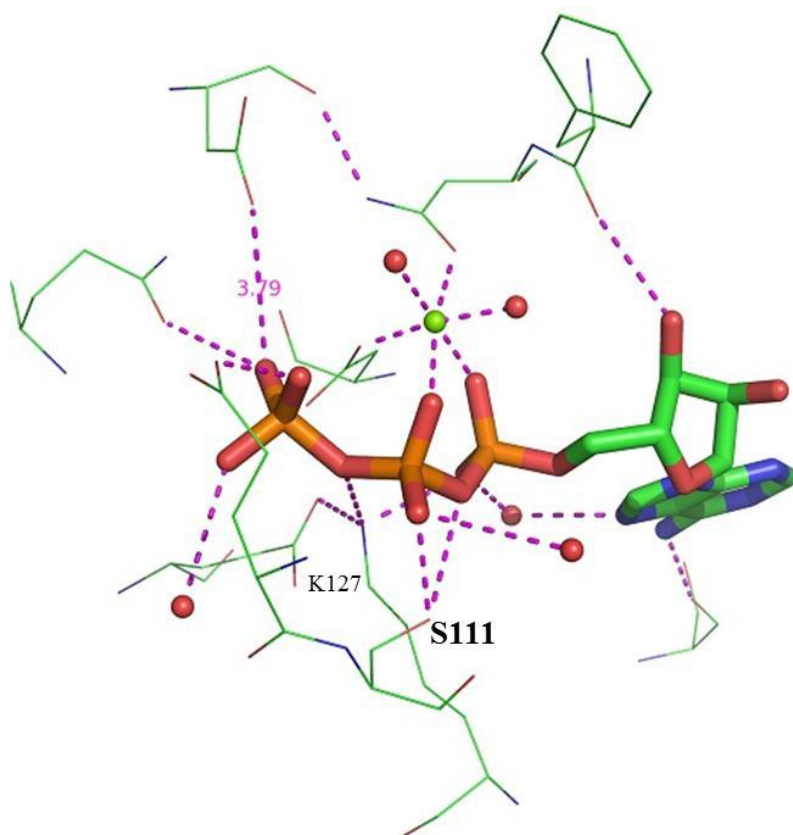
Figure 5.11: Polar Contacts with β Phosphate of ATP

Crystal structure of ctRio2 with 2Fo-Fc map contoured at 1σ. The β phosphate of ATP makes polar contacts with S111 and is coordinated by the Mg ion.
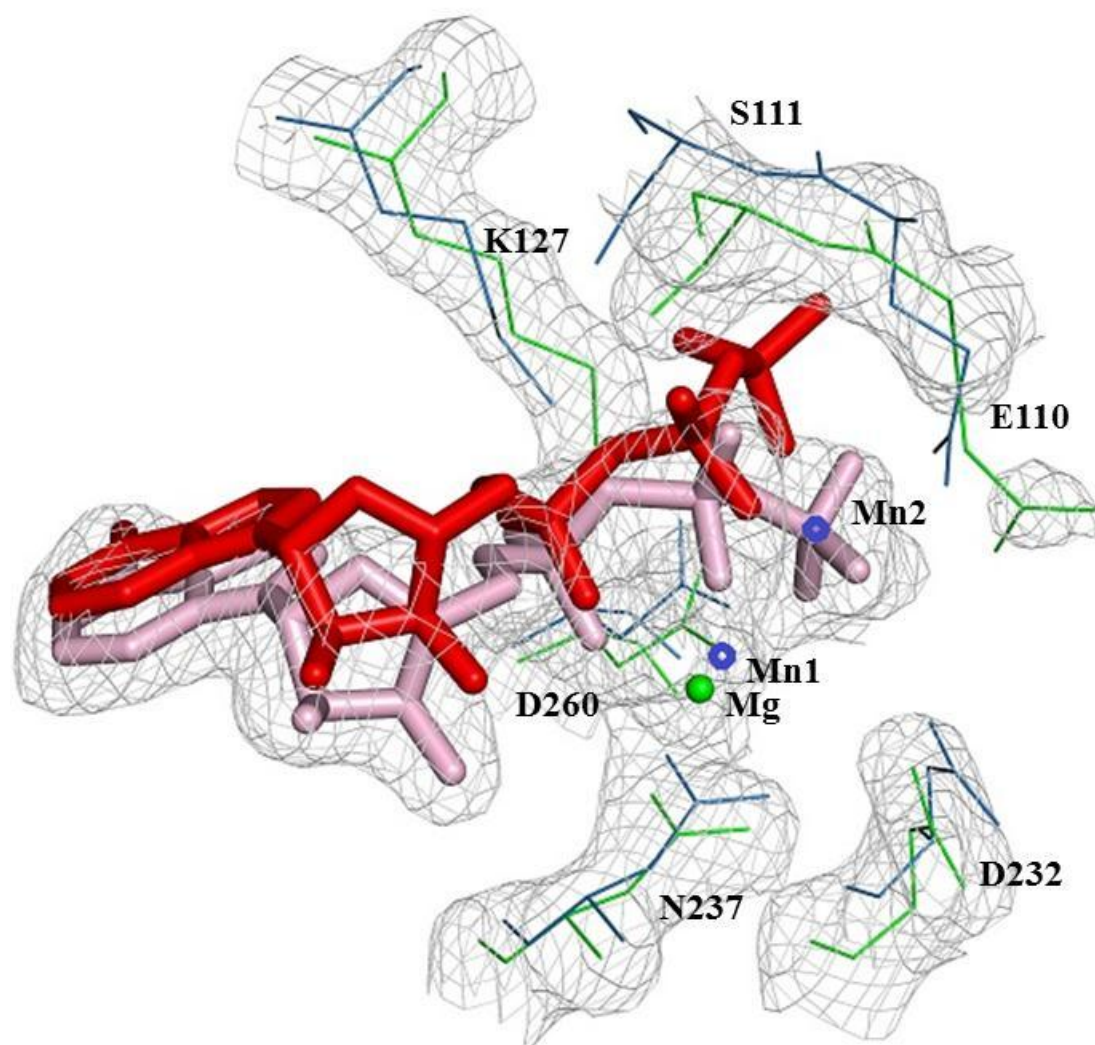
Figure 5.12: Alignment Between Conserved Active Site Resides in AfRio2 and CtRio2

The structures of afRio2and ctRio2are aligned, ctRio2residue numbering is used. Residues for afRio2 are blue and for ctRio2 are green. ATP for afRio2 is red and for ctRio2 is pink. The 2Fo-Fc map for ctRio2 is contoured at 1σ
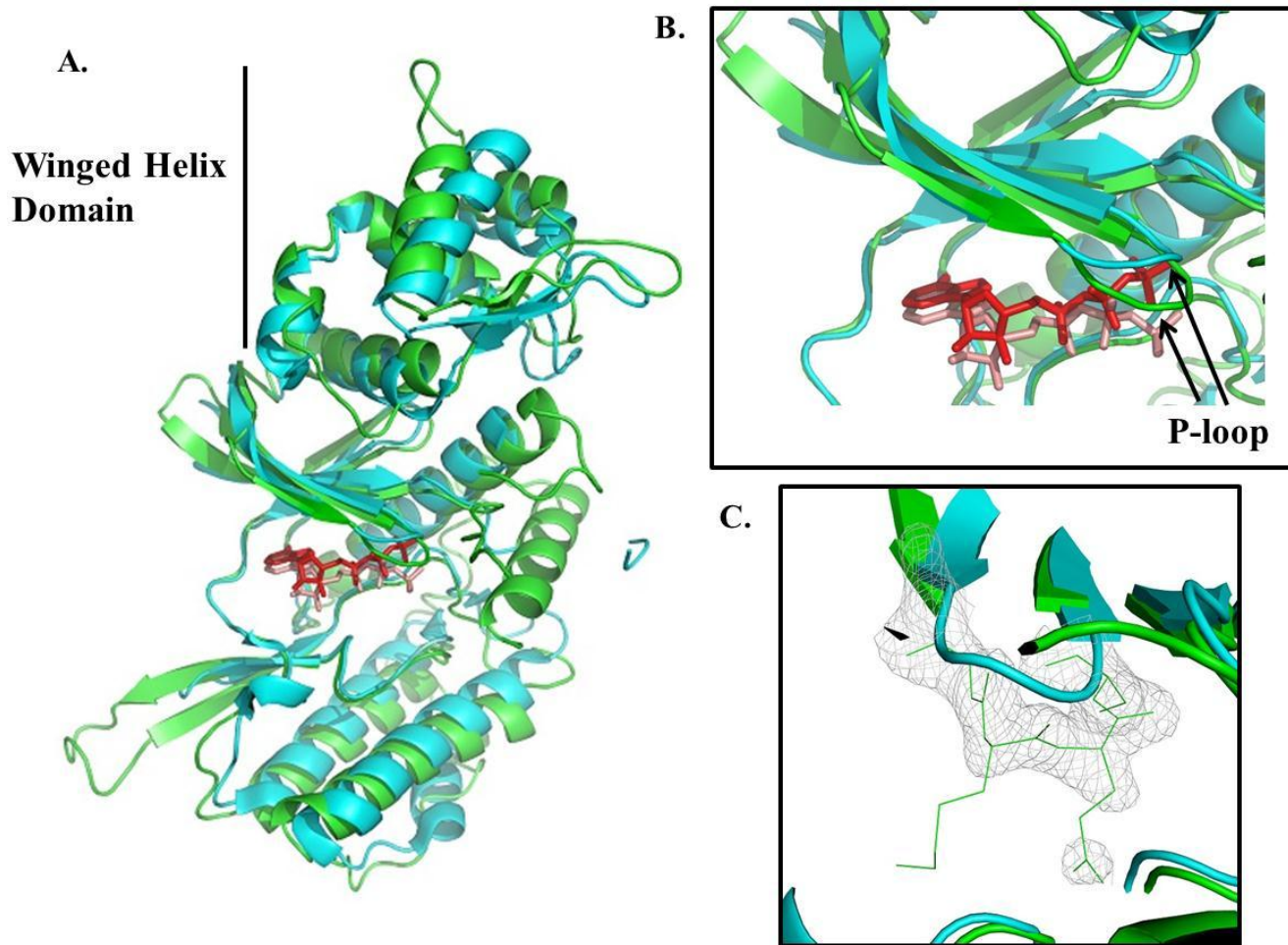
Figure 5.13: Structural Alignment of CtRio2 and AfRio2 and Positioning of the P-loop.

AfRio2(cyan) and ctRio2(green) in complex with ATP are structurally aligned. Rmsd is 2.204A for 217 atoms. ATP for afRio2 is red and for ctRio2 is pink. A. Wing helix domain is indicated. B. Close-up of the P-loop. C. 2Fo-Fc map of the P-loop of ctRio2 is contoured at 1σ.
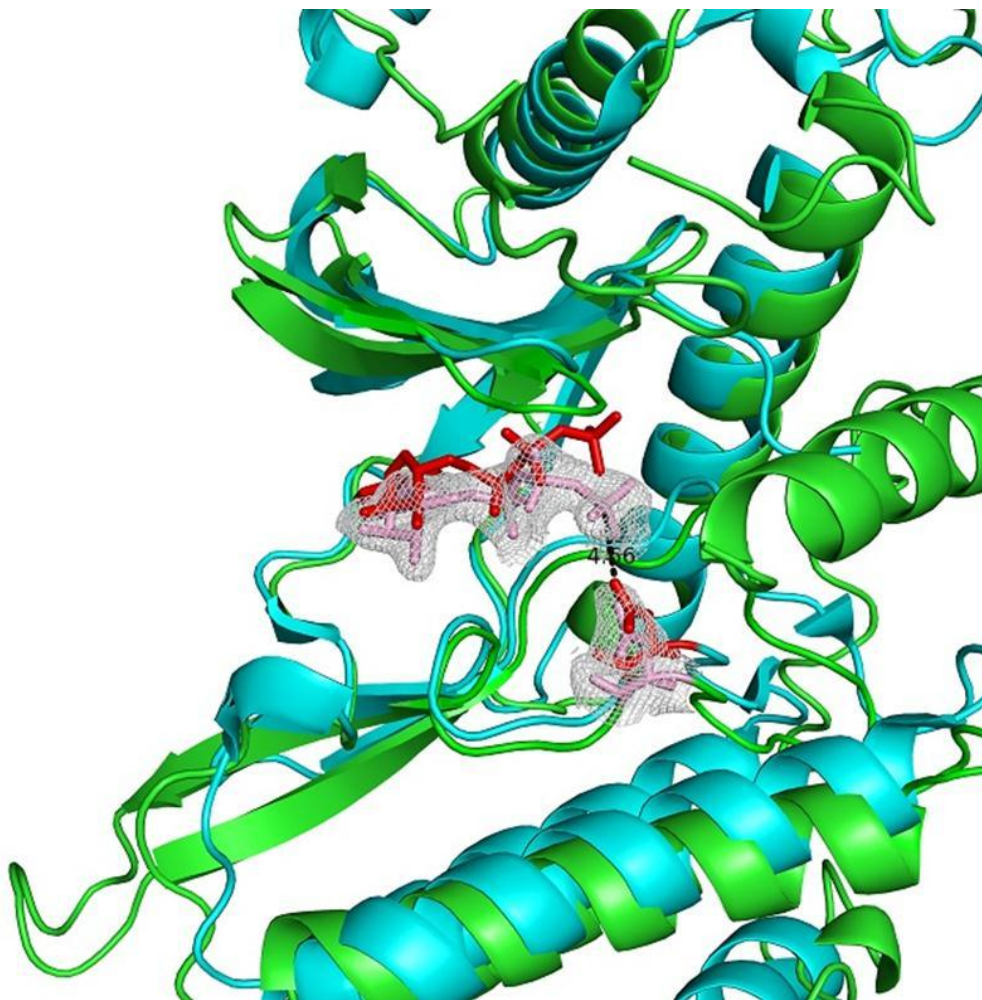
Figure 5.14: Alignment of CtRio2 and AfRio2 and Positioning of the γ Phosphate.

AfRio2(cyan) and ctRio2(green) in complex with ATP are structurally aligned. Rmsd is 2.204A for 217 atoms. ATP for afRio2 is red and for ctRio2 is pink. The γ phosphate of ctRio2 is 4.66 Å from the catalytic D232. 2Fo-Fc map of ATP and D232 of ctRio2 is contoured at 1σ.

Winged Helix Domain

N-Terminal Lobe of the Kinase Domain

C-Terminal Lobe of the Kinase Domain
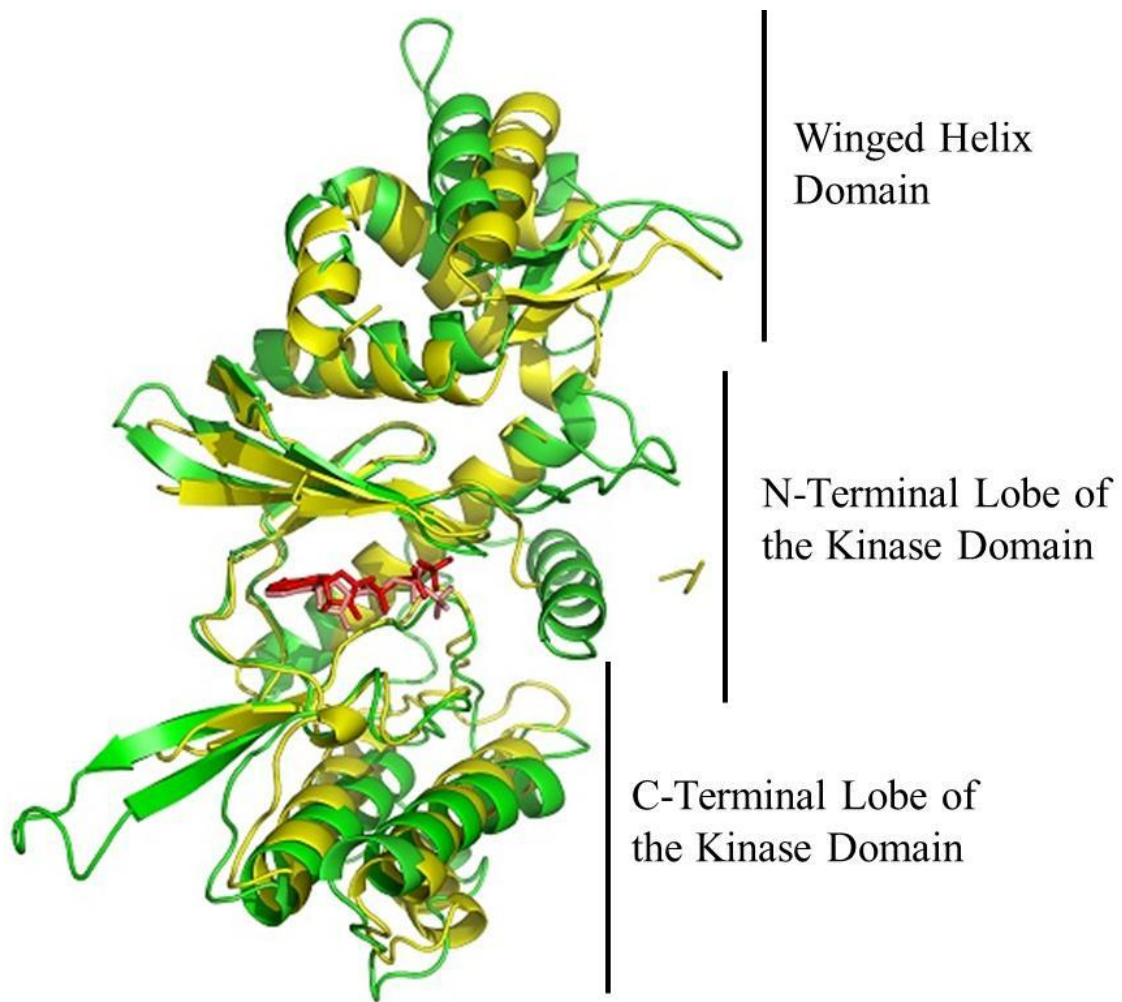
Figure 5.15: Best Fit Alignment Between CtRio2 and AfRio2

AfRio2(yellow) and ctRio2(green) in complex with ATP are structurally aligned using Swiss PDB Viewer. Alignment is for the best fit fragment between the two proteins. ATP for afRio2 is red and for ctRio2 is pink.
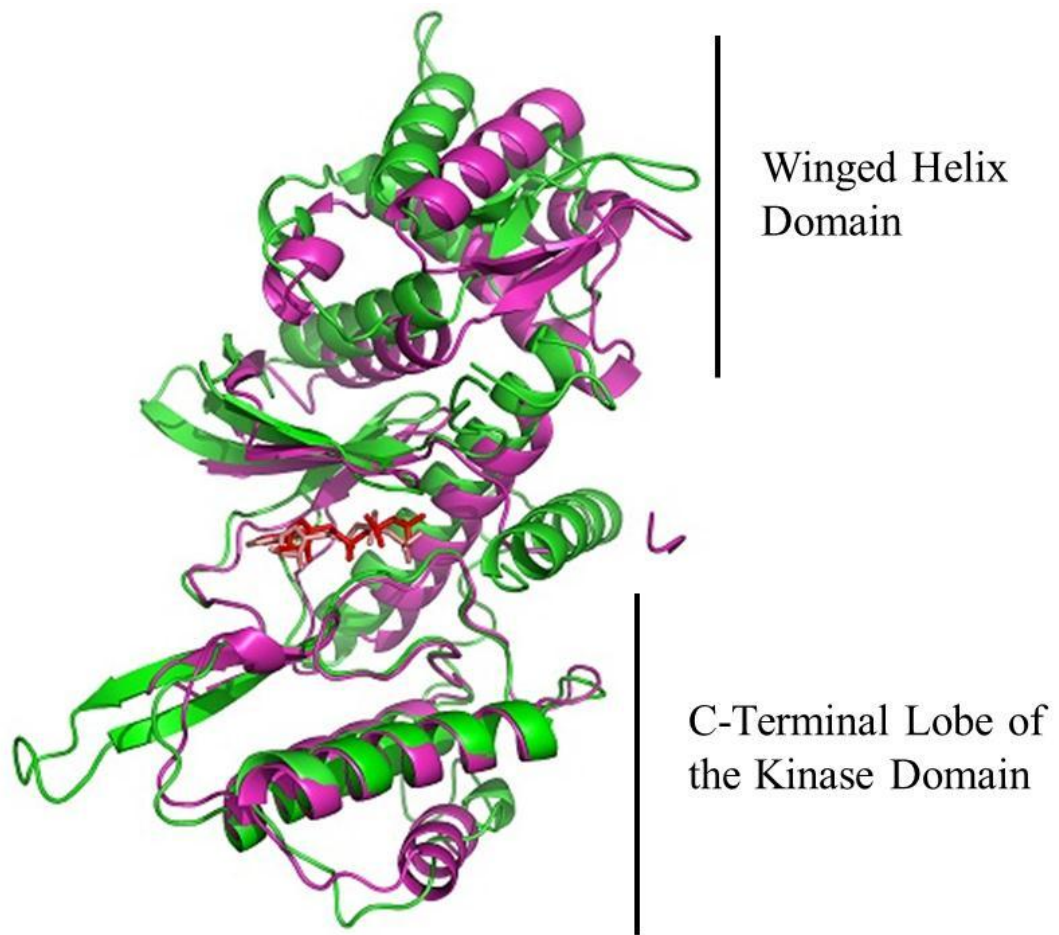
Figure 5.16: C-Terminal Lobe Alignment Between CtRio2 and AfRio2

AfRio2(magenta) and ctRio2(green) in complex with ATP are structurally aligned using Swiss PDB Viewer. Alignment is for the C-terminal lobe between the two proteins. ATP for afRio2 is red and for ctRio2 is pink.

Deconvoluted spectrum
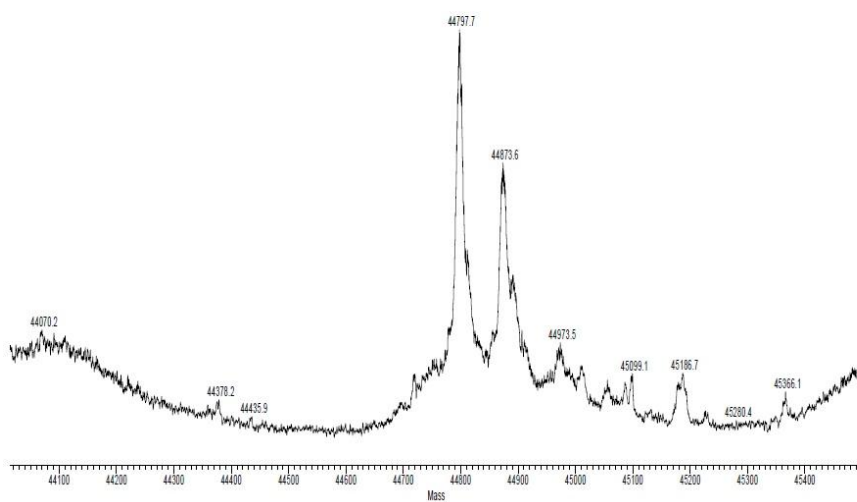
Figure 5.17: Mass Spec of CtRio2

1 mg/mL ctRio2 was incubated with 1mM ATP and 2.5 mM MgCl$_2$ overnight at room temperature. LC/MS was taken of the sample. The expected molecular weight of ctRio2 is 44799kDA. The difference between the two peaks seen in the chromatogram is 75.9 Da or one phosphate group.
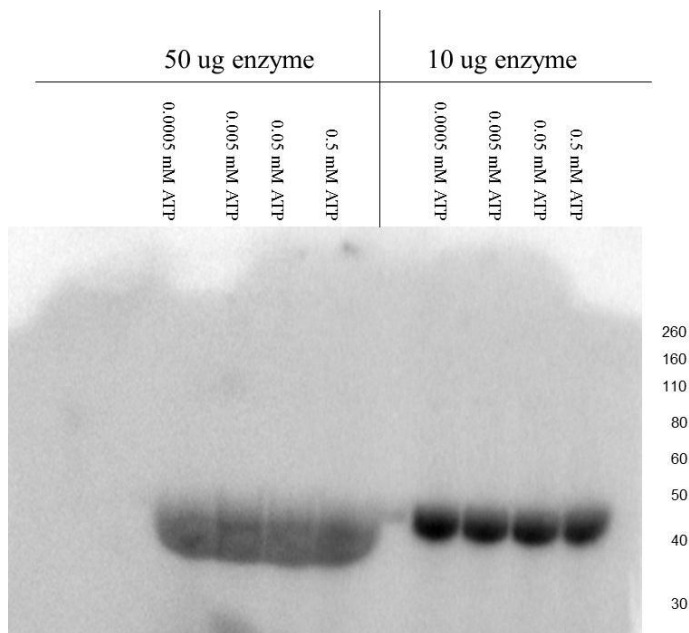
Figure 5.18: Autophosphorylation of CtRio2
10 μg or 50 μg of ctRio2 was incubated with 10 μCi γP^{32} labeled ATP and various concentrations of cold ATP and incubated at 37 °C for one hour. The samples were loaded on a SDS-PAGE gel, dried and exposed overnight on a Phosphor imaging plate.
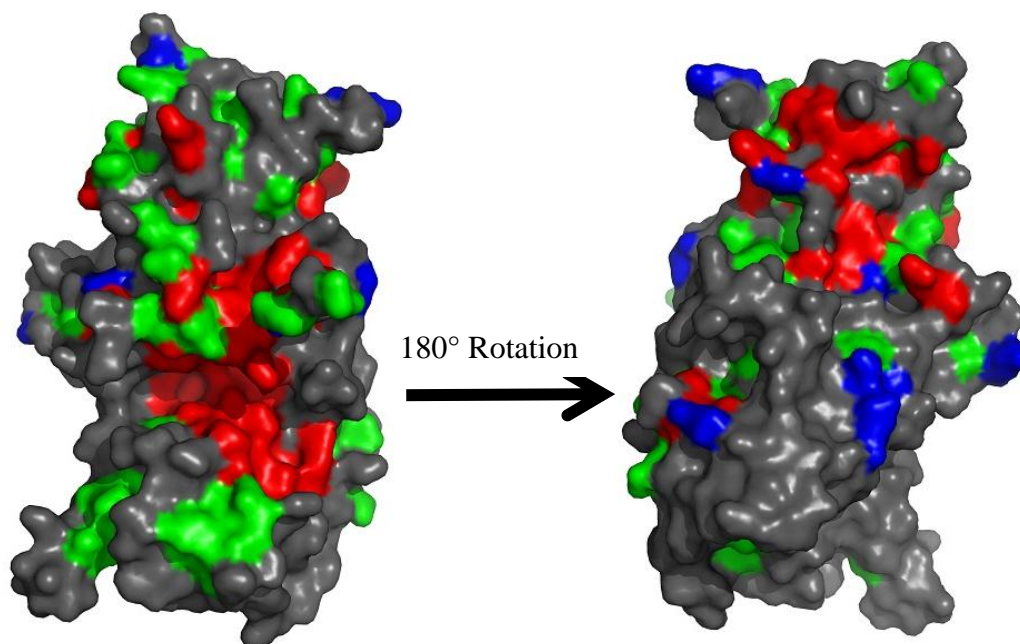
Figure 5.19: Surface Representation of Conserved Resides on CtRio2
A surface representation of ctRio2 is shown with identical, highly similar and weakly similar residues indicated in red, green, and blue colors respectively.

Chapter 6: Role of the Winged Helix Domain in Binding to the Ribosome

6.1: Domain Contribution to Rio2 Ribosome Binding

Previous research has shown that Rio2 binds tightly to the pre-40S ribosome. Rio2p is stably associated with the 43S preribosome in the nucleus and cytoplasm concomitantly with other late 40S processing factors and ribosomal proteins (*35-37*). Sucrose gradients of yeast expressing TAP-tagged Rio2p were performed to analyze ribosome profiles. In sucrose gradients used to analyze ribosome profiles in yeast containing TAP tagged Rio2p, the 40S and 43S fractions were isolated and a TAP tag pull down of Rio2p showed that Rio2p was associated with other late 40S factors including Dim1p, Tsr1p, Enp1p, Hrr25p, Nob1p, and Yor145p. The TAP tag pull down of Rio2p also showed the presence of many small subunit Rps proteins (*35-37*), characterizing Rio2p as a marker for late 40S pre-subunits.

The results of the earlier work done by Hurt and others (*35-37*) were used as the basis for experiments performed in this Chapter. Since we know that afRio2 consists of at least the N-terminal winged helix domain (wHTH) and the C-terminal RIO domain (Figure 1.2), I was interested in determining whether one domain, or the other, or both were involved in binding of Rio2p to the ribosome. To address this question, I used tap-tagged Rio2p constructs containing various domains or point mutations to pull-down and assay ribosome binding.

6.2: Development of an Assay to Determine if Rio2p Constructs can Pull Down the Ribosome

An assay was developed to determine if Rio2p was attaching to the ribosome. The constructs of Rio2p full length (FL Rio2p), Rio2p minus the winged helix domain (Short Rio2p), and the winged helix domain alone (Rio2p WH) were cloned into

plasmids containing a C-terminal TAP tag and a HIS3 gene for histidine production to select for the plasmid. The Short Rio2p consisted of FL Rio2p missing the first 78 amino acids. Rio2p WH consisted of the first 103 amino acids. The TAP tag consisted of C-terminal protein A followed by a TEV cleavage site and a calmodulin binding protein construct (CBP) (Figure 6.1).

The protein A tag allows the TAP tag fusion protein to bind to rabbit IgG beads. The TEV cleavage site allows the protein A tag to be cleaved so that Rio2p can be released from the column. The CBP tag allows for further purification of Rio2p by calmodulin beads. However, in this experiment the TEV cleavage site and the CBP tag were not used since my experiments showed that these were ineffective in purifying Rio2p from the supernatant in sufficiently large quantities. These three constructs were transformed into yeast cells containing GST tagged Rps18a on a separate plasmid containing the URA3 gene, conferring the ability to produce uracil. Rps18a is a constituent of the 40S small ribosomal subunit. The Rio2p constructs were pulled down by their Protein A tag via rabbit IgG beads. Rps18a was pulled down by Rio2p if the 40S preribosome in which Rps18a was incorporated also contained the Rio2p construct. The IgG beads bound to Rio2p were then boiled with SDS loading dye to remove the protein bound to the beads. The proteins extracted into the SDS dye were run on a gel. The GST tag of the pull down small ribosomal protein Rps18a was detected by performing a Western blot using a polyclonal antibody against the GST tag. In addition to binding the GST tag, the rabbit polyclonal antibody bound the protein A portion of the TAP tag. It allowed TAP-tagged Rio2p constructs to be seen on the Western blot in the lysate, supernatant, and pull down elution fractions.

The Western blot was analyzed for the presence of the GST tagged protein to show whether Short Rio2p and FL Rio2p could bind to the ribosome (Figure 6.2). The negative control for this assay was the yeast strain carrying the GST tagged Rps18a plasmid, which was needed to determine whether IgG beads alone could pull down Rps18a. The FL Rio2p was used as the positive control to pull down the 40S pre-ribosome and hence Rps18a.

6.3: Procedures

6.3.1: Transformation of Rio2p Constructs into Yeast Cells Containing GST Tagged Rps18a.

The plasmid containing the GST tagged Rps18a and the URA3 gene pre-transformed into a yeast cell was purchased from Operon (Cat # 166-044-C-10 with name YDR450W). The genotype of the yeast strain was MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0. The Rps18a gene was under the control of the galactose promoter. The Rio2p constructs were cloned into pAG413GAL-ccdB-TAP. This plasmid contained a HIS3 gene. In this plasmid the transcription of the inserted gene is under control of the galactose promoter. The plasmids were transformed into the YDR450W cells using the high efficiency yeast transformation LiAc/SS carrier DNA/PEG method by Geitz, et al. (*55*).

Yeast cells were grown in 5 mL YPAD media (rich yeast media) for 12-16 hours. Then 2.5×10^8 cells were added to 50 mL of pre-warmed 2×YPAD media. Cell titer was determined by taking O.D. at 600 nm. An O.D.600 reading of 0.1 corresponds to 1×10^6 cells/mL. The titer of the inoculation solution should be 5×10^6 cells/ml. The flask was incubated while shaking at 30°C and 200 rpm for about 4 hours or until the cell titer was at least 2×10^7 cells/mL. A 1.0 mL sample of 2 mg/mL salmon sperm DNA was boiled

for five minutes and chilled on an ice bath. The yeast cells were harvested by

centrifugation at 3000g for 5 minutes and resuspended in 25 mL sterile water. The cells

were again centrifuged at 3000g for 5 minutes to pellet the cells. This wash was repeated

with 25 mL sterile water. Then the cells were resuspended in 1.0 mL of sterile water.

The cell suspension was transferred to a 1.5 mL microcentrifuge tube and centrifuged for

30s at 13000g. The supernatant was discarded and the cells were resuspended in 1.0 mL

of sterile water. 100 µl of cell solution containing 10^8 cells were pipetted into 1.5 mL

microcentrifuge tubes, one for each transformation. The cell solution was centrifuged at

13000g for 30s and the supernatant was removed. 240 µl of 50% (w/v) PEG 3350, 35 µl

1.0 M Lithium Acetate, 50 µl 2.0 mg/mL salmon sperm DNA, 34 µl plasmid DNA in

sterile water was added to each cell pellet. The plasmid DNA was approximately 1 µg.

The tubes were vortexed for resuspension and then incubated for 40 minutes at 42ºC.

The tubes were then centrifuged at 13000g for 30s and the supernatant was removed. 1.0

mL of sterile water was pipetted into each transformation tube and the cell pellet was

resuspended. 20 µL and 200 µl of the transformation mixture were plated onto agar

plates made with synthetic complete media lacking uracil and histidine.

6.3.2: Procedure for Pull Down Assay of Rps18a.

The basic procedure for pull down of Rps18a from yeast cells is described in this

section. A yeast culture in 5mL of minimal media minus uracil and histidine was grown

overnight at 30°C. Next day, one mL of this culture at an O.D. 600 of 5.0 was used to

inoculate 2 L of minimal media minus uracil and histidine. The cultures were grown in

flasks at 30°C with shaking at 220 rpm. When the O.D. reached 2.0, the flasks were

removed from the shaker and centrifuged at 4000 rpm for 20 minutes to sediment the

yeast cells. The yeast cells were then washed with 500 mL of cold distilled water. The

cells were aliquoted into 15 mL falcon tubes so that each tube contained 0.5 g of dry

yeast cell pellet. These falcon tubes were frozen at -80°C.

One tube of cells was thawed and resuspended in 10 mL sterile Buffer A (10 mM

K-HEPES, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM PMSF, 2 mM

Benzamidine, ¼ EDTA free protease inhibitor tablet). To Buffer A was added 20 μl

Fisher Brand Optizyme Rnase Inhibitor, Porcine (33,300 U/ml). The suspension was

passed twice through a French Press at 1200 psi. 2 M KCl was added to the lysate to

bring the final concentration to 0.2 M KCl. A 200 μl sample of the lysate was taken and

mixed with 80 μl 5x SDS loading dye and boiled for a Western blot sample. The lysate

was centrifuged at 20.5 Krpm for 30 minutes at 4°C. The supernatant was transferred

into clean 45Ti tubes and centrifuged again at 33.5 Krpm for 1 hour 24 minutes. The

supernatant was then dialyzed at 4°C for two hours against 2 L of Buffer D (20 mM K-

HEPES, 50 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol, 0.5 mM PMSF, and 2

mM Benzamidine). The volume of the supernatant was reduced in half after dialysis. 5

mL of the dialysis buffer was added to the supernatant to return the volume back to 10

mL. A 200 μl sample of the supernatant was taken and mixed with 80 μl 5x SDS loading

dye and boiled for a Western blot sample.

The supernatant was frozen at -80°C. The next day the supernatant was thawed.

A 200 μl sample of the supernatant was taken and mixed with 80 μl 5x SDS loading dye

and boiled for a Western blot sample. To a small Bio-Rad column was added 400 μl of

50% suspension of Rabbit IgG-Agarose beads from Sigma (A2909). The beads were

washed with 5 mL of sterile IPP150 buffer (10 mM Tris-Cl pH 8.0, 150 mM NaCl, and

0.1% NP40). Fifty μl 2 M Tris-Cl pH 8.0 (10 mM final concentration), 200 μl 5 M NaCl

(150 mM final concentration), 100 μl 10% NP40 (0.1% final concentration) was added to

the supernatant. The beads were then added to the 10 mL supernatant in the 15 mL

Falcon tube and shaken on a rocking platform at 4°C for two hours. The bead and

supernatant solution was then transferred back to the Bio-Rad column and the flow-

through was allowed to drain out. 200 μl of the flow-through was retained and mixed

with 80 μl of 5x SDS loading dye and boiled for a Western blot sample. The beads were

washed with 40 mL IPP150 buffer and the last 10 mL of wash was collected. 200 μl of

sample was removed and mixed with 80 μl of 5x SDS loading dye and boiled for a

Western blot sample.

One mL of IPP150 buffer was added to the beads in the Bio-Rad column.   A

suspension was made and it was transferred to a 1.5 mL eppendorf tube. The suspension

was centrifuged at 1000 rpm for 1 minute to settle the beads. The IPP150 buffer was

removed from the top of the beads and 100 μl of 5x loading dye was added to the beads.

The beads were vortexed and then boiled for ten minutes. The beads were vortexed again

and then centrifuged at 13000 rpm for one minute to settle the beads. The top dye was

removed and put into a separate 1.5 mL eppendorf tube and marked eluate.

Polyclonal antibody against the GST tag was used for the Western blotting. The

antibody was purchased from Santa Cruz Biotechnology, Inc. (Cat.# SC-459 HRP). The

antibody was a primary antibody recognizing the GST tag directly, and was conjugated to

horseradish peroxidase enzyme, which produced light in the presence of

chemiluminescent substrate. Dilutions of the samples of lysate, supernatant, wash, and

eluate were run on Invitrogen NuPage 4-12% Bis-Tris Gels, 15 well alongside a sample

of Invitrogen "Novex Sharp Prestained Protein Standards" ladder. The gel was then

transferred to a PVDF membrane (Millipore Immobilon-P membrane Cat No.

lSEQ07850) at 30V for 1 hour.   The ladder was completely transferred to the PVDF

membrane after transfer.   The PVDF membrane was blocked with 1% blocking solution

of powdered skim milk in 1X TBST buffer (25 mM Tris, pH 7.5, 100 mM NaCl, 0.05%

Tween 20). Blocking took place for one hour at room temperature on a shaking platform.

The blocking solution was removed and a 1/200 dilution of antibody in blocking solution

was added to the PVDF membrane. The antibody was incubated with the membrane

while rocking for one hour. The membrane was then rinsed with 40 mL TBST and

washed 4×2 minutes with 40 mL of TBST. Excess TBST was drained by placing one

edge of the membrane on a Kimwipe. Then the membrane was incubated with "Lumi-

Light Western Blotting Substrate" (Cat. No. 12015200001) while rocking for five

minutes.

The membrane was exposed for one minute in a dark box. The antibody was

conjugated to HRP, which in the presence of chemiluminescent substrate produced light.

A photo was taken. A photo was also taken of the protein standards ladder with the

lightbox on.

6.4: Results of First Three Constructs Made in Pull Down Assay

Initially three constructs were cloned—FL Rio2p, Short Rio2p and Rio2p WH.

Samples of the lysate, supernatant, thawed sample, flow-through, wash, and elution were

diluted for each of the constructs before running on a Western blot to maximize the signal

seen for the pull down protein, Rps18a, without overloading the Western blot. GST

tagged Rps18a has a molecular weight of 43kDa, TAP tagged FL Rio2p has a molecular weight of 95kDA and TAP tagged Short Rio2p has a molecular weight of 86kDa.

The Western blot showed that FL Rio2p was more soluble than Short Rio2p, which showed only a faint band in the supernatant fraction (Figure 6.3). No protein eluted in the flow through and wash of FL Rio2p and Short Rio2p, indicating that entire protein in the supernatant became bound to the column. The elution lane showed well-defined bands for the TAP tagged protein in both the Western blots of FL Rio2p and Short Rio2p. Both also had well-defined bands for the GST tagged Rps18a.

Rio2p WH was expressed at low levels in the yeast cell, so no dilutions of the lysate, supernatant, and elution were made. The Western blot shows that Rio2p WH was weakly expressed in the lysate and was slightly soluble in the supernatant as compared to the lysate (Figure 6.4). The elution showed a band for Rio2p WH and no band for the GST tagged Rps18a. The negative control was completely blank.

The experiment showed that the ribosome was equally pulled down with Short Rio2p as with FL Rio2p. Rio2p WH did not pull down GST-tagged Rps18a, perhaps due to low expression of Rio2p WH. The negative control did not show any signal for Rps18a, confirming that the Western blot band for Rps18a was authentic. The negative results with Rio2p WH suggested that the TAP tag alone was not interacting with and pulling down Rps18a.

From the data we can draw the following conclusion: Though the wHTH may make direct contacts with the ribosome, the RIO domain or the C-terminus may make equally strong or stronger contacts with the ribosome. The wHTH does have all of the interactions between Rio2p and the 40S preribosome.

The exterior surface of the wHTH was positively charged in Rio2 suggesting that the wHTH interacted with the negatively charged phosphate backbone of RNA. The closest structures found with the DALI server to wHTH were the transcription factors MarR and SlyA and histone linker protein GH5 (*34*). All of these are known to interact directly with DNA. Crosslinking experiments have shown that Rio2p binds to a small part of the rRNA of the ribosome close to the site D cleavage (*114*). However, it is not known if the wHTH of Rio2p makes these contacts to the rRNA or some other part of the RIO domain binds to the rRNA sequence. The true extent of the rRNA contacts are not known. The crosslinking experiment may represent only the partial binding site of Rio2p to the rRNA. It may be possible that the wHTH does not contact rRNA and may simply interact with the protein complex of the preribosome or have some altogether different function.

6.5: Results of Point Mutations Made in the Rio2p Pull Down Assay

The structure was reexamined to determine what other portions of Rio2p might bind to the rRNA sequence of the 40S preribosome. The structure showed patches of positively charged surface containing conserved positively charged residues. Three conserved positively charged residues were chosen to be mutated in Rio2p. The residues in AF Rio2 corresponded to K102, R253, and R264 (Figure 6.5) and are conserved surface residues among eukaryotes (Figures 6.6).

These conserved residues were mutated to alanine to see whether they were involved in the binding of Rio2p to the rRNA of the ribosome. Three residues (K105A , R271A, and K282A) were mutated in FL Rio2p (FL R271A, FL K282A, and FL K105A) and Short Rio2p (Short R271A, Short K282A, and Short K105A) to determine pull down

of the 40S preribosome in either construct. The fractions of lysate, supernatant, thawed

sample, flow-through, wash, and elution were diluted for each mutated construct to

visualize the pull down of Rps18a. In some cases the fractions had to be diluted

extensively to prevent saturation of the Western blot with the signal for Rps18a. In other

cases the fractions required less dilution because Rps18a was present only in small

amounts.

The Western blot showed that Short K105A was only weakly expressed in the

lysate, requiring no dilution of fractions (Figure 6.7). FL K105A and Short K105A were

not very soluble. They showed a faint band or no band in the supernatant fraction. No

protein eluted in the wash. The elution showed a strong band for the TAP tagged protein

in both the Western blot of the FL K105A and the Short K105A. Only FL K105A

showed a well-defined band for the GST tagged Rps18a.

The Western blot showed that Short R271A was only weakly expressed in the

lysate. It required 1/5 dilution of fractions (Figure 6.80). FL R271A and Short R271A

were equally soluble and showed bands in the supernatant fraction that were equal in

strength to the lysate fraction. No protein eluted in the wash. The elution showed a

strong band for the TAP tagged protein in both the Western blot of the FL R271A and the

Short R271A. Only FL R271A showed a well-defined band for the GST tagged Rps18a.

The Western blot showed that Short K282A was weakly expressed in the lysate,

requiring no dilution of fractions (Figure 6.9). FL K282A was very soluble. It showed a

band in the supernatant fraction that was equal in strength to the band in the lysate

fraction. No protein eluted in the flow through and wash indicating that all the protein in

the lysate did bind. The elution showed a strong band for the TAP tagged protein in both

the Western blot of FL K282A and Short K282A. Both FL K282A and Short K282A showed weak bands for the GST tagged Rps18a.

A Western blot with all supernatants from different constructs was run on the same gel. The supernatant was diluted in order to allow the band of the tagged protein to be seen at about equal levels. A separate Western blot was run with all elutions for different constructs of the GST tagged Rps18a at relatively same dilution as the supernatants. The elutions from different constructs could then be compared and contrasted (Figure 6.10).

FL Rio2p, Short Rio2p, FL K282A, and Short K282A pull down the 40S preribosome. FL K105A pulled down pre-40S ribosome weakly. Short K105A was barely soluble, so it was difficult to determine if it pulls down the pre40S ribosome. However, Short R271A though soluble, does not pull down the ribosome at the dilution chosen to compare with the other constructs. Short Rio2p and Short R271A have similar amounts of protein in the supernatants. However, Short Rio2p pulled down Rps18a while Short R271A does not. In the earlier Western blot, the FL R271A construct did pull down the pre-40S ribosome at 1/5 dilution. This would indicate that FL R271A pulled down the pre-40S ribosome most weakly of all the constructs.

R271 is analogous to R253 in Rio2 from afRio2. This residue is part of α helix F or the second to last helix in the sequence. Helix F is part of the C-terminal lobe of AF Rio2 and contains three other positively charged arginines that are not part of the conserved sequence. In the yeast sequence this helix would have five other positively charged residues. This makes Helix F ideal for making contacts with the negatively charged backbone of rRNA. However, whether it actually interacts with the rRNA is not

known. The point mutation does not appear to prevent Rio2p from folding properly as FL R271A and Short R271A are soluble and show up in the supernatant.

A true test of the proper folding of the protein would be to analyze its serine kinase enzymatic activity. Since AF Rio2 is shown to autophorphorylate, similar activity most likely exists with Rio2p. Though the natural peptide substrates of Rio2p are not known, it can be tested for its ability to autophosphorylate.

6.6: Conclusions of Rio2p TAP Tag Pull Down

A TAP tag pull down assay was performed with different constructs of Rio2p to determine if the winged helix domain was necessary for binding of Rio2p to the ribosome. Without the winged helix domain, Rio2p still bound to the ribosome so it is not absolutely required for ribosomal binding. Three different point mutations were made to FL Rio2p and Short Rio2p to determine what specific residue might interact with the ribosome.   R271 was implicated in binding of Rio2p to the ribosome. R271 is a surface residue strictly conserved among eukaryotes (Figure 6.6). It might be important for the binding of Rio2p to other ribosomal processing factors, to ribosomal proteins, or to the rRNA of the 40S preribosome.

N'	Rio2p	CBP	TEV Cleavage	Protein A	C'

Figure 6.1: TAP tagged Rio2 Construct
A schematic of the C-terminal TAP tag is shown. It consists of calmodulin binding peptide (CBP)
followed by a TEV cleavage site and a protein A tag.
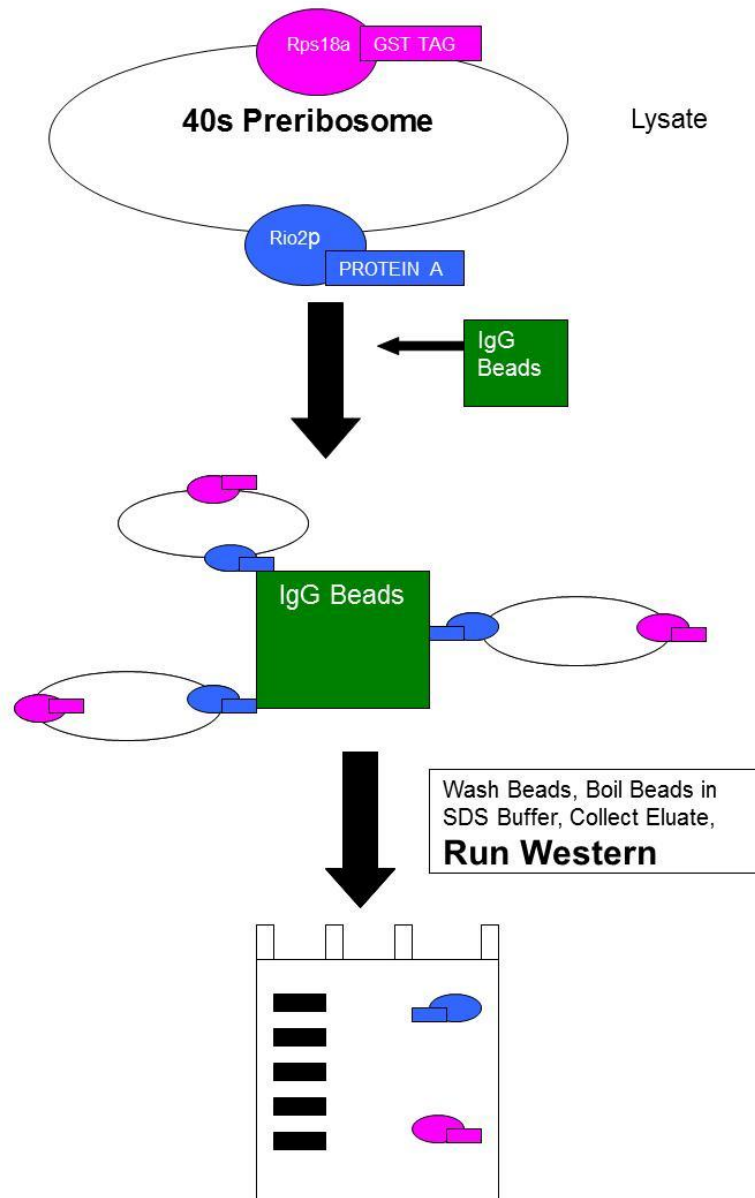
Figure 6.2: Assay for Pull Down of Ribosome
This figure illustrates the pull down assay with Rio2. Both Rps18a with a GST tag (pink) and TAP tagged Rio2 (blue) attach to the ribosome and and are pulled down together with IgG beads(green) which bind to the TAP tag. After the pull down, the IgG beads are boiled in SDS buffer, and the eluate is Western blotted for the GST tag.
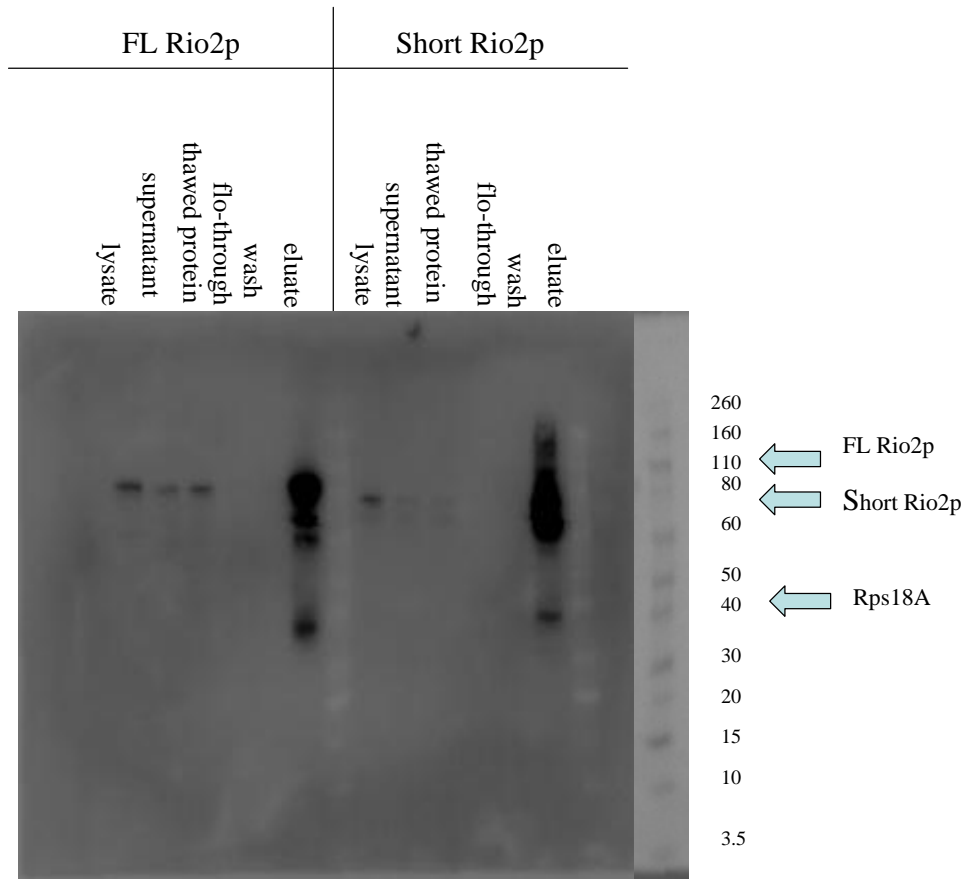
Figure 6.3: Tap Tag Pull Down of Rio2p
Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed (lysate), centrifuged to remove the supernatant, frozen and thawed the next day. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluate"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. This figure represent the fractions of FL Rio2p TAP tag pull down at 1/20 dilutions. This figure also represents the fractions of Short Rio2p TAP tag pull down at 1/20 dilutions.
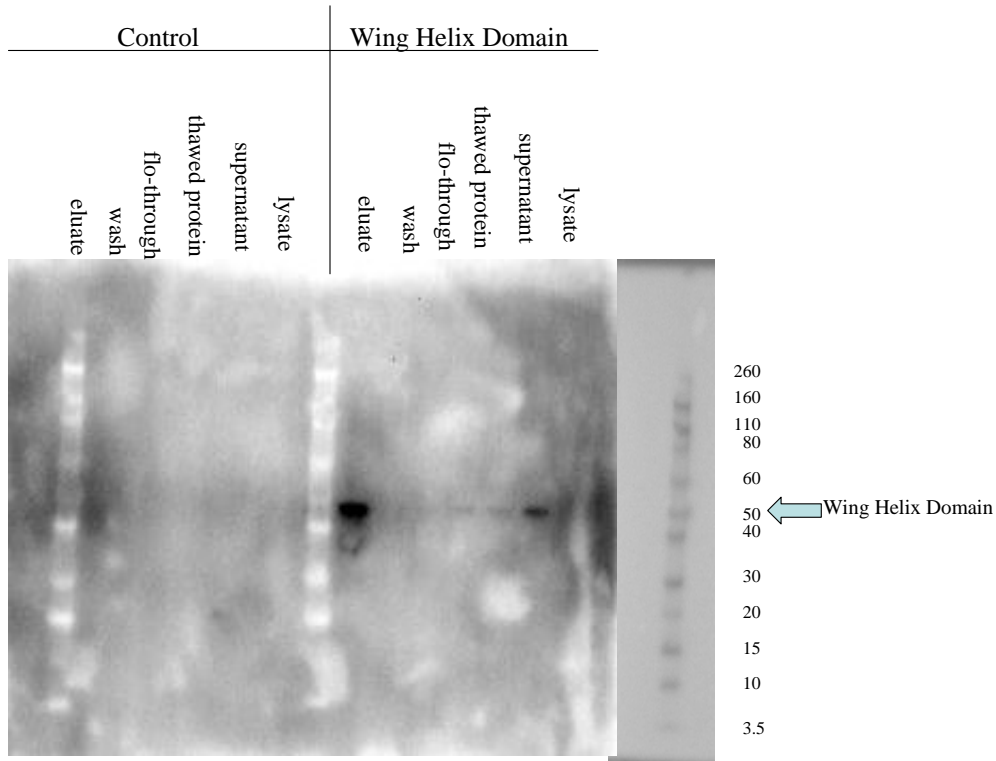
Figure 6.4: TAP Tag Pull Down of Wing Helix Domain and Negative Control
Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed (lysate), centrifuged to remove the supernatant, frozen and thawed the next day. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluate"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. This figure represent the fractions of the Rio2p WH TAP tag pull down at no dilution. This figure also represent the fractions of the negative control pull down at 1/20 dilution.
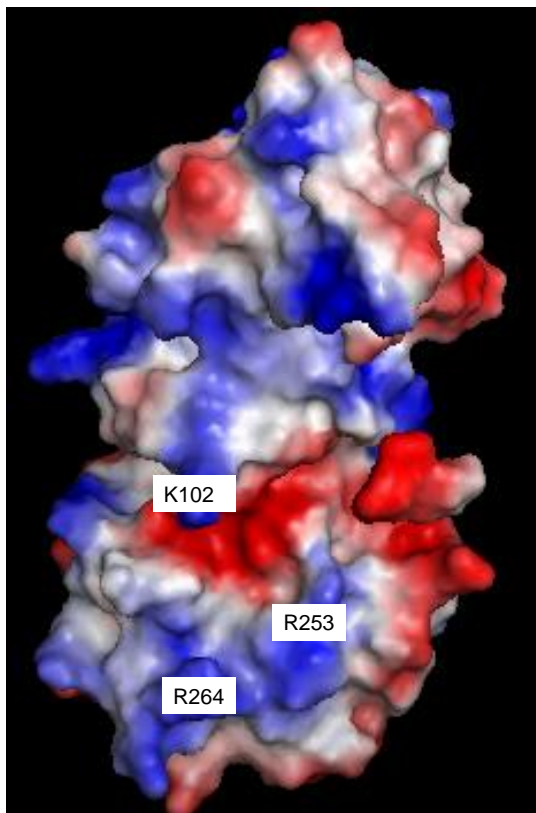
Figure 6.5: Electrostatic Properties of AfRio2

This figures represents the electrostatic properties of the surface of AF Rio2. Blue represent positively charged electron density and red represents negatively charged electron density. The three point mutations K102 (K105A in yeast), R253 (R271A in yeast), and R264A (K282A in yeast), all lie in the positively charged regions of the electron density.

```
              10         20         30         40         50         60
               |          |          |          |          |          |
ddRio2  -MKLDPKALRYLSKDDFRTLVAVEMGMKNHELVPVSLICTIANLKY--GGTKKSIQTLHK
atRio2  -MKLDVNVLRYLSKDDFRVLTAVEMGMRNHEIVPSELVERIACLKH--GGTYKVLKNLLK
rio2p   -MKLDTSHMRYLTTDDFRVLQAVEQGSRSHEVVPTPLIHQISGMRS-QSGTNRAISDLAK
ctRio2  -MKLDTRAMRHLTAEDWRVLTAVEMGSKNHEIVPTPLIEKIARLRGGSSGVHKSIATLAK
drrio2  MGKLNVVILRYLSRDDFRVLTAVEMGMKNHEIVPVSLIASIASLKH--GGCNKVLRELVK
xlrio2  MGKLNVLLLRYLSREDFRVLTAVEMGMKNHEIVPASLIASIASLKH--GGCNKILRELVK
hsrio2  MGKVNVAKLRYMSRDDFRVLTAVEMGMKNHEIVPGSLIASIASLKH--GGCNKVLRELVK
dmrio2  MGKLNVTVLRYLTKEDFRVLTAIEMGMKNHELVPGPLAAAIANLKS--GGVHKLLKELCK

              70         80         90        100        110        120
               |          |          |          |          |          |
ddRio2  FKLLFHD-GRNYDGYKLTYLGYDFLALKTLVSRGVCSYVGNQIGVGKESDIYIVANDDNQ
atRio2  YKLLHHD-SSKYDGFRLTYLGYDFLAIKTLVNRGIFTGVGRQIGVGKESDIFEVAQEDGT
rio2p   LSLISKMRNVKYDGYRLTYNGIDYLALKTMLNRDTVYSVGNTIGVGKESDIYKVSDKNGN
ctRio2  AGLIARMKEAKYDGYRLTYGGLDYLALHTHAARKDVYSVGSRIGVGKESDIMIVADEKGK
drrio2  HKLLAYERSKTVQGYRLNYGGYDYLALKTLSSRDVILSVGNQMGVGKESDIYIVANAEEQ
xlrio2  HKLIAYDRTKTVQGYRLIYGGYDYLALKTLSSREVVTSVGNQMGVGKESDIYIVANEDEN
hsrio2  HKLIAWERTKTVQGYRLTNAGYDYLALKTLSSRQVVESVGNQMGVGKESDIYIVANEEGQ
dmrio2  HKLLAYERGKKYDGYRLTNTGYDYLALKSLTLRGSVSSFGNQIGIGKESNIYVVADEEGT

             130        140        150        160        170        180
               |          |          |          |          |          |
ddRio2  EMVLKLHRLGRVSFKTIKNNRDYLKHRKSAS-WLYLSRLAALKEFAYMKALYENGFPVPT
atRio2  ILAMKLHRLGRTSFRAVKSKRDYLRHRSSFS-WLYLSRLAALKEFAFMKALEEHDFPVPK
rio2p   PRVMKIHRLGRTSFHSVRNNRDYLKKSNQGANWMHLSRLAANKEYQFMSMLYSKGFKVPE
ctRio2  QKVLKIHRLGRISFRTVKANRDYLRNRSTGS-WMHLSRLAAIKEFAFMKALYEEGFPVPE
drrio2  QFALKLHRLGRTSFRNLKNKRDYHKHRNNMS-WLYLSRLSAMKEFAYMKALYDRGFPVPK
xlrio2  QMALKLHRLGRTSFRNLKNKRDYHKHRHKMS-WLYLSRLAAMKEFAYMKALYDTGFPVPK
hsrio2  QFALKLHRLGRTSFRNLKNKRDYHKHRHNVS-WLYLSRLSAMKEFAYMKALYERKFPVPK
dmrio2  PICLKLHRLGRTCFRNVKAKRDYHGRRHKAS-WLYLSRISATREFAYMSALYDRGFPVPK

             190        200        210        220        230        240
               |          |          |          |          |          |
ddRio2  PIDYNRHCIVMSRARGYPLTQIVQLRHPSKVYSDLMNLIVKLASYGLIHGDFNEFNILIN
atRio2  AIDCNRHCVIMVQVK-------QLQNPETIFEKIIGIVVRLAEHGLIHCDFNEFNIMID
rio2p   PFDNSRHIVVMELIEGYPMRRLRKHKNIPKLYSDLMCFIVDLANSGLIHCDFNEFNIMIK
ctRio2  PIAQSRHTIVMSLVDALPMRQVSSVPDPASLYADLIALILRLAKHGLIHGDFNEFNILIR
drrio2  PVDYNRHAVVMELINGYPLCQVREIQDPAALYSEIMELIVKLANHGLIHGDFNEFNLMLD
xlrio2  PHDYNRHAVVMELINGYPLCQIRQLEDPASLYSELMDLIVKLANHGLIHGDFNEFNLMLD
hsrio2  PIDYNRHAVVMELINGYPLCQIHHVEDPASVYDEAMELIVKLANHGLIHGDFNEFNLILD
dmrio2  PIDFNRHCVLMDLVKGWPMTQVHELLDAPQVYDDLMNLIVRLGNSGVIHGDFNEFNLMVT

             250        260        270        280        290        300
               |          |          |          |          |          |
ddRio2  D----EEEITL------IDFPQMVSTSHLNAEMYFDRDVTCIRVFFEKRFNFIGERWPKF
atRio2  D----EEKITM------IDFPQMVSVSHRNAQMYFDRDIECIFKFFRKRFN---------
rio2p   DKLEDENDCGF----VVIDFPQCISIQHQDADYYFQRDVDCIRRFFKKKLK---------
ctRio2  EEKDAEDPSSITLTPIIIDFPQMVSMDHPNAEMYFDRDVQCIKRFFERRFH---------
drrio2  D----NDHVTM------IDFPQMVSTSHINAEWYFDRDVKCIRDFFIKRFNYESELYPTF
xlrio2  E----EDHVTM------IDFPQMVSTSHANAEWYFNRDVKCIRDFFLRRFNYESELYPTF
hsrio2  E----SDHITM------IDFPQMVSTSHPNAEWYFDRDVKCIKDFFMKRFSYESELFPTF
dmrio2  D----AGKPIL------IDFPQMMSTSHENAEFFFERDVNCVREMFRRKFGYESEDYPKF
```

Figure 6.6: Sequence Alignment of Eukaryotic Rio2 and Rio2p
Alignment was performed with the program CLUSTALW using sequences from *Dictyostelium discoideum*
(ddRio2), *Danio rerio* (ddRio2), *Homo sapiens* (hsRio2), *Sacchoromyces cerevisae* (Rio2p), *Arabidopsis
thaliana* (atRio2). *Chaetomium thermophilum* (ctRio2), *Drosophila melanogaster* (dsRio2), and *Xenopus
laevis* (xlRio2). Identical, highly similar, and weakly similar residues are indicated by red, green, and blue
letters respectively. Yellow highlighted residues represent the placements of the K105A, R271A, and
K282A mutations in yeast.

137

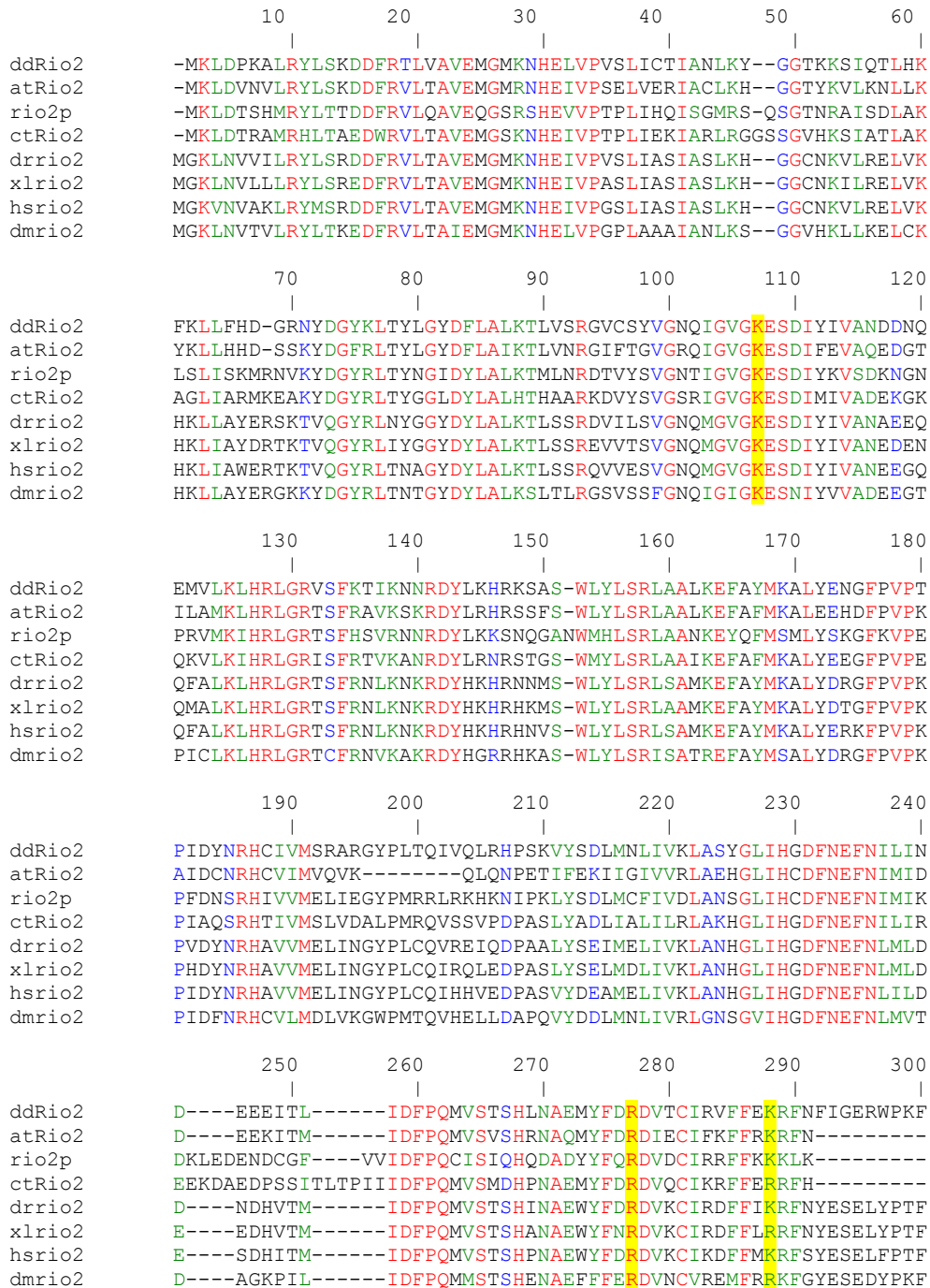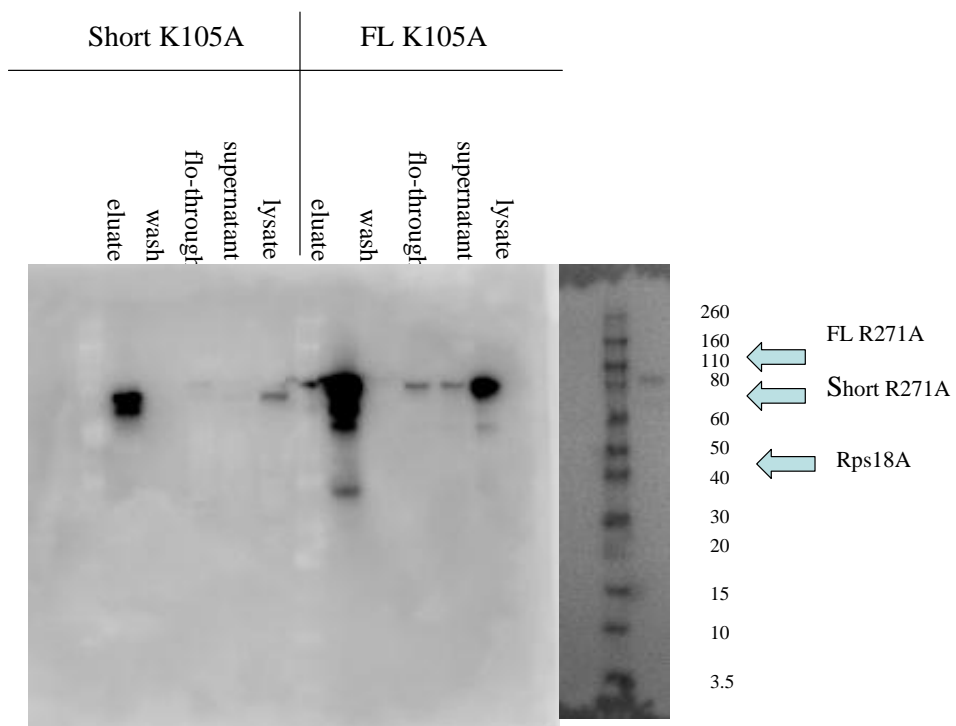**Figure 6.7: TAP Tag Pull Down of K105A Construct**

Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed (lysate), centrifuged to remove the supernatant, frozen and thawed the next day. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluate"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. The figure represent the fractions of FL K105A TAP tag pull down at 1/20 dilution. The figure also represent the fractions of Short K105A TAP tag pull down at no dilution.
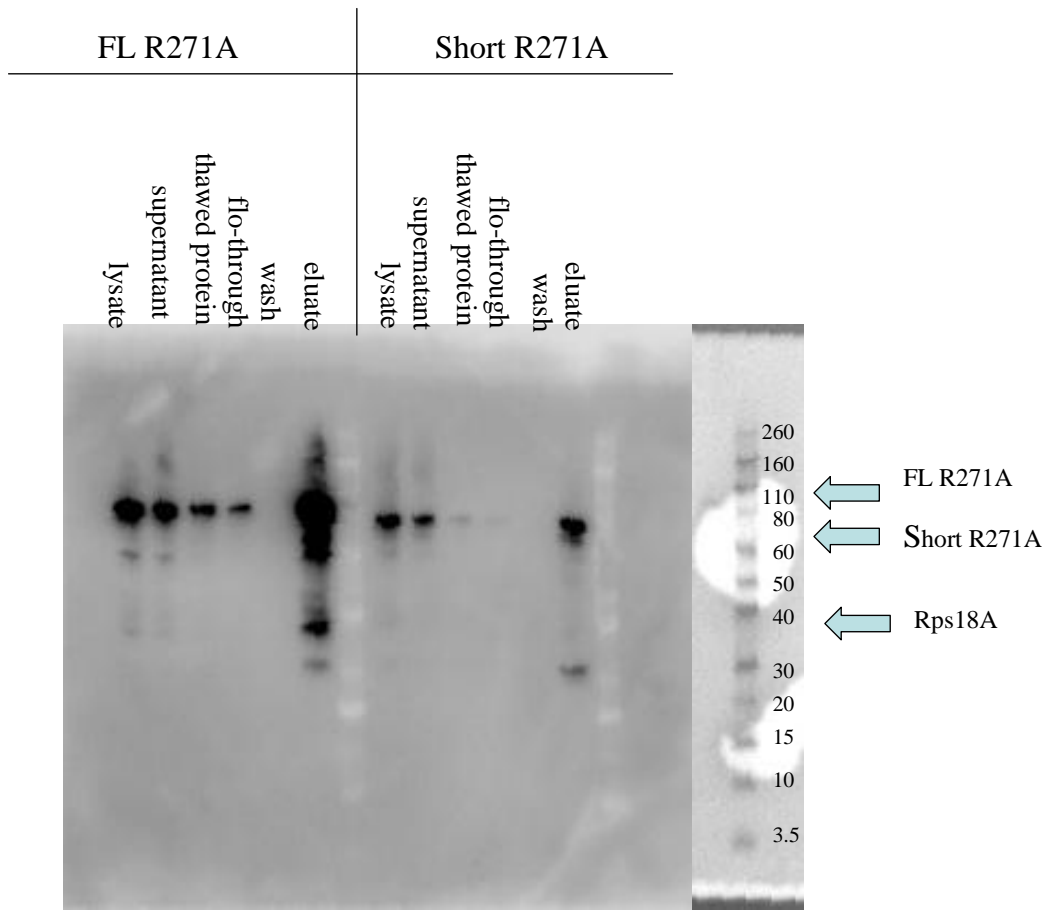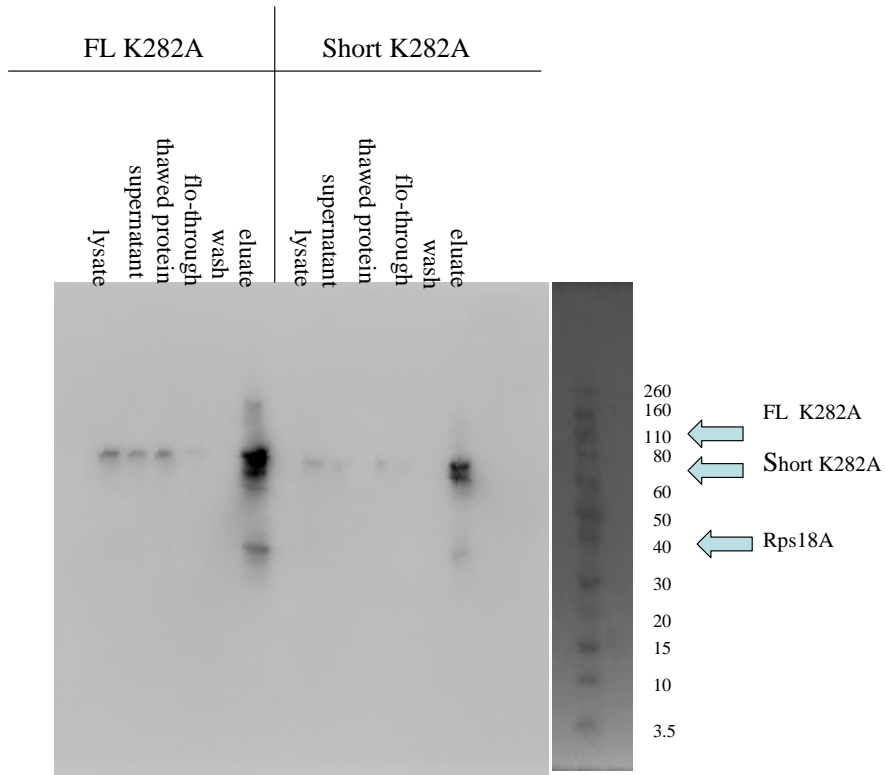
Figure 6.8: TAP Tag Pull Down of R271A Construct
Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed (lysate), centrifuged to remove the supernatant, frozen and thawed the next day. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluate"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. This figure represent the fractions of the Fl R271A TAP tag pull down at 1/5dilutions. This figure also represent the fractions of the Short R271A TAP tag pull down at 1/5 dilutions.
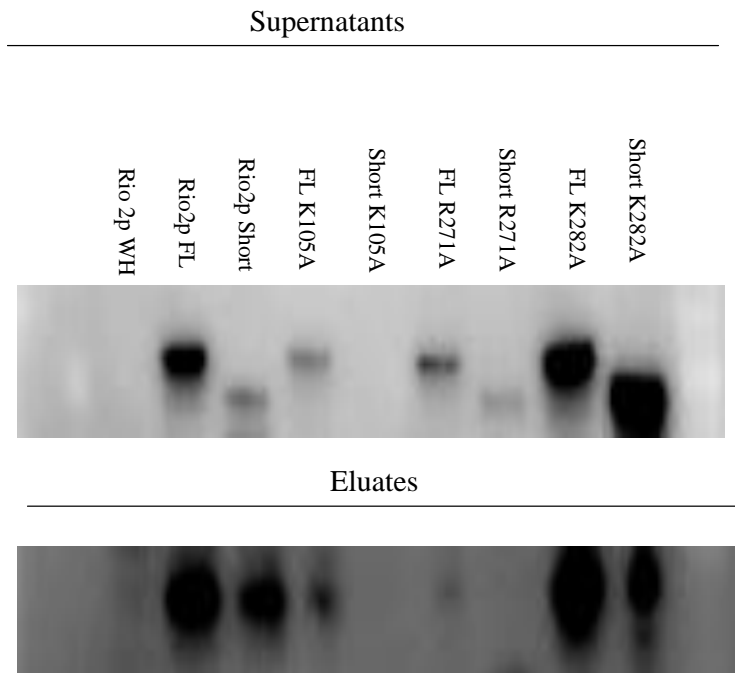
Figure 6.9: TAP Tag Pull Down of K282A Construct
Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed (lysate), centrifuged to remove the supernatant, frozen and thawed the next day. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluate"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. This represent the fractions of the FL K282A TAP tag pull down at 1/20dilutions. This figure also represent the fractions of the Short K282A TAP tag pull down at 1/20 dilutions.

Figure 6.10: TAP Tag Pull Down of All Constructs

Yeast cells containing both GST tagged Rps18a and TAP tagged Rio2p were lysed, centrifuged to remove the supernatant and frozen. The thawed "supernatants" indicated TAP tagged Rio2p. The thawed supernatant was rotated with IgG beads, the IgG beads were collected and the flow-through removed, and the beads washed with binding buffer. The beads were boiled in 5xSDS-loading dye and the dye was removed ("eluates"). The Western blot was probed with polyclonal antibody against the GST tag conjugated to HRP. The eluates indicated GST tagged Rps18a.

The supernatant of the wing helix domain was undiluted, the supernatant of Rio2p FL, the supernatant of Short Rio2p, and the supernatant of FL 105A was diluted 1/5. The supernatant of Short K105A was undiluted. The supernatant of FL R271A was diluted 1/5. The supernatant of Short R271A was undiluted. The supernatant of Fl K282A was diluted 1/5. The supernatant of Short K105A was undiluted.

The eluate of the wing helix domain was undiluted. The eluate of Rio2p , the supernatant of Short Rio2p, and the eluate of FL 105A was diluted 1/20. The eluate of Short K105A was undiluted. The eluate of FL R271A was diluted 1/20. The eluate of Short R271A was diluted 1/5. The eluate of Fl K282A was diluted 1/20. The eluate of Short K105A was undiluted.

Chapter 7: Crystal Structure of MphR(A) with and without the Drug Erythromycin

7.1: Statement of Authorization

This research was originally published under the title Structure and Function of the Macrolide Biosensor Protein, MphR(A), with and without Erythromycin in *Journal of Molecular Biology* (2009; 387:1250-1260). The scientists who collaborated on the study were: Jianting Zheng, Vatsala Sagar, Adam Smolinsky, Chase Bourke, Nicole LaRonde-LeBlanc, and T. Ashton Cropp.

7.2: Overview

Dr. Jianting Zheng and I were co-first authors of the paper on the crystal structure of MphR(A) and my role in the work we did together was as follows: I supervised Dr. Jianting Zheng and worked closely with him to set up crystal trays. I then did all the computer work to determine and refine the structure of MphR(A). My supervisor Dr. Nicole LaRonde-LeBlanc guided me in this endeavor. The paper also describes mutagenesis done by Dr. Jianting Zheng.

7.3: Introduction to MphR(A)

Macrolide antibiotics are used to treat many bacterial infections. Some macrolide antibiotics include erythromycin, azithromycin, and clarithromycin (Figure 7.1). However, bacteria have developed resistance to these antibiotics by various mechanisms, including modifying the drug target, using efflux pumps, and modifying the drug itself (*56-58*). Our study examined the repressor of the mphA gene whose transcribed gene product phosphorylates the 2' hydroxyl group on the desosamine sugar of the antibiotic

(*59*). This modification prevents the drug from interacting with the drug target—the ribosome.

MphR(A) is regulated by the presence of erythromycin (*60*). When erythromycin is present, it can no longer bind to the promoter region of the gene cassette containing mphA, mrx, and MphR(A) and transcription of mph(A) is de-repressed. The MphR(A) binding site completely overlaps that of the PmphA promoter (*60*) and is 2173 basepairs upstream of its own start codon. The MphR(A) binding site is 35 base pairs in length. Thus MphR(A) has a negative feedback loop of its own expression (*60*). This gene cassette has been found in many clinical and environmental sources and as such is mobile (*61, 62*).

MphR(A) is predicted to be part of the TetR class of transcriptional regulators based on the sequence of the N-terminal helix turn helix motif (*63*). The dimeric motif is responsible for binding of Tet repressors to the promoter region. Thus it is predicted for MphR(A) that its similar N-terminal motif is responsible for binding to the promoter region.   While the C-terminal domains of the MphR(A) and TetR classes of repressors differ in sequence, they both function to bind their respective ligand antibiotics.

MphR(A) can bind to various different ligands of different sizes including 12-14-, and 16-membered macrolactone antibiotics. MphR(A) binds to these ligands irrespective of the ligand's function. MphR(A) can bind erythromycin, azithromycin, methymycin, and josamycin. While these ligands all contain a macrocyclic lactone and a desosamine sugar, they differ in aspects of size and shape. This brings us to the focus of this study, which was to determine the structure of MphR(A) in order to know how it can bind ligands of different sizes. Once this is known, it is possible that MphR(A) can be

engineered to bind to and detect ligands other than the lactones to which it is known to bind.

The gene for MphR(A) and the section of promoter to which it binds have also been utilized for many screens. It has been fused to other genes to control expression of those genes in mammalian tissue culture and mice (*63, 65*). The same fusion has been employed *in vitro* to determine the presence of erythromycin in milk (*66, 67*). These fusions have also been used to determine the presence of macrolide antibiotics in Streptomyces fermentation extracts (*68*). Thus in future, this type of gene system could be used to quickly and accurately screen soil and culture extracts for macrolide antibiotics .

Towards this end the structure of MphR (with and without the ligand erythromycin) was determined. The protein was expressed as a selenomethionine derivative and its liganded structure was determined by the single wavelength anomalous diffraction (SAD) method. The apo structure was then determined by molecular replacement.

7.4: Experimental Procedures

7.4.1: Crystallization

In the crystallization trials, 10% glycerol was added to the MphR(A) solutions to stabilize the protein. The final concentration of protein screened was 10 mg/ml. For ligand screens, 10 mM erythromycin was added. The first screens for crystal hits were performed using the sitting drop vapor diffusion method and encompassed 480 different conditions at three protein-to-well solution ratios. Those ratios included one-to-one protein-to-well solution, two-to-one protein-to-well solution, and one-to-two protein-to-

well solution. Crystals were optimized by mixing 2 µl of protein solution with 1 µl of well solution in hanging drops. The apo form of MphR(A) crystallized in a solution of 18-22% (w/v) polyethylene glycol 3350, 0.2 M ammonium acetate, and 0.1 M Bis-Tris pH 5.5 MphR(A) bound to erythromycin crystallized in 35-40% (w/v) polyethylene glycol 3350, 0.2 M MgCl2, and Bis-Tris, pH 5.5. Crystals were flash frozen in well solution containing 20% glycerol.

7.4.2: Data Collection, Phasing and Refinement

Data was processed using HKL2000(*46*), and the MphR(A)-erythromycin structure was determined using the program HKL2MAP to locate the selenium atoms and calculate the phases from selenium SAD. The map was put through heavy-atom position refinement, density modification, and model building using autoSHARP (*69*). Additional model building was performed using Coot (*70*). Molecular replacement with the ligand-bound MphR(A) was used to find the phases for the apo MphR(A) structure using MolRep. Refmac (*50*) was used for refinement. TLS refinement was used by setting each monomer as a separate TLS group (*71*).

7.5: Structure of MphR(A) with Erythromycin

The first structures were determined with the liganded MphR(A) because it was easier to crystallize. The first crystals appeared after three days. Selenomethionione was incorporated in the protein to determine phases by SAD. Crystal and refinement statistics are shown in Table 7.1. Each asymmetric unit contained a dimer. Electron density was visible for residues 5-189 in one subunit and 6-189 in the second subunit with each subunit containing 194 residues. There was clear electron density for erythromycin A binding in a central cavity in each subunit. We found that MphR(A) contains nine α-

helices per monomer (Figure 7.2) as characteristic of the structure of proteins in the same class—TetR (72), EthR (73), and QacR (74), etc. The α-helices of MphR(A) encompass residues 9-22 (α1), 30-37 (α2), 41-48(α3), 51-72 (α4), 80-90 (α5), 99-111 (α6), 118-133 (α7), 145-160 (α8), and 165-181(α9). Structural alignment of the coordinates using the Secondary Structure Matching (SSM) (75) search algorithm revealed the closest homology (rmsd of 2.49Å for 156 matched Cα atoms in the monomer) to the regulator, IcaR from *Staphylococcus epidermidis (76)* (Protein Data Bank entry 2ZCM), thought the sequence identity was only 14.8%. The purpose of IcaR is thought to be "antibiotic sensing." However, IcaR recognizes a different class of antibiotics, gentamicins, and is also important in controlling biofilm formation. Structurally similar proteins that align their tertiary structure with MphR(A) include TetR (3.26 Å, 135 matched C^α atoms), EthR (2.98 Å, 147 matched C^α atoms), and QacR (2.76 Å, 144 matched C^α atoms).

The N-terminal residues of MphR(A) are responsible for binding to the DNA promoter region. Residues 8-50 form a helix-turn-helix motif. The two helix-turn-helix motifs (HTH) that are part of the dimer are separated by a distance of 46.3Å as measured from the two C^α atoms of L44. L44 is a conserved residue of the TetR family. The same distance between N termini is seen in QacR and Act—44 and 46Å distances respectively. Dissociation of MphR(A) from the promoter, which likely occurs upon ligand binding, is due to the large increase in distance between the N-termini. The portion of the HTH motif that comes into contact with the promoter contains hydrophobic and conserved positively charged residues to bind to the negatively charged phosphate backbone. The residues include K35, R41, and R51.

Each monomer of MphR(A) binds one molecule of erythromycin. Since erythromycin is largely hydrophobic so is the binding pocket for it. The cavity is spanned by residues from helices α4-α8 and includes the dimeric interface of the other monomer. The pocket is large enough to accommodate the entire ligand, including the desosamine and cladinose sugar residues. To enter the central cavity the ligand has to traverse a short tunnel (Figure 7.2). The cavity is lined with 17 residues of mainly hydrophobic nature that make close contact with the erythromycin. The side chains are within 4Å of erythromycin (Figure 7.3). The cavity also contains a small portion of positively charged residues to counteract a negatively charged portion on erythromycin. These residues consist of R122, N123, and H147. Hydrogen bonds are formed between N123 and the C-9 and C-11 oxygen atoms of erthryomycin, H147 and the C-9 carbonyl, and R122 and the C1-carbonyl. A water-mediated reaction involving the C-12 hydroxyl group of erythromycin and the backbone carbonyl of A151 and the hydroxyl of Y103 further stabilizes negative charges on erythromycin. A chloride ion was placed at a position of high electron density that was greater than 10 σ in the difference density map. The chloride ion interacts between the desosamine moiety of the erythromycin and the backbone amide of N94. Only one of the monomers shows a hydrogen bond between K21 and the erythromycin cladinose O-6 and the 2'hydroxyl, and between S106 and the C-1carbonyl. The 2' hydroxyl of the desosamine residue of erythromycin is the target for MphA phosphotransferase protein that conveys resistance of the organism to erythromycin. The Ɛ-nitrogen of K21 is located 4.6Å from this hydroxyl group in the second monomer. Thus the phosphorylated product would be stabilized by this

interaction between its 2'-O-phosphate and the positively charged NH_3 group, leading to the hypothesis that the phosphorylated erythromycin is a better ligand for MphR(A).

7.6: Structural Effects of Ligand Binding

Once the liganded structure of MphR(A) was determined, the apo form could be solved by molecular replacement. Clear electron density was seen for residues 8-94, and 101-190 for one monomer, and for residues 6-95 and 99-186 for the second monomer. The electron density was too weak to model residues 95-200 for the first monomer and residues 96-98 for the second monomer. Both erythromycin-bound and apo structures of MphR(A) were aligned as dimers. Their overlap had an rmsd of 0.86Å for 298 C^α atoms as determined by least squares method. When one monomer was aligned the rmsd changed to 1.028Å for 173 atoms. This occurs because when the rmsd is greater than 4Å, it is excluded as an outlier. Thus the rmsd between atoms when the dimer is overlapped with a dimer contains more of these outliers than the rmsd calculated when one monomer is overlapped with the dimer. It appears as if a large shift between monomers occurs upon ligand binding such that the distance between N-termini in the apo form becomes shorter. This distance between the two C^α atoms of L44 is shorter and is only 42.9Å (Figure 7.4). This shorter distance allows for correct positioning of the N-termini of the repressor on the promoter for optimal binding. When erythromycin is bound, this distance increases and the hold on the promoter is lost, resulting in de-repression. The apo form of MphR(A) also contains a larger cavity volume than the erythromycin-bound MphR(A). As calculated by VOIDOO, the two cavities in the liganded-dimer are 1183 and 1234Å^3 and in apo MphR(A) dimer are 2273 and 1756Å^3. Thus the cavity closes around the bound ligand, and brings the N-terminal HTH closer towards the cavity,

increasing the distance between termini. The changes in structure can be accounted for by introduction of new bonds between MphR(A) and the ligand. Hydrophobic residues T17 and L20 on helix 1, in the apo form, are more spaced out. When erythromycin binds, these residues converge to make contacts with the hydrophobic region on the cladinose sugar residue. Also on helix 1 are residues E14 and R61 that converge to form a strong ionic interaction when erythromycin is bound.

7.7: Discussion

The TetR family of proteins is known for its role in antibiotic resistance. However, these proteins may also have a role in secondary metabolism. Some TetR family members have a role in secondary metabolism and as such are activated by metabolic intermediates and precursor ligands (*77*). In fact, when a BLAST-P search is done with MphR(A), it reveals homologues that are part of the TetR family members including those from Myxococcus, Salinispora, and Streptomyces species. 10% of the genome of these species is involved in secondary metabolism. Alignment shows that these homologues might function similarly as they share many conserved residues (Figure 7.5). Importantly, they share a hydrophobic residue that in MphR(A) is L20 and is involved in the interaction of helix one with the binding pocket of erythromycin. In addition proteins from Myxococcus and Rhodopseudomonas both have the anionic side chain of E14 (or D14) and the cationic side chain of R61 that form an ionic interaction between helices 1 and 4 when the ligand is bound. Conserved residues include R122 and N123 that in MphR(A) form favorable contacts with erythromycin. It may be that proteins from these organisms recognize the same ligands as MphR(A). Thus these proteins may have a dual function to protect the organism from environmental toxins and to regulate secondary metabolism.

In addition to binding erythromycin, MphR(A) binds to larger ligands with extra sugar residues. Those sugar residues might hang outside the entrance tunnel since the binding pocket snugly fits erythromycin with no extra room. MphR(A) does not make any polar contacts with the desosamine sugar of erythromycin. It probably does not make direct contacts to sugar molecules. The sugar molecules are not critical to ligand recognition. The polar contacts to the oxygen atoms at C-1, C-9, and C-11 suggest these must be the critical points of anchoring the ligand to the binding site. If the phosphorylated erythromycin does indeed serve as a better ligand for MphR(A), it would create a cascade effect for activation of the operon.

Mutations to MphR(A) might be able to allow a different ligand to bind to MphR(A) or improve the sensitivity of MphR(A) for its natural ligands. Such changes could lead to amplifying the MphR(A) response to the ligand creating a more pronounced activation of the operon. Including multiple copies of the MphR(A) binding site might slow the response to ligands and give a positive signal for ligands when large concentrations are present. Alternatively, the promoter could be modified to create a cascade response to detect smaller amounts of ligand. Thus the MphR(A) regulatory system can be modified to create responses to different ligands, and to different concentrations of ligands (*78, 79*).
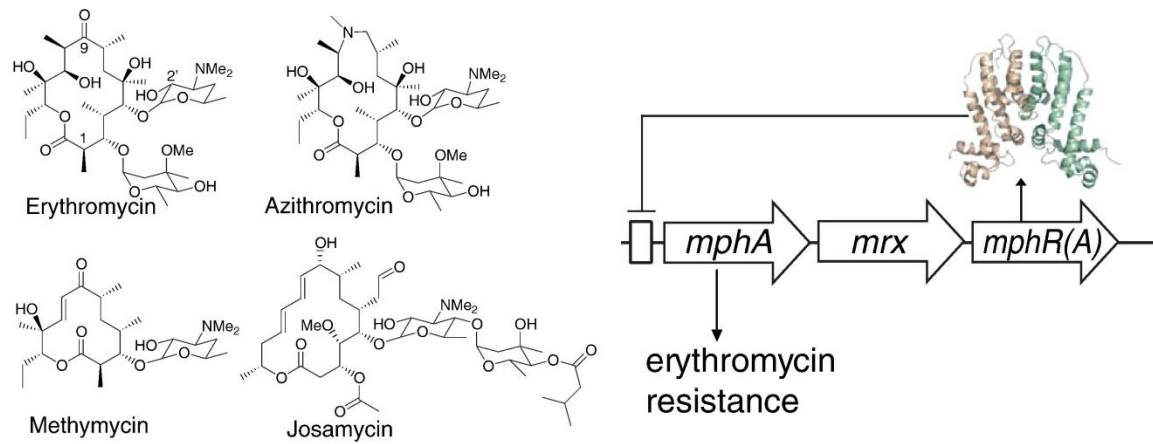
Figure 7.1: MphR(A) Ligands (*64*)

The structures of known ligands of MphR(A) and order of genes. MphA is the erythromycin 2'-phosphotransferase. The function of Mrx is not known.
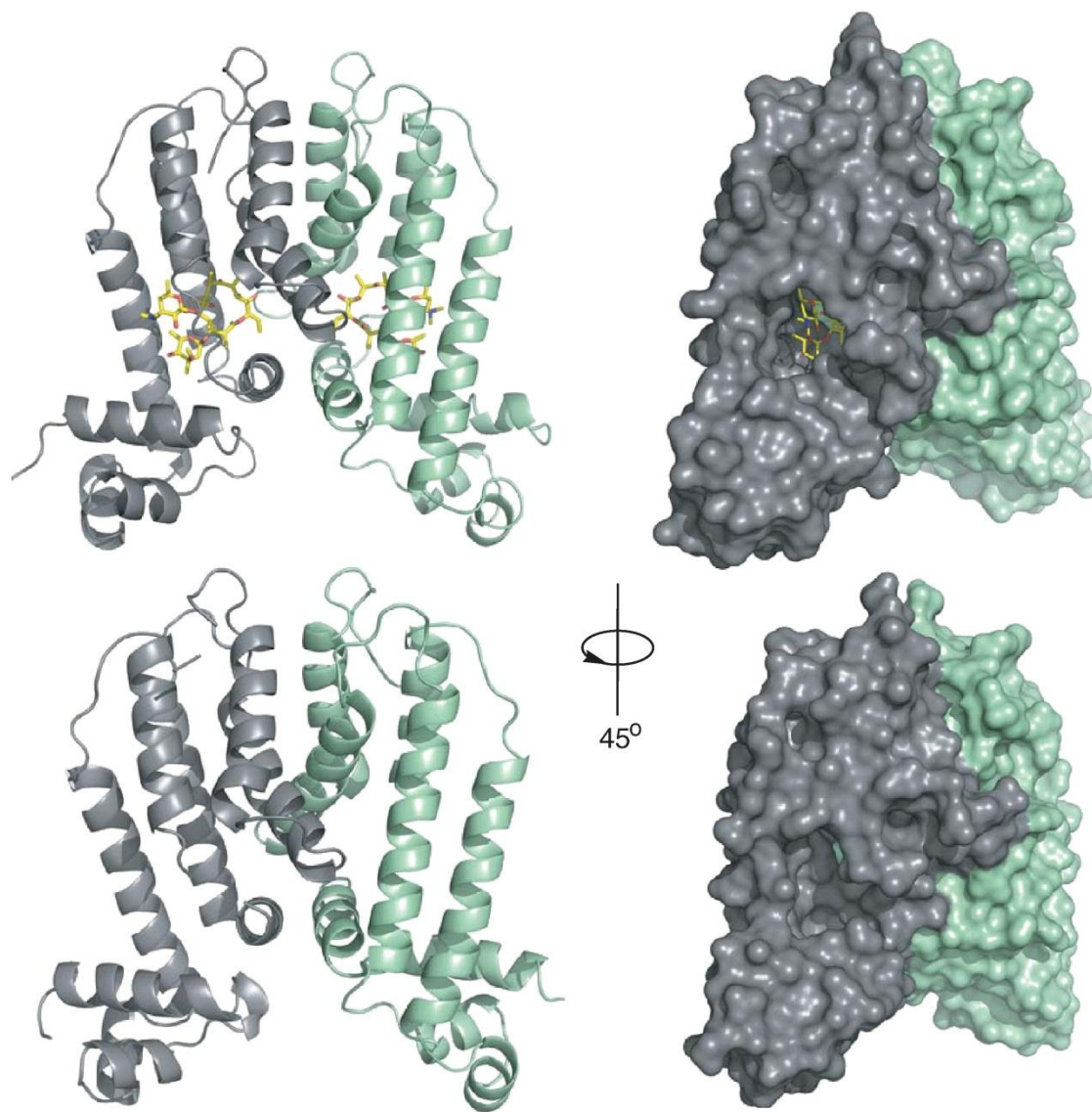
Figure 7.2: Model of MphR(A) (*64*)

This figure shows MphR(A) with and without erythromycin. Space filling models show the tunnel erythromycin has to travel to enter the ligand binding site.

Table 7:1: Data collection and refinement statistics

	MphR(A)–SeMet	MphR(A)–erythromycin	MphR(A)–APO
Protein Data Bank code		3FRQ	3G56
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	44.4, 61.1, 134.7	44.3, 61.2, 134.6	37.4, 113.8, 43.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 93.9, 90.0
Molecules per asymmetric unit	2	2	2
Wavelength (Å)	0.9795	0.9795	1.000
Resolution (Å)	50–2.0	50–1.76	50–2.1
R_{sym} (last shell)	0.070 (0.321)	0.064 (0.419)	0.040 (0.467)
$I/\sigma I$ (last shell)	51.2 (8.1)	30.0 (4.6)	30.2 (2.1)
Completeness (%) (last shell)	99.7 (97.4)	99.5 (97.9)	98.5 (90.7)
Redundancy (last shell)	7.3 (7.0)	7.8 (7.4)	4.1 (3.6)
No. of reflections	45,208	35,182	20,802
Refinement			
Resolution (Å)		30–1.76	57–2.1
R_{work}/R_{free} (%)		18.1/22.1	22.0/28.4
No. of protein residues in monomers A/B/C/D/E/F/G/H		184/183	176/179
No. of non-protein atoms			
Ligand		102	0
Solvent		273	84
Mean *B*-factors (Å^2)		30.0	48.1
RMSD			
Bond lengths (Å)		0.017	0.015
Bond angles (°)		1.569	1.419
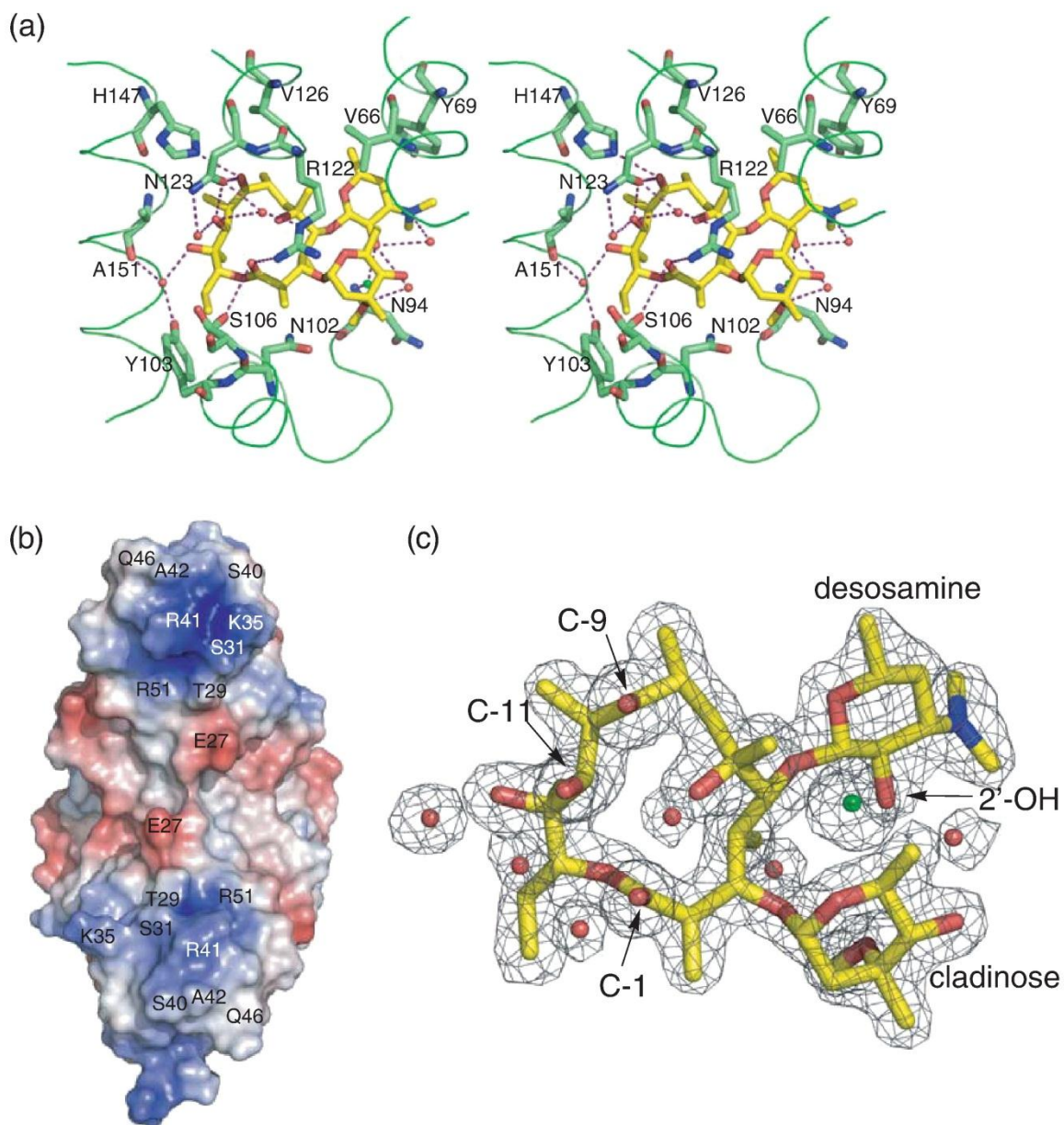
Figure 7.3: Interactions of MphR(A) (*64*)

(a) The figure is a stereo image of the ligand binding site of MphR(A) with erythromycin. Polar contacts within 4Å are shown as dotted lines. (b) The surface charge of Mphr(A) as seen from the bottom HTH. Surface resiudes are labeled. (c) The 2F$_o$-F$_c$ electron density map of the erythromycin bound to the ligand binding site at 1σ density is shown. Water molecules are depicted as red spheres and the chloride ion as a green sphere.
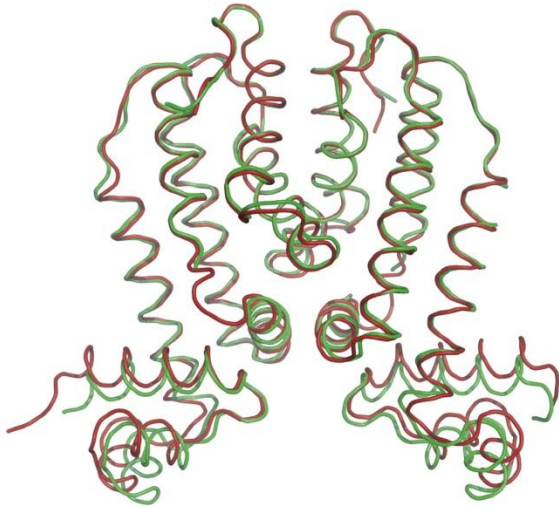
Figure 7.4: The Backbone Structures of MphR(A) With and Without Erythromycin Superimposed (*64*)
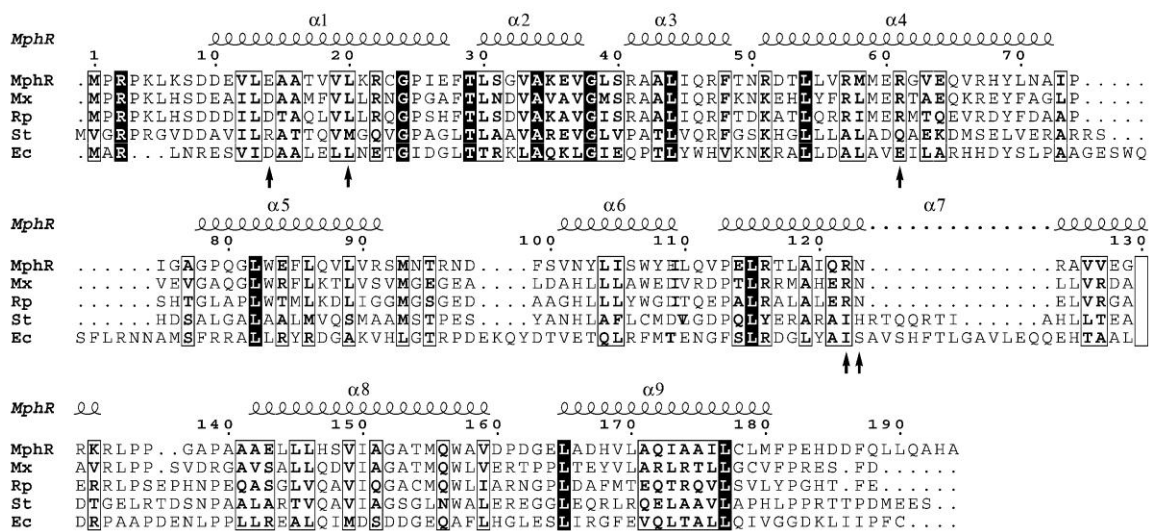
Figure 7.5: CLUSTALW alignment of homologues of MphR (*64*).
Mx, *Myxococcus Xanthus* (accession no. YP_633146); Rp, *Rhodopseudomonas palustris* (accession no. YP_ABE4088); St, *Salinispora tropica* (accession no. YP_001159116); Ec, TetR from *E.coli.* Black boxes represent complete identity, open boxes represent similar residues, and arrows indicate the residues mentioned in the paper (E14, L20, R61, R122, and N123).

Chapter 8: Crystal Structure of N-Acetyl-L-glutamate Synthase from *Neisseria gonorrhoeae*

8.1: Overview

This chapter contains work published along with Dr. Dashuang Shi, Dr. Norma Allewell and Dr.Mendel Tuchman. As author of this dissertation manuscript, it is my responsibility to discuss those portions of the project I completed and what was contributed by other authors. Cloning, purification, and crystallization of *Neisseria gonorrhoeae* N-acetylgluatamate synthase (NAGS) was done solely by me. I also completed the activity assays. Computer manipulations of the x-ray diffraction data to determine the actual structure of *N. gonorrhoeae* NAGS was done by Dr. Dashuang Shi.

8.2: Statement of Authorization

This research was originally published in The Journal of Biological Chemistry by Dashuang Shi, Vatsala Sagar, Zhongmin Jin, Xiaolin Yu, Ljubica Caldovic, Hiroki Morizono, Norma M. Allewell, and Mendel Tuchman. The paper was titled: The Crystal Structure of N-Acetly-L-glutamate Synthase from *Neisseria gonorrhoeae* Provides Insights into Mechansims of Catalysis and Regulation. *The Journal of Biological Chemistry.* 2008; 283(11): 7176-7184. © the American Society for Biochemistry and Molecular Biology.

8.3: Introduction

The first step in the synthesis of arginine is the acetylation of glutamate by N-acetylgluatamate synthase (NAGS) to create N-acetlyglutamate (NAG). NAG is then phosphorylated by NAG kinase (NAGK) to become NAG phosphate and then transformed into N-acetylornithine. In organisms such as fungi, plants, and many

bacteria, NAG can also be generated by transferring the acetyl group from N-acetyl ornithine to glutamate by ornithine acetyl transferase (OAT) (*80, 81*). If organisms possess OAT they mainly use this enzyme to produce NAG in which process feedback inhibition by arginine usually targets NAGK and, to a much lesser extent, NAGS. However, if organisms do not possess OAT, then arginine mainly inhibits NAGS and does not at all inhibit NAGK (*82*). *Neisseria gonorrhoeae*, in addition to having NAGS, contains an OAT that can use both acetyl CoA and N-acetylornithine as substrates to transfer the acetyl group (*83*). However, NAG and OAT have little sequence similarity (*82*).

In mammals NAG acts as an obligatory allosteric activator of carbamylphosphate syntheatase I in the urea cycle (*84*). Contrary to microorganisms, arginine is used by mammals to activate NAGS. Without the production of NAG by NAGS, hyperammonemia ensues due to a dearth of carbamylphosphate synthetase I that causes a block in ureagenesis (*85*).

While enzymes of the arginine cycle usually show similar sequence similarity across phyla, the NAGS enzyme differs among mammals, fungi, and bacteria (*86-88*). It is thought that the modern day NAGS was generated from the fusion of two genes containing NAGK and N-acetyltranferase (NAT). Support for this hypothesis comes from the existence of enzymes that are bifunctional as kinases and acetyl transferases (*88*). The existence of enzymes that are considerably shortened in some bacteria also favors this hypothesis (*89*).

While several NAGK and many NAT related structures have been solved, no NAGS structure from any organism had been solved until now (*90-95*). This is probably

because NAGS proteins are usually unstable and not crystallizable (*94*). The first look at a three-dimensional structure of a NAGS protein comes from the crystal structure of NAGS from *Neisseria gonorrhoeae*. The structure reveals two domains that are connected but fold independently. The domains are an N-terminal amino acid kinase domain and a C-terminal GCN5-related NAT domain. Two structures have been solved with Acetyl Coenzyme A (AcCoA) bound and with both CoA and NAG bound. The structures reveal aspects of the catalytic process and the allosteric effects of arginine.

8.4: Experimental Procedures

8.4.1: PCR of *N. gonnorhoeae* NAGS

PCR of the argA gene was done using *N. gonnorhoeae* genomic DNA (ATCC 53420). The primers used were CGG<u>CATATG</u>AACGCGCCCGACAGCTTTGT and CTT<u>CTCGAG</u>ATATCAGCGGTGCAGGCGAC. The PCR product was cloned into a pET28a expression vector (Novagen). The gene was sequenced and contained three point variations that resulted in V312I, D336N, and P427S amino acid substitutions. These three amino acids are found in the NAT domain sequence of *Neisseria meningitides* NAGS (accession codes NP_274872, AAF42210 or CAM07850). The construct that was cloned is the AAK domain of *N. gonorrhoeae* and the NAT domain of *N. meningitides*.

8.4.2: Protein Purification and Crystallization of NAGS

Protein was expressed in Eschrichia coli BL21(DE3) cells (Invitrogen). The protein was first purified on a Ni affinity column and subsequently on a DEAE column (GE Healthcare). The protein was concentrated to 10 mg/ml. It was incubated for 30 minutes with 23 mM AcCoA or AcCoA with 100 mM L-glutamate. CsCl was then

added to 100 mM of the concentrate. The hanging drop method was used for crystal formation. 2 µl of well solution containing 100 mM CsCl, 6% polyethylene glycol 2250 and 100 mM sodium citrate at pH 5.8 was mixed with 2 µl of the protein mixture to make the hanging drop. The selenomethione- (SeMet) substituted protein was made using Overnight Auto-induction System 2 (Novagen) (*94, 95*). The same procedure was used to make the SeMet crystals. AcCoA and glutamate formed the reaction product CoA and NAG when incubated and these products crystallized with the protein.

8.4.3: Activity of NAGS

NAGS activity was measured in a solution containing 50 mM Tris-HCl buffer, pH 8.5, 100 mM NaCl, 10 mM L-glutamate, and 2.5 mM AcCoA and 0.16 ug of enzyme in 100 µl (*88*). The reaction was incubated at 30°C for five minutes and terminated with 100 µl of 30% trichloroacetic acid. The amount of NAG produced was determined using liquid chromatography-mass spectroscopy as described (*86, 87*). NAGK activity was measured colorimetrically. 1.0 µg of enzyme was mixed in 100 µl 100 mM Tris-HCl, pH 8.5, 100 mM NaCl, 100 mM NAG, 20 mM ATP, 40 mM $MgCl_2$ and 400mM hydroxylamine at room temperature for 20 minutes. The reaction was terminated and the color was developed by adding 100 µl of 5% FeCl3, 8% trichloracetic acid, and 0.3 M HCl. The absorbance of product was measured at 540nm. Inhibition of NAGS activity was measured over 0.01 mM -10 mM L-arginine concentrations.

8.5: Enzymatic Activity

The enzyme had a specific NAGS activity of 29.1 µmol/min/mg protein. The activity is about 10 times lower than the NAGS activity of NAGS/K enzyme from *Xanthomonas campestris* and 4 times lower than NAGS from *E. coli (96)*.

N. gonorrhoeae NAGS activity is inhibited by arginine with 50% inhibition with 0.15mM arginine. The enzyme has no NAGK activity, in line with the fact that *N. gonorrhoeae* contains a separate gene encoding NAGK (accession number YP_207961).

8.6: Linkage of the AAK and NAT Domains

The monomeric *N. gonorrhoeae* NAGS consists of two domains that are linked together by three amino acids (10Å) (Figure 8.3). The N-terminal Amino Acid Kinase (AAK) domain and the C-terminal GCN5-related NAT domain share no other monomeric contacts. The AAK domain consists of an eight-stranded parallel β sheet bracketed by α helices on both sides. Proteins with similar folds include NAGK (*92, 93*), carbamate kinase (*98*), carbamate kinase-like CPS (*99*), UMP kinase (*100, 101*), aspartokinase (*102*), and 5-glutamate kinase (*103*). The NAGK structure of *Thermotoga maritima*, and *Pseudomonas aeruginosa*, when overlayed with the AAK domain of *N. gonorrhoeae*, have a root mean square deviation of 1.7 and 1.6Å for the 238 and 234 equivalent C^{α} atoms. The amino acid sequence of the AAK domain from *N. gonorrhoeae* has 25.9 and 27.4% sequence identity to *T. maritima* and *P. aeruinosa* NAGK, respectively. Differences between organisms occur in the N-terminal lobe, which is responsible for binding different substrates and making dimer contacts. The C-terminal lobe, since it is responsible for binding the same ATP molecule, is similar (*101, 103*). Among different organisms NAGK mainly varies in the β3-β4 loop. Different conformations lead to either opening or closing of the active site. For the AAK domain of *N. gonorrhoeae* the active site is closed similar to the state of *Mycobacterium tuberculosis* NAGK without substrate.

The C-terminal NAT domain on the other hand contains a central anti-parallel β sheets surrounded by α helices on both sides. The β sheets are divided to form a V shape with AcCoA binding inside the V. On the N-terminus of the V is a four-stranded anti-parallel β sheet surrounded by three α helices. On the C-terminus of the V is a three-stranded anti-parallel β sheet surrounded by two α helices. At the closure of the V, the β strands share four hydrogen bonds between two parallel strands (β22 and β23). The bulge caused by A355 and C356 causes the V to split. According to a DALI data base search, the NAT domain of *N. gonorrhoeae* is most similar to the GCN5 histone acetyltransferase (Protein Data Bank code 1M1D) with a z score of 16.4 (*104*). Overlaying the two structures gives a root mean square deviation of 1.7Å for 121 C^α atoms.

8.7: NAGS Monomers Form a Hexameric Ring

Crystals of NAGS from *N. gonorrhoeae* have one molecule in the asymmetric unit. The crystals have a three-fold crystallographic symmetry axis that aligns with the molecular three-fold axis, and a two-fold crystallographic symmetry axis that aligns with the molecular two-fold axis. Thus it can be said that the hexamer consists of either three dimers or two trimers.

The AAK domain of one monomer makes contacts with the NAT domain of a neighboring monomer. Two stacked trimers form a ring with outer diameter 100Å and inner diameter 20Å and a height of 110Å (Figure 8.4A). The width of the inside ring is smaller than the 26-32Å arginine sensitive NAGK structures, which change ring size depending on whether arginine is bound or not (*92*).

When the three dimensional structure of the hexamer is viewed perpendicularly to the 3-fold axis, the AAK domains form a zigzag stripe across the middle of the molecule. The NAT domains are at the top and bottom of this stripe (Figure 8.4A). Each AAK domain makes contacts with two other AAK domains. The hexamers in the crystal are packed together along the c axis to form long tubes of alternating AAK and NAT layers.

8.8: Interactions Across the Dimer Interface

The AAK domain can make two types of dimer contacts with other AAK domains along the zigzag stripe across the middle of the molecule (Figure 8.4, B and C). The first is the dimer contact between K1 and K4 (Figure 8.4C). The contacts occur amongst residues from strand 5, helix 4, helix 5, and the loop between strands 11 and 12 in the N-terminal lobes of both subunits. The contacts are planar with α helices flanking strand 5. Strand 5 makes antiparallel β strand contacts with strand 5 from the neighboring subunit and transforms the 8-stranded β sheet from each monomer into an extended 16-stranded β sheet. The interactions that occur are mainly the formation of hydrogen bonds including the main chain atoms of L125, S127, and N129. The α helices interact principally through hydrophilic contacts including the side chains of Q96, R103, S104, E107, N129, Q161 and H175. There are also some hydrophic interactions involving V126, F160, and A164 in the dimer interface. The strength of the many hydrogen bonds and hydrophic contacts suggests that the crystal structure represents the natural dimeric interface.

Another dimer contact occurs between K1 and K5 and involves interlacing helices (Figure 8.4B). The helix bends at conserved P16 and is divided into an N-terminal helix and a C-terminal helix. The N-terminal helix is 15 residues long and amphipathic with charges on one side and hydrophobic residues on the other side. The C-terminal helix is

five residues long and thought to be responsible for binding arginine at a place similar to that where ariginine binds in NAGK enzyme (*92*). Y16, which aligns with Y15 and Y21 in the sequence of *T. maritima* and *P. aeruginosa* NAGK, is in this C-terminal helix. F7, F11, A14, I18, and H10 form hydrophobic contacts across the N-terminal helices creating a two-fold dimer axis. Since the N-terminal helix is amphipathic, it is lined with R12, R19, and R22 on the other side and these residues also make contacts between K1 and K5. Hexamer formation is likely to be reliant on the N-terminal helix. This helix is probably responsible for the arginine regulatory response and formation of the hexamer is likely to be key to the arginine response. Thus hexamer formation probably also occurs in nature (*92*).

The dimeric contacts between K1-K4 and K1-K5 occupy a buried surface area of 1206 and 1340 Å^2. This indicates that the hexamer is most likely to occur both in solution and in the crystal structure.

8.9: The Interaction of AAK and NAT Domains

Though the AAK domain and the NAT domain of a monomer do not interact, the AAK domain and the NAT domain of adjacent monomers do interact (Figure 8.4D). These contacts are polar and involve the side chains of R134, D140, D143, R151, K152 and the main chain oxygen or nitrogen atoms of I136, V138, and G141 from the AAK domain interacting with N393, T394, N426, and H428 from the NAT domain. The AcCoA bound to the NAT domain also makes contacts with the AAK domain via atoms O2*, O3*, and OAQ and R134, R151, and K152 from the AAK domain (Figure 8.4D).

8.10: The AAK Domain Cannot Catalyze a Reaction

Though NAGK proteins and AAK proteins share significant sequence similarity (20-27%), the AAK domain lacks several key residues that would allow it to bind ATP and NAG and enable phosphorylation of NAG. K8 is one such residue that is conserved in all NAGK sequences including *X. campestris* NAGS/K and in the AAK domain of *N. gonorrhoeae* NAGS is changed to G29. Furthermore, the side chain of M186 occupies the place where the K8 side chain would lie. All NAGK enzymes have a conserved N160-A161-D162 (*E. coli* NAGK numbering), but the AAK domain contains D185-M186-V187 in its place. Finally the important K217 is replaced by L241 in the AAK domain. These replacements would suggest that the NAGK domain of *N. gonorrhoeae* NAGS is inactive (*88*).

8.11: The Catalytic Mechanism

The densities of AcCoA, CoA, and NAG delineate the active site of the NAT domain (Figures 8.1 and 8.2A). AcCoA and CoA bind to the NAT domain using the loop-joining strands 22 and 12 to engage the pyrophosphate group. The sequence N364-E365-G366-G367-Y368-G369 in the loop agrees with the (R/N)-Xaa-Xaa-G-Xaa-(G/A) sequence for binding AcCoA in the GCN5-related NAT superfamily members (*105*). However, as previously mentioned, the side chains of R134, R151, and K152 from the adjacent AAK domain also take part in binding AcCoA (Figure 8.4D). The adenosine ring is anchored via hydrogen bonds to the side chain of E397. AcCoA bends about 180° at the pyrophosphate group so that the CoA chain is parallel to the adenosine ring. It also bends 90° at the CBO atom of pantetheine to bury the acetyl group in the core of the NAT domain. The acetyl group is bound in a hydrophobic pocket that is lined by the side chains of A355, L357, and F399. The acetyl carbonyl hydrogen atom bonds to the main

chain nitrogen atoms of C356 and L357 and positions the acetyl group towards the opening of the V created by the B bulge of A355-C356 (Figure 8.1). The same hydrogen bonding between the acetyl carbonyl group and the NH group of the residue aligning with L357 have been shown in other GCN5-related NAT proteins (*106*). What is missing in the case of the acetyl group is a conserved tyrosine that interacts with the sulfur or carbonyl atom of AcCoA in other NAT proteins (*91, 107*). S392 is the closest alternative but its side chain oxygen atom is 3.7Å away from the sulfur atom of AcCoA. However, S392 does form a hydrogen bond with the sulfur atom of CoA (3.3Å). The side chain of S392 appears to form a hydrogen bond with the main chain nitrogen atom of T395, preventing its locomotion. The pantetheine moiety of AcCoA and CoA form hydrogen bonds to the NH group of V359 (NH) and carbonyl group of L357, similar to the hydrogen bonding in an anti-parallel β-sheet. The carbonyl atoms of AcCoA and CoA make hydrogen bonds with the side chains of T395 and Q364.

Since the acetyl group of AcCoA is buried deep within the protein, the glutamate molecule must enter the active site from a passageway at the other end of the protein. This passageway allows AcCoA access to the solvent. The ternary complex of AcCoA and glutamate shows that glutamate or the NAGS carboxyl group form bonds with R316, R425, S427, and the main chain nitrogen atoms of C356 and L314. The hydrophobic side chains of L314 and L391 make interactions with the methylene groups in the side chain of glutamate. The acetyl carbonyl atom of the NAG hydrogen bonds to the main chain nitrogen atoms of C356 and L357. The side chain of A416 has a fixed position due to hydrogen bonding with the side chain of conserved E353 and it also may contribute to the positive charge stabililizing the negative caroboxyl group of L-glutamate or NAG.

Compared to the crystal structure with bound AcCoA, where both CoA and NAG are bound, the side chain of R316 swings 8Å closer to the active site (Figure 8.2A). This movement might be critical for catalysis.

There are two schools of thought on how NAGS enzymes accomplish catalysis. The first is the two-step ping-pong method in which the acetyl group of AcCoA is first transferred to a cysteine group on the enzyme (*108-110*). Then the L-glutamate binds and the acetyl group is transferred to it. According to the crystal structure, the potential acceptor of the acetyl group, C356, is 8.4Å away from the sulfur atom of AcCoA making covalent attachment less likely. The other hypothesis is that the active site accommodates both substrates simultaneously and that the acetyl group is directly transferred from AcCoA to L-glutamate. Most members of the GCN5-related NAT family use this mechanism of acetyl group transfer (*90*). A possible scenario for catalysis can be modeled based on the two solved structures (Figure 8.2B). In this scenario the α-amino nitrogen atom of L-glutamate is 2.5Å from the carbon atom of the acetyl group. The nitrogen atom is positioned to attack the carbonyl atom directly to form a S-configured tetrahedral transition state. The carbonyl atom of the acetyl group would be hydrogen bonded to the main chain nitrogen atoms of C356 and L357 imparting a slight positive charge to the carbon atom. The products of CoA and NAG form and the sulfur atom of CoA moves 0.9Å towards the side chain of Ser392 for protonation before release. The plane of the acetyl group of NAG rotates almost 90° from the acetyl group of AcCoA. At the same time, the side chain of R316 moves away to allow room for NAG to leave the active site. The main chain carbonyl oxygen atom of L391 hydrogen bonds to the amino nitrogen atom of NAG and is close to the sulfur atom of CoA. This

positioning may help proton transfer from the amino group of glutamate to the sulfur atom of CoA.

8.12: Classes of NAGS Protein

There are four different known types of NAGS. The first class is the "classical" *E.coli*-like NAGS. The second class is the bifunctional vertebrate-like NAGS-K. The third class is the "short" NAGS. The fourth class is the argH-argA fusion NAGS (*111*). The first two classes consist of an AAK domain and a NAT domain comprised of approximately 440 amino acids (*87*). From the current crystal structure, it can be deduced that the AAK domain plays a regulatory role and the NAT domain plays a catalytic role. This is consistent with the fact that the AAK domain can be removed and the NAT domain will still be functional, and that the AAK domain is not functional in all NAGS. Sometimes the AAK domain may be part of an argH encoded argininosuccinate lyase (*111*). Contrary to prevalent thought, this crystal structure shows that the AAK domain does not contribute to glutamate binding (*92, 112*).

Aspartate transcarbamoylase is similar to NAGS as follows: the catalytic trimer is active alone and when combined with inactive dihydroorotase, or regulatory units, or even when fused to other proteins such as carbamoylphosphate synthetase and dihyroorotase (*113*).
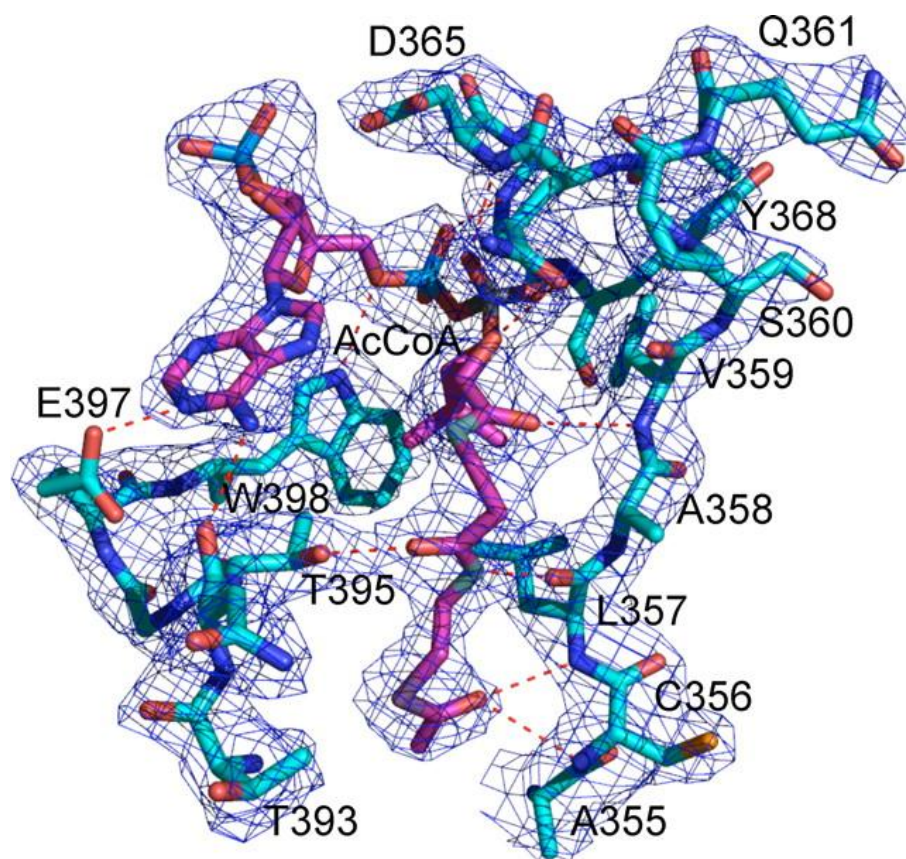
Figure 8.1: AcCoA Binding Site (*97*)

The $2F_o$-F_c electron density map is shown for the AcCoA binding site at 1.0 σ contour. The carbon atoms for AcCoA are represented by pink sticks. The carbon atoms for the protein are represented by blue sticks and hydrogen bonds between AcCoA and the protein are seen as red dashed lines.
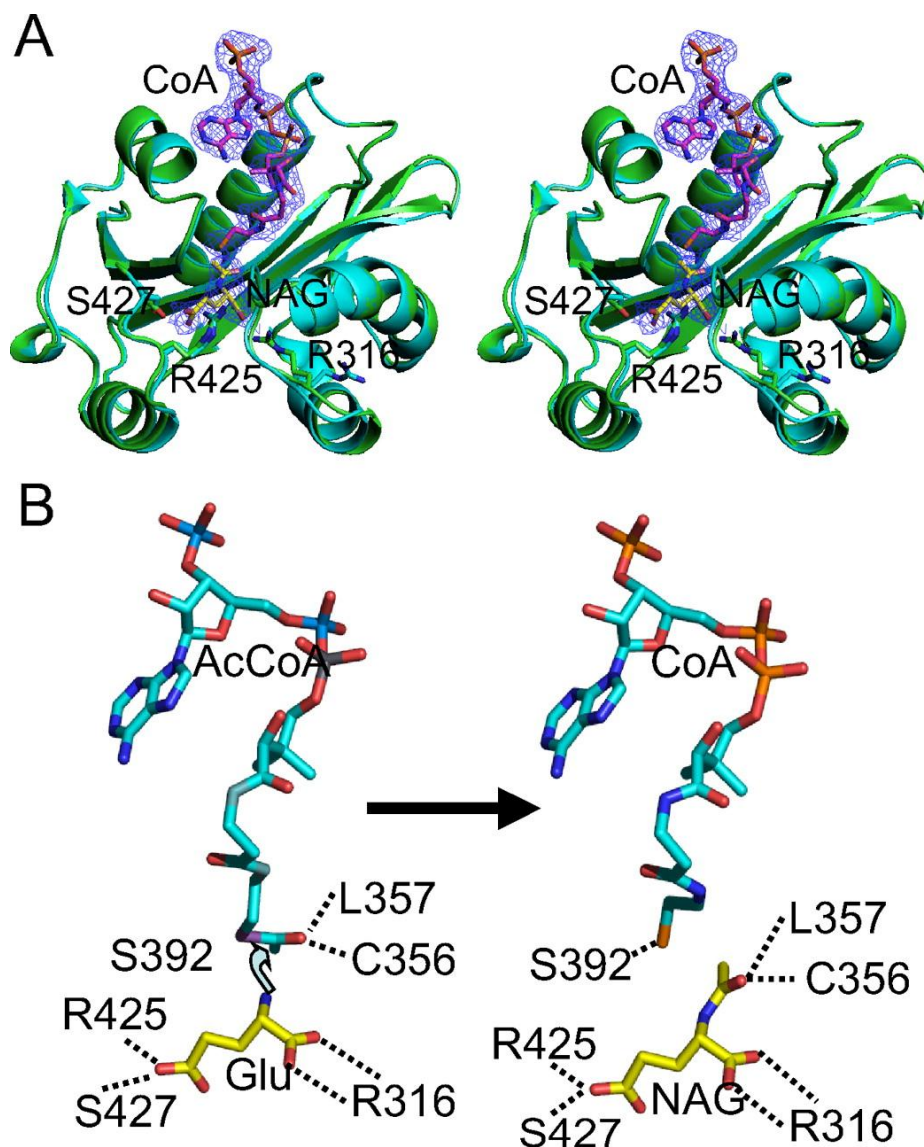
Figure 8.2: Ribbon Diagram of the the NAT domain for the AcCoA and CoA and NAG complexed structures and the reaction model. (*97*)

(a) A stereo diagram of the structures of NAGS with AcCoA in light blue and NAGS with CoA and NAG in green are superimposed. The bound CoA and NAG are shown as an $2F_o$-F_c electron density map at 1.0 σ contour. The carbon atoms of AcCoA are shown as pink sticks while the carbon atoms of NAG are shown as yellow sticks. The side chains of R316 and R425 are shown as sticks because they are involved in NAG or glutamate binding. (b) The reaction model for the attack of the nitrogen of the α amino group of glutamate on the carbon atom of the acetyl group. The nitrogen and carbon atoms are 2.5Å away and poised to react.
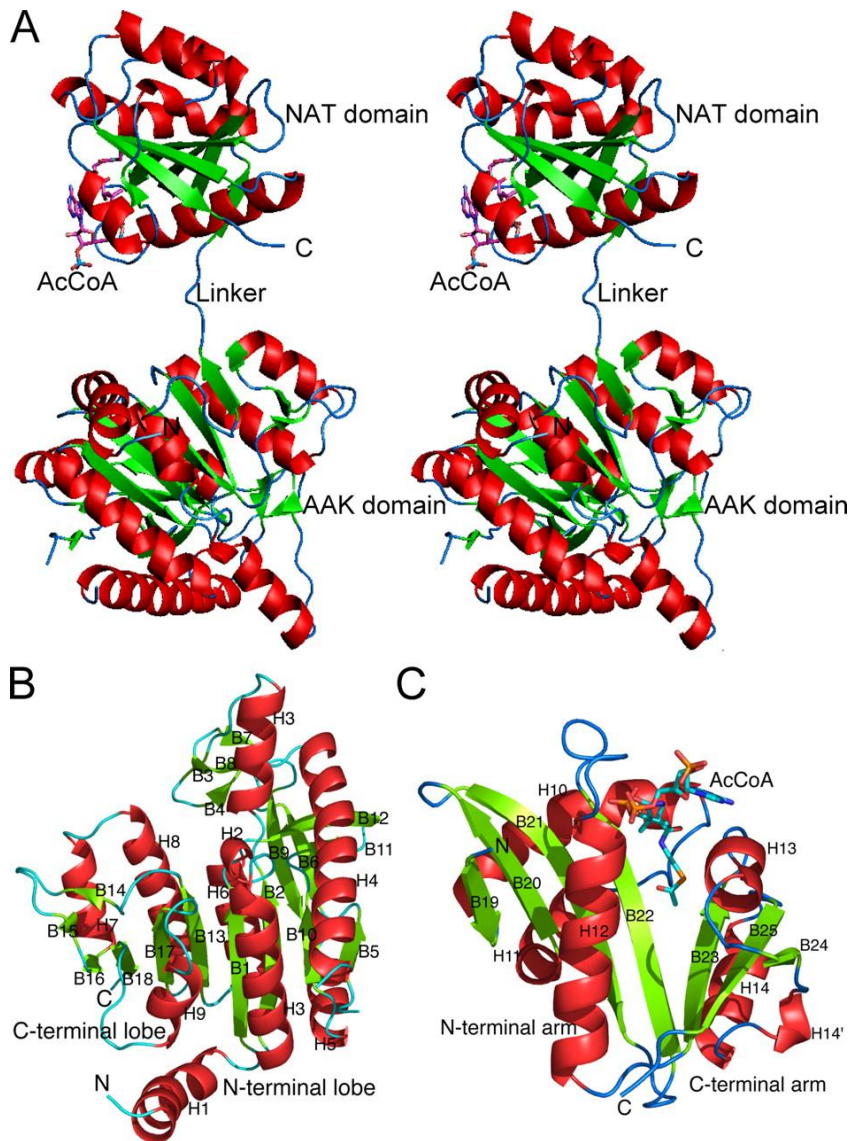
Figure 8.3: Ribbon diagram for a NAGS monomer(A, shown in stereo) and its AAK(B) and NAT (C) domains. (*97*)

The green arrows show β sheets and the direction of the β sheets and α helices are in red. AcCoA is shown as sticks.
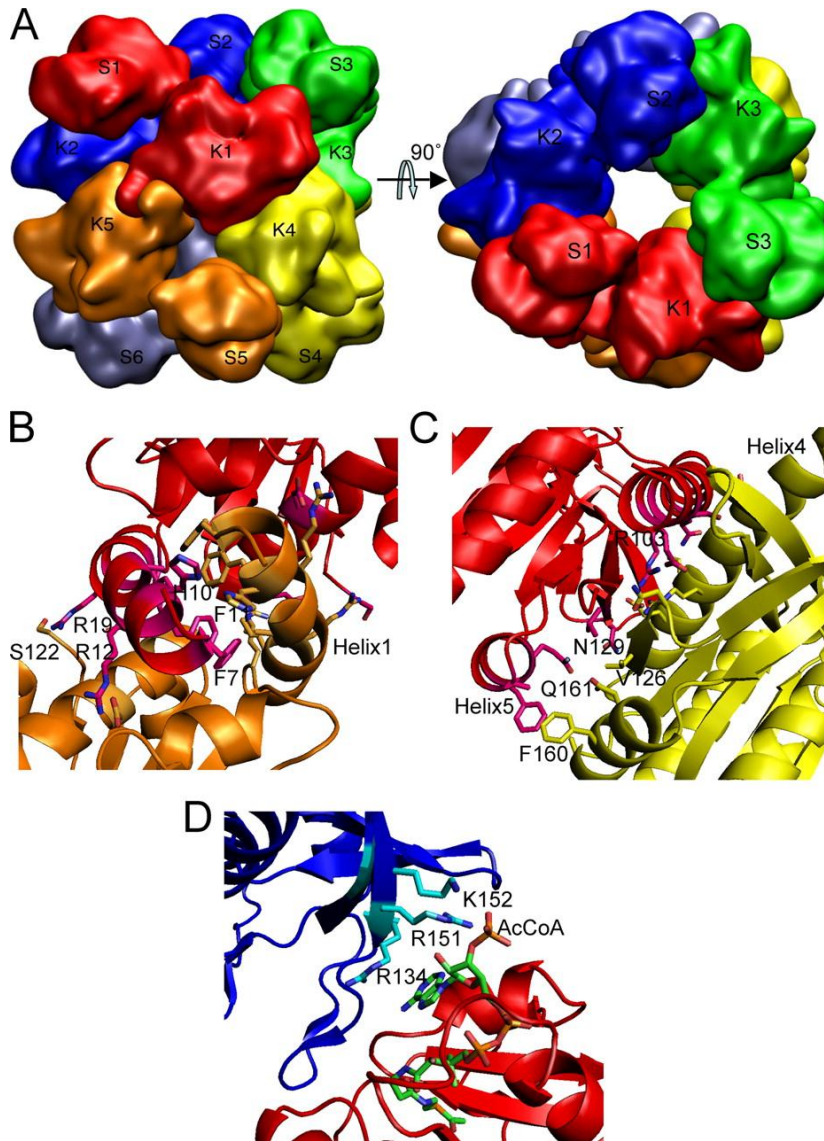
Figure 8.4: Molecular hexamer of *N. gonorrhoeae* NAGS and its interfaces. (*97*)
(a) Quaternary interactions between the subunits are shown. (b) Dimer interface between the AAK domain of K1 and K5 is shown. (c) Dimer interface between the AAK domain of K1 and K4 is shown. (d) Trimer interface between the AAK domain of K2 and the NAT domain of S1 is shown. Each subunit is shown in a different color. The side chains of residues involved in the interaction are shown as sticks.

Chapter 9: Future Studies

Now that the crystal structure of a eukaryotic Rio2 kinase is known, we can begin to examine the role of residues that appear important to the function of Rio2. Recent papers have been published by Karbstein and colleagues that elucidate the binding site of Rio2 in the ribosome and human Rio2's potential binding interactions with other ribosomal 43S processing factors (*115, 116*). In-vitro binding assays with ribosomal processing factors and ribosomal proteins indicate yeast Rio2 seems to directly interact with Rps15, Rps5, Pno1, Tsr1, Dim1, Nob1, and indirectly interacts with Ltv1 and Enp1(*116*). In addition, cryoEM studies were done of purified pre40S ribosomal particles pulled down by TAP tagged Rio2 in which various ribosomal processing factors were depleted or deleted to determine the exact positioning of these assembly factors in the ribosome (*115*). Future studies of Rio2 might involve docking the eukaryotic structure of Rio2 into the electron density seen by cryoEM and observing whether conserved surface residues on ctRio2 might be involved in binding to other ribosomal processing factors or binding to the rRNA or proteins of the ribosome. The surface residues can be mutated and TAP tagged Rio2 can be pulled down to view if it remains associated with the ribosome and with all the known ribosomal processing factors. Sucrose gradients can be run to observe whether mutated TAP tagged Rio2 is associating with the pre 40S ribosomal particle. In vitro experiments can be performed to determine if purified Rio2 with various point mutations can bind to MBP tagged ribosomal processing factors as done previously by Karbstein (*116*). Northern blots can be performed to determine if the point mutations have any affect on rRNA processing of 20S rRNA to 18S rRNA.

The dependence of rRNA processing on the catalytic activity of Rio2 can be addressed. Based on the crystal structure, mutations can be made to Rio2 that do not affect binding of ATP but affect catalysis. A Northern blot can be run to witness the effect of the mutation on 20S rRNA processing. Similarly TAP tagged Rio2 with the catalytic mutation can be pulled down to determine if catalytic activity has any effect on binding to the ribosome or association of ribosomal assembly factors.

The future direction of the studies of the wing helix domain and Rio2p would be to dock ctRio2 into the cryoEM of the pre40S particle. I would look for support of the data thus collected that extensive contacts besides the wing helix domain are made between Rio2 and the ribosomal particle. I would also look for the position of R271 to suggest how it interacts with the ribosome or ribosomal processing factors. The mutation of R271 on ribosomal binding could be tested in vivo by Northern blots to see if it affects ribosomal rRNA processing. It can also be tested by sucrose gradient to support that mutated Rio2p is not associating with the ribosome. The effect of the mutation of R271 on folding of Rio2p can be tested in vivo by CD of purified Rio2p or tested by proof of autophosphorylation of the mutated Rio2p.

Cancer cells exhibit enhanced ribosomal production as discussed in Chapter 1. Inhibitors of human Rio2 might be potential anticancer drugs. The study of inhibitors of yeast Rio2 is also advantageous because most of the studies of ribosome biogenesis have been in yeast. Since I have purified human Rio2 and yeast Rio2 the next step would be to study the binding of various kinase inhibitors with a thermo fluorescence assay to determine melting point. In this assay, purified protein is incubated with the inhibitor and then mixed with fluorescent dye which binds to hydrophobic residues. The sample is put

in a thermocycler which increments the temperature slowly while monitoring flourescence. When the protein become denatured, a spike in fluorescence is seen. The inflection point of the fluorescent change marks the melting temperature of the protein. The binding of the kinase inhibitors to Rio2 can be determined by an increase in melting temperature. The enzyme kinetics of human Rio2 and yeast Rio2 can be studied by autophoshphorylation rate with and without inhibitors. When a potential inhibitor for human Rio2 has been identified, it can be cocrystallized with ctRio2 to observe how it binds and whether it can be used as a drug specific to Rio2.

Crystallization of ctRio2 will help us make advances to the understanding of eukaryotic ribosomal biogenesis. The presence of Rio2 is necessary for maturation of the 43S preribosome and cleavage of the 20S rRNA to18S rRNA (*21-23*). At least in vitro, Rio2 makes direct or indirect contacts with many of the 43S preribsomal processing factors and might be critical to their function (*116*). As of now, the catalytic mechanism of Rio2 is not known, nor are its substrates or how it is important for cleavage of the 20S rRNA. Understanding of how Rio2 structure relates to function is the next step of study. Understanding ribosomal biogenesis is also critical to the field of cancer therapeutics. Cancer cells have upregulated ribosomal biogenesis because of their rapid proliferation and need for increased protein production for the new daughter cells (*1-4*). Ribosomal biogenesis is also correlated to cell cycle control and loss of cell cycle control in cancer cells (*1-4*).

BIBLIOGRAPHY

1. Derenzini, M., Montanaro, L., Chillà, A., Tosti, E., Vici, M., Barbieri, S., Govoni, M., Mazzini, G., and Treré, D. (2005) Key role of the achievement of an appropriate ribosomal RNA complement for G1-S phase transition in H4-II-E-C3 rat hepatoma cells, *J Cell Physiol 202*, 483-491.
2. Moss, T. (2004) At the crossroads of growth control; making ribosomal RNA, *Curr Opin Genet Dev 14*, 210-217.
3. Ruggero, D., and Pandolfi, P. P. (2003) Does the ribosome translate cancer?, *Nat Rev Cancer 3*, 179-192.
4. Thomas, G. (2000) An encore for ribosome biogenesis in the control of cell proliferation, *Nat Cell Biol 2*, E71-72.
5. Nazar, R. N. (2004) Ribosomal RNA processing and ribosome biogenesis in eukaryotes, *IUBMB Life 56*, 457-465.
6. Derenzini, M., Trerè, D., Pession, A., Montanaro, L., Sirri, V., and Ochs, R. L. (1998) Nucleolar function and size in cancer cells, *Am J Pathol 152*, 1291-1297.
7. Derenzini, M., Trerè, D., Pession, A., Govoni, M., Sirri, V., and Chieco, P. (2000) Nucleolar size indicates the rapidity of cell proliferation in cancer tissues, *J Pathol 191*, 181-186.
8. Kamath, R. V., Thor, A. D., Wang, C., Edgerton, S. M., Slusarczyk, A., Leary, D. J., Wang, J., Wiley, E. L., Jovanovic, B., Wu, Q., Nayar, R., Kovarik, P., Shi, F., and Huang, S. (2005) Perinucleolar compartment prevalence has an independent prognostic value for breast cancer, *Cancer Res 65*, 246-253.
9. McLean, I. W., and Gamel, J. W. (1988) Prediction of metastasis of uveal melanoma: comparison of morphometric determination of nucleolar size and spectrophotometric determination of DNA, *Invest Ophthalmol Vis Sci 29*, 507-511.
10. Warner, J. R. (1999) The economics of ribosome biosynthesis in yeast, *Trends Biochem Sci 24*, 437-440.
11. Granneman, S., and Baserga, S. J. (2004) Ribosome biogenesis: of knobs and RNA processing, *Exp Cell Res 296*, 43-50.
12. Kihm, A. J., Hershey, J. C., Haystead, T. A., Madsen, C. S., and Owens, G. K. (1998) Phosphorylation of the rRNA transcription factor upstream binding factor promotes its association with TATA binding protein, *Proc Natl Acad Sci U S A 95*, 14816-14820.
13. Voit, R., and Grummt, I. (2001) Phosphorylation of UBF at serine 388 is required for interaction with RNA polymerase I and activation of rDNA transcription, *Proc Natl Acad Sci U S A 98*, 13631-13636.
14. Voit, R., Hoffmann, M., and Grummt, I. (1999) Phosphorylation by G1-specific cdk-cyclin complexes activates the nucleolar transcription factor UBF, *EMBO J 18*, 1891-1899.
15. Voit, R., Kuhn, A., Sander, E. E., and Grummt, I. (1995) Activation of mammalian ribosomal gene transcription requires phosphorylation of the nucleolar transcription factor UBF, *Nucleic Acids Res 23*, 2593-2599.

16. Ciarmatori, S., Scott, P. H., Sutcliffe, J. E., McLees, A., Alzuherri, H. M., Dannenberg, J. H., te Riele, H., Grummt, I., Voit, R., and White, R. J. (2001) Overlapping functions of the pRb family in the regulation of rRNA synthesis, *Mol Cell Biol 21*, 5806-5814.

17. Treré, D., Ceccarelli, C., Montanaro, L., Tosti, E., and Derenzini, M. (2004) Nucleolar size and activity are related to pRb and p53 status in human breast cancer, *J Histochem Cytochem 52*, 1601-1607.

18. Voit, R., Schäfer, K., and Grummt, I. (1997) Mechanism of repression of RNA polymerase I transcription by the retinoblastoma protein, *Mol Cell Biol 17*, 4230-4237.

19. Cavanaugh, A. H., Hempel, W. M., Taylor, L. J., Rogalsky, V., Todorov, G., and Rothblum, L. I. (1995) Activity of RNA polymerase I transcription factor UBF blocked by Rb gene product, *Nature 374*, 177-180.

20. Kaye, F. J., Kratzke, R. A., Gerster, J. L., and Horowitz, J. M. (1990) A single amino acid substitution results in a retinoblastoma protein defective in phosphorylation and oncoprotein binding, *Proc Natl Acad Sci U S A 87*, 6922-6926.

21. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

22. Rouquette, J., Choesmel, V., and Gleizes, P. E. (2005) Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells, *EMBO J 24*, 2862-2872.

23. Vanrobays, E., Gelugne, J. P., Gleizes, P. E., and Caizergues-Ferrer, M. (2003) Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae, *Mol Cell Biol 23*, 2083-2095.

24. Dragon, F., Gallagher, J. E., Compagnone-Post, P. A., Mitchell, B. M., Porwancher, K. A., Wehner, K. A., Wormsley, S., Settlage, R. E., Shabanowitz, J., Osheim, Y., Beyer, A. L., Hunt, D. F., and Baserga, S. J. (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis, *Nature 417*, 967-970.

25. Fatica, A., and Tollervey, D. (2002) Making ribosomes, *Curr Opin Cell Biol 14*, 313-318.

26. Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schäfer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A. C., and Hurt, E. (2002) 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors, *Mol Cell 10*, 105-115.

27. Bossemeyer, D. (1995) Protein kinases--structure and function, *FEBS Lett 369*, 57-61.

28. Hanks, S. K., Quinn, A. M., and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains, *Science 241*, 42-52.

29. Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification, *FASEB J 9*, 576-596.

30. Krupa, A., and Srinivasan, N. (2002) Lipopolysaccharide phosphorylating enzymes encoded in the genomes of Gram-negative bacteria are related to the eukaryotic protein kinases, *Protein Sci 11*, 1580-1584.

31. Leonard, C. J., Aravind, L., and Koonin, E. V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily, *Genome research 8*, 1038-1047.

32. LaRonde-LeBlanc, N., Guszczynski, T., Copeland, T., and Wlodawer, A. (2005) Autophosphorylation of Archaeoglobus fulgidus Rio2 and crystal structures of its nucleotide-metal ion complexes, *Febs Journal 272*, 2800-2810.

33. LaRonde-LeBlanc, N., Guszczynski, T., Copeland, T., and Wlodawer, A. (2005) Structure and activity of the atypical serine kinase Rio1, *Febs Journal 272*, 3698-3713.

34. LaRonde-LeBlanc, N., and Wlodawer, A. (2004) Crystal structure of A-fulgidus Rio2 defines a new family of serine protein kinases, *Structure 12*, 1585-1594.

35. Gavin, A. C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes, *Nature 415*, 141-147.

36. Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sørensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry, *Nature 415*, 180-183.

37. Schafer, T., Strauss, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes, *Embo Journal 22*, 1370-1380.

38. Geerlings, T. H., Faber, A. W., Bister, M. D., Vos, J. C., and Raue, H. A. (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae, *Journal of Biological Chemistry 278*, 22537-22545.

39. Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard,

P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002) Functional profiling of the Saccharomyces cerevisiae genome, *Nature 418*, 387-391.

40.   Moy, T. I., and Silver, P. A. (1999) Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins, *Genes Dev 13*, 2118-2133.

41.   Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5, *Nature 364*, 412-420.

42.   Gajiwala, K. S., and Burley, S. K. (2000) Winged helix proteins, *Current opinion in structural biology 10*, 110-116.

43.   Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A., and Head, J. F. (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution, *Nature structural biology 8*, 710-714.

44.   Wolberger, C., and Campbell, R. (2000) New perch for the winged helix, *Nature structural biology 7*, 261-262.

45.   Angermayr, M., and Bandlow, W. (2002) RIO1, an extraordinary novel protein kinase, *FEBS Lett 524*, 31-36.

46.   Otwinoski, Z., Minor, W., and Charles W. Carter, J. (1997) Processing of X-ray diffraction data collected in oscillation mode., *Methods of Enzymology*, 307-326.

47.   McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *Journal of applied crystallography 40*, 658-674.

48.   Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta crystallographica. Section D, Biological crystallography 66*, 213-221.

49.   Adams, P. D., Afonine, P. V., Grosse-Kunstleve, R. W., Read, R. J., Richardson, J. S., Richardson, D. C., and Terwilliger, T. C. (2009) Recent developments in phasing and structure refinement for macromolecular crystallography, *Current opinion in structural biology 19*, 566-572.

50.   Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta crystallographica. Section D, Biological crystallography 53*, 240-255.

51.   Delano, W. (2002) The Pymol molecular graphics system, *Delano Scientific*.

52.   Zemp, I., Wild, T., O'Donohue, M. F., Wandrey, F., Widmann, B., Gleizes, P. E., and Kutay, U. (2009) Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2, *The Journal of cell biology 185*, 1167-1180.

53. Chen, L., Jiao, Z. H., Zheng, L. S., Zhang, Y. Y., Xie, S. T., Wang, Z. X., and Wu, J. W. (2009) Structural insight into the autoinhibition mechanism of AMP-activated protein kinase, *Nature 459*, 1146-1149.

54. Hu, S. H., Parker, M. W., Lei, J. Y., Wilce, M. C., Benian, G. M., and Kemp, B. E. (1994) Insights into autoregulation from the crystal structure of twitchin kinase, *Nature 369*, 581-584.

55. Gietz, R. D., and Schiestl, R. H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method, *Nature protocols 2*, 31-34.

56. Leclercq, R., and Courvalin, P. (1991) Intrinsic and unusual resistance to macrolide, lincosamide, and streptogramin antibiotics in bacteria, *Antimicrobial agents and chemotherapy 35*, 1273-1276.

57. Ross, J. I., Eady, E. A., Cove, J. H., Cunliffe, W. J., Baumberg, S., and Wootton, J. C. (1990) Inducible erythromycin resistance in staphylococci is encoded by a member of the ATP-binding transport super-gene family, *Molecular microbiology 4*, 1207-1214.

58. Weisblum, B. (1995) Erythromycin resistance by ribosome modification, *Antimicrobial agents and chemotherapy 39*, 577-585.

59. Noguchi, N., Emura, A., Matsuyama, H., O'Hara, K., Sasatsu, M., and Kono, M. (1995) Nucleotide sequence and characterization of erythromycin resistance determinant that encodes macrolide 2'-phosphotransferase I in Escherichia coli, *Antimicrobial agents and chemotherapy 39*, 2359-2363.

60. Noguchi, N., Takada, K., Katayama, J., Emura, A., and Sasatsu, M. (2000) Regulation of transcription of the mph(A) gene for macrolide 2'-phosphotransferase I in Escherichia coli: characterization of the regulatory gene mphR(A), *J Bacteriol 182*, 5052-5058.

61. Szczepanowski, R., Krahn, I., Bohn, N., Puhler, A., and Schluter, A. (2007) Novel macrolide resistance module carried by the IncP-1beta resistance plasmid pRSB111, isolated from a wastewater treatment plant, *Antimicrobial agents and chemotherapy 51*, 673-678.

62. Szczepanowski, R., Krahn, I., Linke, B., Goesmann, A., Puhler, A., and Schluter, A. (2004) Antibiotic multiresistance plasmid pRSB101 isolated from a wastewater treatment plant is related to plasmids residing in phytopathogenic bacteria and carries eight different resistance determinants including a multidrug transport system, *Microbiology 150*, 3613-3630.

63. Ramos, J. L., Martinez-Bueno, M., Molina-Henares, A. J., Teran, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R., and Tobes, R. (2005) The TetR family of transcriptional repressors, *Microbiol Mol Biol Rev 69*, 326-356.

64. Zheng, J., Sagar, V., Smolinsky, A., Bourke, C., LaRonde-LeBlanc, N., and Cropp, T. A. (2009) Structure and function of the macrolide biosensor protein, MphR(A), with and without erythromycin, *J Mol Biol 387*, 1250-1260.

65. Weber, W., Fux, C., Daoud-el Baba, M., Keller, B., Weber, C. C., Kramer, B. P., Heinzen, C., Aubel, D., Bailey, J. E., and Fussenegger, M. (2002) Macrolide-based transgene control in mammalian cells and mice, *Nature biotechnology 20*, 901-907.

66. Link, N., Weber, W., and Fussenegger, M. (2007) A novel generic dipstick-based technology for rapid and precise detection of tetracycline, streptogramin and macrolide antibiotics in food samples, *Journal of biotechnology 128*, 668-680.

67. Weber, C. C., Link, N., Fux, C., Zisch, A. H., Weber, W., and Fussenegger, M. (2005) Broad-spectrum protein biosensors for class-specific detection of antibiotics, *Biotechnology and bioengineering 89*, 9-17.

68. Mohrle, V., Stadler, M., and Eberz, G. (2007) Biosensor-guided screening for macrolides, *Analytical and bioanalytical chemistry 388*, 1117-1125.

69. Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) Automated structure solution with autoSHARP, *Methods Mol Biol 364*, 215-230.

70. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta crystallographica. Section D, Biological crystallography 60*, 2126-2132.

71. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement, *Acta crystallographica. Section D, Biological crystallography 57*, 122-133.

72. Hinrichs, W., Kisker, C., Duvel, M., Muller, A., Tovar, K., Hillen, W., and Saenger, W. (1994) Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance, *Science 264*, 418-420.

73. Dover, L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatituri, V., Besra, G. S., and Futterer, K. (2004) Crystal structure of the TetR/CamR family repressor Mycobacterium tuberculosis EthR implicated in ethionamide resistance, *J Mol Biol 340*, 1095-1105.

74. Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2001) Structural mechanisms of QacR induction and multidrug recognition, *Science 294*, 2158-2163.

75. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta crystallographica. Section D, Biological crystallography 60*, 2256-2268.

76. Jeng, W. Y., Ko, T. P., Liu, C. I., Guo, R. T., Liu, C. L., Shr, H. L., and Wang, A. H. (2008) Crystal structure of IcaR, a repressor of the TetR family implicated in biofilm formation in Staphylococcus epidermidis, *Nucleic Acids Res 36*, 1567-1577.

77. Willems, A. R., Tahlan, K., Taguchi, T., Zhang, K., Lee, Z. Z., Ichinose, K., Junop, M. S., and Nodwell, J. R. (2008) Crystal structures of the Streptomyces coelicolor TetR-like protein ActR alone and in complex with actinorhodin or the actinorhodin biosynthetic precursor (S)-DNPA, *J Mol Biol 376*, 1377-1387.

78. Anderson, J. C., Voigt, C. A., and Arkin, A. P. (2007) Environmental signal integration by a modular AND gate, *Molecular systems biology 3*, 133.

79. Sayut, D. J., Kambam, P. K., and Sun, L. (2007) Engineering and applications of genetic circuits, *Molecular bioSystems 3*, 835-840.

80. Cunin, R., Glansdorff, N., Piérard, A., and Stalon, V. (1986) Biosynthesis and metabolism of arginine in bacteria, *Microbiol Rev 50*, 314-352.

81. Slocum, R. D. (2005) Genes, enzymes and regulation of arginine biosynthesis in plants, *Plant Physiol Biochem 43*, 729-745.

82. Caldovic, L., and Tuchman, M. (2003) N-acetylglutamate and its changing role through evolution, *Biochem J 372*, 279-290.

83. Martin, P. R., and Mulks, M. H. (1992) Sequence analysis and complementation studies of the argJ gene encoding ornithine acetyltransferase from Neisseria gonorrhoeae, *J Bacteriol 174*, 2694-2701.

84. GRISOLIA, S., and COHEN, P. P. (1953) Catalytic rôle of of glutamate derivatives in citrulline biosynthesis, *J Biol Chem 204*, 753-757.

85. Bachmann, C., Colombo, J. P., and Jaggi, K. (1982) N-acetylglutamate synthetase (NAGS) deficiency: diagnosis, clinical observations and treatment, *Adv Exp Med Biol 153*, 39-45.

86. Caldovic, L., Morizono, H., Gracia Panglao, M., Gallegos, R., Yu, X., Shi, D., Malamy, M. H., Allewell, N. M., and Tuchman, M. (2002) Cloning and expression of the human N-acetylglutamate synthase gene, *Biochem Biophys Res Commun 299*, 581-586.

87. Caldovic, L., Morizono, H., Yu, X., Thompson, M., Shi, D., Gallegos, R., Allewell, N. M., Malamy, M. H., and Tuchman, M. (2002) Identification, cloning and expression of the mouse N-acetylglutamate synthase gene, *Biochem J 364*, 825-831.

88. Qu, Q., Morizono, H., Shi, D., Tuchman, M., and Caldovic, L. (2007) A novel bifunctional N-acetylglutamate synthase-kinase from Xanthomonas campestris that is closely related to mammalian N-acetylglutamate synthase, *BMC Biochem 8*, 4.

89. Errey, J. C., and Blanchard, J. S. (2005) Functional characterization of a novel ArgA from Mycobacterium tuberculosis, *J Bacteriol 187*, 3039-3044.

90. Dyda, F., Klein, D. C., and Hickman, A. B. (2000) GCN5-related N-acetyltransferases: a structural overview, *Annu Rev Biophys Biomol Struct 29*, 81-103.

91. He, H., Ding, Y., Bartlam, M., Sun, F., Le, Y., Qin, X., Tang, H., Zhang, R., Joachimiak, A., Liu, J., Zhao, N., and Rao, Z. (2003) Crystal structure of tabtoxin resistance protein complexed with acetyl coenzyme A reveals the mechanism for beta-lactam acetylation, *J Mol Biol 325*, 1019-1030.

92. Ramon-Maiques, S., Fernandez-Murga, M. L., Gil-Ortiz, F., Vagin, A., Fita, I., and Rubio, V. (2006) Structural bases of feed-back control of arginine biosynthesis, revealed by the structures of two hexameric N-acetylglutamate kinases, from Thermotoga maritima and Pseudomonas aeruginosa, *J Mol Biol 356*, 695-713.

93. Ramón-Maiques, S., Marina, A., Gil-Ortiz, F., Fita, I., and Rubio, V. (2002) Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis, *Structure 10*, 329-342.

94. Shi, D., Caldovic, L., Jin, Z., Yu, X., Qu, Q., Roth, L., Morizono, H., Hathout, Y., Allewell, N. M., and Tuchman, M. (2006) Expression, crystallization and preliminary crystallographic studies of a novel bifunctional N-acetylglutamate synthase/kinase from Xanthomonas campestris homologous to vertebrate N-acetylglutamate synthase, *Acta Crystallogr Sect F Struct Biol Cryst Commun 62*, 1218-1222.

95.   Shi, D., Yu, X., Roth, L., Morizono, H., Hathout, Y., Allewell, N. M., and
        Tuchman, M. (2005) Expression, purification, crystallization and preliminary X-
        ray crystallographic studies of a novel acetylcitrulline deacetylase from
        Xanthomonas campestris, *Acta Crystallogr Sect F Struct Biol Cryst Commun 61*,
        676-679.
96.   Marvil, D. K., and Leisinger, T. (1977) N-acetylglutamate synthase of Escherichia
        coli: purification, characterization, and molecular properties, *J Biol Chem 252*,
        3295-3303.
97.   Shi, D., Sagar, V., Jin, Z., Yu, X., Caldovic, L., Morizono, H., Allewell, N. M.,
        and Tuchman, M. (2008) The crystal structure of N-acetyl-L-glutamate synthase
        from Neisseria gonorrhoeae provides insights into mechanisms of catalysis and
        regulation, *J Biol Chem 283*, 7176-7184.
98.   Marina, A., Alzari, P. M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I., and
        Rubio, V. (1999) Carbamate kinase: New structural machinery for making
        carbamoyl phosphate, the common precursor of pyrimidines and arginine, *Protein
        Sci 8*, 934-940.
99.   Ramón-Maiques, S., Marina, A., Uriarte, M., Fita, I., and Rubio, V. (2000) The
        1.5 A resolution crystal structure of the carbamate kinase-like carbamoyl
        phosphate synthetase from the hyperthermophilic Archaeon pyrococcus furiosus,
        bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and
        provides insight into substrate binding and stability in carbamate kinases, *J Mol
        Biol 299*, 463-476.
100.   Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Barzu, O., and Gilles, A.
        M. (2005) Structure of Escherichia coli UMP kinase differs from that of other
        nucleoside monophosphate kinases and sheds new light on enzyme regulation, *J
        Biol Chem 280*, 25533-25540.
101.   Marco-Marín, C., Gil-Ortiz, F., and Rubio, V. (2005) The crystal structure of
        Pyrococcus furiosus UMP kinase provides insight into catalysis and regulation in
        microbial pyrimidine nucleotide biosynthesis, *J Mol Biol 352*, 438-454.
102.   Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006)
        Structures of R- and T-state Escherichia coli aspartokinase III. Mechanisms of the
        allosteric transition and inhibition by lysine, *J Biol Chem 281*, 31544-31552.
103.   Marco-Marín, C., Gil-Ortiz, F., Pérez-Arellano, I., Cervera, J., Fita, I., and Rubio,
        V. (2007) A novel two-domain architecture within the amino acid kinase enzyme
        family revealed by the crystal structure of Escherichia coli glutamate 5-kinase, *J
        Mol Biol 367*, 1431-1446.
104.   Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of
        distance matrices, *J Mol Biol 233*, 123-138.
105.   Neuwald, A. F., and Landsman, D. (1997) GCN5-related histone N-
        acetyltransferases belong to a diverse superfamily that includes the yeast SPT10
        protein, *Trends Biochem Sci 22*, 154-155.
106.   Marmorstein, R., and Roth, S. Y. (2001) Histone acetyltransferases: function,
        structure, and catalysis, *Curr Opin Genet Dev 11*, 155-161.
107.   Vetting, M. W., Roderick, S. L., Yu, M., and Blanchard, J. S. (2003) Crystal
        structure of mycothiol synthase (Rv0819) from Mycobacterium tuberculosis

shows structural homology to the GNAT family of N-acetyltransferases, *Protein Sci 12*, 1954-1959.

108. Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E., and Noble, M. E. (2000) Structure of arylamine N-acetyltransferase reveals a catalytic triad, *Nature structural biology 7*, 560-564.

109. Wong, L. J., and Wong, S. S. (1983) Kinetic mechanism of the reaction catalyzed by nuclear histone acetyltransferase from calf thymus, *Biochemistry 22*, 4637-4641.

110. Yan, Y., Harper, S., Speicher, D. W., and Marmorstein, R. (2002) The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate, *Nature structural biology 9*, 862-869.

111. Xu, Y., Glansdorff, N., and Labedan, B. (2006) Bioinformatic analysis of an unusual gene-enzyme relationship in the arginine biosynthetic pathway among marine gamma proteobacteria: implications concerning the formation of N-acetylated intermediates in prokaryotes, *BMC Genomics 7*, 4.

112. Xu, Y., Labedan, B., and Glansdorff, N. (2007) Surprising arginine biosynthesis: a reappraisal of the enzymology and evolution of the pathway in microorganisms, *Microbiol Mol Biol Rev 71*, 36-47.

113. Labedan, B., Xu, Y., Naumoff, D. G., and Glansdorff, N. (2004) Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase, *Mol Biol Evol 21*, 364-373.

114. Granneman S, Petfalski E, Swiatkowska A, Tollervey D. (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA-protein cross-linking, *EMBO J.,* 29(12), 2026-36.

115. Strunk BS, Loucks CR, Su M, Vashisth H, Cheng S, Schilling J, Brooks CL 3rd, Karbstein K, Skiniotis G. (2011) Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates, Science., 333(6048), 1449-53.

116. Campbell MG, Karbstein K. (2011) Protein-protein interactions within late pre-40S ribosomes., PLoS One., 6(1):e16194.