UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Characterization of the Archaebacterial Ortholog of**

**the Processing Body Subunit Winnebago/Growl**

A thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

in

Biology

by

Amir Motamedi

Committee in charge:

Professor James E. Wilhelm, Chair
Professor Randolph Y. Hampton
Professor Partho Ghosh

2010

This thesis of Amir Motamedi is approved and it is acceptable in quality and form from publication on microfilm and electronically:

_____

_____

_____
Chair

University of California, San Diego

2010

## DEDICATION

This work is dedicated to my Mom, Dad, and Aunt for their continuous and tremendous love and support. Your enormous strength and encouragement empowers me as I face every journey in my life.

Thank you and I love you!

## EPIGRAPH

*Human beings are members of a whole,*
*In creation of one essence and soul.*
*If one member is afflicted with pain,*
*Other members uneasy will remain.*
*If you have no sympathy for human pain,*
*The name of human you cannot retain.*

Persian poet **Sa'di of Shiraz**

## TABLE OF CONTENTS

## LIST OF ABBREVIATIONS

RNP……Ribonucleoprotein
*nos*……*nanos*
*osk*……*oskar*
Tral……Traler Hitch
Exu……Exuperantia
Yps…… Ypsilon Schachtel
RRM……RNA recognition motif
Winn……Winnebago
MTHF……Methenyltetrahydrofolate
MTHFS……Methenyltetrahydrofolate synthetase
5-FTHF ……5-formyltetrahydrofolate
AMP……Ampicillin
LB……Luria Broth
□-ME …… beta-mercaptoethanol
PI……Protease inhibitor
Ni-NTA …… nickel-nitrilotriacetic acid
SDS-PAGE..... sodium dodecyl sulfate polyacrylamide gel electrophoresis
IEC……Ion exchange chromatography
pI……Isoelectric point
FA……..Folinic acid

**List of Figures, Tables, Diagrams and Flowcharts**

**ACKNOWLEDGEMENTS**

I would like to greatly thank Dr. James Wilhelm for welcoming me to his laboratory and allowing my growth in carrying out independent scientific research. He has encouraged my aspirations, and helped me in my research projects throughout the past two years, particularly in the past year. I would also like to thank Dr. Andrew Shiau with whom I initially worked and learned a great deal about basic molecular techniques in the Wilhelm lab.

Moreover, I would like to thank the members of the Wilhelm lab, Risa Maruyama Broyer, Elena Monfort, Brian Sato, and Charlie Noree who helped me throughout this journey. They not only helped me troubleshoot countless experiments gone awry, but also shared with me their kindness, wisdom, and friendship. In addition, I would like to show appreciation toward all the help that I received from Philip Jiang.

Most importantly, I would not have achieved my goals thus far had it not been for the encouragement, support, and the teachings of my family. Their unconditional love, always available ears, strength, and knowledge have provided me with the ability to set my goals high and achieve them one by one. Thus, here is thank you to Mom, Zahra Aminilari,

, Dad, Hamid Motamedi , my brothers Morteza and Majid Motamedi, and my aunt and her husband, Marjan and Alan Loghman.

I also want to thank Afshin Salehi for his friendship and support.

ABSTRACT OF THE THESIS

**Characterization of the Archaebacterial Ortholog of the Processing Body Subunit Winnebago/Growl**

by

Amir Motamedi

Master of Science in Biology

University of California, San Diego, 2009

Professor James E. Wilhelm, Chair

Because the transcriptional machinery is inactive throughout late stages of oogenesis and early stages of embryogenesis, the initial development of the embryo is driven by post-transcriptional regulation of maternally deposited mRNA. In *Drosophila*, this post-transcriptional regulation occurs at specialized sites called P- bodies. While P bodies

contain proteins required for mRNA degradation, translational control, and microRNA processing, the complete composition of the P body is unknown. Recent work on the biochemistry of the P body has identified a novel protein of unknown function called Winnebago. Winnebago contains an atypical conserved ATPase motif homologous with the cyclo-ligase human methenyltetrahydrofolate synthetase (MTHFS) and a conserved RNA recognition motif suggesting that Winnebago may be a novel RNA modifying enzyme. Consistent with this, the substrate binding site of the Winnebago and its archaebacterial homologs are not conserved with the human MTHFS, indicating that Winnebago has a unique class of substrate. In this study, Swinn, the Winnebago homolog from *Sulfolobus,* was purified and it was demonstrated that it is an ATPase. We also found that Swinn does not use the normal substrates of MTHFS, arguing that it acts on a novel class of cyclo-ligase substrates. This work lays the groundwork for future studies to identify the substrate of the Winnebago family of ATPases and their role in RNA regulation.

# INTRODUCTION

## Maternal mRNA Regulation and Stability

In metazoans, transcription is silent throughout late oogenesis and during the early stages of embryogenesis. As a result, development during these stages is dictated by a large pool of maternal mRNA deposited in the embryo during early oogenesis [1]. However, the stockpiling of these maternal transcripts in the developing oocyte poses a problem for oogenesis, since the accumulation of these mRNAs would normally cause a corresponding increase in the proteins encoded by those transcripts. As a result, the developing oocyte translationally regulates the pool of maternally deposited mRNAs. This post-translational regulation occurs primarily by translationally repressing the maternal transcripts through a combination of transcript deadenylation and the recruitment of inhibitors of translation to particular messages via sequence specific RNA binding proteins [2] . While the dominant mode of post-transcriptional regulation of maternal messages is via translational regulation, translational regulation is often combined with mRNA localization and regulated mRNA stability to drive the events of early embryogenesis.

*Drosophila* has been a particularly powerful system for understanding how these different forms of regulation are combined. It

has been shown that asymmetric localization of several maternal mRNA during oogenesis in *Drosophila* is essential for normal development of embryo and establishment of the germ line [1]. One of the maternal mRNAs involved in proper embryonic patterning in *Drosophila* is *oskar (osk)* mRNA. During early stages of oogenesis, *osk* mRNA is localized to the posterior pole of developing oocyte, where it is translationally activated and results in formation of OSK protein gradient. This posterior localization is accomplished though a combination of mRNA localization and translational control. First, *osk* translation is repressed by binding to an RNA-binding protein called Bruno. This would result in recruitment of Cup protein which is involved in both localization and translational repression of *oskar* mRNA [3, 4]. Following *osk* localization to the posterior, the translational repression of osk is relieved by an unknown mechanism that leads to localized translation of OSK at the posterior pole.

Posterior development is also dependent on the posterior localization and translation of *nanos (nos)* mRNA, another maternal mRNA. Unlike *osk* mRNA the generation of a NOS gradient is dependent on the regulation of both mRNA stability and translation. This is done though deadenylation and subsequent destabilization and decay of *nos* mRNA throughout the oocyte except at the posterior pole where deadenylation is inhibited though activity of post-transcriptional

regulatory factor. Deadenylation occurs by binding of Smaug protein to the *nanos* 3' UTR resulting in recruitment of the deadenylating complex CCR4-NOT. However, at the posterior pole Oskar protein inhibits binding of Smaug to the *nanos* mRNA resulting in stabilization and translation of Nanos protein [5].

## Maternal RNP Complexes

Since maternal messages are often subject to multiple forms of post-transcriptional control (i.e. translation, localization, and stability), in order to drive the events of early development, the identification and characterization of the proteins that mediate these forms of regulation has been a subject of intense investigation in multiple developmental systems. These studies have led to the discovery of large ribonucleoprotein (RNP) complexes, variously called P bodies, P granules, or sponge bodies. Interestingly, while these complexes have acquired a variety of names throughout the years, they are all composed of components that are highly conserved across a variety of species including *Drosophila* [6] (Table 1). Functional studies of the proteins involved in the P bodies indicates that these large complexes are not only involved in mRNA degradation, but also they participate in mRNA localization, repression, and storage [7]. For example, Trailer hitch (Tral),

one of the most conserved components of P bodies, is involved in mRNA localization and translational control. Me31B protein is involved in translational repression of maternal mRNA, which prevents unlocalized protein from ectopically being translated. Additionally, in many cases these component work together on the same transcript to allow allows for a correct localization and pattering of mRNA and its resulting protein, leading to a correct development of embryo. For example, the correct pattering of Oskar protein is achieved by the activity of many components of RNP complex. For instance, Exuperantia (Exu) protein along with Ypsilon Schachtel (YPS) protein are involved in localization of *osk* mRNA, prior to its translation [8]. Cup is involved in both localization and repression of *osk* mRNA. Such precise mechanism of localization, prior to translation of mRNA, is essential for the normal development of the embryo. Previous studies have shown that ectopic translation of maternal mRNA results in abnormal embryonic patterning and lethality of the embro [1]. Interconnectivity of functions of components of P-bodies and their tight post-transcriptional regulation of maternal mRNA raises the question that there might be additional modes of mRNA regulation that remain to be discovered.

## Identification of Winnebago

The first step in discovery of additional modes of mRNA regulation by RNP complexes is to identify all of the components of the maternal RNP complex. Previous work in *Drosophila* focused on purifying maternal RNP complexes and characterizing the functional roles of their subunits. In these studies, Exu protein was used as a biochemical handle for immunopercipitation technique in order to identify the components of these large complexes. Following the process of immunopercipitation of Exu and mass spectrometry of pulled down particles, at least 10 proteins were identified along with Exu in the RNP complex. One of the proteins in association with this complex is a 59.74 KDa previously uncharacterized protein named Growl/Winnebago.

## Domains of Winnebago

In addition to co-purifying with maternal RNP complexes, Winnebago also co-localizes with other RNP subunits of the P-bodies in the oocyte of *Drosophila* egg chamber [9]. The primary protein sequence of Winnebago also shows two interesting conserved domains (Figure 1). Consistent with the association of Winnebago with RNP complexes, the carboxyl terminal domain contains an RNA recognition motif (RRM), which

is an RNA binding motif. Winnebago also possesses a domain homologous to human methenyltetrahydrofolate-synthetase (MTHFS).

MTHFS is a cyclo-ligase. This enzyme catalyzes the formation of a carbon nitrogen ring by converting 5-formyltetrahydrofolate (5-FTHF) into 5-10 methenyltetra-hydrofolate. The proposed chemical reaction for this enzyme utilizes ATP and Mg2+. The mechanism of reaction starts by a nucleophilic attack of 5-formyl oxygen on the γ-phosphate of ATP. The result is formation of an N5-iminium phosphate intermediate which itself undergoes a cyclization reaction via nucleophilic attack by N10, resulting in formation of phosphoimidazolidine tetrahedral intermediate. Finally, the irreversible process of phosphate elimination follows, which results in generation of methenyltetrahydrofolate product (figure 2).

Sequence comparisons of the catalytic domain of MTHFS and Winnebago shows that the catalytic residues involve APTase activity are conserved, suggesting that Winnebago may perform a similar or related chemical reaction on RNA.

One of the interesting features of MTHFS is that it is an ATPase that lacks a canonical ATP binding motif. Typically, most of the ATPase enzymes belong to "ATP-binding cassette transporter" (ABC-transporter) superfamily. This class of ATPase has two common ATPase motifs. First, is the "Walker A" or "P-loop" motif which is consist of two β-sheets and six α

helices with the consensus sequence of GXXGXGK-S/T, where X is any amino acid. The second common motif of this superfamily is known as "Walker B" motif with consensus sequence of R/K-X(7-8)- ΦΦΦΦD, where Φ is a hydrophobic residue (Figure 3). MTHFS carries an unusual ATPase motif with a consensus sequence of GXGXG. This sequence seems to be conserved between MTHFS, Winnebago, and its archaebacterial homologs (Figure 4). This sequence analysis suggests that Winnebago is a novel RNA regulatory/modifying enzyme that uses ATP hydrolysis as part of its function.

**Experimental approach**

In this study, we focus on characterizing the catalytic activity of Winnebago homologs. We start by identifying, expressing and purifying a soluble Winnebago homolog. Then, we attempt to determine the potential substrate of the purified protein by obtaining the crystal structure of the protein. Lastly, we focus on confirming and measuring the extent of ATP hydrolysis activity of the purified protein by performing a $K_m$ and $K_{cat}$ analysis on the protein.

**Table 1: Conserved component of P bodies across species**

This table shows that many protein components of P bodies are conserved between *Drosophila, Homo sapiens, Xenopus, C. elegans*, and *Yeast*. These components play different roles such as localization, repression, and translation of mRNA.

| *Drosophila* | *Human* | *Xenopus* | *C. Elegans* | *Yeast* | Functions |
|---|---|---|---|---|---|
| **Me31B** | RCK | Xp54 | CGH-1 | Dhh1 | Repression |
| **Cup** | 4ET | Maskin | Spn2 | - | Localization<br>Repression |
| **YPS** | YB1 | FRGY2 | CEY-2, 3, 4 | - | Localization |
| **Exu** | - | - | - | - | Localization |
| **Tral** | RAP55 | - | Car-1 | SCD6 | Localization<br>Translation |

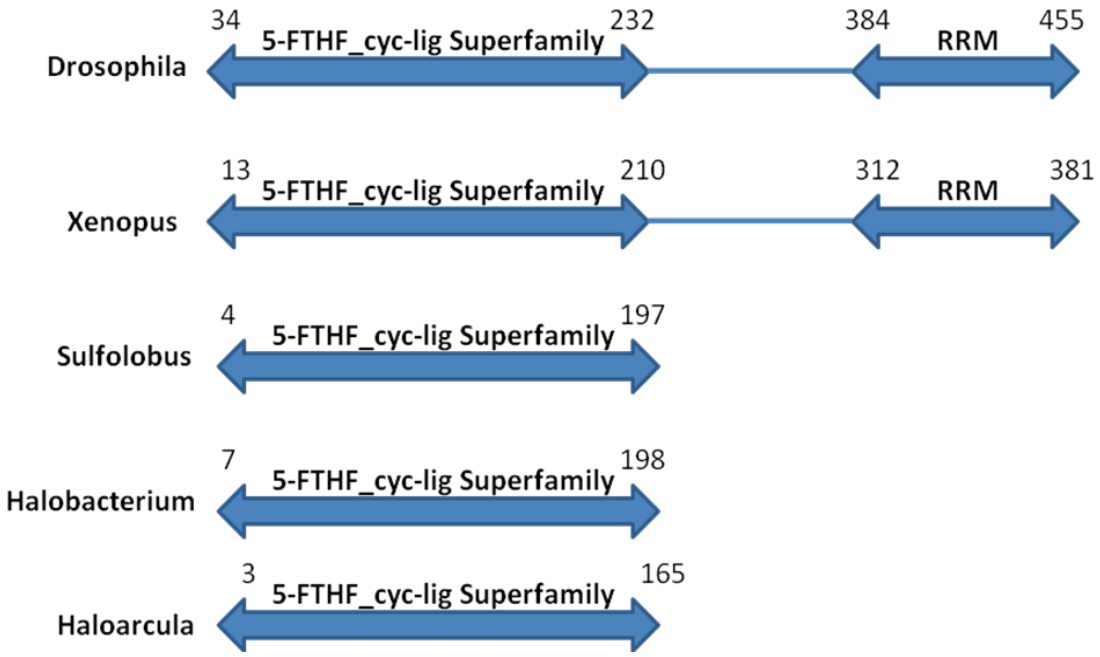



**Figure 1: *Drosophila* Winnebago has two conserved domains: RNA Recognition Motif and 5-formyltetrahydrofolate cyclo-ligase/5-methenyltetrahydrofolate synthetase motif**

This figure indicates that cyclo-ligase domain is conserved between *Drosophila* Winnebago and the homologous protein from eukaryotes and archaebacteria. The RNA Recognition Motif (RRM) is, however, only conserved in eukaryotes.

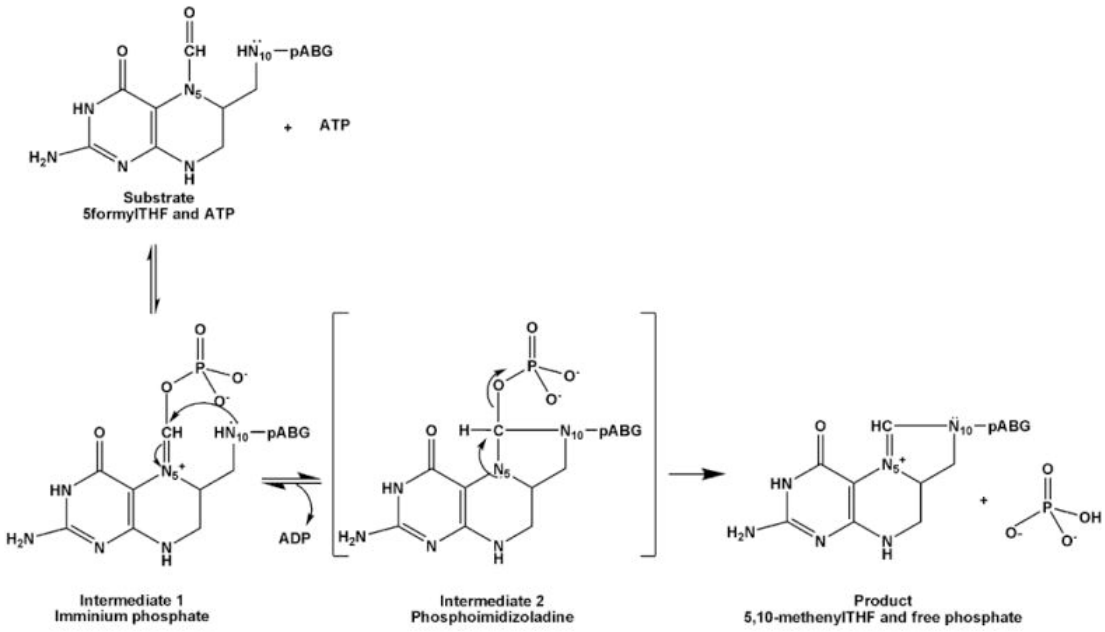



**Figure 2: Chemical reaction of human MTHFS**

Human MTHFS utilizes ATP and $Mg^{2+}$ in order to convert 5-FTHF into 5-10-MTHF, through a nucleophilic attack of 5-formyl oxygen on the γ-phosphate, resulting in formation of an N5-iminium phosphate intermediate. This intermediate is unstable and undergoes a cyclization reaction via nucleophilic attack by N10, resulting in formation of phosphoimidazolidine tetrahedral intermediate. Finally, the phosphate gets eliminated and 5-10-MTHF product will be generated.

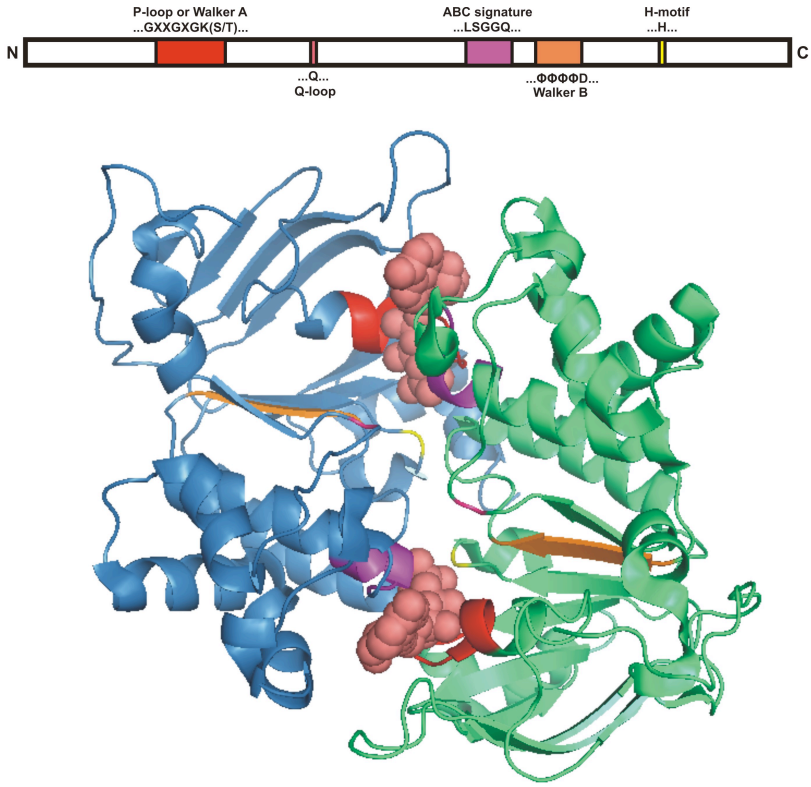



**Figure 3: Walker A and Walker B motif**

This figure shows the typical ATPase "Walker A" or "P-loop" motif with the consensus sequence of GXXGXGK-S/T, where X is any amino acid. Walker A motif consists of two β-sheets and six α helices. Walker B motif is another typical ATPase motif with the consensus sequence of R/K-X(7-8)- ΦΦΦΦD, where Φ is a hydrophobic residue.

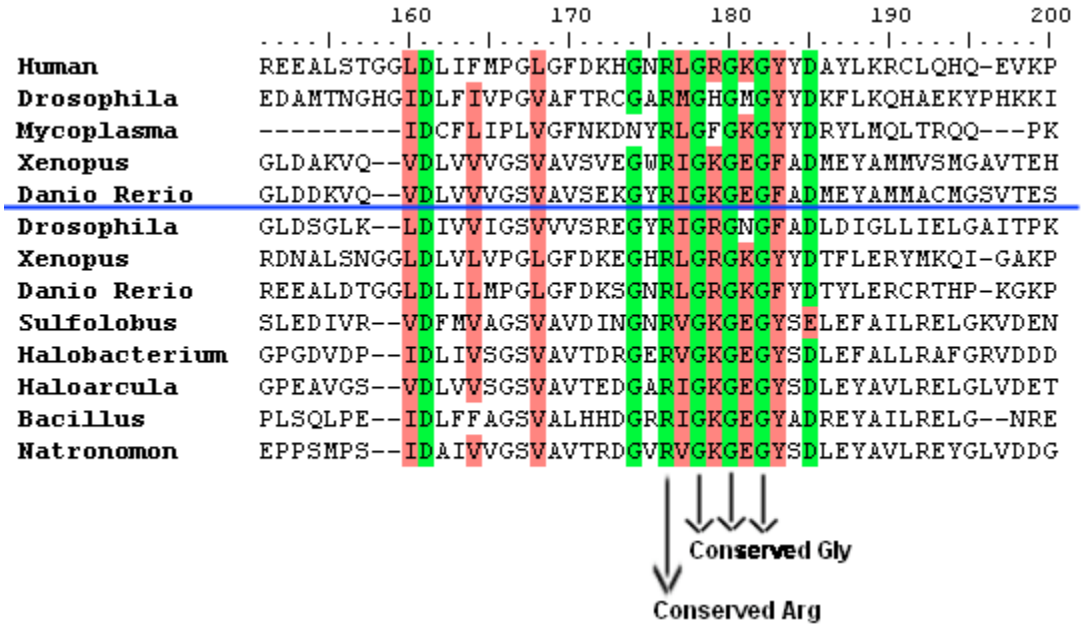



**Figure 4: Conserved residues of atypical ATPase of MTHFS between different species**

This figure shows that the indicated glycine and arginine residues inside the ATP binding pocket of the protein is absolutely conserved between different species containing this protein. The protein sequences above the blue line are from true 5-FTHF cyclo ligase protein and the ones below the blue line are from 5-FTHF cyclo ligase- homologous proteins. Some organisms have both versions. Green shading indicates the absolute identity, and red shading indicates the similarity of the residues.

# MATERIALS AND METHODS

## Protein expression

The open reading frame from cDNA sequence of Winnebago homologs from *Danio Rerio*, *Xenopus*, *Bacillus*, *Natronomonas*, *Haloarcula*, *Halobacterium*, and *Sulfolobus* species were sub-cloned into pET151/D-TOPO vector using Topo cloning system from Invitrogen kit. The cloned vectors were transformed into the *E. coli* BL21 (DE3) cells. Following transformation, the cells were incubated overnight at 37°C on an agar plate containing ampicillin (Amp), in order to obtain colonies. Each colony was inoculated overnight at 37°C in Luria Broth (LB) medium containing 100µg/ml Amp. The overnight culture was diluted to 0.05 OD at □=600nm. The diluted culture was subsequently grown to 0.2 OD and then transferred to the growth temperatures of 19 °C, 27°C , or 37°C until they reached 0.5- 0.7 OD. At this OD, protein expression was induced by adding 500□M Isopropyl □-D-1-thiogalactopyranoside to the growing culture and allowing the culture to express protein for 4 hours at their corresponding temperature. The culture was then spun down at 4000rpm for 15 minutes at 4°C. The supernatant was discarded and the pelleted cells were frozen using liquid nitrogen and stored at -80°C.

**First step of protein purification: Metal Chromatography**

In this step, the expressed protein was purified using the His-Tag at the N-terminus of the protein. The frozen cell pellet from 1.5 mL cell culture was thawed out in 40mL of metal chromatography lysing buffer containing 50mM Tris-Cl pH 8.0, 250mM NaCl, 14.3mM beta-mercaptoethanol (□-ME), and cocktail protease inhibitor (PI). The cells were lysed using a micro-fluidizer. The lysed cells were centrifuged at 15000rpm for 30 minutes at 4°C. The pellet was discarded and the supernatant was used protein purification using metal chromatography. The supernatant was bound in batch to 2ml of pre-equilibrated nickel-nitrilotriacetic acid (Ni-NTA)Agarose beads. The wash buffer of metal chromatography containing 25mM Tris-Cl pH 8.2, 150mM NaCl, and 14.3mM □-ME was used for equilibration of the beads. The binding process was allowed for 15 minutes at 4 °C. Subsequently, the beads were separated from the supernatant by a spinning process at 2000rpm for 5 minutes at 4 °C. The beads were washed three times using 25mL wash buffer. The purified protein was eluted from the beads four times using 5mL elution buffer containing 150mM NaCl, 250mM imidazole pH 8.0, and 14.3mM □-ME. The eluted protein were stored on ice and used for further purification using ion exchange chromatography. Samples of solutions from each step of metal chromatography were analyzed using SDS-

polyacrylamide gels (SDS-PAGE gel) and the existence of protein was confirmed using Coomassie blue staining.

**Second step of purification: Ion Exchange Chromatography**

In this step of purification, Q fast flow sepharose beads, an anion exchanger, were used for purification of Winnebago homolog from *Haloarcula* and *Halobacterium*, which were named Arcwinn, and Bacwinn, with pI of 4.57, and 4.62, respectively. SP fast flow sepharose beads, a cation exchanger, were used for purification of Winnebago homolog from *Sulfolobus* specie, which was named Swinn with pI of 8.96. The eluted Swinn protein from metal chromatography was diluted 1:5 using dilution buffer containing 25mM HEPES pH 7.2, and 14.3mM □-ME; and the Arcwinn and Bacwinn eluted protein from metal chromatography step was diluted 1:3 using dilution buffer containing 25mM Tris pH 8, and 14.3mM □-ME. Each diluted solution was applied to a corresponding pre-equilibrated 600µl of unpacked beads in a column. Wash buffer was used for pre-equilibration of each of the beads. The wash buffer for Swinn contained 25mM HEPES pH 7.2, 50mM NaCl, and 14.3mM □-ME. The wash buffer for Arcwinn and Bacwinn contained 25mM Tris pH 8, 50mM NaCl, and 14.3mM □-ME. After collecting the flow through, each column was washed with 6mL of wash buffer. Subsequently, the protein

was eluted from each column 10 times using 300□l of corresponding elution buffer. In order to optimize the elution condition, the protein was eluted by varying the salt concentration in the elution buffer. The optimal elution buffer for the Swinn protein was found to be 25mM HEPES pH 7.2, 300mM NaCl, and 14.3mM □-ME, and the optimal elution buffer for Arcwinn and Bacwinn was found to be 25mM Tris pH 8, 300mM NaCl, and 14.3mM □-ME. The collected protein was run on the SDS-PAGE gel and the purity and existence of the corresponding protein was confirmed using Coomassie blue staining technique.

**His-Tag cleavage followed by metal chromatography:**

This process was performed using 1mL of Ni-NTA agarose beads, pre-equilibrated in the corresponding elution buffer from ion exchange chromatography step, which was also used as a wash buffer for the step. The cleavage process was performed using 2.5U proTEV protease per 100µg of protein. The Swinn and Bacwinn protein were successfully cleaved, however, the Arcwinn protein was not cleaved. The cleaved protein was then reapplied to the Ni-NTA column using previously described metal chromatography procedure. The flow through containing the cleaved protein was collected and analyzed by SDS-PAGE

gel in order to confirm the success of process and purity of the obtained protein.

**Heat treatment**

The process of heat treatment was performed at 50°C, and 60°C for 30 minutes, following by a rapid cool down of the protein solution on ice. The heat treated solution was spun down at 14000rpm for 30 minutes at 4°C in order to eliminate the precipitated contaminant protein. The obtained solution was run on the SDS-PAGE gel and the success of purification was examined using Coomassie blue staining.

**Crystallography**

The process of crystallography was performed using predefined reagent kits such as CSHT, JCSG, The ComPAS, The MPD, and The Protein Complexes. During this process over six thousand different conditions were screened. These conditions included protein incubated with different nucleotide such as 5mM ATP, ATPγS, AMP-PNP, AMP-PCP, ADP, and ADP $AlF_3$ at both room temperature and 4°C. For each condition1µl of wells buffer was added to 1ul of purified protein incubated with appropriate nucleotide overnight. The mixed buffer and sample were placed in a sub-well and the wells were sealed.

**Enzymology**

In order to determine if the Swinn protein can hydrolyze ATP we used the "Promega ADP-Glo Kinase Assay" assay to identify the optimal reaction conditions of the protein, as well as to determine the Km and Kcat of the protein. Since the true substrate of the Swinn protein has not been identified yet, our experiments focused on the determining the basal ATPase activity level of the Swinn protein in the absence of its substrate. This assay detects the amount of ADP produced by the protein during the ATP hydrolysis reaction by eliminating all the remaining ATP substrate of the reaction after the ATP hydrolysis activity, and then converting the ADP product of the hydrolysis reaction into ATP. The newly synthesized ATP is then used in a luciferase/luciferin reaction to generate luminescence light. The amount of light generated correlates to the amount of ADP generated in the ATP hydrolysis reaction.

In our initial studies of Swinn protein, we sought to determine the existence of ATPase activity for the solution of purified Swinn protein. In the ATPase activity assay the ATP concentration was set to 100uM, and a 1:3 serial dilution of Swinn protein was performed from 0 mg/ml to 1 mg/ml over 12 points. The reaction condition was set to 50mM HEPES 7.2, 150mM NaCl, 1mM $MgCl_2$, and the reaction was performed at 50°C for one hour.

Subsequently, the Promega ADP-Glo Kinase kit was used to detect the level of ATPase activity.

After establishing that our protein had ATPase activity, we sought to use the heat stability of the Swinn protein to demonstrate that the ATPase activity was not due to a contaminating *E. coli* ATPase. In order to do so, the same experiment was performed with mutated, and heat-treated version of Swinn protein.

In order to determine the optimal reaction condition of Swinn protein, the effect of different reaction temperature was examined at 50°C and 60°C. Subsequently, the effect of different pH was examined using 50mM MES pH 6, 50mM HEPES pH 7 and 8, and 50mM Tris pH 9. Additionally, the effect of different $MgCl_2$ concentration at pH 7.2 was examined.

The Km of the Swinn protein was calculated by performing an experiment using 50mM MES at pH 6, and 1mM $MgCl_2$ at 50°C. The protein concentration was set to 0.5mg/ml. The experiment was performed for different ATP concentration of 300 µM, 250 µM, 200 µM, 100 µM, 75 µM, 50 µM, 25 µM, and 10 µM. Samples of reaction solution was taken at different time points from 0 minute to 1 hour during the reaction and the amount of ADP produced at each time point was calculated using a standard curve for the reaction.

Lastly, the experiment was performed in presence of folinic acid (FA) and 5,10 MTHF. In this experiment, the ATP concentration was set to 100 μM, and the FA and MTHF concentration was set to 250 μM. The reaction was performed at 50°C at 50mM HEPES pH 7.2.

# RESULTS

## Identification of the substrate for Winnebago

Since the number of candidate substrates is large, we have pursued a two pronged strategy that uses both enzymological and structural approaches to identify the natural substrates of Winnebago. This section outlines our structural studies of Winnebago in order to identify its substrate binding pocket as a first step towards defining the characteristics of Winnebago substrates.

MTHFS is an ATPase that lacks a canonical ATP binding motif. Previous work on the MTHFS crystal structure has identified GXGXG as the key residues that form the ATP binding pocket and mediate ATP hydrolysis (Figure 5). A multiple sequence alignment between Winnebago and MTHFS enzymes revealed that GXGXG residues of the ATP binding site were conserved between MTHFS enzymes and Winnebago family members, arguing that Winnebago is an atypical ATPase (Figure 4). In addition to this motif, all of the Winnebago and MTHFS family members have a conserved arginine residue. Previous studies on the crystal structure of MTHFS of *Mycoplasmas* revealed that this arginine is located within the ATP binding pocket and directly interacts with γ-phosphate of the ATP. The hydrolysis reaction happens by glycine residues wrapping

around the ATP molecule while the positively charged arginine residue stabilizes the negatively charged γ-phosphate. Following this interaction, the ATP molecule is hydrolyzed, and the high energy γ-phosphate is transferred onto the 5-formyl oxygen of ring open FTHF. This results in formation of a high-energy intermediate which makes the cyclo-ligase reaction energetically favorable. Therefore, MTHFS essentially acts as a kinase for 5- formyl tetrahydrofolate. The fact that both the GXGXGX motif and hydrophilic arginine residue inside the hydrophobic binding pocket of ATP are highly conserved among Winnebago family members as well as MTHFS family members suggest that Winnebago performs the same ATP hydrolysis/substrate phosphorylation reaction as MTHFS.

Our analysis of primary sequence of these homologous proteins also identified a second region of homology among MTHFS family members that is absent from the Winnebago family of enzymes. This set of motifs constitute the MTHFS binding site for 5-formyltetrahydrofolate. The lack of conservation of the substrate binding site suggests that while MTFHS and Winnebago both share the same motifs required for ATP hydrolysis, they act on different substrates. However, the fact that they both consist of similar atypical kinase-like phosphotransferase residues suggests that Winnebago catalyzes a cyclo-ligase reaction on a substrate distinct from that used by MTHFS. In order to gain further insights into the nature of the

Winnebago substrate, we sought to obtain the crystal structure of the Winnebago, and attempted to model the potential substrate that fits in the binding pocket.

**Purification of Winnebago homologs**

The process of crystallography of an enzyme requires an abundant source of pure, soluble protein. Therefore, we attempted to express full-length *Drosophila* Winnebago in *E. coli* BL21 (DE3) cells. However, full-length *Drosophila* Winnebago was completely insoluble when expressed in *E. coli*. In order to overcome this obstacle, we identified Winnebago orthologs from several different species including archaebacteria species. However it is possible that Winnebago homologs may act on different substrates than *Drosophila* Winnebago, Therefore, in order to ensure that the crystal structure of the homologous protein gives us the insight into the substrate of Winnebago, we selected our potential candidates for protein expression based on two different criteria. First was the candidate's identity to Winnebago (Flowchart 1 &2). Our candidate homologs ranged between 30% to 35% identity to Winnebago protein. Secondly, we considered homologous proteins from sources in which the protein was more likely to be soluble. Considering the second criteria, we decided that recombinant proteins from archaebacterial species would provide us

with the most soluble and stable protein due to the fact that these organisms grow in harsh environmental conditions. Using the set criteria, we selected Winnebago homologs from 6 different species. The process of protein expression was performed in *E. coli* BL21 (DE3) cells at 37°C, 27°C, and 19°C in order to find the expression temperature at which the most soluble and stable protein would be obtained. Out of the 6 homologs, three were insoluble when expressed in *E. coli* BL21 (DE3) cells (flowchart 1& 2). Of the remaining three soluble proteins, the Winnebago homolog from *Sulfolobus* expressed poorly in *E. coli* BL21 (DE3) cells. We suspected that the lack of a strong expression could be due to the fact that the recombinant protein is toxic to the cell. Consequently, we attempted to express the *Sulfolobus* homolog in different strains of *E. coli* BL21 (DE3) cells that has been altered for improved resistance to the expression of toxic proteins. The results indicated that *Sulfolobus* homolog is expressed the best in C41(DE3) at 37°C. The other two homologs from *Haloarcula* and *Halobacterium* species showed a good level of recombinant protein expression in *E. coli* BL21 (DE3) cells at 37°C.

Following the process of protein expression, we developed a purification strategy for each of our three soluble proteins. The first step in the purification of all three proteins was to express His-tagged versions of each protein following by metal chromatography of expressed protein.

The results show that about 80% of expressed protein from *Sulfolobus*, 50% of expressed protein from *Haloarcula*, and 50% of expressed protein from *Halobacterium* was recovered. However, these protein samples also contained a large number of co-purifying contaminating proteins as assayed by Coomassie Blue staining of the SDS-PAGE gel of the samples (Figure 6).

In order to obtain purer protein, metal chromatography step was followed by an ion exchange chromatography (IEC) step. Since each of our three proteins has a different calculated isoelectric point (pI), the binding buffer, the elution buffer, and the IEC resin used were chosen to take advantage of the pI of each protein (Flowchart 3). The protein was eluted from the IEC column using buffers with different salt concentrations in order to identify the elution condition with the least amount of contaminant protein and the greatest amount of intended protein gets eluted from the column (Figure 7 and 8). It was found that the optimal elution condition is at 300mM NaCl for all three proteins. The introduction of an IEC step after metal chromatography significantly improved the purity of our protein samples. However, there were still several contaminating proteins that could interfere with our crystallographic and enzymological studies.

As a result, we decided to take advantage of the fact that *Sulfolobus solfataricus* is an extremophile archaebacteria, which naturally grows at volcanic spring at 75-80°C. Therefore, the proteins from this species are thermally stable. In order to precipitate and eliminate the co-purified contaminant *E. coli* BL21 (DE3) proteins, a 60°C heat treatment followed by a rapid cool down of the protein solution was performed on the eluted protein form IEC step. Since the purified protein uses ATP as a substrate, the process of heat treatment was also performed in presence of nucleotides in order to allow protein to assume its correct tertiary structure after refolding. Analysis of the purified protein by SDS-PAGE and Coomassie Blue staining revealed that the addition of a heat treatment step significantly reduced the amount of contaminating proteins in our samples (figure 9). This result also indicated that the treatment process results in the same level of purification regardless of presence of nucleotides. It is important to mention that the heat treatment step was an optional step, which was not performed on all purified samples used for crystallography, due to potential adverse effect on the folding and activity of the Winnebago homolog protein.

The last step of purification was another optional step to further purify the recombinant protein. In order to eliminate additional contaminant proteins, we exploited the fact that the His-tag can be

separated from the expressed protein by cleavage with TEV protease (Diagram 1). This allowed us to cleave the His-tag from our expressed protein and re-run the sample over a metal chromatography column. Since the cleaved protein should flow through the column while all of the contaminants should rebind to the resin, we expected that this step would significantly improve the purity of our protein samples.

While the cleavage of His-tag was successful for recombinant proteins from *Sulfolobus* and *Halobacterium*, the His-tag was not cleavable from recombinant protein from *Haloarcula*. Since the failure of the His-tag to be accessible to TEV protease is often a sign of partial misfolding of the protein, we focused our subsequent efforts on the *Sulfolobus* and *Halobacterium* version of Winnebago. An examination of the metal chromatography column flow-through for the cleaved protein revealed that the addition of the cleavage step resulted into a purer protein by eliminating co-purified *E. coli* protein visible on the gel in elution column (Figure 10). The purified recombinant proteins were subsequently used for crystallography and enzymology studies.

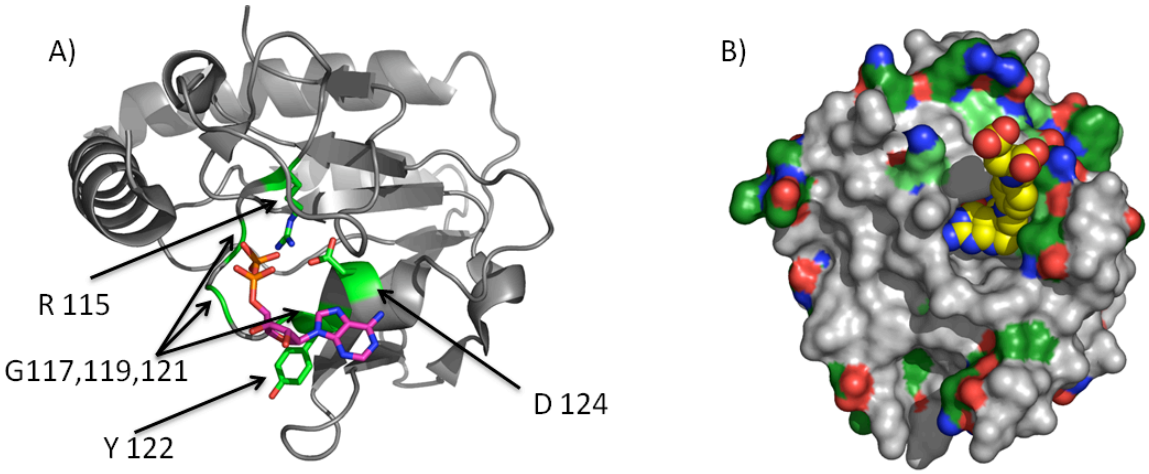
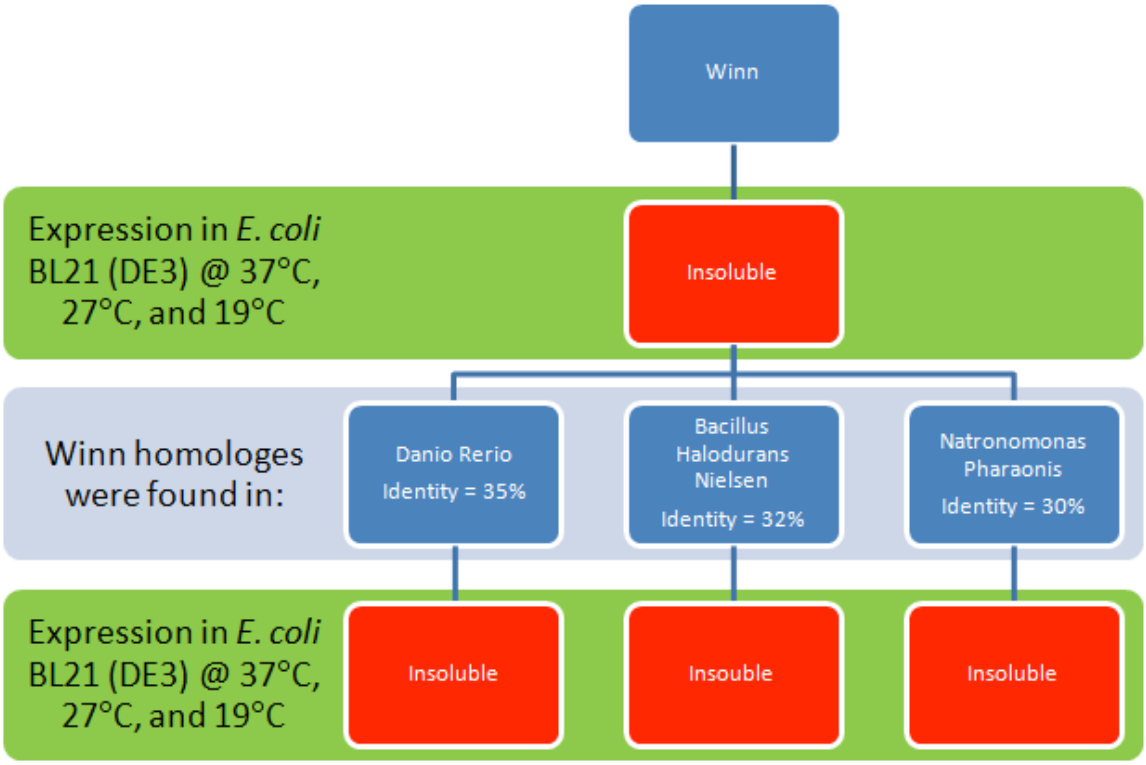


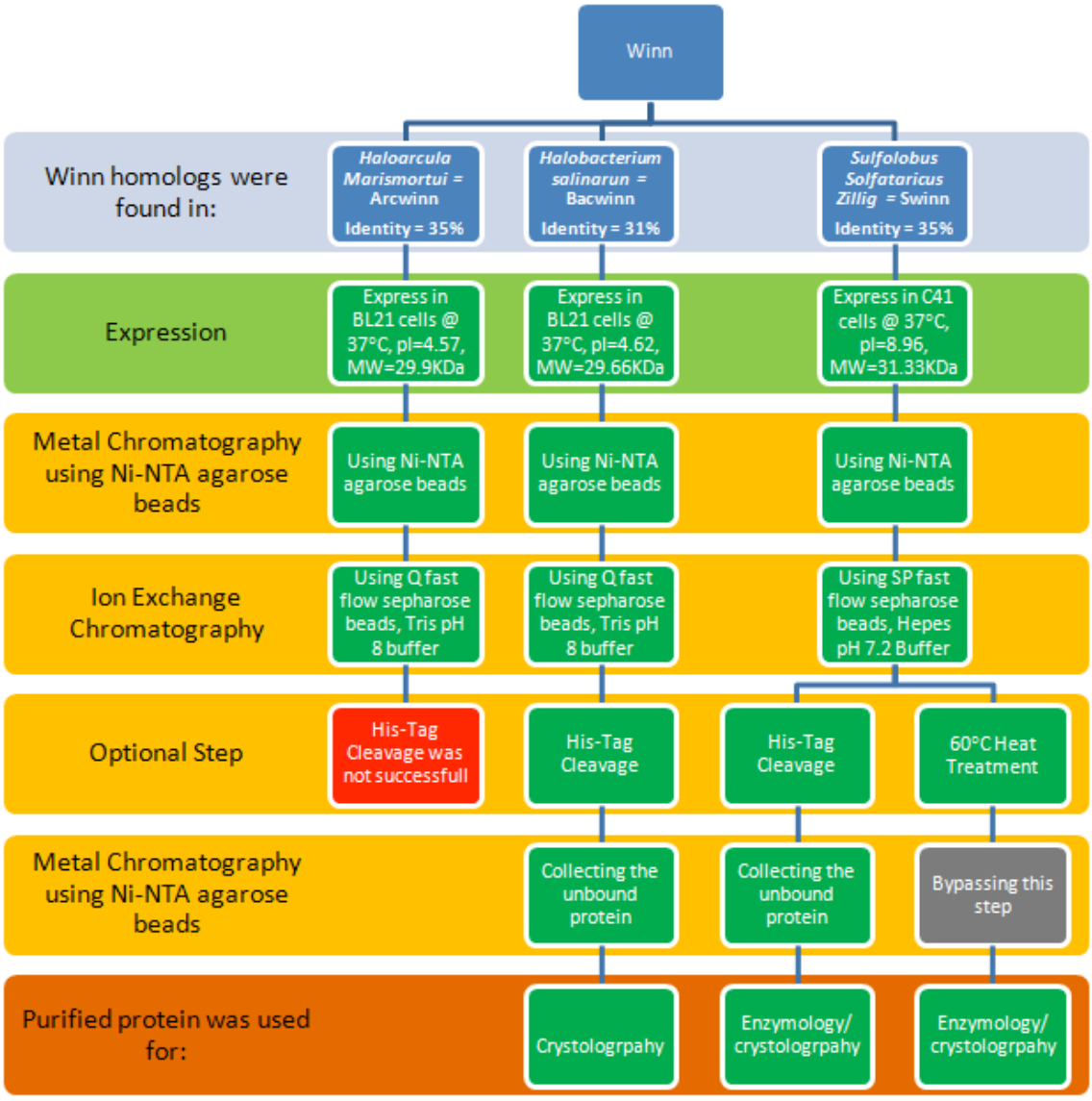

**Figure 5: Crystal structure of *Mycoplasma* MTHFS**

Figure A shows the ribbon diagram of *Mycoplasma* MTHFS along with the location of the three conserved glycine and one conserved arginine residues. Figure B shows the space-filling model of the same protein along with the binding pocket for the folate substrate. The folate substrate is represented in yellow. The conserved residues are represented in light green and the similar residues are represented in dark green. This figure shows that the residues around the substrate-binding pocket are not conserved (residues in gray).




**Flowchart 1: Identification of different source of soluble Winn homologs**

This flowchart shows the process of identification and expression of different protein, in order to find a soluble recombinant Winn homolog.




**Flowchart 2: Identification, and purification of different soluble Winn homologs**

This flowchart shows the process of identification, expression and purification of different soluble recombinant Winn homolog.

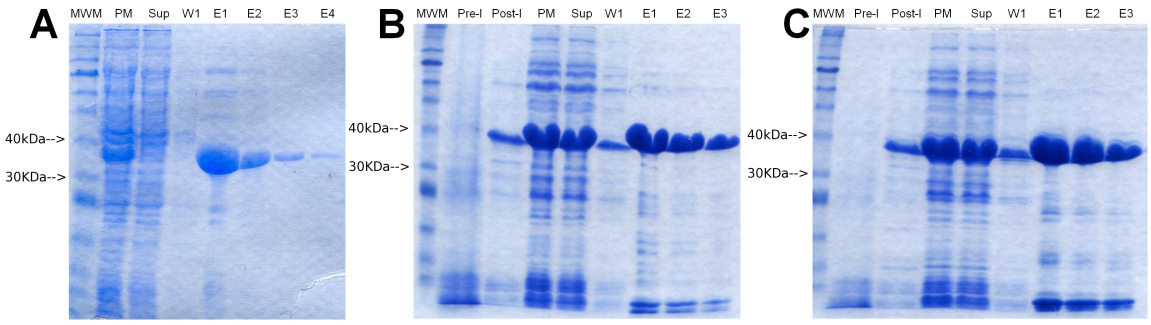
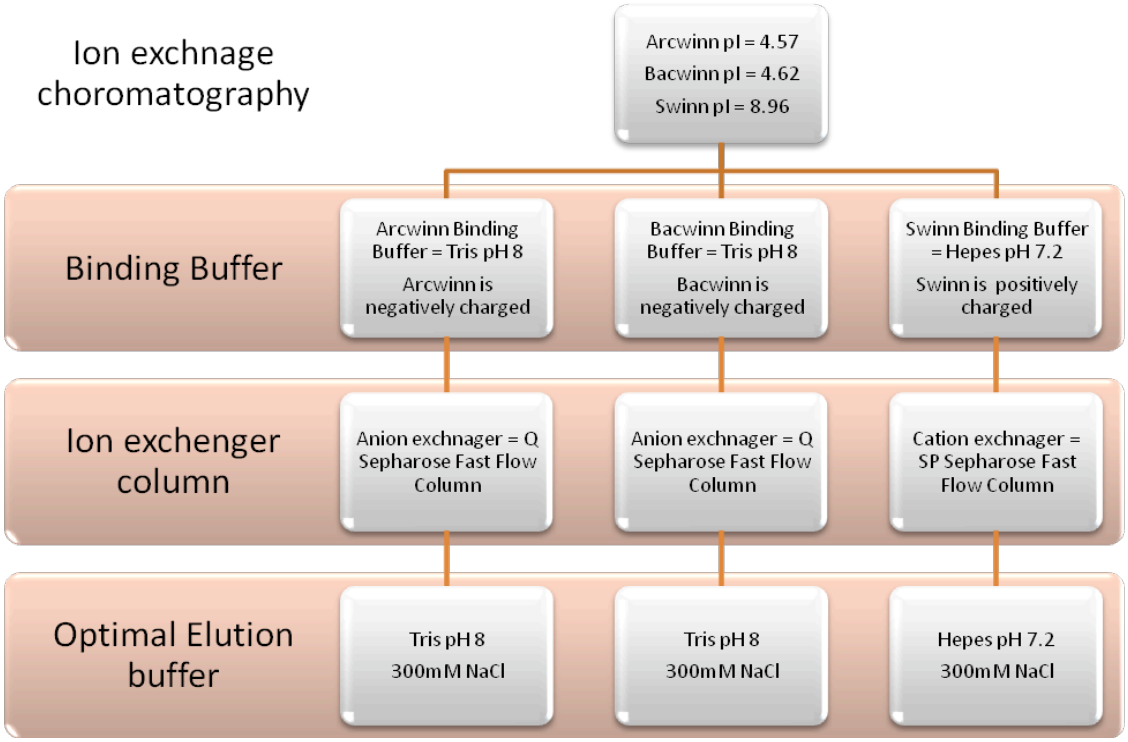



**Figure 6: Metal chromatography of Swinn, Arcwinn, and Bacwinn eliminated a considerable amount of contaminant protein**

This figure show the results of metal chromatography of the three proteins after running the samples on SDS-PAGE gel and visualizing the protein using Coomassie blue technique. Using the metal chromatography, we were able to eliminate a considerable amount of contaminant protein while recovering more than 50% of our recombinant protein. MWM = Molecular weight marker, Pre-I = pre induction of protein, Post-I= post induction of protein, PM = Pre metal chromatography, Sup = Unbound supernatant of protein-beads mixture, W1 = first wash, E1-4 = represents the eluted protein. Figure A is metal chromatography of Swinn protein. Figure B is metal chromatography of Arcwinn. Figure C is metal chromatography of Bacwinn protein.




**Flowchart 3: Process of ion exchange chromatography of Swinn, Arcwinn, and Bacwinn**

This flowchart shows the buffer, type of beads, and salt concentration of elution buffer used for ion exchange chromatography.

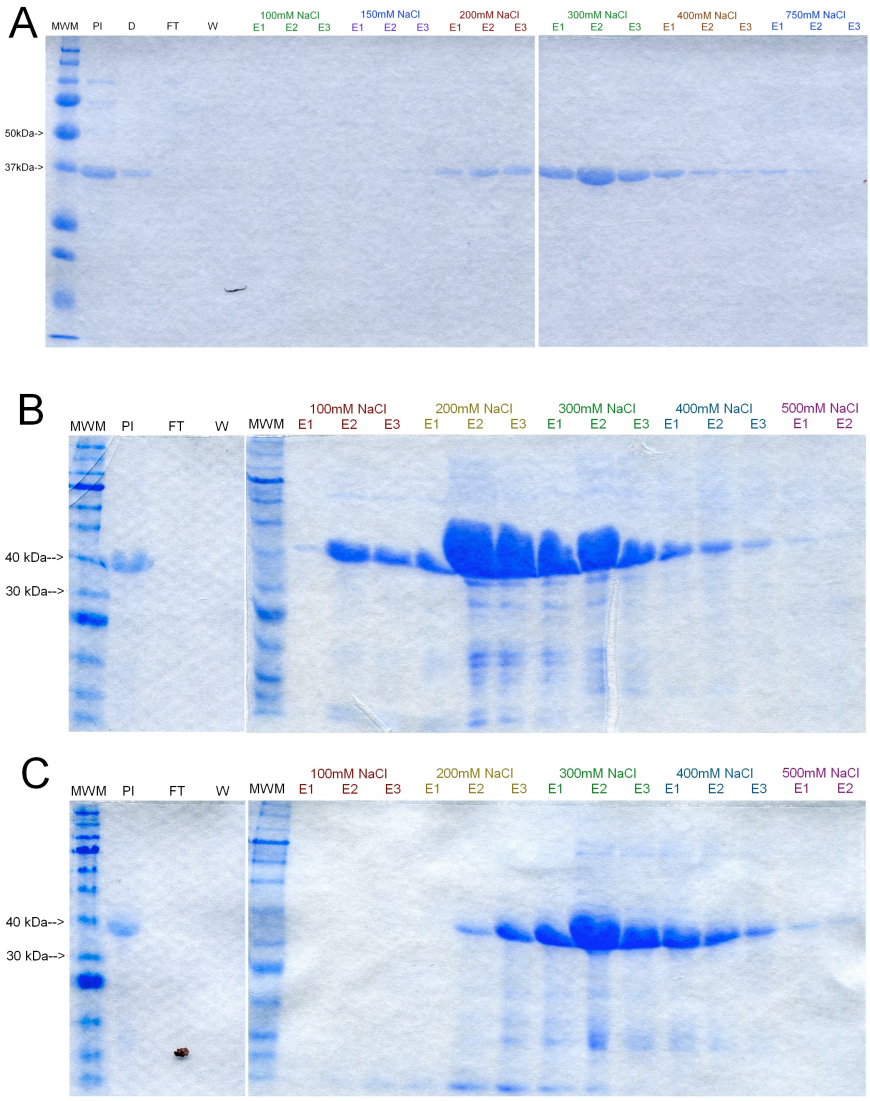



**Figure 7: Optimal salt concentration for elution of Swinn, Arcwinn, and Bacwinn in ion exchange chromatography was found to be 300mM NaCl**

This figure shows that the optimal salt concentration of the elution buffer at which most of the intended protein gets eluted off of ion exchange chromatography column is at 300mM NaCl. The purity was visualized using Coomassie blue technique. MWM = Molecular weight marker, PI = Pre ion exchange chromatography, FT = Unbound flow through from the column, W = wash, E1-3 = represents the eluted protein using different salt concentration. Figure A represents Swinn protein. Figure B represents Arcwinn. Figure C represents Bacwinn protein.

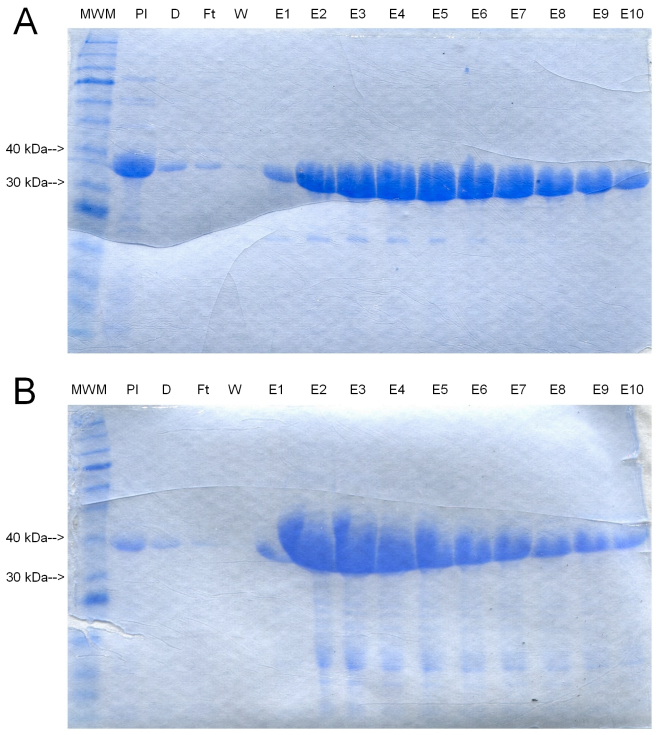



**Figure 8: Ion exchange chromatography of Swinn and Bacwinn purified protein into a higher extent**

This figure represents the purified protein after ion exchange chromatography step. The protein is visualized using Coomassie blue technique. MWM = Molecular weight marker, PI = Pre ion exchange chromatography, D = diluted protein, FT = Unbound flow through from the column, W = wash, E1-10 = represents the eluted protein using elution buffer with 300mM NaCl. Figure A represents ion exchange chromatography of Swinn protein. Figure B represents ion exchange chromatography of Bacwinn.

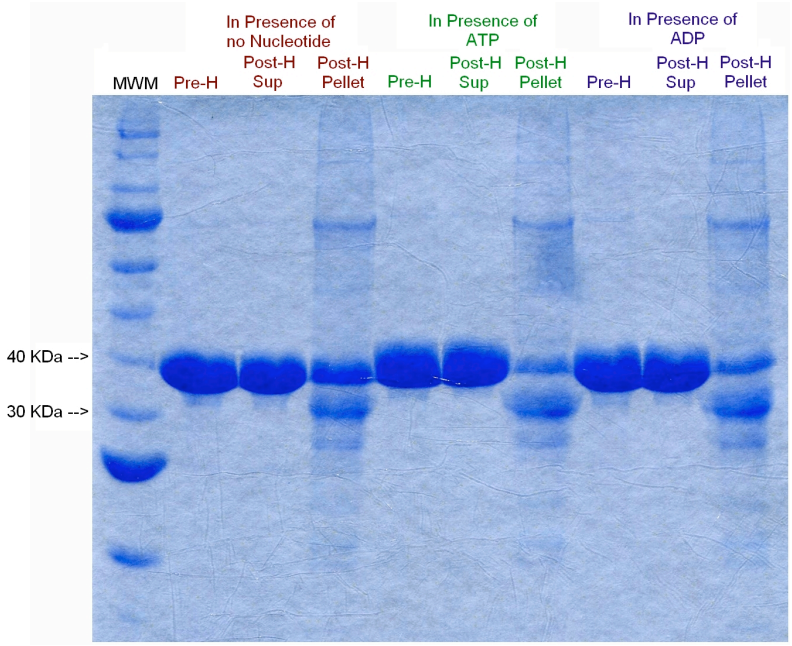



**Figure 9: Heat treatment of Swinn in presents of different nucleotides eliminated many of the non heat resistant contaminant protein**

This figure shows that the heat treatment purification step was able to eliminate the co-purified *E. coli* contaminant protein resulting in a very pure Swinn. Heat treatment was performed in presence of ATP and ADP; however, this figure indicates that there is no difference in the purity of the protein after heat treatment in presence of different nucleotides. Pre-H = Pre heat treated protein, Post-H sup = supernatant of heat treated protein, Post-H pellet = pellet and aggregates of heat treated protein.

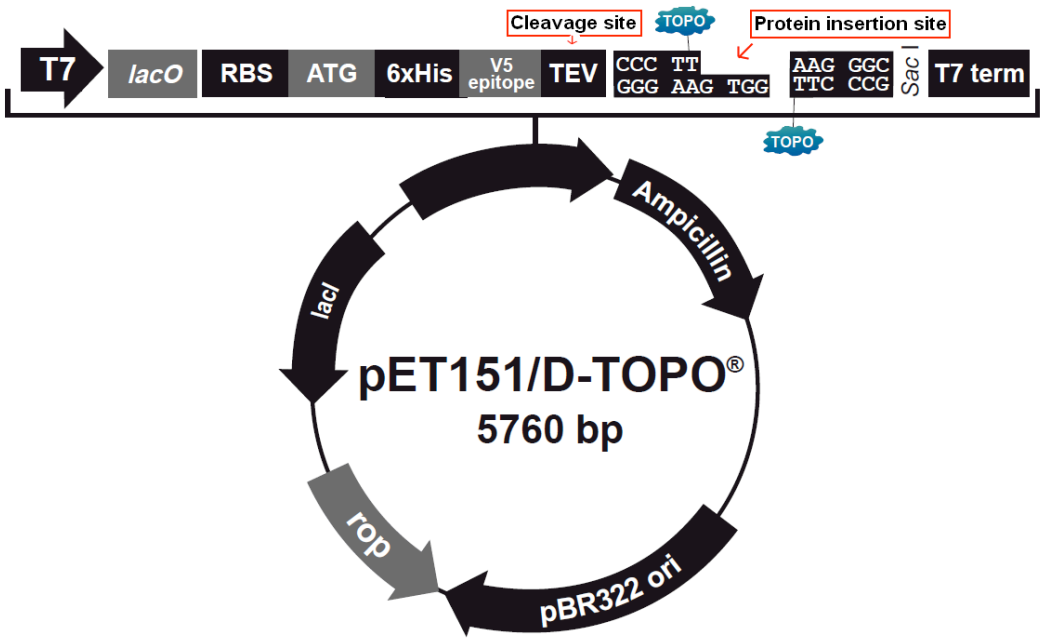



**Diagram 1: TEV cleavage site**

This diagram shows that the N terminus His-Tag can be cleave from the expressed protein using a proTEV enzyme which cuts the sequence at TEV sequence shown on the diagram.

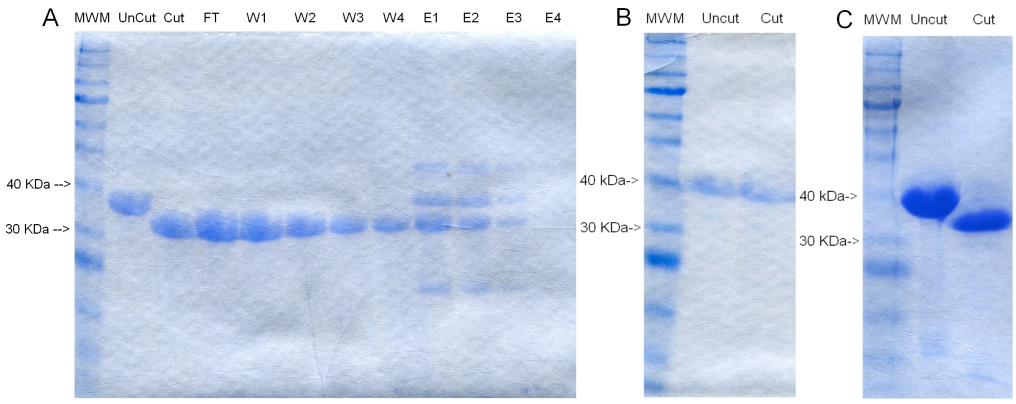



**Figure 10: His-tag cleavage and subsequent metal chromatography eliminated nickel interacting contaminant protein**

Figure A shows that the cleavage of His-tag was successful, and the process resulted in elimination of some of the contaminant protein as they are visible in elution lane. The flow through and wash lanes contain the His-tag cleaved Swinn protein since it does not bind to the nickel column due to the lack of His-tag. MWM = Molecular weight marker, Uncut = uncleaved protein, Cut = His-tag cleaved protein, FT = unbound flow through from the column, W1-3 = wash, E1-3 = elution. Figure B shows the result of cleavage from Arcwinn protein. The figure indicates that the His-tag is not cleaved. This is the sign of potential miss folding of the protein. Figure C shows the result of cleavage of Bacwinn protein. The His-tag for this protein was successfully cleaved.

**Crystallography**

The crystal structure of a protein can give useful information as to what substrate can potentially bind to the substrate binding site of a protein. Even though the exact substrate of the protein cannot be identified from the crystal structure, the number of potential substrate can be reduced having useful information about the possible size and the charge of the substrate.

The process of crystallography requires an abundant source of soluble protein. As discussed in the previous section, the recombinant *Drosophila* Winnebago protein is insoluble. Therefore, we used soluble Swinn and Bacwinn protein with an identity of 35% and 31% to Winnebago, respectively. These two proteins were purified using different techniques in order to achieve the most pure protein. For example, the His-tag was cleaved for both Swinn and Bacwinn and these proteins were run on the Nickel affinity column. Additionally, Swinn was heat treated to eliminate the co-purified *E. coli* protein. However, since heat treatment and protease treatment might partially denature our expressed protein, we tested all possible combinations of uncleaved, cleaved, heat-treated, and non-heat-treated protein in our crystallographic trials.

During crystallographic studies we tested over six thousand different conditions were screened at both room temperature and 4°C using

predefined reagent kits such as CSHT, JCSG, The ComPAS, The MPD, and The Protein Complexes . Each condition had a unique mixture of pH, salts, polymers and organics molecule in order to promote crystal growth. Since the protein is an ATPase, the conditions were examined in presence and absence of nucleotides. Nucleotides could potentially promote formation of crystals of this ATPase protein. We attempted to crystallize the protein in presence of 5mM ATP, ATP□S, AMP-PNP, AMP-PCP, ADP, and ADP-AlF3. After the initial screen, potential crystals were observed for non-hydrolysable ATP conditions in presence of Jeffamine M-600. However further analysis using no-protein control condition suggested that the obtained crystals were nucleotide crystals. Additionally, potential protein crystals were recently observed in two additional wells. The condition for the first crystal is composed of 1µl of 10mg/ml His-tag removed Swinn incubated with 5mM ADP and 5mM $MgCl_2$ added to 1µl of 0.1M HEPES pH 7.5 containing 10% w/v Polyethylene glycol 8,000 and 8% v/v Ethylene glycol (Figure 11). The condition for the second crystal is composed of 1µl of 10mg/ml His-tag removed Swinn incubated with 5mM ADP, 5mM $AlCl_3$ , and 20mM NaF added to 1µl of 0.1M MES monohydrate pH 6.5 containing 12% w/v Polyethylene glycol 20,000. The obtained crystals will be sent for crystallography and obtaining the diffraction pattern in order to solve the crystal structure of the protein.

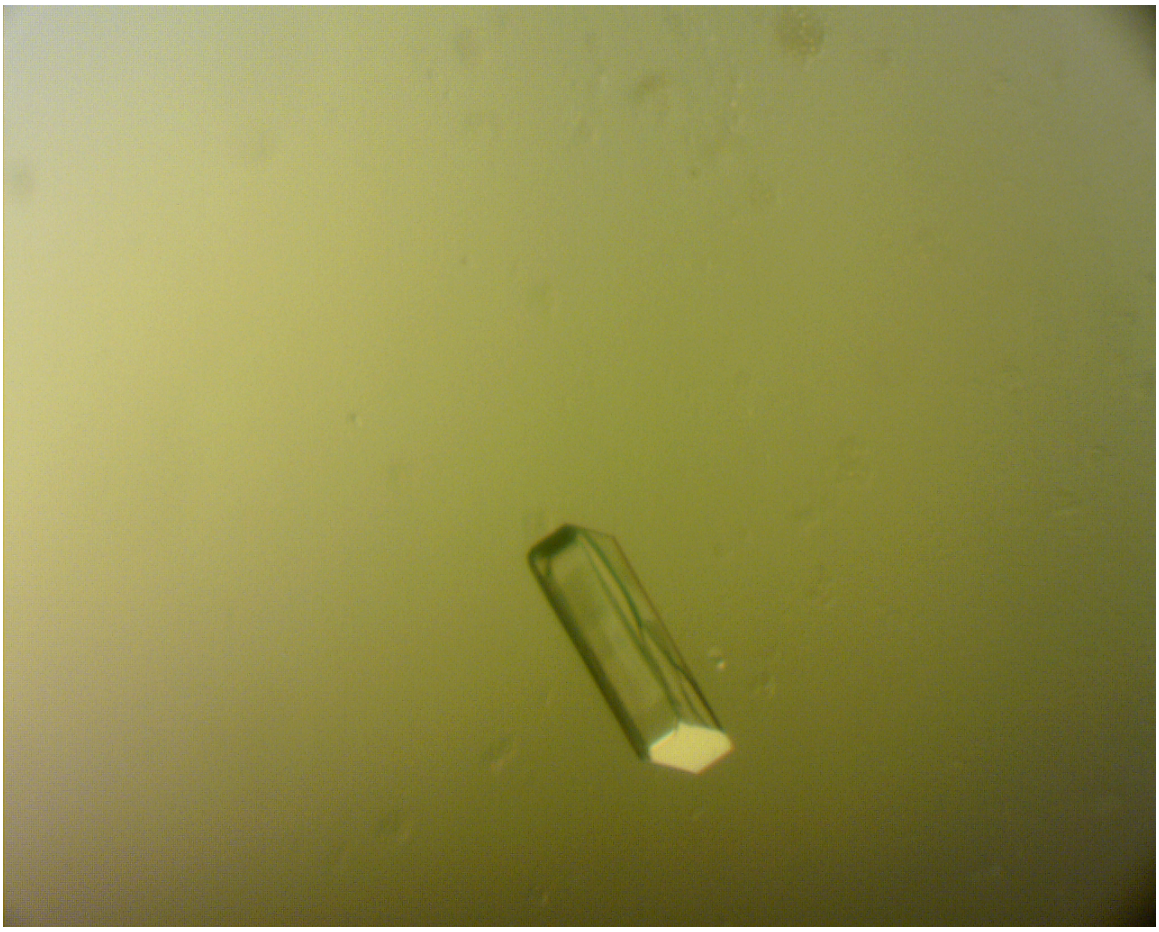


**Figure 11: His-Tag cleaved Swinn crystal**

This figure shows the crystal obtained for Swinn protein. The condition under which the crystal was obtained is composed of 1µl of 10mg/ml His-tag removed Swinn incubated with 5mM ADP and 5mM $MgCl_2$ added to 1µl of 0.1M HEPES pH 7.5 containing 10% w/v Polyethylene glycol 8,000 and 8% v/v Ethylene glycol.

**Enzymology**

Analysis of the primary sequence of Winnebago and its homologs revealed that it possessed a conserved motif that is used in MTHFS to bind and hydrolyze ATP. This suggests that Winnebago utilizes the high-energy phosphate of ATP as a source of energy in order to perform its hypothesized cyclo-ligase activity. The first step in testing this hypothesis is to demonstrate that Winnebago and/or its homologs are ATPases. Since the correct substrate of the cyclo-ligase domain has not been identified, our approach has been to focus on measuring the basal ATP hydrolysis rate of Winnebago. As part of these studies we have purified Winnebago homologs and demonstrated that they are ATPases that use the atypical ATPase motif present in MTHFS. We have also measured the $K_m$ and $K_{cat}$ of the protein.

**Detection Assay**

Measuring and characterizing the enzymatic activity of a protein requires a pure source of relatively stable protein and a reliable detection assay for measuring the activity of the enzyme. Of the Winnebago homologs that we characterized, the Winnebago homolog (Swinn) from *Sulfolobus* archaebacteria was the most pure and stable recombinant protein. Therefore, we focused our attention determining if Swinn had

ATPase activity. For these studies we utilized "the Promega ADP-Glo™ Kinase Assay" for detecting the ATP hydrolysis activity of Swinn. This assay uses a two-step process for measuring the ATPase activity by detecting the amount of ADP produced during the enzymatic assay. The first step involves addition of ADP-Glo™ Reagent to the reaction assay, which terminates the ATPase reaction and depletes all the non-hydrolyzed ATP remaining in the reaction solution. Then, the Kinase Detection Reagent is added to the mix, which results in conversion of the ADP produced during the enzymatic reaction to ATP. The newly synthesized ATP is measured using a luciferase/luciferin reaction, resulting in production of luminescence light. The amount of luminescence light generated correlates to the amount of ADP generated in ATPase assay, indicating the existence of ATPase activity. Finally, the kinetics of ATPase activity of the enzyme can be measured using a standard curve for converting the amount of luminescence light generated to the amount of ADP generated.

**Functionality of purified Swinn**

In order to determine whether our purified recombinant Swinn protein had ATPase activity, we performed an in vitro ATPase reaction assay using both his tagged Swinn and Swinn where the tag had been

cleaved. This is done to analyze the effect of cutting the His-Tag on the functionality of the protein. In this assay, the concentration of the ATP as a substrate was set to a constant value of 100uM. The reaction was performed at 50°C for one hour, due to the fact that *Sulfolobus* naturally grows at volcanic spring at 75-80°C. Figure 12 indicates the existence of an ATPase activity for both purified version of recombinant Swinn. However, the results indicate that the cleaved version possesses a lower level of ATPase activity. This suggests that the process of His-Tag cleavage might have an effect on the tertiary structure of the recombinant protein or that a portion of the ATPase activity is due to contaminating proteins. Since we were concerned about the effects of cleaving the His Tag on the structure of the protein, the cleaved version of the protein, even though slightly purer, was not used for the rest of the enzymatic analysis.

**Mutations in the putative ATPase motif block Swinn ATP hydrolysis**

One concern with our analysis of Swinn ATPase activity is that the basal level of ATP hydrolysis that we observe may be due to a contaminating ATPase.. In order to exclude this possibility, we pursued two different strategies.

First, we attempted to generate a mutant version of recombinant Swinn, one that has the same structure, external charge on the surface,

and binding capability to ATP but does not hydrolyze ATP. The idea behind this strategy is that if the ATPase activity belongs to a co-purified *E. coli* protein, then a Swinn mutant that is expressed and purified in the same exact way as the wild type Swinn should exhibit the same level of contaminating ATP hydrolysis activity. Since there is no crystal structure of a Winnebago family member, we based our mutagenesis on previous structural studies of MTHFS protein (Figure 5). The crystal structure of MTHFS shows that ATP, in its binding pocket, is wrapped around by three conserved glycine residues, and the binding is stabilized by direct interaction of the negatively charged ATP with a conserved positively charged arginine residue. The fact that all of these four residues are conserved is an indication of the importance of these residues in the ATP hydrolysis activity of this motif. Since the ATPase motif in MTHFS is also present in Swinn ,(Figure 4), we were able to design two different point mutants which should lack the ability to hydrolyze ATP – R175K and R175M. The R175K and R175M point mutants of Swinn were constructed, expressed, and purified using the same purification strategy that was used to purify wild type recombinant Swinn. Analysis of the ATPase activity of both mutants (Figure 13) indicates a drop of about 60% in ATP hydrolysis activity of both point mutants as compared to the wild type Swinn when measured at the highest protein concentration. These results suggest that

a large fraction of the ATP hydrolysis activity that we observe is dues to the Swinn protein.

However, since the design of our mutants was guided by structural studies of a distantly related protein, it was possible that the failure of our mutants to completely eliminate ATP hydrolysis activity was due to differences in the MTHFS and Swinn structure. In order to provide additional evidence for Swinn ATPase activity, we exploited the fact that Swinn is a thermostable enzyme. The assumption in this strategy is that since *E. coli* is not a thermophilic bacterium, its endogenous proteins are not thermally stable and prone to denaturation by heat treatment. If the observed ATPase activity is due to a co-purifying *E. coli* protein, then a heat-treated sample of purified protein should exhibit a reduction or loss of ATPase activity. In order to test this, we heat-treated a subset of purified Swinn to 50°C and 60°C following by a rapid cool down of the solution on ice. Assays of the ATPase activity of both heat-treated and non-heat-treated Swinn showed that there was no significant difference in ATPase activity between the two Swinn samples, indicating that the ATP hydrolysis activity is a genuine function of Swinn (Figure 14). While this was a strong argument supporting that the ATPase activity we observe is due to Swinn, it was still formally possible that a heat-stable E. coli ATPase could be responsible for our results. In order to address this possibility, we combined

our mutant analysis with our heat treatment to determine if our mutations might make Swinn heat labile. To test this possibility, each of the point mutants and wild type protein was heat treated at either 50°C or 60°C. While wild type Swinn displayed the same ATPase activity whether it was heat treated or not , when both of the mutant Swinn proteins were exposed to a 60°C heat treatment, they lost about 60% of their ATPase activity as compared to non-heat-treated version at the highest concentration of protein (Figure 15). Thus, the Swinn ATPase activity is heat stable unless Swinn protein has been mutated – a result that strongly argues that the observed ATP hydrolysis activity is due to Swinn enzymatic function. Additionally, we observed no change in enzymatic activity of 50°C heat-treated mutant comparing to non-heat-treated mutant. These results indicate that both Swinn mutants are temperature sensitive mutations for ATPase activity.

**Optimizing the ATPase assay**

The first step in determining the $K_m$ and $K_{cat}$ of Swinn is to optimize the conditions for its ATPase activity. In order to optimize the reaction conditions, we varied temperature, pH, and concentration of Mg to identify the conditions where Swinn has the highest rate of ATP hydrolysis.

First, we compared the activity of the protein at two temperatures of 37°C and 50°C. Figure 16 shows that Swinn ATP hydrolysis increase by about 20% at 50°C comparing to 37°C. Thus, 50°C was chosen as our standard temperature for the assay. Then, we attempted to optimize the pH at which the reaction is performed. We performed the enzymatic assay at pH 5, 6, 7, 8, and 9. It was found that the highest level of activity is at pH 6 (Figure 17). The increase in activity from pH 7 to pH 6 at highest protein concentration was about 20%. We were not able to test the activity at lower pH than 6 since lower pH would interfere with our ATP detection assay. Finally, we assayed the effect of different concentration of Mg on the level of ATP hydrolysis by Swinn at pH 6 and pH 7.2. The results indicated that at both pH 6 and pH 7.2, the ATP hydrolysis activity increases by about 22% when the concentration of Mg is reduced from 1mM to 100μM Mg (Figure 18). Additionally, the level of activity at pH 7.2 was compared for different concentration of Mg and Mn, however, there was no significant difference in ATP hydrolysis activity between reactions with Mg, or Mn (Figure 19). Thus, we set the optimal condition at 100μM Mg.

**Measuring the $K_m$ and $K_{cat}$**

Using our optimal reaction conditions of 50°C, pH 6, with 100μM concentration of Mg we measured the $K_m$ and $K_{cat}$ of Swinn. For these measurements, the concentration of protein was set to 0.5 mg/ml since this is an average protein concentration at which a detectable rate of ATPase can be observed for Swinn protein. The reaction condition was set to 50°C, in pH 6. Even though the optimal Mg concentrations was determined to be at 100μM, we used 1mM Mg concentration in order to provide enough Mg in the reaction mix to support the higher concentration of ATP. The reaction was performed over a course of 1 hour using different concentrations of ATP ranging from 0μM ATP to 300μM ATP. Samples of each reaction were taken at intervals throughout the 1 hour reaction in order to measure the reaction rate at each concentration. The rates of the reactions were then plotted versus their corresponding substrate concentration, and the $V_{max}$, $K_m$, and $K_{cat}$ of the protein was calculated (Figure 20 ). From these experiments we determined that Swinn has a $K_m$ of 33.55μM and a $K_{cat}$ of .044 $min^{-1}$.

**Testing the substrate and product of MTHFS on Swinn ATP hydrolysis activity**

Comparing the primary sequence of Swinn to MTHFS revealed that the ATP binding site in MTHFS is conserved in Swinn. However, the residues

involved in substrate binding are not conserved between these two proteins, suggesting that they potentially work on different substrates. In order to confirm that Swinn is not just a MTHFS and has its own unique substrates, we tested whether Swinn could use MTHFS substrates. For most ATPases, the basal level of ATP hydrolysis is greatly stimulated in the presence of its substrate and decreased in presence of product of the reaction. MTHFS is a cyclo-ligase enzyme, which works on Folinic Acid (FA) and catalyzes the formation of carbon-nitrogen ring, resulting in production of 5,10-Methenyltetrahydrofolate (MTHF). If Swinn were a MTHFS we would predict that Folinic acid would stimulate Swinn ATPase activity and MTHF would inhibit ATPase activity. Measurements of the Swinn ATPase activity in presence of 250μM FA. or 250 μM MTHF found that there was no significant difference in the Swinn ATPase activity when either FA or MTHF were present in the reaction (Figure 21). These results suggest that FA is not the correct substrate of Swinn.

.

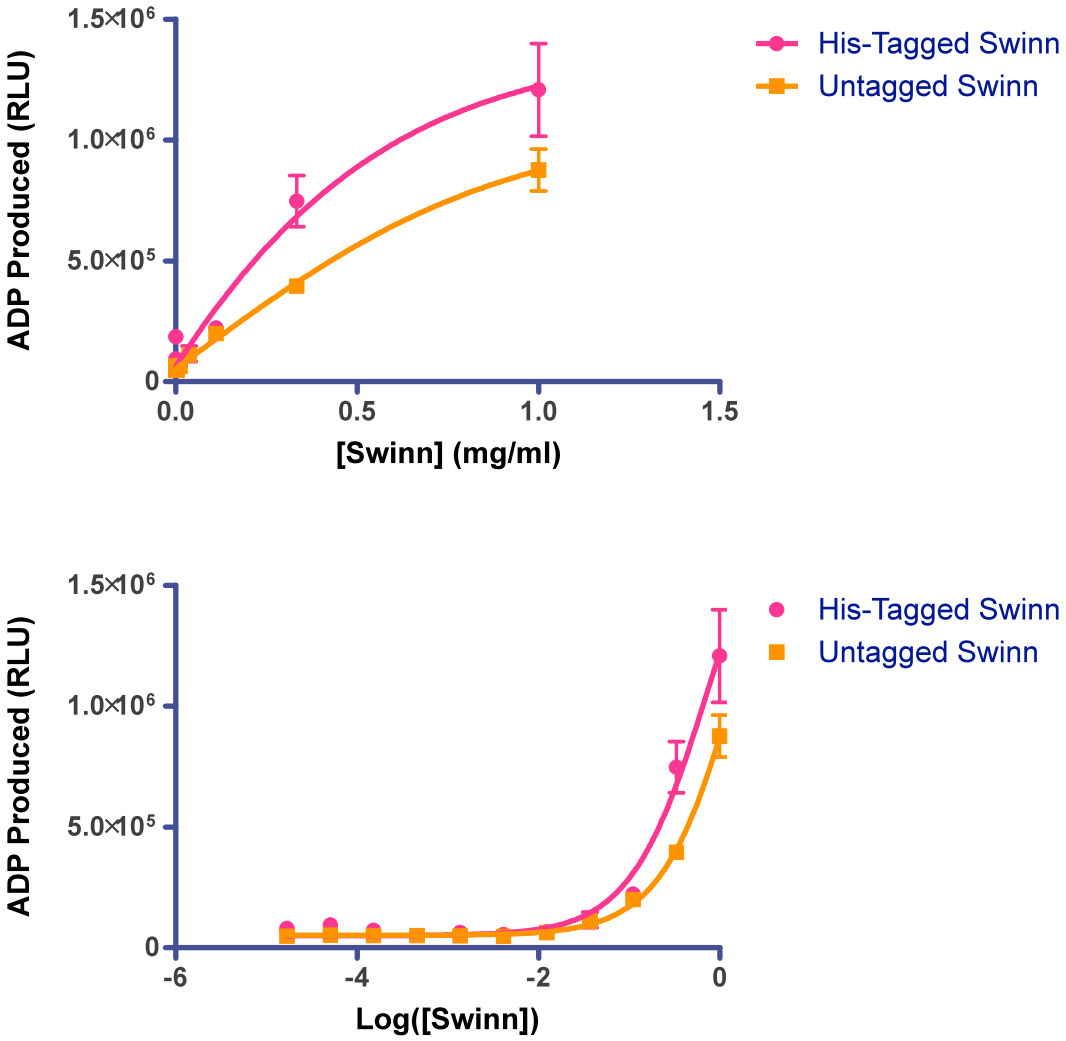



**Figure 12: Purified Swinn shows an ATP hydrolysis activity**

This figure indicates that there an ATP hydrolysis activity for both sample of uncleaved and His-tag cleaved protein solution. The uncleaved protein solution exhibits a higher level of ATPase activity. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale). The RLU unit is a luminescence unit representing the level of ATP hydrolysis activity.

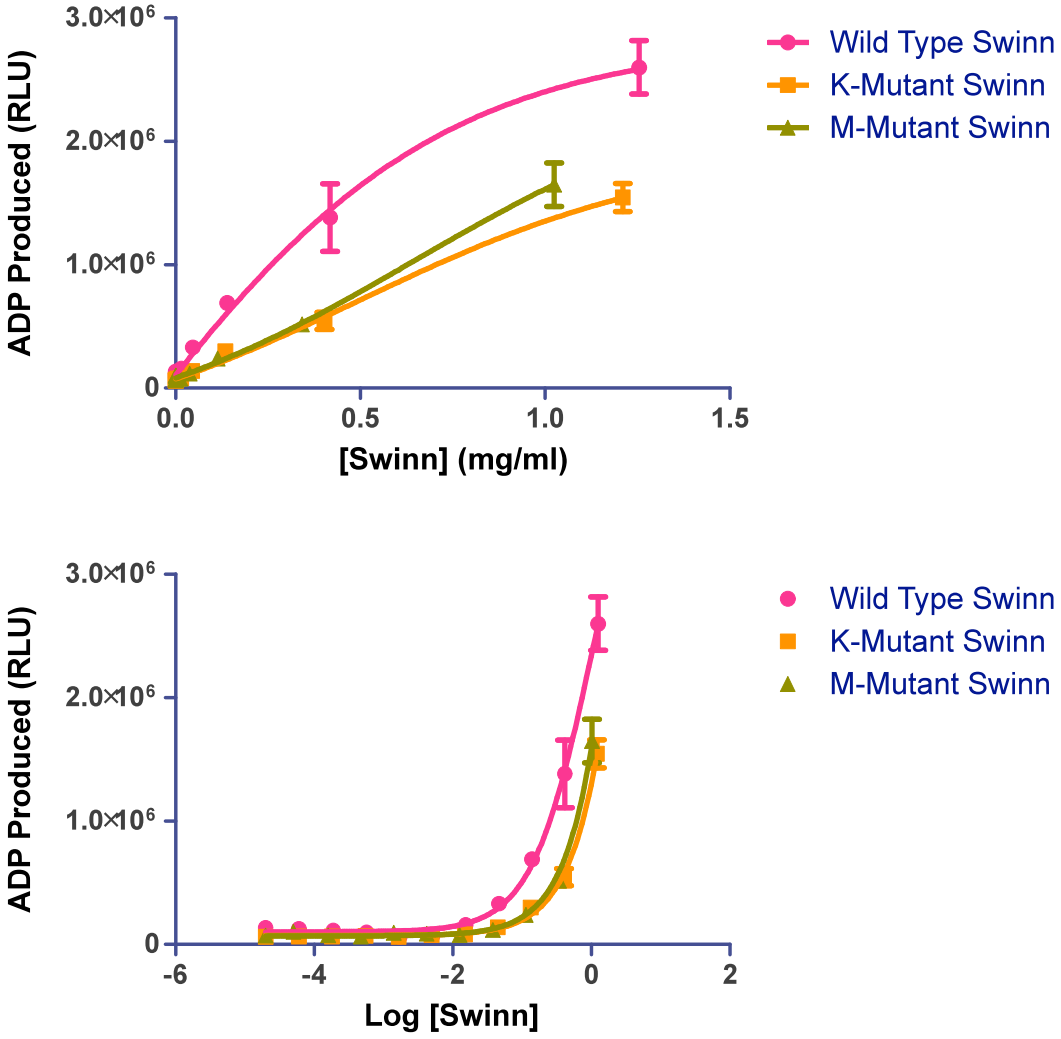



**Figure 13: Observed ATPase activity belong to Swinn due to reduction in ATP hydrolysis ability of Swinn by introduction of point mutation of arginine residue**

This figure shows that the level of ATPase activity drops by 60% after mutating the conserved arginine residue inside the ATPase domain of Swinn. This result indicates that ATP hydrolysis is a genuine function of Swinn and it is not due to some contaminant *E. coli* protein. K-mutant Swinn is a R175K mutation and M-mutant Swinn is a R175M mutation. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

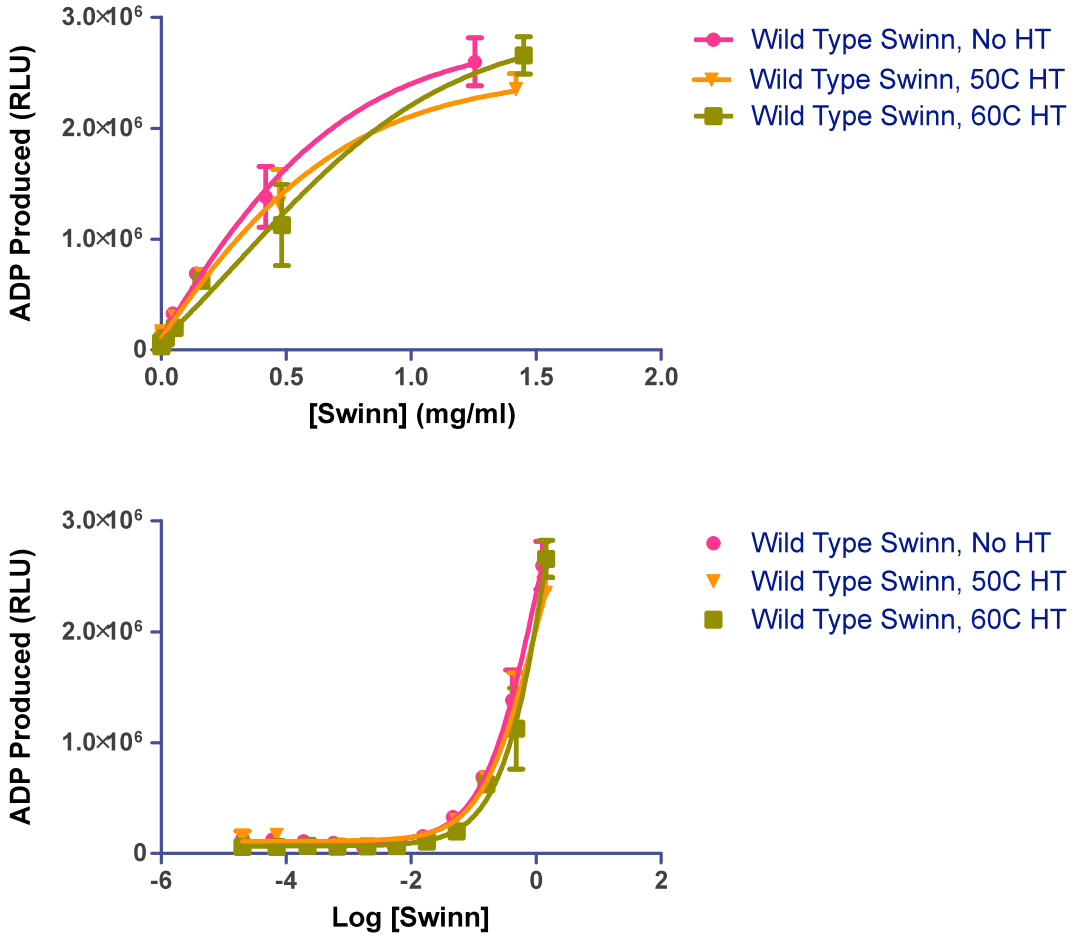



**Figure 14: The heat treatment of thermostable Swinn protein shows that the observed ATPase activity belongs to Swinn protein**

This figure indicates that the level of ATP hydrolysis activity is maintained after heat treatment of solutions of wild type Swinn protein. Since the Swinn is a heat resistance protein, and assuming that contaminant *E. coli* protein are not heat resistance, then this results is a sufficient evidence of genuinity of Swinn ATP hydrolysis activity. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

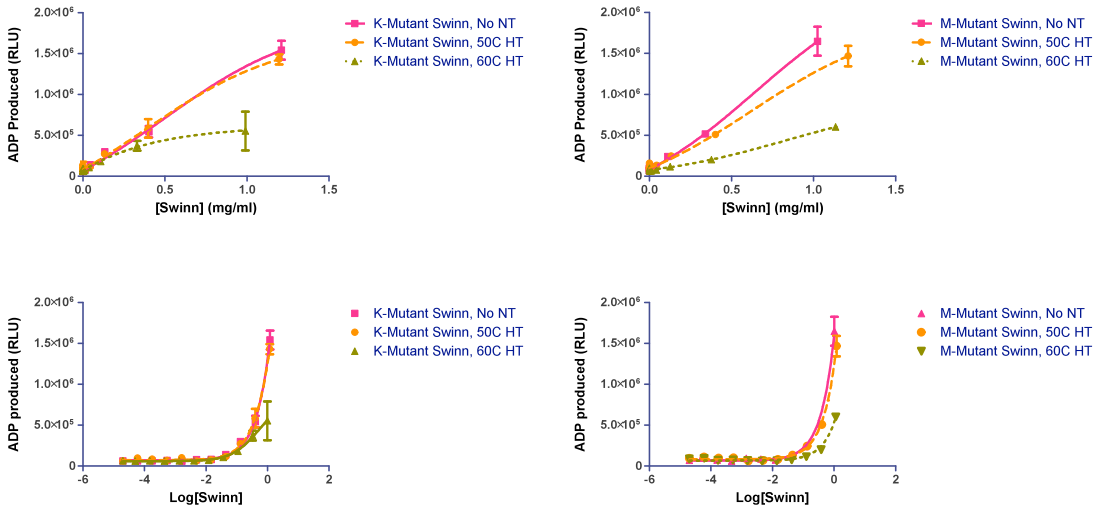



**Figure 15: Heat treatment of the mutant Swinn indicates that the observed ATPase activity belongs to Swinn protein**

This figure shows a drop in the level of ATP hydrolysis activity of the point mutants after 60°C heat treatment. This figure eliminates the possibility of a co-purified heat resistant *E. coli* protein being responsible for the ATPase activity. This is due to the fact that the 60°C heat treatment of wild type did not show any loss of activity (Figure 15). The two figures on the left are representative of the same experiment, drawn on different scales (Normal and log scale). The two figures on the right are representative of the same experiment, drawn on different scales (Normal and log scale).

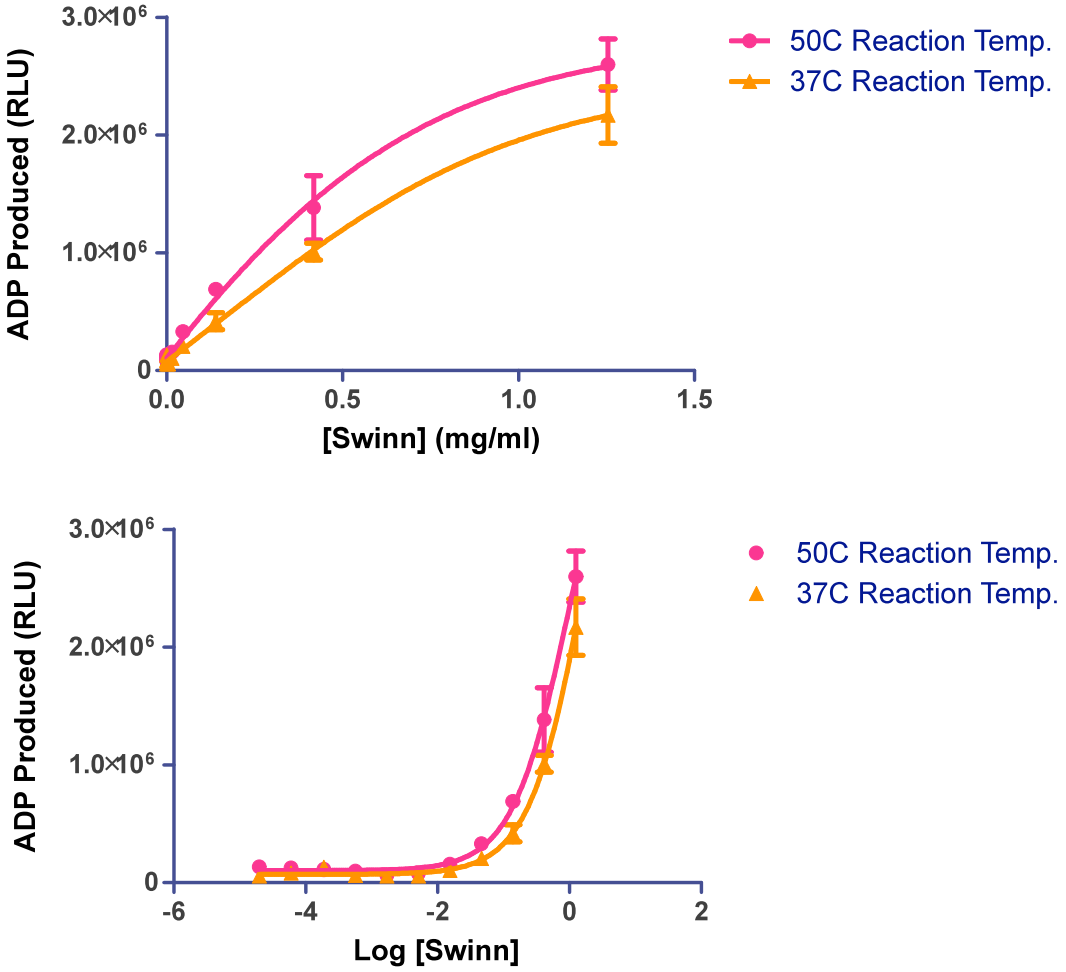



**Figure 16: The optimal reaction temperature for Swinn is 50°C**

This figure shows that the ATP hydrolysis activity of Swinn protein increases by 20% at 50°C comparing to 37°C. Thus, the optimal reaction condition for Swinn was found to be at 50°C. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

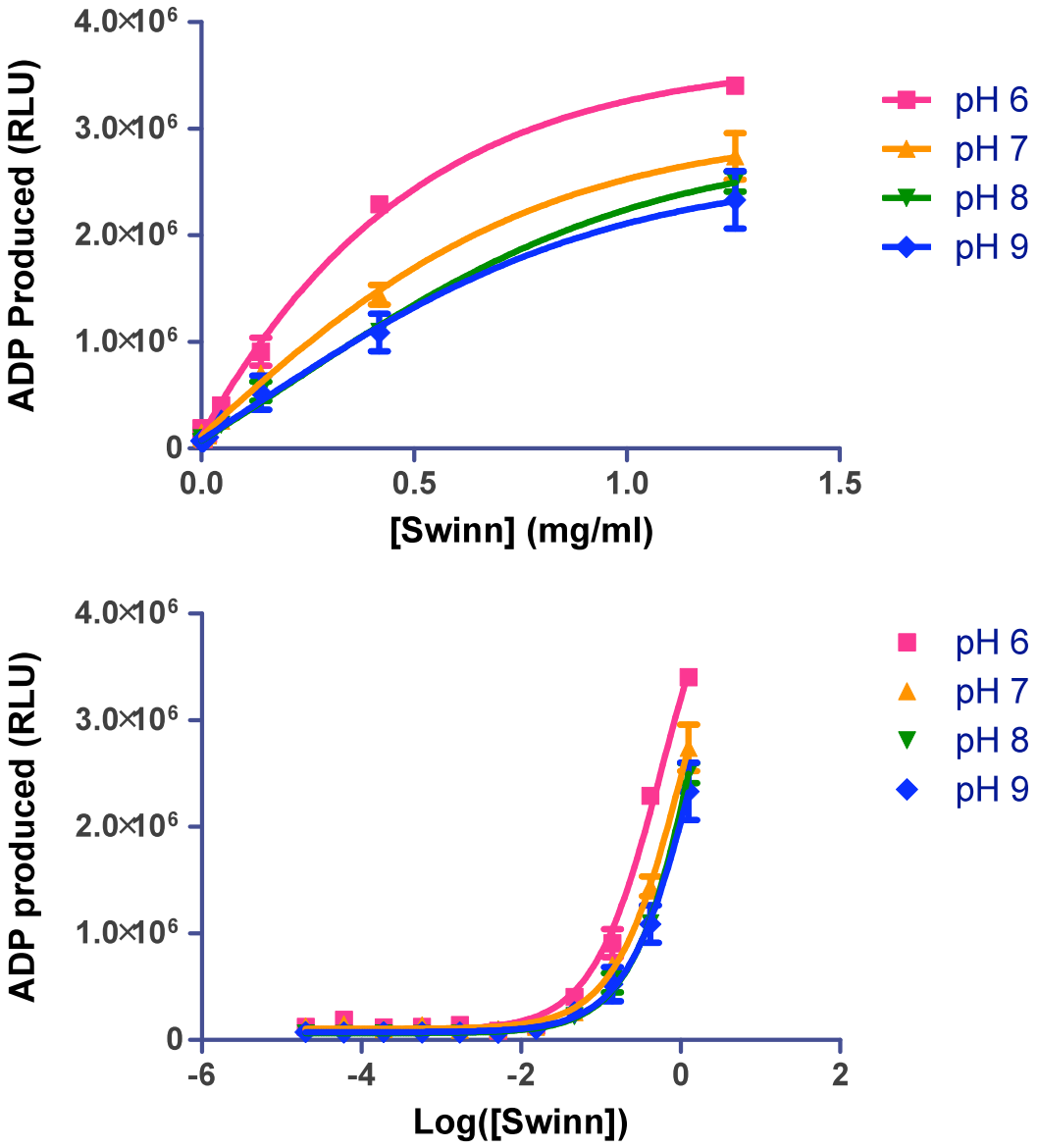



**Figure 17: The optimal pH for Swinn activity is at pH 6**

In this experiment the effect of different pH was examined using 50mM MES pH 6, 50mM HEPES pH 7 and 8, and 50mM Tris pH 9. This figure shows that the ATP hydrolysis activity of Swinn protein increases by 20% from pH 7 to pH 6. Thus, the optimal reaction condition for Swinn was found to be at pH 6. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

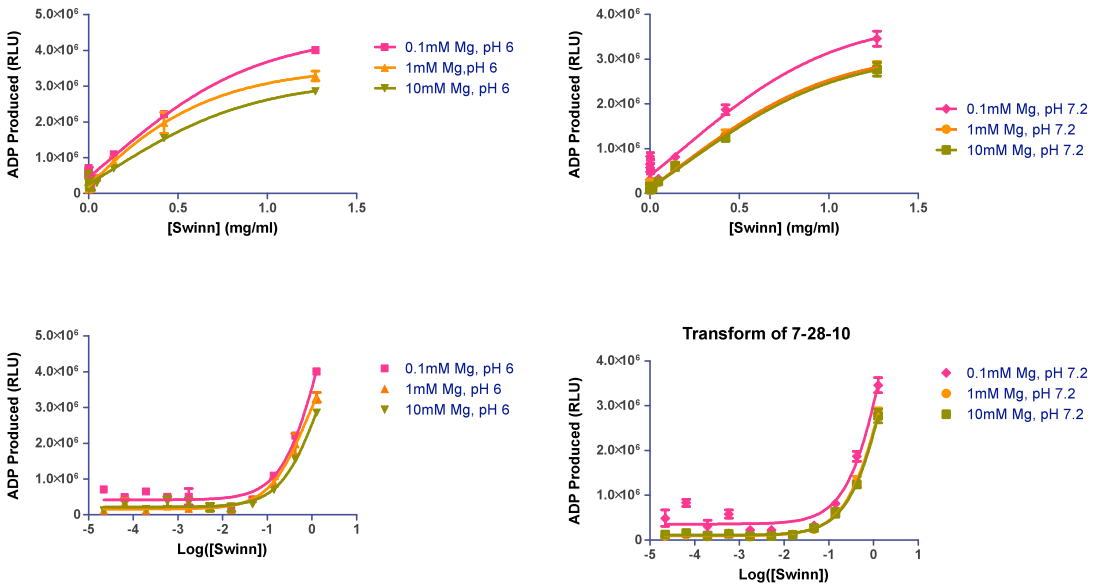



**Figure 18: The optimal [Mg] for Swinn activity is at [Mg] = 100μM**

This figure shows that the ATP hydrolysis activity of Swinn protein increases by 22% when the concentration of Mg is reduced from 1mM to 100μM Mg at both pH 6 and pH 7.2. Thus, the optimal Mg concentration for the assay was found to be at 100μM. The two figures on the left are representative of the same experiment, drawn on different scales (Normal and log scale). The two figures on the right are representative of the same experiment, drawn on different scales (Normal and log scale).

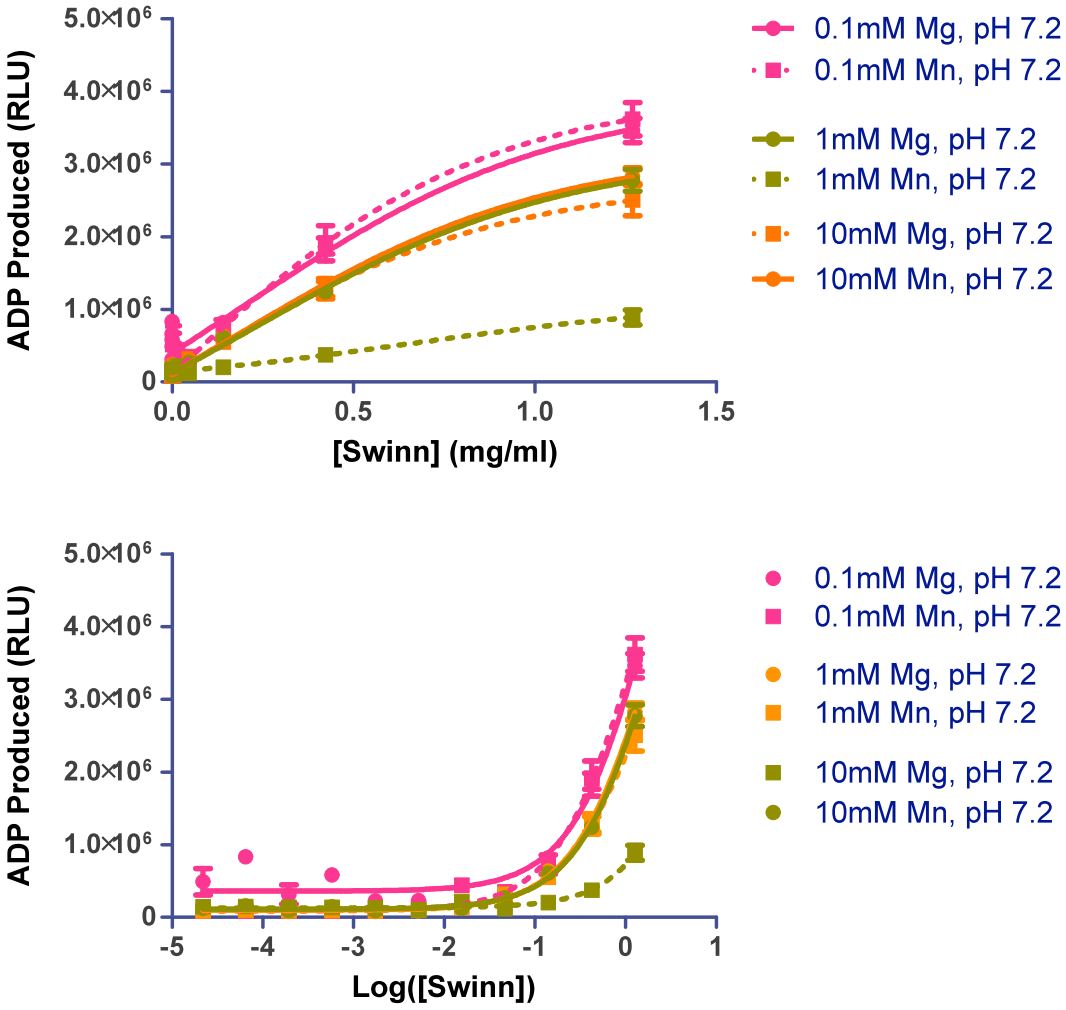



**Figure 19: Mn does not improve the ATPase activity of Swinn comparing to Mg**

This figure indicates that there is no significant difference in ATP hydrolysis activity between reaction condition with Mg and Mn at their corresponding100μM and 1mM concentration. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

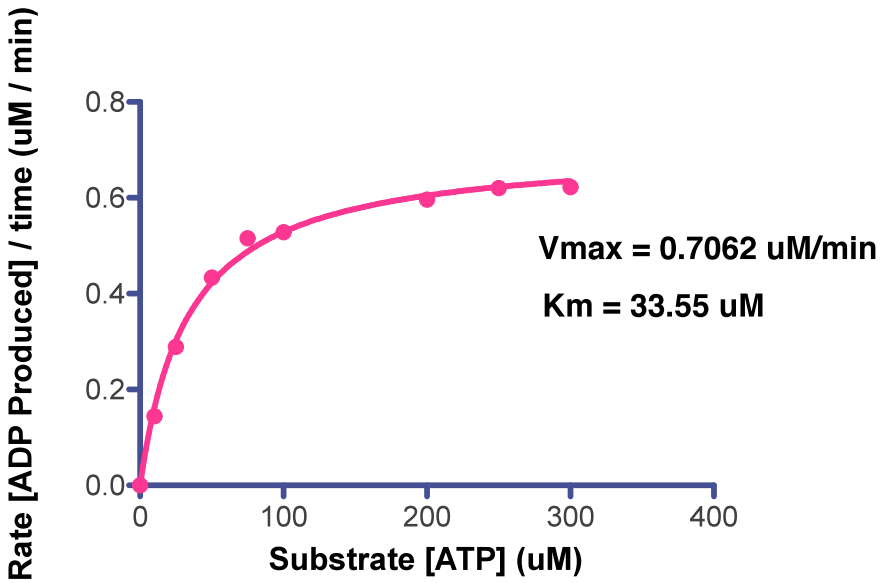



**Figure 20: Swinn has a Km of 33.55uM**

This figure shows the $K_m$ and $V_{max}$ of the Swinn protein. The $K_{cat}$ of Swinn was calculated to be .044 $min^{-1}$.

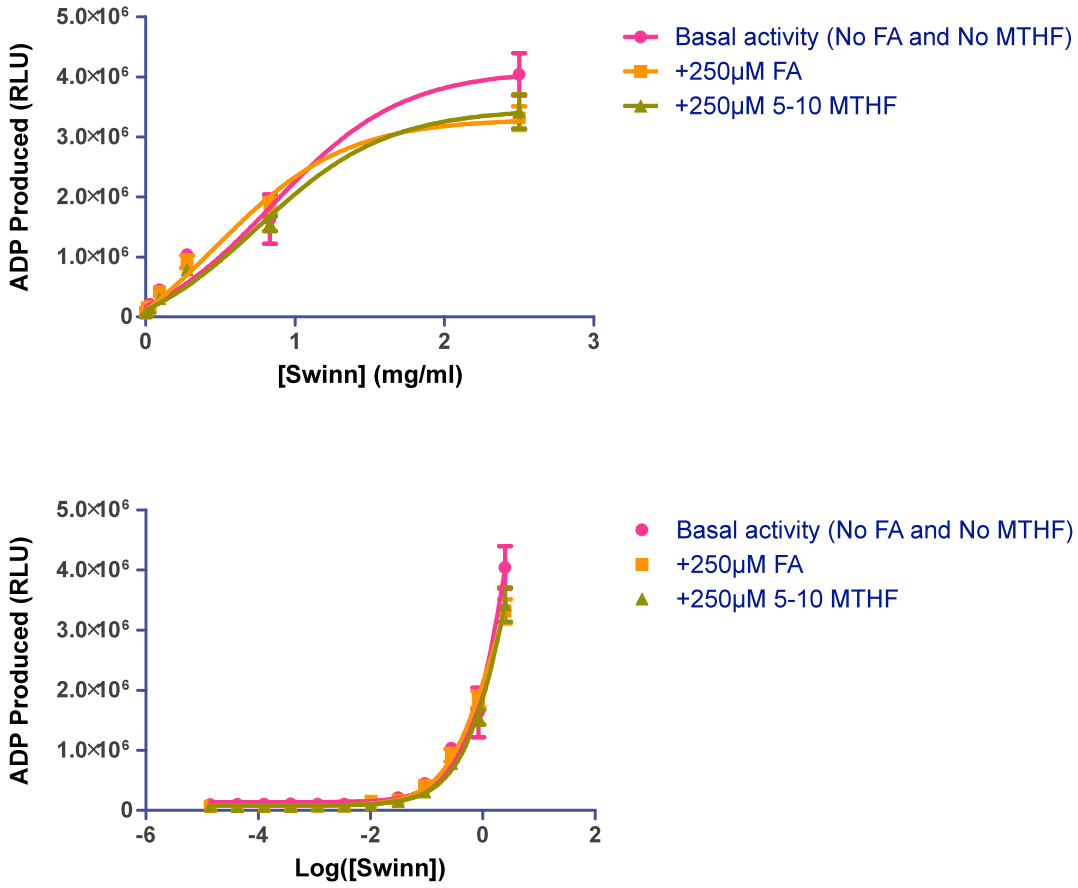



**Figure 21: Swinn does not work on the substrate of MTHFS**

In this experiment, folinic Acid as the substrate of MTHFS and 5,10-Methenyltetrahydrofolate as product of MTHFS was tested to analyze their effect on the ATP hydrolysis activity of Swinn. The results indicated that there is no significant difference between in the level of ATPase activity between these three experiments. These results suggest that FA is not the correct substrate of Swinn. Both of the above figures belong to the same experiment, drawn on different scales (Normal and log scale).

# DISCUSSION

## Summary of our results

Our study focused on characterizing the enzymatic activity of Winn protein. However, since the recombinant *Drosophila* Winn protein is not soluble, we used a homologous protein to Winn from an archaebacterial source, called Swinn. Our experiments confirmed a basal ATPase level for Swinn in absence of its true substrate. Additionally, we confirmed that folinic acid is not a substrate of Swinn, indicating that Swinn has a unique class of substrate.

In order to obtain structural information as to what substrate could potentially bind to the Swinn protein, we attempted to crystallize the protein. These experiments identified two conditions where Swinn can form protein crystals. The analysis of diffraction pattern of these two crystals is under investigation. Our hope is that by obtaining the correct three-dimensional structure of the protein, we can leverage the information about the substrate-binding pocket to narrow the list of potential substrates of the protein.

**Future experiments and our hypothesis for potential substrate**

While our work has focused on defining the enzymatic activity of the Winn family, it will be necessary to relate that enzymatic function back to its in vivo role. Previous genetic work on Winnebago using both over-expression and RNAi has found that it plays a critical role in muscle development [10]. Analysis of a hypomorphic allele of Winnebago found that Winnebago is required for testis development. The results indicated that flies with hypomorphic allele of Winnebago are male sterile. Unfortunately, it remains unclear why loss of Winnebago affects both the testis and muscle. Future studies directed at using the biochemical insights generated by our work to define the molecular defects responsible for both the testes and muscle phenotypes will help close the gap between the genetic and biochemical studies of Winn.

One interesting hint as to how to close this gap comes from the fact that both testis and muscles contain large numbers of mitochondria. Since oxidative phosphorylation by mitochondria generates a large amount of reactive oxygen species, mutations that compromise the response to reactive oxygen species often cause defects in both muscle and testes similar to those seen in Winnebago mutants. Given the fact that Winn is an RNP complex component, this raised the question of whether or not there was a mechanistic link between reactive oxygen species, mRNA

regulation, and Winn function. Since a mature mRNA can undergo damages reactive oxygen species, just like DNA, we propose that Winn is a candidate RNA repair enzyme that can utilize its cyclo-ligase activity to repair RNA bases whose rings have been opened by oxidative damage. This model is attractive since it combines our understanding of Winn enzymology, its role as an RNP component, and the requirement for Winn in muscle and testes into a simple testable model. Future work will be focused on testing whether Winn has the ability to use oxidatively damaged RNA as a substrate. In particular we will focus on oxidative damage to the 5' cap (7-methylguanosine cap) of mRNA since it is particularly susceptible to opening of its carbon nitrogen ring (Figure 22) and since the stereochemistry required to repair an oxidatively damaged 5'cap is identical to that required to close the ring of 5-formyltetrahrofolate (folinic acid) (Figure 23).

While the idea that Winn is an RNA repair enzyme is highly speculative, our combination of biochemical and structural studies also establishes a framework pursuing an unbiased screen for Winn substrates as well as testing substrates suggested by our crystallographic studies. Together these approaches hold great promise in allowing us to define the biochemical role of Winn in regulating RNA and understanding how that regulation is used during *Drosophila* development.

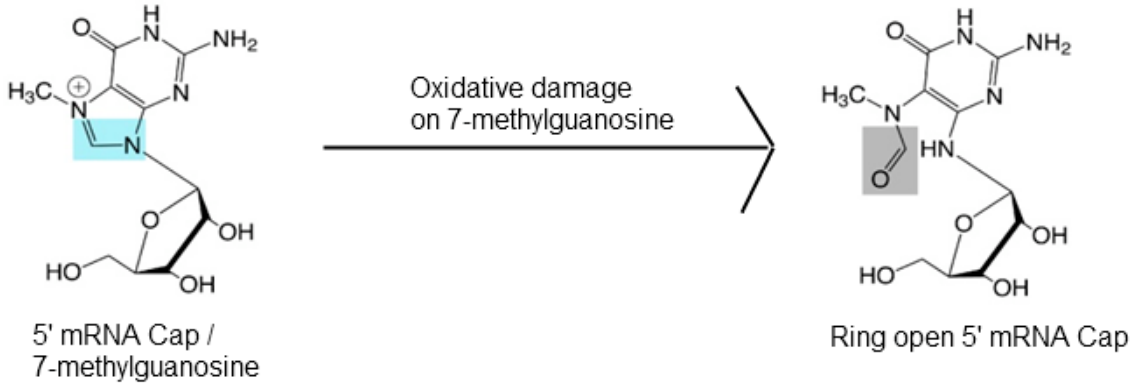



**Figure 22: Oxidative damage to the 5' cap (7-methylguanosine cap)**

Oxidative damage due to activity of free oxygen radicals can potentially damage the 5' cap (7-methylguanosine) of an mRNA molecule and results in formation of a ring open non-functional mRNA molecule.

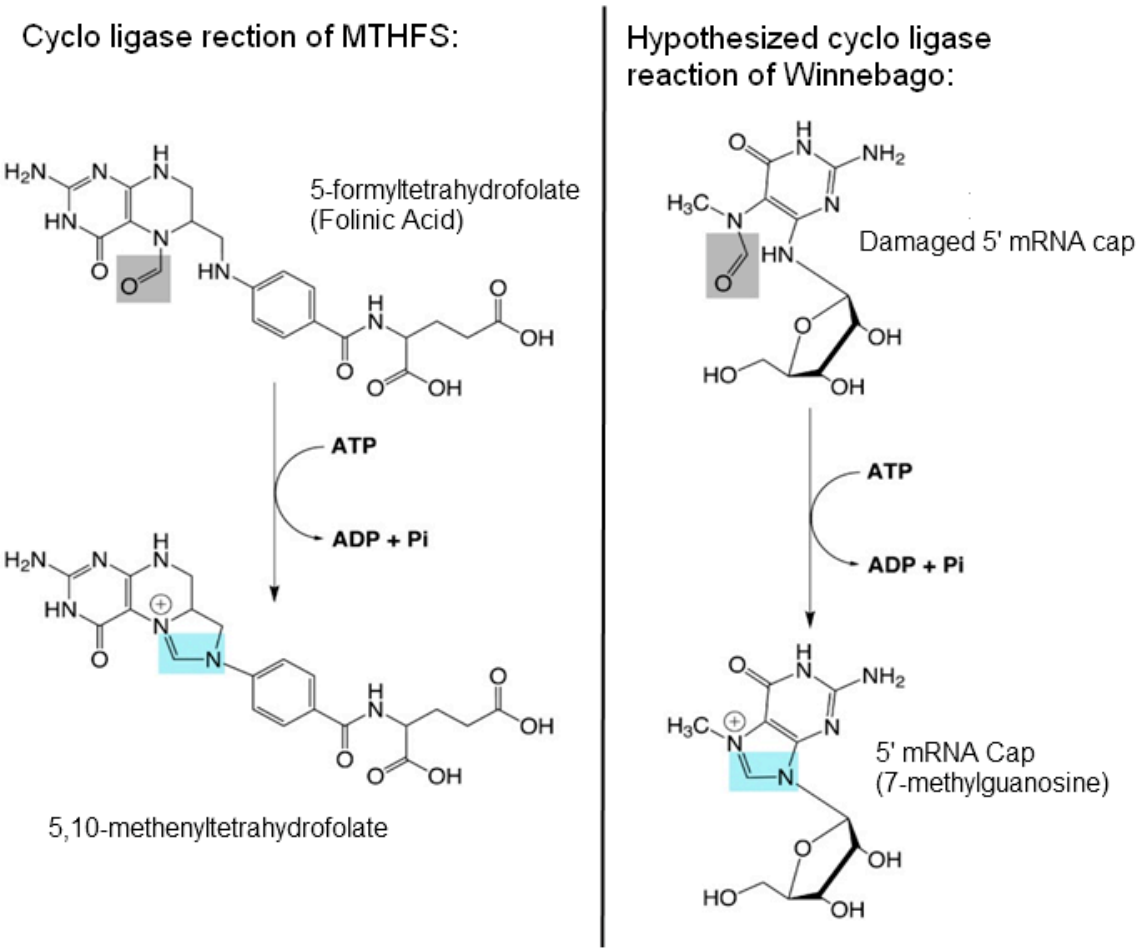



**Figure 23: Hypothesized cyclo-ligase reaction of Winnebago**

The left figure shows the known cyclo-ligase reaction of MTHFS. Due to the similarities between the damaged ring open 5' cap and the open ring of folinic acid, we hypothesize that Winnebago performs a similar chemical reaction in order to fix the damaged 5' mRNA cap.

# APPENDIX

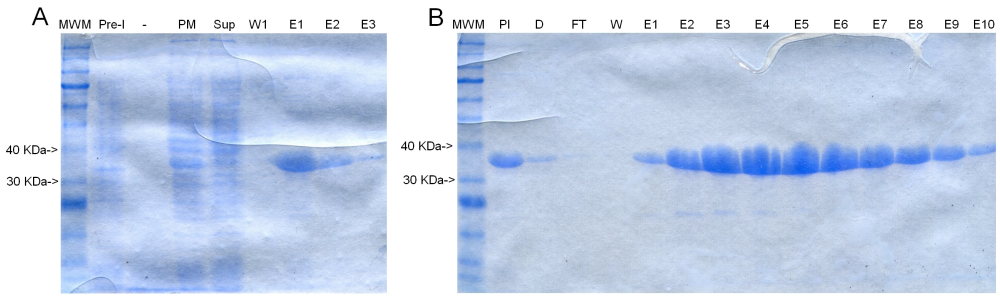



**Figure A.1: Metal chromatography and ion exchange chromatography of K mutant**

Figure A shows the result of metal chromatography of R175K mutant. Figure B shows the result of ion exchange chromatography of R175K mutant. The proteins were visualized using Coomassie blue technique. MWM = Molecular weight marker, Pre-I = pre induction of protein, PM = Pre metal chromatography, Sup = Unbound supernatant of protein-beads mixture, W1 = first wash, E1-3 = represents the eluted protein. PI = Pre ion exchange chromatography, D = diluted protein, FT = Unbound flow through from the column, W = wash, E1-10 = represents the eluted protein using elution buffer with 300mM NaCl.

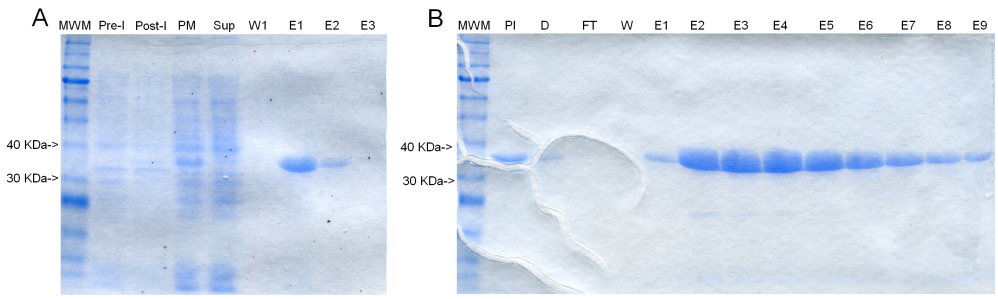



**Figure A.2: Metal chromatography and ion exchange chromatography of M mutant**

Figure A shows the result of metal chromatography of R175M mutant. Figure B shows the result of ion exchange chromatography of R175M mutant. The proteins were visualized using Coomassie blue technique. MWM = Molecular weight marker, Pre-I = pre induction of protein, PM = Pre metal chromatography, Sup = Unbound supernatant of protein-beads mixture, W1 = first wash, E1-3 = represents the eluted protein. PI = Pre ion exchange chromatography, D = diluted protein, FT = Unbound flow through from the column, W = wash, E1-10 = represents the eluted protein using elution buffer with 300mM NaCl.

# REFERENCES


1. Castagnetti, S., et al., Control of oskar mRNA translation by Bruno in a novel cell-free system from Drosophila ovaries. Development, 2000. 127(5): p. 1063-8.

2. Osborne, H.B., et al., Post-transcriptional regulation in Xenopus embryos: role and targets of EDEN-BP. Biochem Soc Trans, 2005. 33(Pt 6): p. 1541-3.

3. Kim-Ha, J., K. Kerr, and P.M. Macdonald, Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. Cell, 1995. 81(3): p. 403-12.

4. Wilhelm, J.E., et al., Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. J Cell Biol, 2003. 163(6): p. 1197-204.

5. Zaessinger, S., I. Busseau, and M. Simonelig, Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development, 2006. 133(22): p. 4573-83.

6. Boag, P.R., A. Nakamura, and T.K. Blackwell, A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in C. elegans. Development, 2005. 132(22): p. 4975-86.

7. Sheth, U. and R. Parker, Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science, 2003. 300(5620): p. 805-8.

8. Mansfield, J.H., J.E. Wilhelm, and T. Hazelrigg, Ypsilon Schachtel, a Drosophila Y-box protein, acts antagonistically to Orb in the oskar mRNA localization and translation pathway. Development, 2002. 129(1): p. 197-209.

9. Snee, M.J. and P.M. Macdonald, Dynamic organization and plasticity of sponge bodies. Dev Dyn, 2009. 238(4): p. 918-30.

10. Staudt, N., et al., Gain-of-function screen for genes that affect Drosophila muscle pattern formation. PLoS Genet, 2005. 1(4): p. e55.