STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE
ESSENTIAL RNA HELICASE MTR4

by

Ryan N. Jackson

A dissertation submitted in partial fulfillment
of the requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
Biochemistry

Approved:

Dr. Sean J. Johnson  Dr. Lance C. Seefeldt
Major Professor  Committee Member

Dr. Joan M. Hevel  Dr. Christopher P. Hill
Committee Member  Committee Member

Dr. Alvan C. Hengge  Dr. Mark R. McLellan
Committee Member  Vice President for Research and
                   Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012
ABSTRACT

Structural and Functional Characterization of the Essential RNA Helicase Mtr4

by

Ryan N Jackson, Doctor of Philosophy
Utah State University, 2012

Major Professor: Dr. Sean J. Johnson
Department: Chemistry and Biochemistry

The essential protein Mtr4 is a conserved Ski2-like RNA helicase that maintains the integrity of nuclear RNA by promoting the 3’ end decay of a wide variety of RNA substrates. Mtr4 activates the multi-protein exosome in RNA processing, surveillance, and turnover pathways by unwinding secondary structure and/or displacing associated proteins from RNA substrates. While Mtr4 may be able to promote decay independently, it is often associated with large multi-protein assemblies. Specifically, Mtr4 is the largest member of the TRAMP (Trf4/Air2/Mtr4 polyadenylation) complex which targets a plethora of RNA substrates for degradation by appending them with small (~5nt) poly(A) tails via the polymerase activity of Trf4. Mtr4 preferentially binds and unwinds RNAs with short poly(A) tails. Notably, the mechanism by which Mtr4 recognizes the length and identity of the RNA 3’ end is coupled to the modulation of poly(A) polymerase activity of Trf4. The lack of structural data for Mtr4 and associated complexes severely limits the understanding of Mtr4 function. Particularly, it is unclear
how Mtr4 senses RNA features, acts on RNA substrates, delivers RNA substrates to the exosome, and assembles into larger protein complexes. Presented here is the x-ray crystal structure of Mtr4 combined with detailed structural and biochemical analysis of the enzyme. The structure reveals that Mtr4 contains a four domain helicase core that is conserved in other RNA helicases and a unique arch-like RNA binding domain that is required for the \textit{in vivo} processing of 5.8S rRNA. Furthermore, kinetic and \textit{in vivo} analysis of conserved residues implicated in the poly(A) sensing mechanism demonstrates that ratchet helix residues regulate RNA unwinding and impact RNA sequence specificity. A comparison of the apo Mtr4 structure with the RNA/ADP bound structure (determined elsewhere) provides a view of the range of motion that individual domains of Mtr4 adopt upon substrate binding as well as the possible conformations that occur during RNA translocation. These studies provide an important framework for understanding the fundamental role of Mtr4 in exosome-mediated RNA decay, and more broadly describe common themes in architecture and function of the Ski2-like helicase family.
PUBLIC ABSTRACT

Structural and Functional Characterization of the

Essential RNA Helicase Mtr4

To insure the integrity of nuclear RNA, the eukaryotic cell employs surveillance systems that identify and degrade RNAs that are detrimental or unneeded. The failure of RNA surveillance systems can lead to neurodegenerative disease states and cancer. The essential RNA helicase Mtr4 is required for the degradation and processing of several nuclear RNAs. To further the understanding of RNA surveillance and processing in eukaryotes, Ryan Jackson of the Department of Chemistry and Biochemistry at Utah State University has determined the molecular structure of Mtr4 and has used this structure to interrogate Mtr4 function biochemically. The structure revealed that Mtr4 contains a four domain helicase core, and a novel domain required for 5.8S ribosomal RNA processing. Furthermore, it was discovered that Mtr4 possesses a ratchet helix that contains conserved residues required for sensing specific substrate characteristics and unwinding activity. This five year, $300,000 project has significantly increased the understanding of Mtr4 structure and function and is currently being used as a platform for future studies aimed at dissecting the mechanisms of RNA surveillance.
ACKNOWLEDGMENTS

It is overwhelming to think of all the people who have helped, encouraged, directed, and mentored me as I have traveled the road of scientific research and of late pursued a Ph.D. in Biochemistry. I would first like to thank the many faculty members of the Chemistry and Biochemistry Department that have inspired me through their teaching and enthusiasm for science. As an undergraduate I was privileged to take courses from Drs. Lisa Berreau, Lance Seefeldt, Joan Hevel, and Scott Ensign. I greatly enjoyed the passion that these instructors have for teaching. Their example helped entice me towards a career in academics. Furthermore, the quality of their teaching reflects the enthusiasm of the Department and University as a whole. I certainly owe Dr. Joseph K.K. Li a tremendous thank you for setting me on the path of research by allowing me to work as an undergraduate and master’s student in his virology lab. Many of the techniques and attitudes I learned from him have helped me in my career.

My supervisory committee, Drs. Joan Hevel, Lance Seefeldt, Alvan Hengge, and Chris Hill, have been incredibly supportive. Their counsel and advice have been a great asset to my training and I hope to be able to continue to receive their mentorship in the future as I pursue a career in academics.

I would like to shout out a big thank you to the members past and present of the Johnson Lab. It has been a lot of fun, yet I am impressed at the dedication our little community has had to hard work and to each other. You have been great friends to me and we have accomplished some great science together. Specifically I would like to thank Sandy Oveson, Chris Bitner, Bradley Hintze, Kristopher Lange, Kelby Bosshardt,
Amy Crandall, Megi Rexhepaj, Anna Lytle, Lacy Taylor, Lindsey Lott, and Jeremy Bakelar for your friendship and examples.

About six years ago Dr. Sean Johnson asked me if I would like to pursue a Ph.D in his lab. I was honored and thrilled to attempt protein x-ray crystallography at the start and it has been quite a ride to get to the finish. Sean has been a wonderful mentor and example throughout the years and has provided me with a solid foundation for future success. I am happy to know that he will continue to give me guidance and direction as I move forward in life.

Lastly I would like to thank my parents, Neal and Shelley Jackson, my children Daniel, Russell, Lydia, Ethan, and Alex, and especially my beautiful, trusting, and very patient wife, Amber. I especially want to thank Amber for her support and strength throughout these last few months and weeks. It is hard to believe that we are finished with such a great task. I could never have done this without you.

Thank you everyone!

Ryan N. Jackson

This work was supported by NSF grant MCB 0952920.
To my children Daniel, Russell, Lydia, Ethan and Alex I dedicate this dissertation.
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>PUBLIC ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>1.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>A COMPREHENSIVE COMPILATION OF METHODS USED TO CHARACTERIZE MTR4</td>
<td>29</td>
</tr>
<tr>
<td>3.</td>
<td>THE CRYSTAL STRUCTURE OF MTR4 REVEALS A NOVEL ARCH DOMAIN REQUIRED FOR rRNA PROCESSING</td>
<td>57</td>
</tr>
<tr>
<td>4.</td>
<td>CONSERVED FEATURES IN THE MTR4 RATCHET HELIX REGULATE RNA HELICASE ACTIVITY</td>
<td>98</td>
</tr>
<tr>
<td>5.</td>
<td>THE SK12-LIKE RNA HELICASE STRUCTURES: COMMON THEMES AND COMPLEX ASSEMBLIES</td>
<td>123</td>
</tr>
<tr>
<td>6.</td>
<td>SUMMARY AND FUTURE DIRECTIONS</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>APPENDICES</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td>APPENDIX A</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>APPENDIX B</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>CURRICULUM VITAE</td>
<td>188</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Buffers used for purification of Mtr4 protein</td>
<td>38</td>
</tr>
<tr>
<td>2.2</td>
<td>Crystallization conditions of Mtr4</td>
<td>42</td>
</tr>
<tr>
<td>2.3</td>
<td>T7 purification buffers</td>
<td>50</td>
</tr>
<tr>
<td>2.4</td>
<td>Data collection, phasing and refinement statistics (SAD)</td>
<td>51</td>
</tr>
<tr>
<td>3.1</td>
<td>Data collection, phasing and refinement statistics (SAD)</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Unwinding and dissociation constants of Mtr4 constructs</td>
<td>108</td>
</tr>
<tr>
<td>5.1</td>
<td>X-ray crystal structures of Ski2-like helicases</td>
<td>128</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Many exosome variants and exosome activating complexes exist in the cell</td>
<td>5</td>
</tr>
<tr>
<td>1-2</td>
<td>Activation of the RNA exosome by the TRAMP complex</td>
<td>8</td>
</tr>
<tr>
<td>2-1</td>
<td>Structure determination of ATCase</td>
<td>52</td>
</tr>
<tr>
<td>3-1</td>
<td>Structural overview of Mtr4</td>
<td>64</td>
</tr>
<tr>
<td>3-2</td>
<td>Architecture of Domains 1, 2, and 3 in Mtr4</td>
<td>66</td>
</tr>
<tr>
<td>3-3</td>
<td>Domain 3a – the “arch” domain</td>
<td>68</td>
</tr>
<tr>
<td>3-4</td>
<td>Comparison of the Mtr4 fist with the L14e ribosomal protein</td>
<td>70</td>
</tr>
<tr>
<td>3-5</td>
<td>RNA binding model</td>
<td>73</td>
</tr>
<tr>
<td>3-6</td>
<td>Surface conservation</td>
<td>76</td>
</tr>
<tr>
<td>3-7</td>
<td>Construction of the Mtr4-archless mutant</td>
<td>78</td>
</tr>
<tr>
<td>3-8</td>
<td>The arch domain of Mtr4 is required for RNA processing and degradation, but not for helicase or ATPase activity</td>
<td>80</td>
</tr>
<tr>
<td>3-9</td>
<td>Model for 5.8S rRNA processing</td>
<td>82</td>
</tr>
<tr>
<td>4-1</td>
<td>Conserved ratchet helix residues interact with nucleic acid</td>
<td>104</td>
</tr>
<tr>
<td>4-2</td>
<td>Unwinding assays of ratchet helix point mutants</td>
<td>107</td>
</tr>
<tr>
<td>4-3</td>
<td>RNA-dependent ATPase enhancement</td>
<td>110</td>
</tr>
<tr>
<td>4-4</td>
<td>Growth complementation of an Mtr4 knockout strain by ratchet helix mutants</td>
<td>111</td>
</tr>
<tr>
<td>5-1</td>
<td>Members of the Ski2-like helicase family share a common four domain helicase core</td>
<td>129</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5-2</td>
<td>Substrate binding induces conformational changes in Mtr4</td>
<td>137</td>
</tr>
<tr>
<td>5-3</td>
<td>Exosome activating complexes are mediated by Mtr4 and Ski2</td>
<td>142</td>
</tr>
<tr>
<td>6-1</td>
<td>The core and the fist of Mtr4 bind structured RNA</td>
<td>161</td>
</tr>
<tr>
<td>6-2</td>
<td>Mtr4 protein crystal growth in the presence of nucleic acid</td>
<td>163</td>
</tr>
<tr>
<td>6-3</td>
<td>The ATPase activity of Mtr4 helicase core mutants</td>
<td>165</td>
</tr>
<tr>
<td>6-4</td>
<td>Probing the substrate sensing mechanism of Mtr4</td>
<td>168</td>
</tr>
<tr>
<td>A-1</td>
<td>Alignment of Mtr4 and Ski2 sequences</td>
<td>173</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

**Functional RNA levels are regulated by processing and degradation machineries**

The steady-state abundance of mature RNA is directly linked to gene expression and cell function (Munsky et al., 2012). In response to cellular signals and environmental stimuli the cell regulates the copy number of protein coding mRNA transcripts available to translational processes in an adaptable and controlled manner, allowing for quick and reproducible alteration of gene expression. A variety of processes that enhance transcription and/or RNA processing increase functional mRNA copy number, whereas degradation pathways are utilized to turnover mRNA that is no longer needed or non-functional (Schaeffer et al., 2011; Schoenberg and Maquat, 2012). Interestingly, only a small portion of the human genome codes for proteins; however, most of the genome is transcribed into RNA (Bickel and Morris, 2006; Struhl, 2007; Willingham and Gingeras, 2006). Many of these non-coding RNAs, including, tRNAs, rRNAs, snRNAs, and snoRNAs, are also heavily involved in the regulation of gene expression (Collins, 2011; Coy and Vasiljeva, 2011; Zhang et al., 2012).

Mature RNA levels, coding and non-coding, are tightly regulated by many processing events that are required to yield correctly folded and fully functional RNA (Bernstein and Toth, 2012). When a functional RNA is no longer needed, or if it is incorrectly processed, it is degraded by RNA turnover machineries (Fasken and Corbett, 2009; Schaeffer et al., 2011; Wolin et al., 2012). Furthermore, RNA surveillance mechanisms are dedicated to the degradation of aberrantly transcribed or modified RNAs.
The RNA exosome is a central player in RNA processing, turnover, and surveillance

In eukaryotes a central component to processing, turnover and surveillance pathways is the RNA exosome (Allman et al., 1999b; Briggs et al., 1998; Butler, 2002; Lebreton and Seraphin, 2008; Lykke-Andersen et al., 2009; Mitchell et al., 1996; Mitchell et al., 1997). The exosome is located in the nucleolus, nucleus, and cytoplasm, where it acts on virtually all RNAs that require 3’end ribonucleolyis (Lubas et al., 2011; Lykke-Andersen et al., 2011). In the cytoplasm the exosome is utilized in general mRNA turnover (Anderson and Parker, 1998; Schaeffer et al., 2011; van Hoof et al., 2000b), and in the surveillance and removal of mRNAs that have been aberrantly transported to the cytoplasm or are incorrectly processed (Doma and Parker, 2007; Hilleren and Parker, 2003; Isken and Maquat, 2007; Mitchell and Tollervey, 2003). Targets of the cytoplasmic exosome also include RNAs with AU-rich elements and the byproducts of RNA interference (Gherzi et al., 2004; Mukherjee et al., 2002; Orban and Izaurrealde, 2005).
Furthermore, the cytoplasmic exosome is proposed to degrade invading viral RNA in innate immune responses (Schaeffer et al., 2011; Toh et al., 1978; Widner and Wickner, 1993).

The nuclear exosome is required for the degradation of a plethora of RNA species. These include byproducts of RNA processing events, incorrectly processed or hypomodified RNAs (Bousquet-Antonelli et al., 2000; Hilleren et al., 2001; Kadaba et al., 2004; Milligan et al., 2005; Torchet et al., 2002; Wang et al., 2008b), RNAs that fail to fold correctly or associate properly within ribonucleic protein complexes (RNPs) (Assenholt et al., 2008; Libri et al., 2002; Rougemeille et al., 2007), and several species that arise from “pervasive” transcription (Belostotsky, 2009; Chen et al., 2001; Lubas et al., 2011; Neil et al., 2009; Preker et al., 2008; van Bakel and Hughes, 2009; Wyers et al., 2005). Furthermore the exosome is required for the 3’-end processing of rRNAs, snRNAs, and snoRNAs (Allmang et al., 1999a; de la Cruz et al., 1998; Suzuki et al., 2001; van Hoof et al., 2000a) and general nuclear mRNA turnover (Bousquet-Antonelli et al., 2000; Guo et al., 2009).

**The eukaryotic exosome resembles archaeal and bacterial RNA degrading complexes**

The eukaryotic RNA exosome is composed of an invariable nine subunit core that forms a ring-like structure with a central channel. Six of these proteins (Rrp41, Rrp42, Rrp43, Rrp45, Rrp46, and Mtr3) contain RNase PH domains and assemble into a hexameric structure that is capped by three proteins (Rrp4, Rrp40, and Csl4) that contain RNA binding S1 and KH domains (Januszyk and Lima, 2011; Liu et al., 2006; Lykke-
Andersen et al., 2011; Malet et al., 2010). The exosome core architecture is reminiscent of analogous RNA degrading complexes found in the other domains of life; namely the archaeal exosome, bacterial RNase PH, and the polynucleotide phosphorylase (PNPase) of the bacterial and organellar degradosome (Borowski et al., 2010; Lin-Chao et al., 2007; Lykke-Andersen et al., 2009). Contrary to what is observed in the archaeal, bacterial, and organellar machineries, the human and yeast RNA exosome core is catalytically inactive due to changes in crucial residues of the RNase PH domain (Lykke-Andersen et al., 2009). Catalytic ribonuclease activity is acquired through association with 3’-5’ exonucleases Rrp44/Dis3 and Rrp6 (Bonneau et al., 2009; Briggs et al., 1998; Butler and Mitchell, 2011; Lorentzen et al., 2008; Mamolen and Andrulis, 2009). Rrp44/Dis3 also contains endonuclease activity that is provided by an accessory PIN domain (Lebreton et al., 2008; Mamolen et al., 2010; Schaeffer et al., 2009; Schneider et al., 2009).

Depending on the organism and cellular localization, the association dynamics of the exosome core with Rrp44/Dis3 and Rrp6 or their respective homologs can vary (Figure 1-1) (Cristodero et al., 2008; Liu et al., 2006; Lykke-Andersen et al., 2011). For example in budding yeast the nuclear exosome associates with both Rrp6 and Rrp44/Dis3 whereas the cytosolic exosome only associates with Rrp44 (Allmang et al., 1999b; Liu et al., 2006; Schaeffer et al., 2011). In humans a different type of association exists. The Rrp6 homolog (hRRP6) is observed in association with every known exosome variant, whereas the two Rrp44/Dis3 homologs, one that possesses both exo and endonuclease activities (hDIS3) and another that contains only exonuclease activity (hDIS3L), do not
Figure 1-1. Many exosome variants and exosome activating complexes exist in the cell. The core exosome (grey), associates with different nucleases (pacmans) and interacts with different activating complexes depending on cellular localization and organism. In yeast (top) the core associates with Rrp44 and Rrp6 in the nucleus and only Rrp44 in the cytoplasm. In humans (bottom) the core exosome associates with human Rrp6 (hRrp6) throughout the cell and associates with different Rrp44-like (hDIS3, hDIS3L) nucleases dependent on localization. Several diverse exosome activating complexes reside in different cellular localizations. They include the TRAMP5 (nucleolus), TRAMP4 (nucleus) and SKI (cytoplasm) complexes in yeast and the human version of TRAMP (nucleolus), and SKI (cytoplasm) as well as the recently discovered NEXT (nucleus) complex. Notably the Ski2-like helicases Mtr4 and Ski2 are central components of exosome activating assemblies.
associate with the nucleolar exosome (Lubas et al., 2011; Staals et al., 2010; Tomecki et al., 2010). It is unclear why different exonucleases are associated with the exosome in a localization dependent manner. However, it may be due to specific characteristics of targeted RNAs in those areas.

**The exosome requires cofactors for *in vivo* activity**

Nuclease complexes including the RNA exosome commonly require RNA binding and modifying enzyme cofactors to effectively recruit and shape substrates in a manner that makes them accessible to hydrolytic active sites (Houseley and Tollervey, 2009; Lin-Chao et al., 2007; Lykke-Andersen et al., 2009). For example, the hexameric PNPase of the *E. coli* degradasome associates directly with a DEAD-box RNA helicase RhlB and possibly with the poly(A) polymerase PAP1 (Py et al., 1994; Py et al., 1996; Raynal and Carpousis, 1999; Xu and Cohen, 1995). The helicase activity of RhlB is thought to unwind secondary structure and/or displace associated proteins from complex RNA substrates, while the addition of non-structured oligo(A) tails by PAP1 is proposed to provide a landing pad for PNPase binding and nucleolytic attack (Lykke-Andersen et al., 2009). The RNA exosome utilizes several cofactors and protein complexes for *in vivo* activation and substrate recruitment (Figure 1-1) (reviewed in (Anderson and Wang, 2009; Houseley and Tollervey, 2009; Lebreton and Seraphin, 2008; Lykke-Andersen et al., 2009; Schmid and Jensen, 2008)). These include large complexes that contain diverse enzymatic activities such as the Ccr4-not complex, the Nrd1-Nab3 complex, the TRAMP complex, the FFC, and the NEXT complex as well as independent cofactors that
associate directly with the nucleases like Rrp47 and Mpp6 (Baker et al., 2012; Butler and Mitchell, 2011; Collart and Panasenko, 2012; Norbury, 2011). Of all the many cofactors involved the pathways of nuclear exosome activation, the Ski2-like RNA helicases Mtr4 and Ski2 appear to be the gatekeepers of degradation.

**Mtr4 is a member of several exosome activating complexes**

Mtr4 is a conserved and essential Ski2-like RNA helicase that is a major activator of the nuclear exosome (de la Cruz et al., 1998; Liang et al., 1996). Although Mtr4 may stimulate the exosome independently (LaCava et al., 2005), it is a major component of several exosome activating complexes (Figure 1-1). In budding yeast, Mtr4 is the largest protein of the TRAMP4 (Trf4/Air2/Mtr4/Polyadenylation) complex (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005), which tags RNA substrates for degradation by appending small (~5nt) poly(A) tails through the combined activities of the RNA binding zinc knuckle Air2 protein and the poly(A) polymerase Trf4 (Vanacova et al., 2005). The delivery of tagged substrates to the exosome for decay requires Mtr4, which when mutated or reduced causes a buildup of polyadenylated RNAs (Liang et al., 1996; Vanacova et al., 2005). TRAMP5 (Trf5/Air1/Mtr4 Polyadenylation), an analogous yeast complex, similarly targets RNAs for decay, but localizes predominantly to the nucleolus (Dez et al., 2006; Dez et al., 2007; Houseley and Tollervey, 2006; Rougemaille et al., 2007; San Paolo et al., 2009; Wery et al., 2009). Notably the cytosolic homolog of Mtr4, Ski2, forms a quaternary complex, called the Ski-complex, composed of the RNA helicase Ski2, with two Ski8 proteins containing a WD repeat fold and a tetratricopeptide
repeat protein Ski3 (Brown et al., 2000; Schaeffer et al., 2011; Synowsky and Heck, 2008). While Ski2 highly resembles Mtr4 and likely presumes a similar function in

**Figure 1-2.** Activation of the RNA exosome by the TRAMP complex. The TRAMP complex acts on RNA substrates (represented by the line) by first distinguishing a non-functional RNA from functional RNAs. RNA substrates are tagged for degradation via the polyadenylation activity of Trf4 and are delivered to the exosome for 3’ end decay via the unwinding activity of Mtr4.
exosome activation, the activities of Ski8 and Ski3 are not equivalent to Trf4 and Air2, demonstrating that distinct cofactor complexes are utilized by the cell depending on localization (see Figure 1-1).

Likewise, other organisms display variation in Mtr4 containing exosome activating complexes in a location dependent manner. Human Mtr4 associates with a homologous TRAMP complex that resides in the nucleolus as well as the NEXT (Nuclear Exosome Targeting) complex which is found exclusively in the nucleus (Figure 1-1) (Lubas et al., 2011; Norbury, 2011). It is unclear why such a variety of complexes are utilized to activate the exosome, however the rationale may be related to differences in the RNA substrates being surveyed and the specific characteristics expressed by distinct exosome complexes that are also differentiated based on cellular localization (Lubas et al., 2011). Notably, the association of a Ski2-like RNA helicase (Mtr4 or Ski2) is a common feature observed in almost all known exosome activating complexes. It is proposed that the helicase activity of Mtr4 promotes exosomal decay by unwinding secondary structure and displacing bound proteins from RNA transcripts (Figure 1-2) (Anderson and Wang, 2009; Houseley and Tollervey, 2009; Lykke-Andersen et al., 2009).

**Structural information regarding Mtr4 and related exosome cofactors is lacking**

When the work described in this dissertation began a large amount of structural and biochemical information was available regarding the RNA exosome core and associated nucleases (Butler, 2002; Januszyk and Lima, 2011; Lykke-Andersen et al.,
2009). However, there was little to no structural information regarding Mtr4, Ski2, or any TRAMP complex components. The structural information available was limited to amino acid sequence alignments with proteins with known structures. The sequence of Mtr4 demonstrated the presence of the RecA-like domains 1 and 2 that contain the canonical helicase motifs required for ATPase and helicase activity (de la Cruz et al., 1998; Liang et al., 1996). These regions were predicted to align with the RecA folds of the Ski2-like helicase Hel308 for which there were x-ray crystal structures available (Buttner et al., 2007; Oyama et al., 2009; Richards et al., 2008; Zhang et al., 2008). The remaining C-terminal sequence of Mtr4, including the region described as a DSHCT (Dob1/Ski2/He1Y/C-terminal) domain (Staub et al., 2004), did not align with any known structural feature. The lack of structural data for Mtr4 left many unanswered questions regarding the function of Mtr4 and the exosome activating complexes with which it associates. These ranged from basic questions regarding the general structure of the protein, and the contributions of the C-terminus to Mtr4 activity, to more complex questions involving the structural association of Mtr4 with protein complexes and the exosome for targeting and delivery of RNA for decay. In particular, it was unclear how RNA substrates that required 3’ ribonucleolysis were distinguished from functional RNAs when the various substrates have no common feature in sequence, secondary structure or associated proteins, and are derived from all three RNA polymerases (I, II, and III).

In an effort to shed light on these unanswered questions this work has aimed to structurally and functionally characterize the RNA helicase Mtr4. Our efforts to
determine the apo x-ray crystal structure of Mtr4 and follow up biochemical analyses have significantly contributed to the understanding of Mtr4 function within exosome-mediated decay pathways. To provide a more in depth understanding of what information was available at the time this work was started the next section will describe the background history and function of Mtr4. Following the Mtr4 background section a description of how this work has impacted the understanding of Mtr4 structure and function will be discussed followed by a description of the summary and future directions of this work.

**History and function of Mtr4**

Mtr4 was first discovered by the Tartakoff lab in a yeast genetic screen for proteins involved in the transport of RNAs out of the nucleus (Liang et al., 1996). A mutagenic strain mtr4-1 (C942Y) caused a build-up of nuclear poly(A) RNA, and it was assumed at the time that the mutation precluded mRNA transport. Thus, Mtr4 stands for mRNA transport 4. It was also determined that Mtr4 is a Ski2-like RNA helicase essential for yeast growth. Mtr4 was concurrently identified as an essential protein by the Tollervey lab in a genetic screen probing for mutations that could be rescued by overexpression of the elongation initiation factor 4b (e14b) (de la Cruz et al., 1998). Mtr4 was given the alias Dob1 or dependent on b 1, because when mutated, yeast viability could only be acquired through the overexpression of e14B. The Tollervey study was the first of many to show that Mtr4/Dob1 is required for the correct processing and degradation of RNA through exosome activation (de la Cruz et al., 1998). A few years later, Mtr4 was discovered to be a member of the TRAMP complex that is involved in the
exosome-mediated degradation of a plethora of RNA substrates (LaCava et al., 2005; Vanacova et al., 2005; Wolin et al., 2012; Wyers et al., 2005). The discovery of TRAMP explained the initial observation of poly(A) RNA buildup upon Mtr4 depletion (Liang et al., 1996). The knock down or mutation of Mtr4 caused the buildup of several RNA species with poly(A) tails, demonstrating that Mtr4 is a critical link between the substrate targeting poly(A) activity of Trf4/5 and exosomal ribonucleolysis (Houseley and Tollervey, 2006). Furthermore, Mtr4 is required for maintaining correct gene expression. In fission yeast TRAMP has been shown to silence aberrantly transcribed genes from heterochromatin and telomere regions (Bayne et al., 2007; Buhler et al., 2007; Houseley et al., 2007; Wang et al., 2008a). In Neurospora Crassa under the alias of FRH (Frequency Related Helicase) Mtr4 forms a stoichometric 1:2 FRH-FRQ complex (FFC) that negatively regulates the mRNA levels of the frequency allele and other genes by targeting transcribed mRNA to the exosome (Baker et al., 2012; Cheng et al., 2005; Guo et al., 2009; Shi et al., 2010). Also, in Arabidopsis thaliana the Mtr4 homolog HEN2 was shown to be involved in controlling reproductive organ identity, and it is proposed to function by regulating mRNA levels (Western et al., 2002).

It was initially unclear if Mtr4 was a bona fide helicase. Genetic studies in budding yeast demonstrated that mutation of some of the canonical helicase motifs in Mtr4 caused dominant negative growth defects (Bernstein et al., 2006). Expression of some motif mutations in an Mtr4 deletion strain caused lethality or slow growth (Bernstein et al., 2006; Liang et al., 1996; Wang et al., 2008b). However, other mutations of the helicase core motifs had little to no effect in vivo. Further
characterization of recombinant Mtr4 from *S. cerevisiae* demonstrated that Mtr4 possesses 3’-5’ unwinding activity, RNA dependent ATPase activity and a preference for poly(A) RNA (Bernstein et al., 2008; Wang et al., 2008b). Moreover, mutations that diminished Mtr4 unwinding *in vitro* caused a buildup of hypomodified tRNA\(^{1}_{\text{met}}* in vivo*, suggesting that the Mtr4 unwinding activity is important for exosome-mediated RNA degradation (Wang et al., 2008b). Although these initial studies provided some characterization of Mtr4 as a helicase and identified what substrates it acted upon, many questions regarding Mtr4 function remained unanswered. In particular it was unclear how Mtr4 identified substrates, interacted with other proteins, and what role the C-terminal region contributed to function.

**Structural studies of Mtr4**

The primary aim of the work described in this dissertation was to better understand the function of Mtr4 using structural and biochemical analysis. One of the major contributions of the work was the determination of the x-ray crystal structure of apo Mtr4 (described in Chapter 3). We showed that Mtr4 contains a helicase core composed of canonical RecA folds (domains 1 and 2), a winged helix domain (domain 3), and a helical bundle domain (domain 4) observed in related Ski2-like helicases and more distant DEAH-box helicases (Buttnner et al., 2007; Halbach et al., 2012; He et al., 2011; Oyama et al., 2009; Pena et al., 2009; Richards et al., 2008; Walbott et al., 2010; Zhang et al., 2009; Zhang et al., 2008). The discovery of the structural conservation via amino acid sequence analysis was precluded by a large 265 amino acid insertion. The insertion extends above the helicase core with two groups of antiparallel coiled coils that
terminate in a globular fist. Structure comparison search engines revealed that the insertion was a novel domain (Holm and Rosenstrom, 2010; Krissinel and Henrick, 2004). Because of the striking similarity in appearance to Delicate Arch, which is located in Utah, we coined the this insertion the “arch domain.” We showed that the arch domain is unique to Mtr4 and Ski2 proteins, is not required for RNA unwinding or ATPase activity, but is required for proper 5.8S rRNA processing in vivo. Concurrently, another group described a second Mtr4 structure bound to a 5 poly(A) oligo and ADP. They also showed that the arch domain binds structured RNA substrates (tRNA$_{\text{met}}$) and is not necessary for TRAMP formation (Weir et al., 2010).

Because the apo and substrate bound Mtr4 structures were published at approximately the same time, the initial publications did not include their structural comparison. In Chapter 5 of this dissertation we compare the structures of apo and substrate bound Mtr4 in a comprehensive review of Ski2-like helicases. We show that conformational differences in domain positions exist in the substrate bound and apo states. We also show that the arms of the arch domain are fairly rigid but the fist appears to be able to conform to a variety of positions. Furthermore, conformational differences of key residues that are likely involved in the ATPase dependent unwinding activity of Mtr4 within the core were identified. Several mutations of these residues have been made and initially tested for unwinding and ATPase activity. The results of these investigations are not yet published and are further discussed in the future directions section of Chapter 6.
The initial structure determination of Mtr4 and the discovery of the arch domain provided a strong framework for understanding Mtr4 function. However, many questions regarding Mtr4 function remain unanswered for which x-ray crystal structures of Mtr4 in complex with other proteins and/or structured RNAs would provide much needed direction towards resolution. Attempts at the crystallization of several structured RNA and DNA substrates with Mtr4 and related protein constructs that include the arch domain and the fist have been made. Conditions in which Mtr4 requires the presence of nucleic acid to crystallize have been discovered. However, to date no crystals grown with nucleic acid have diffracted to an acceptable degree. An explanation of the crystallization methods used to obtain these potential Mtr4-nucleic acid co-crystals is described in Chapter 2. Mtr4 can form many complexes with other proteins and, although significant attempts have been made towards obtaining crystals of full length TRAMP4 from S. cerevisiae, a vast amount of crystallization trials remain to be explored. Chapter 6 delineates what future experiments should be pursued in an attempt to obtain Mtr4 crystal structures bound to structured nucleic acid and/or associated proteins.

The poly(A) tail is central to TRAMP activity

After the determination of the x-ray crystal structures of Mtr4, several publications further characterized the functional activity of Mtr4 and role of the poly(A) tail in TRAMP function (Bernstein et al., 2010; Jia et al., 2011; Jia et al., 2012; Wlotzka et al., 2011). A paper from the Toth lab indentified the minimal binding site for Mtr4 (~5nt) and showed that Mtr4 has two distinct modes of RNA binding (Bernstein et al., 2010). Furthermore, it was shown that the preference for poly(A) over random RNA
sequence is heightened when only one Mtr4 binding site is available and that sensitivity to sequence is dependent on the identity of nucleotide present. These data fit nicely with work from the Jankowsky lab that showed that the mechanism by which Mtr4 senses the 4-5 adenosine 3’ tail is central to TRAMP function (Jia et al., 2011; Jia et al., 2012). Within a TRAMP context, Mtr4 recognizes the number of adenosines located at the 3’ end of RNA transcripts, and when less than 4 adenosines are present, Mtr4 enhances Trf4 activity. Once 4 or more adenosines accumulate, Mtr4 restricts Trf4 polymerase activity effectively limiting the length of the poly(A) tail to ~5nt (Jia et al., 2011). Intriguingly, RNA substrates with a 3’ overhang of 5 adenosines are unwound at a faster rate than substrates carrying a longer poly(A) overhang or a non-(A) sequence of similar length (Jia et al., 2012). These data demonstrate that Mtr4 and Trf4/Air2 act synergistically to modify the 3’ end of RNA substrates in order to obtain optimal TRAMP function. Central to this optimization is the ability of Mtr4 to sense the length and identity of the 3’ tail of RNA substrates. Notably, Mtr4 translocates across various different lengths and sequences of RNA substrates after sensing the poly(A) tail, suggesting a mechanism in which Mtr4 switches from a sensing conformation to a translocating conformation.

We used the x-ray crystal structures of Mtr4 and conservation analysis to identify residues along the RNA binding path likely to sense identity and sequence of RNA substrates. Several conserved residues along the ratchet helix of domain 4 were identified, mutated and analyzed both in vitro and in vivo. The results of these studies are described in detail within Chapter 4 and Chapter 6 of this dissertation. We show that conserved ratchet helix residues R1030 and E1033 are required for proper helicase
function \textit{in vitro} and \textit{in vivo} and may be involved in poly(A) tail sensing. Interestingly, each residue has a distinct role within the unwinding mechanism. When R1030 is mutated to an alanine, unwinding activity decreases. Conversely, when E1033 is mutated to an alanine, the unwinding activity improves and a greater maximum unwinding rate constant is observed. However, the observed rate constant enhancement is at the expense of substrate functional affinity. Not only does the E1033A mutant Mtr4 have less functional affinity for RNA substrates, it also differs from the wild-type enzyme in failing to enhance ATPase activity in a substrate dependent. The expression of the ratchet mutations \textit{in vivo} demonstrate that they are important for normal growth but are not essential for viability. The double mutant of these residues does not compound the observed growth phenotype, suggesting that both residues act in the same pathway that is disrupted by mutation. These studies have provided an initial analysis of key residues likely to be involved in the poly(A) binding pathway from which future studies will be generated.

\textbf{Conclusion}

The work described in this dissertation has made a significant impact in the field of RNA surveillance and quality control. Mtr4 is a major activator of the nuclear RNA exosome that is necessary for the processing and degradation of a myriad of RNAs. The determination of the first x-ray crystal structure of Mtr4 revealed the presence of the arch domain that binds RNA and is required for 5.8S rRNA processing. It was also shown that the four domain arrangement that makes up the helicase core of Mtr4 is a conserved feature observed throughout the Ski2-like and DEAH-box helicase families. Kinetic
studies demonstrated that conserved ratchet helix residues of domain 4 that lie along the RNA binding path are important for Mtr4 unwinding and may be involved in the poly(A) sensing mechanism central to TRAMP function. The in-depth structural analysis of apo and RNA/ADP bound Mtr4 structures as well as the preliminary biochemical investigation of Mtr4 core residue point mutants provides a strong foundation for future studies regarding Mtr4 function.

References


Anderson, J.S. and Parker, R.P. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. EMBO J, 17, 1497-1506.


de la Cruz, J., Kressler, D., Tollervey, D. and Linder, P. (1998) Dbl1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J*, 17, 1128-1140.


Suzuki, N., Noguchi, E., Nakashima, N., Oki, M., Ohba, T., Tartakoff, A., Ohishi, M. and Nishimoto, T. (2001) The *Saccharomyces cerevisiae* small GTPase, Gsp1p/Ran, is involved in 3' processing of 7S-to-5.8S rRNA and in degradation of the excised 5'-A0 fragment of 35S pre-rRNA, both of which are carried out by the exosome. *Genetics*, **158**, 613-625.


CHAPTER 2
A COMPREHENSIVE COMPILATION OF METHODS
USED TO CHARACTERIZE MTR4

Abstract

Several methodologies were adapted to determine the structure and function of Mtr4, including molecular cloning, protein purification, protein crystallization, structure solution and structure modeling strategies. Established ATPase, Helicase and Electrophoretic Mobility Shift Assay (EMSA) techniques were modified to explore the biochemical parameters of Mtr4 and Mtr4 mutants. An integral part of the biochemical assays and many crystallization screens was the presence of RNA. An in vitro RNA transcription protocol was developed that utilizes Johnson Lab T7 polymerase and transcription buffers. Herein is a comprehensive overview of the methods and techniques used to probe Mtr4 structure and function, written in greater detail than that found in the methods section of subsequent chapters, to facilitate the successful repetition of the experiments described in this dissertation. Methodologies not covered in subsequent chapters are addressed here. Moreover, a description of the methods employed to solve the structure of small cubic crystals that grew in the presence of the N-terminus of Ski2 construct is described.

Introduction

Recombinant Mtr4 was cloned expressed and purified within a BL21 DE3 codon plus E. coli expression system before I began to work on this project. However, no protein crystals had been obtained. My efforts began with optimization of the protein
purification process, which included switching from a Tris buffer system to HEPES. Contemporaneously, Mtr4 surface entropy reduction mutants (Cooper et al., 2007) were made in an effort to aid Mtr4 crystallization. Purification of the surface entropy Mtr4 mutant (E47A/E48A/K49A) in HEPES buffer allowed for crystallization. When purified in HEPES the wild-type protein also crystallized in the same conditions, indicating that the HEPES buffer provided the parameter that yielded crystallization. Thus, purification of Mtr4 and Mtr4 mutants is commonly performed in HEPES buffer. However, the stability of Mtr4 in Tris buffers appears to be equivalent to that observed in HEPES and other buffers such as Tris, NaPO₄ and MOPS have been used successfully in biochemical assays probing Mtr4 function (Bernstein et al., 2008; Jia et al., 2011; Wang et al., 2008). Notably an experiment by Bernstein and colleagues determined that Mtr4 was most stable at pH 9.5 in CHES buffer (Bernstein et al., 2008). Rapid purification of Mtr4 within a 24 hour period provided the best yield. Below, under Mtr4 expression and purification strategies, are the buffers, columns and schedules utilized to express and purify Mtr4.

Mtr4 at times was difficult to mutate and several different strategies were used to successfully modify Mtr4 constructs. These included the use of PCR, restriction digestion and ligation strategies, a modified version of the QuikChange protocol from Agilent Technologies and a PCR based mutation method that proved to be the only strategy that worked to mutate certain locations on the Mtr4 construct (Liu and Naismith, 2008). All three of these methods are described under the heading Mtr4 and Mtr4 Mutant construct design.

Mtr4 first crystallized in a Hampton SaltRX robot screen (Hampton Research).
Subsequent hand-trays successfully produced Mtr4 crystals at Utah State and the conditions were further optimized to obtain crystals that diffracted to 3.4 Å. A variety of Mtr4 constructs in combination with appropriate substrates that include ADP, AMP-PNP, RNA and DNA were screened in crystallization trials. Several crystal hits of Mtr4 in the presence of nucleic acid were observed. However, to date none of these putative co-crystals has diffracted to a resolution better than 10 Å. Under the heading Mtr4 crystallization, the conditions and protein preparations that have yielded crystals are described in detail.

The biochemical analysis of Mtr4 required the production of in vitro transcribed RNA. Initially the Ampliscribe T7-Flash kit from Epicentre was used to produce the 42-mer RNAs used as the bottom strand in the helicase assay described in Chapter 3. However, crystallization trials of Mtr4 with structured RNAs, such as tRNA and rRNA, required that in vitro transcription be done large scale. The cost of commercial enzymes and buffers precluded their use for in vitro transcription that would provide the quantities of RNA desired for crystallization trials. To obtain RNA for crystallization and biochemical analysis, T7 polymerase, buffers and NTPs were made and/or prepared within the Johnson lab for large scale in vitro transcription. Under the heading in vitro transcription, are described the buffers and protocols associated with T7 expression and purification, large scale in vitro transcription and RNA purification.

The biochemical characterization of Mtr4 was performed initially by Bernstein and colleagues (Bernstein et al., 2008) and Wang and colleagues (Wang et al., 2008). These seminal studies used helicase and ATPase assays to confirm that Mtr4 has RNA-
dependent ATPase activity and ATP-dependent 3'-5' duplex RNA unwinding activity. Further studies explored the function of Mtr4 independent of TRAMP (Bernstein et al., 2010; Jia et al., 2011) and within a TRAMP context (Jia et al., 2011). The protocols developed by these groups were utilized and slightly modified in this work to explore the activity of the Mtr4-archless and Ratchet helix mutants. Also, EMSA techniques were used to demonstrate preliminary RNA binding and to determine dissociation constants with RNA substrates. The assays are fairly straightforward and are described in full detail in Chapters 3 and 4.

The structure solution for Mtr4 was performed using single wavelength anomalous dispersion methods on Selenomethionine derived Mtr4 crystals. Several techniques were used for the modeling and refinement of Mtr4 and they are described in full detail in Chapter 3.

A second protein structure was solved using SAD techniques during the course of this work in an attempt to determine the crystal structure of the N-terminus of the Ski2 RNA helicase. The N-terminus of Ski2 was successfully cloned, expressed and purified. However, purified protein contained a small fraction (~5%) of contaminating proteins that could be observed on a Coomassie stained gel. Crystallization trials yielded small cubic crystals that were difficult to reproduce. Eventually, seeding methods allowed for the reproduction of some small crystals that were used to collect native and heavy metal soaked x-ray diffraction data sets. The cloning, expression, purification and crystallization methods of the N-terminus of Ski2 are explained in the document entitled “ATCase purification crystallization and structure solution” located in the lab drive of the
Johnson lab. Within this chapter is described the heavy atom soaking, data collection, data processing and SAD strategies used to determine the structure of the E. coli protein ATCase, which was a contaminant in the N-terminus of Ski2 protein preparation.

The Consurf server (Ashkenazy et al., 2010) scores conservation according to an algorithm that considers phylogenetics into its calculation by weighting the conservation of residues in more distantly related organisms more heavily. The ConSurf server was used to calculate the conservation scores of Mtr4 and related Ski2 like helicases. Although the process of calculating a conservation score is clearly explained on the ConSurf server webpage, the exact details of how the conservation scores described in Chapters 3 and 4 were obtained are described in this chapter.

This chapter describes in detail the methods use to structurally and functionally characterize Mtr4 as well as the methodologies developed to produce structured RNAs. It is anticipated that this comprehensive methods chapter will be of great use to future researchers studying helicase function or determining protein x-ray crystal structures.

**Methods**

**Mtr4 and Mtr4 mutant construct design**

As described in the methods section of Chapter 3, the full length Mtr4 DNA sequence from *S. cerevisiae* was inserted into a pET 151-D-topo *E. coli* expression vector (Invitrogen) with an N-terminal, TEV-protease cleavable, hexahistidine-tag. Removal of the his-tag by TEV cleavage leaves six amino acids N-terminal to the native sequence. The insertion of Mtr4 DNA into the pET 151-D-topo vector was performed by following
the protocol in the pET 151-D-topo manual. Several truncated Mtr4 vectors were also made in the same fashion. The pET Directional TOPO® system employs an elegant cloning procedure that utilizes a topoisomerase to cut and ligate a complementary sequence into the pET 151-D-topo vector. It requires a CACC nucleotide sequence to be designed into the PCR forward primer in frame with the first codon of the protein or protein truncation for which you are generating. The CACC anneals with a complementary 3’ overhang located on the TOPO vector and is covalently linked to the vector by the topoisomerase.

The use of restriction digestion enzymes and ligases was commonly employed to cut and paste DNA from PCR products and/or existing vectors into Mtr4 expression vectors. In particular the Mtr4 - archless mutant was constructed in this fashion by replacing protein residues 615-878 with a 4 residue ELST linker. This was done by first cloning the C-terminal truncation Mtr4 1-615 into the pET 151-D-topo vector as described above with a SalI restriction enzyme site at the 3’ end (archived as SJJ014 and SJJ035). A 5’ SalI restriction enzyme site was incorporated using PCR into the Mtr4 879-1073 DNA sequence and was cloned directly into pET 151-D-topo (archived as SJJ029 and SJJ034). The truncated Mtr4 vectors were transformed into E. coli DH5α cells, amplified and then purified using the Qiagen QIAprep Miniprep plasmid purification kit. Fermentas restriction enzymes and protocols were used to cut and ligate the two constructs together, utilizing the SalI restriction site to make the Mtr4-archless construct (archived as SJJ040). A similar procedure was used to remove a 30 residue loop within Mtr4 to make the Mtr4 loopless construct (archived as SJJ042 ). Plasmids were
transformed, amplified, purified and sequenced before archiving. A wide variety of Mtr4 N and C terminal truncations have been constructed in this manner. Sequence information for these constructs is available at the Johnson lab pymol wiki webpage and the plasmids are archived at -20 and -80°C.

A modified version of the QuikChange Site-Directed Mutagenesis System protocol (Agilent Technologies) was used to make a majority of Mtr4 point mutants. We did not use the buffer or polymerase provided in the QuikChange kit to make the mutations. However, the optimal forward and reverse mutagenesis primers are designed using QuikChange Primer Design Program for the desired mutation. The site-directed mutagenesis protocol uses Johnson Lab made PFU polymerase (see Johnson Lab Protocols for purification strategy) and 10X PFU buffer (Batey and Kieft, 2007) that consists of 200 mM HEPES pH 8.6, 10 mM (NH4)2SO4, 100 mM KCl, 0.1% Triton X-100 and 20 mM MgSO4. Also, PFU polymerase buffers from commercial sources such as Stratagene's Herculase 10X buffer has been used successfully. As an example of this method, the following is a detailed description of how the ratchet helix mutants were made.

The primers for the ratchet helix mutants R1033A, R1033F and R1033W were designed using the Agilent Technologies QuikChange webpage calculator. Mutagenesis reactions were performed in 50 μl final volumes with 1X Herculase buffer (Agilent Technologies), 125 ng each of forward and reverse primers, 10 mM of dNTPs and 1 μl of diluted PFU polymerase (1:300) and included 5ng, 35 ng or 50 ng of DNA template for each desired mutation resulting in 9 mutagenesis reactions total. The PCR parameters
were set as follows: 2 minutes at 95°C, followed by 20 cycles of 50 seconds at 95°C, 50 seconds at 53°C and 11 minutes at 68°C, finished with a final extension of 68°C for 11 minutes. After the mutagenesis reactions were finished, digestion of original plasmid by dpn1 restriction enzyme was performed for 1 hour at 37°C. Transformations were performed into chemically competent DH5α cells using volumes of 2μl, 4μl and 5μl for each PCR reaction resulting in 27 total transformations. Cells were plated onto LB plates containing ampicillin to select for the pET plasmid. Colonies were only observed for PCR mutagenesis reactions that used at least 35 ng of template DNA. Interestingly, the E1033A mutant only produced one colony total from the 35 ng template PCR reaction and the 4 μl transformation. The E1033F mutagenesis reaction produced colonies with the 35 and 50 ng PCR reactions while E1033W only produced colonies with the 50 ng template mutagenesis reaction. The common theme between all these mutagenesis reactions is that 35-50 ng of DNA template should be used in the mutagenesis reaction and the use of several different template concentrations in the initial mutagenesis reactions can improve the odds of a positive outcome. Additionally, point mutagenesis of Mtr4 using 35 ng of template has worked several times for other members of the Johnson lab.

Several cloning strategies were tried without success to mutate conserved residues in the arch domain of Mtr4 and the aspartate and glutamate of motif II in the Mtr4 helicase core. A modified site directed mutagenesis protocol (Liu and Naismith, 2008) that uses a PCR based method to amplify the desired mutation was utilized to successfully modify the Mtr4 construct in these difficult regions. Primers were designed
that contained an overlapping and non-overlapping region. The non-overlapping region allowed the primers to bridge the nicks found on newly polymerized DNA and discouraged primer dimerization. The primers were designed so that the non-overlapping region had a melting temperature (Tm) 5-10°C greater than the overlapping region, which promoted annealing to template strands in the non-overlapping region. The final PCR reactions contained 1X PFU Turbo buffer, 10-15 ng of Mtr4 pET 151-D-topo vector, 1 μM of the forward and reverse primers, 8 % DMSO and 1.2 μl of PFU Turbo in a 50 μl reaction. The PCR parameters were set as 1 cycle at 95°C for 5 minutes, followed by 18 cycles of 95°C for 1 minute, 5°C below the Tm of the non-overlapping region for 1 minute and 68°C for 13 minutes 30 seconds, finished with a 68°C extension for 30 minutes followed by storage at 4°C.

Mtr4 Protein Expression and Purification Strategies

Mtr4 protein of S. cerevisiae was recombinantly expressed in an E. coli BL21(DE3) codon+RIL cell line (Agilent Technologies). Protein expression was induced using an autoinduction protocol (Studier, 2005). Autoinduction yielded a greater amount of protein per media volume (15-20 mg/L) than traditional IPTG induction methods. Protein expression of Mtr4 was executed in 2.5 L baffled Erlenmeyer flasks. A volume of 500 ml of autoinduction media prepared with appropriate antibiotics for protein expression (chloramphenicol and ampicillin) was added to each flask. Growths were inoculated with a small amount (~10 μl) of frozen glycerol stock prepared following protocols described in Studier et al. Flasks were allowed to grow at 37°C at ~350 rpm in
a standing shaker for 5-6 hours, after which the flasks were moved to a room temperature floor shaker and set to a speed of 230 rpm. The room temperature shaker has a greater radius (2 inches) and provided more aeration for the cells. Robust Mtr4 expression was observed when cells were grown for at least 24 hours after transfer to room temperature. Often cells were grown up to 36 hours after transfer and no adverse effects in protein solubility or activity were observed. The growth at room temperature allowed for greater Mtr4 solubility than that observed for growths performed at 37°C. Cells were harvested by centrifugation with a Sorval centrifuge and a SLC-4000 rotor at 8000 rpm and were stored at -80°C until cell lysis.

Cell lysis was performed by first manually breaking up the frozen cell pellet with a stainless steel spatula followed by incubation in lysis buffer (Table 2.1) to which lysozyme, 10 mM imidazole and protease inhibitors were previously added. The final concentration of protease inhibitors in the lysis solution was 2X of aprotinin (1 µg/ml), pepstatin (1.4 µg/ml) and leupeptin (1 µg/ml) and 1X concentration of PMSF (25 µg/ml). Lysozyme was added to obtain a final concentration of 100 mg of protein for every 30 ml of lysis buffer. Lysis buffer was added at a ratio of 2 mls of buffer for every 1 gram of cell pellet. The lysis buffer and crushed pellet were mixed thoroughly with the spatula. Cells were incubated in lysozyme buffer until the solution became very viscous. The solution was then sonicated on ice at a power of 6 and a 80% pulse of 30 seconds with 30 second intervals. The sonicated cell pellet solution was clarified by centrifugation at 20,000 rpms for 30 minutes.

The clarified lysate was incubated on Qiagen or Fisher brand Ni-affinity resin for
at least one hour after which bound Mtr4 was washed and eluted off the resin. Each Ni wash solution contained a concentration of 10 mM of imidazole. The wash volumes were 500 mls of lysis buffer, 100 mls of buffer B and 250 mls of buffer A (see Table 2.1). Mtr4 protein was eluted in 200 ml of Buffer A with 250 mM of imidazole and 100 ml of Buffer A with 500 mM imidazole. All lysis, wash, elution and purification buffers contained 5% glycerol, 50 mM HEPES pH 7.5, and 2 mM beta-mercaptoethanol. Salt concentration varied in the buffers used in the chromatography procedures and are listed in Table 2.1. Following elution off of Ni resin, SDS-PAGE was performed to identify the abundance of the protein in each fraction. Fisher resin commonly binds a contaminating protein that runs at about 25 kDa on SDS PAGE and Qiagen resin binds a 70 kDa contaminating protein. Incubation of the clarified lysate on Ni resin for more than 3 hours resulted in a large quantity of these and other contaminating proteins sticking even after wash steps of substantial volume. By minimizing the incubation time on Ni resin to 1 hour a large portion of the observed contamination was avoided.

Further purification with Heparin affinity, Q and Gel Filtration chromatography is required after Ni-affinity to remove several contaminating proteins away from Mtr4 in

**Table 2.1** Buffers used for purification of Mtr4 protein

<table>
<thead>
<tr>
<th>Buffer name</th>
<th>Lysis Buffer</th>
<th>Buffer A</th>
<th>Buffer B</th>
<th>Buffer Q</th>
<th>Sizing Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycerol</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
</tr>
<tr>
<td>HEPES pH 7.5</td>
<td>50 mM</td>
<td>50 mM</td>
<td>50 mM</td>
<td>50 mM</td>
<td>50 mM</td>
</tr>
<tr>
<td>β-ME</td>
<td>2 mM</td>
<td>2 mM</td>
<td>2 mM</td>
<td>2 mM</td>
<td>2 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>500 mM</td>
<td>160 mM</td>
<td>1 M</td>
<td>50 mM</td>
<td>160 mM</td>
</tr>
<tr>
<td>Volume used</td>
<td>550 mL</td>
<td>1 L</td>
<td>500 mL</td>
<td>500 mL</td>
<td>1 L</td>
</tr>
</tbody>
</table>
preparation for crystallization and biochemical analysis. The Ni resin elutions were loaded directly onto a 5 ml heparin column that was previously prepared by washing with 50 ml of Buffer B and 50 ml of Buffer A. Several contaminating proteins including those that are commonly pulled down by Qiagen and Fisher Ni resins flowed directly through the heparin resin. The column was washed with Buffer A until the 280 absorbance stayed flat for at least 5 mls. A 100 ml gradient from 0 to 100% buffer B was performed using an AKTA purifier or AKTA prime chromatography system. Mtr4 eluted at approximately 30 mS/cm. The peak is fairly broad and requires concentrating before the use of a gel filtration column.

In an effort to concentrate the protein, the 2.5 ml collected fractions from the heparin elution peak were pooled and desalted into Buffer Q using a desalting column. Buffer Q contains 50 mM NaCl allowing Mtr4 to associate with the Q resin. The desalted Mtr4 was then loaded onto a 5 ml Q column with Buffer Q and then eluted in a steep 35 ml Buffer B gradient of 0 to 100%. Mtr4 quickly elutes off the Q column, effectively concentrating the protein into about 8-12 mls.

The 2 ml fractions of the Q column were pooled and loaded into an injection loop of the AKTA prime/purifier system to which a 320 ml Superdex 200 26/60 column is attached. The gel filtration column was washed with one column volume of Sizing buffer and then the concentrated Mtr4 was loaded onto the column. The protein elutes off the Superdex 26/60 column at about 180 mls after injection. The final buffer conditions of purified protein were those described above plus 160 mM NaCl. The protein was then pooled and concentrated to at least 10 mg/ml. For biochemical assays and/or
crystallization trials Mtr4 can be concentrated to as much as 80 mg/ml. The Nanodrop (Thermofisher) spectrophotometer was used to determine the final concentration of the protein using an extinction coefficient calculated using the ProtParam software package found on the ExPasy webpage.

Purified Mtr4 protein was stored by first diluting the concentrated protein 1:1 in sizing buffer that contained 55% glycerol, bringing the final glycerol concentration up to 30% before freezing. Mtr4 protein in 30% glycerol was frozen for storage by pipetting 20 µl beads onto liquid Nitrogen. After freezing the beads are collected and stored at -80°C.

To purify the Mtr4 protein in roughly 24 hours, the buffers need to be made (Table 2.1) and Ni resin regenerated the previous day. It takes approximately 12 hours to get the protein prep from cell lysis to the injection loop of the sizing column. Time can be saved by washing the Heparin and Q columns either the day before the prep or during the centrifugation of the lysate or 1 hour incubation on Ni resin. It is very helpful to have access to two AKTA systems, primarily for preparation of the desalting column while the protein is eluting off of the Heparin column. The desalting and subsequent concentration of Mtr4 on the Q column make the 12 hour period from lysis to sizing injection possible. Mtr4 has also been successfully prepved with full activity in a 48 hour period. This timeline usually includes a break after elution off of the Heparin column. The protein is stored at 4°C until the following day when the prep can be continued.

Mtr4 Crystallization

Crystallization of Mtr4 was performed using standard vapor diffusion methods.
The crystals that allowed for structure solution were grown in sitting drop trays at 4°C in 2.4 M ammonium phosphate dibasic, and 0.5 M Tris-base pH 8.5 at a 1:2 protein:well drop ratio, using 10 mg/ml his-tagged Mtr4 protein. Although TEV-cleaved protein is amenable to crystallization in the same conditions, the crystals grow in a distorted manner and there was no improvement in the diffraction of the crystals. In order to provide phasing data for structure solution, selenomethionine-substituted (Se-met) Mtr4 was expressed in E. coli using a modified autoinduction protocol (Studier, 2005). Se-met purification and crystallization was the same as described for the native protein.

Due to the modest diffraction of the original Mtr4 crystals, many efforts were made to improve the resolution of data through optimization of Mtr4 crystal growth. Several variables were probed including the addition of nucleotides and/or other additives, seeding strategies, growth temperature and 80/20 ratios with established robot screens. For an extensive description of all crystallization trials see the Jackson crystal log located in the Johnson lab. Mtr4 crystals grew in the presence of various additives to the original buffer, and it appears that addition of PEG causes the crystals to grow larger in size and may slightly improve the diffraction properties of the crystal. Although the efforts to improve diffraction were significant, by no means is the optimization complete and it is possible that further optimization could yield better diffracting crystals. It was discovered that Mtr4 will grow at room temperature and 13 °C when seeded from a preexisting crystal. To seed, a small Mtr4 crystal grown at 4°C was crushed in 10 μl of well solution buffer until no large crystal chunks could be identified using the microscope. This 10 μl sample was then diluted from 100-5000 fold in well solution
before addition to crystallization drops at a 0.5 μl volume.

Several other crystallization conditions were discovered in robot screens for Mtr4 that required the presence of DNA or RNA oligonucleotides. However, these crystals proved difficult to work with because some could not be reproduced and others did not diffract well or were difficult to place into cryo solution. Included in Table 2.2 are the three Mtr4 crystal hits that occurred in the presence of nucleic-acid that were reproduced in hand trays. The optimization of the crystallization parameters in the presence of RNA/DNA oligos was minimal and it is hoped that this table can be a reference for others.

<table>
<thead>
<tr>
<th>Table 2.2. Crystallization conditions of Mtr4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construct</strong></td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Mtr4 full</td>
</tr>
<tr>
<td>Loopless</td>
</tr>
<tr>
<td>Mtr4 full</td>
</tr>
<tr>
<td>Mtr4 full</td>
</tr>
<tr>
<td>Mtr4 full</td>
</tr>
<tr>
<td>Mtr4 Loopless</td>
</tr>
<tr>
<td>Δ 74-Mtr4 Loopless</td>
</tr>
<tr>
<td>Δ 74-Mtr4 Loopless</td>
</tr>
</tbody>
</table>

*The RNA was prepared by adding nuclease free water to an oligo purchased from IDT, all other oligos were resuspended in Mtr4 concentrator flow through.*
seeking to pursue co-crystal structures of Mtr4 with nucleic-acid. The poly(A) oligos used in these trials were purchased from IDT. The 37 nt DNA oligo that forms a hairpin structure meant to mimic the acceptor stem of a tRNA was also purchased from IDT, however the RNA version of the tRNA hairpin was made using the large scale in vitro transcription reaction (described below). The DNA sequence used was 5’-AGCGCCGCTCGGAAAAACCGAGCGGCCTACCA AAA A-3’ (the portions of the sequence that hydrogen bonds to form the hairpin are underlined). The RNA hairpin was very similar with the sequence 5-GCGCCGCUCGGAAAAACCGAGCGGCGCUACCA AAA A-3’.

The small oligos were prepared by resuspending precipitated samples with the flow-through from Mtr4 post-gel filtration concentration. Samples were usually incubated at a 1:1.5 protein:nucleic-acid ratio. Larger oligos such as tRNA\textsubscript{met} were produced using the in vitro transcription reaction described below. Although several crystallization trials have been performed with tRNA\textsubscript{met}, a condition that produces crystal growth with an Mtr4 construct has yet to be discovered.

**Cryosolvent and cryocooling of Mtr4 crystals**

Two forms of cryosolvents were commonly used for freezing Mtr4 apo crystals. The first cryosolvent included glycerol and the second dextrose. Mtr4 crystals were stepped up into well solution with 20% glycerol in increments of 5% glycerol with sitting times of 30 seconds to 2 minutes. Direct plunging of crystals into well solution with 20% glycerol did not impact diffraction for better or worse. Thus, crystals were commonly moved directly into cryosolvent that contained 20% glycerol, 2.4 M ammonium hydrogen
phosphate dibasic, 0.5 M Tris pH 8.5. The second method for creating cryo-solvent was used to avoid the production of phosphate crystals that occurred when buffer containing ammonium phosphate in the presence of glycerol. This method is found on the Hampton Research website and is executed by adding 7.5 mg and 15 mg of dextrose to 50 µl aliquots of well solution. This effectively creates a 15% and 30% w/v dextrose solution. The crystals are stepped up first into the 15% and then into the 30% solutions to obtain cryosolvency. X-ray diffraction of Mtr4 apo crystals at room temperature demonstrated improved diffraction over those frozen in cryo suggesting that there is room for optimization. However, the improvement in resolution was not substantial.

**Data Collection, Structure Determination, Modeling, and Refinement of Mtr4**

The methodologies used to collect the crystallographic data, solve the structure, model atom positions and refine the model of Mtr4 are described in thorough detail in Chapter 3 of this dissertation.

**Sequence alignment and conservation analysis**

The ConSurf server was used to determine sequence conservation in a manner that considered phylogenetic relationships into the calculation. Sequence alignments and conservation scores shown in Chapters 3 and 4 were calculated with ConSurf (Landau et al., 2005). In Chapter 3, the arch sequence from *S. cerevisiae* Mtr4 was used as a query and 108 unique Mtr4 sequences were identified for analysis. The conservation scoring observed in Figures 3-3, 3-5 and 3-6 come from this analysis. A similar search with the putative arch sequence from *S. cerevisiae* Ski2 retrieved 42 unique sequences from which
the color scheme observed for Ski2 conservation in Figures 3-3 and 3-6 was derived.

In Chapter 3 and Chapter 4 alignments were also made with model species using Clustal W. The program ESPRIPT (Gouet et al., 1999) was used to calculate conservation scoring of the full length alignment of Mtr4 and Ski2 sequences from model species displayed in Appendix A and referred to in Chapter 3. The Conservation scores calculated by Consurf in Chapter 4 were determined from Clustal W alignments of sequences of Mtr4 and other Ski2-like helicases from model species (Larkin et al., 2007).

**Biochemical analysis of Mtr4**

The ATPase, RNA unwinding and RNA binding activities of Mtr4 and Mtr4 mutants were analyzed using methodologies previously developed to probe Mtr4 and TRAMP function (Bernstein et al., 2008; Bernstein et al., 2010; Jia et al., 2012; Wang et al., 2008). In Chapter 3 modified versions of the colorometric malachite green assay described in Bernstein et al (Bernstein et al., 2008) and the unwinding described in Wang et al (Wang et al., 2008) were used to show that removal of the arch domain did not severely impact the fundamental enzymatic activities of Mtr4. Chapter 4 probes the activities of Mtr4-ratchet helix mutants with buffers and substrates that were optimized by the Jankowsky lab (Jia et al., 2012) with an emphasis on unwinding activity. Slight modifications were made to the unwinding buffer from a 40 mM NaH$_2$PO$_4$ (pH 6.0) buffer system to a 40mM MOPS (pH 6.5). Salt concentrations were modified slightly between the two versions of the assay. In that the primary analysis of Chapter 4 was to determine the effect of substrate on Mtr4 activity the same buffer used for unwinding was also used in the malachite green and RNA binding assays. The full details regarding
buffers and substrate concentrations are described in methods sections of Chapters 3 and 4.

*Preparation of RNA substrates for unwinding assays*

For the initial unwinding assays, the 40 mer RNAs used in Chapter 3 for 3’ overhangs were *in vitro* transcribed using the Ampliscribe T7 RNA Transcription Kit (Epicentre). The 16 nt RNA that was displaced in all the unwinding assays as well as the 22 oligomer RNAs used for overhangs in Chapter 4 were ordered from IDT. The 16 nt RNA was radiolabeled using ATP, [γ32P]-6000Ci/mmol and Kinase Max 5’ end labeling kit (Ambion). The final 5’ labeling reaction contained 100 pmols of RNA, 1X Kinase Buffer (Ambion), 1U/μl of ribolock (Fermentas), 1 μl of T4 kinase and 25 pmols of [γ32P]-6000Ci/mmol ATP. The reaction was incubated for 2-5 hours at 37°C after which the kinase was inactivated by heating the reaction to 95°C for 10 minutes in a sand heat block. To produce labeled duplex RNA the solution was spun down and the overhang RNA strand was added at a 1.6:1 (overhang:top strand) ratio. The solution was again heated to 95°C for 5 minutes after which it was allowed to cool to room temperature. Duplex RNA was successfully made by simply allowing the solution to cool on the bench top. However, slower cooling method was also commonly employed in which the temperature to the sand heat block was turned off and the solution was spun down every 10 minutes in the centrifuge until the temperature reached 60°C at which point the reaction was moved to the bench top.

To purify the duplex RNA away from the excess radiolabeled ATP and reaction components, the annealing reaction was combined with helicase loading dye and loaded
onto a native 15% 1X TBE polyacrylamide gel. The gel was run at 100 V for 60-90 minutes. The gel was then removed wrapped in cellophane and exposed to x-ray film to determine the location of the radiolabeled duplex RNA. The film was developed in the dark room using developer and fixer purchased from Sigma. Using the developed film as a guide, the RNA was cut out of the gel as a polyacrylamide slab and placed in 300 µl of extraction buffer (0.5 M ammonium acetate, 0.1 mM EDTA, 0.1% SDS) for 24 hours. The RNA that diffused into the ammonium acetate buffer was precipitated by adding 3 volumes of ice cold ethanol and 3 µl of 20 mg/ml glycogen. The solution was incubated for 15–30 minutes at −80°C followed by centrifugation at 4°C for ~15 minutes. The supernatant was removed and the pellet of precipitated RNA was allowed to dry. RNA was then re-suspended in an appropriate volume of autoclaved nuclease-free water.

**In vitro transcription of RNA**

A large scale *in vitro* transcription reaction was developed to make structured RNAs like tRNA and rRNA at the quantities required for crystallization trials and biochemical analysis. The reaction was also used to produce radiolabeled RNA substrates with [α-³²P] nucleotides. PCR templates were generated using PFU polymerase and associated made in the Johnson lab (protocols are located in the Johnson lab). T7 RNA polymerase was expressed and purified in the Johnson lab (described below). All NTPs were made as individual stock solutions of 100 µM for which pH was adjusted to 7.4 and filtered before use. Transcription buffer was modified from those previously described (Batey and Kieft, 2007; Kieft and Batey, 2004) and the 10X buffer consisted of 400 mM Tris pH 8.0, 240 mM MgCl₂, 10 mM Spermidine, 50 mM DTT and 0.1% Triton
X-100. The final concentrations for a 1 ml transcription reaction were transcription buffer 1X, PCR product 8% total volume, NTPs 4 mM each, GMP 16 mM and T7 polymerase. Reactions were incubated at 42°C in a water bath for 3 hours. The transcribed RNA was then purified using polyacrylamide gel extraction techniques or gel filtration.

**T7 polymerase protein expression and purification**

The PAR1219 plasmid (Sigma) was transformed into BL21 DE3 codon plus cells from which glycerol stocks were made. Autoinduction media (Studier, 2005) volumes of 500 mL were grown in baffled 2.5L flasks at 37°C in a standing shaker at 350 rpm for 24 hours after which flasks were moved to a room temperature at 230 rpm for 5 more hours. Cells were pelleted at 5,000 rpm in the Sorval SLC-4000 rotor and frozen at -80°C until lysis.

The T7 polymerase purification procedure was adapted from an existing protocol (Li et al., 1999). Protease inhibitors and lysozyme was added to T7 lysis buffer in the same manner described for the lysis buffer of Mtr4. Likewise, the manual breaking of the pellet, incubation with lysis solution, subsequent sonication and clarification were all performed as was done for Mtr4. A second centrifugation step was performed with the clarified lysate using the ultra-centrifuge at 45,000 rpm for 1 hour. The ultra-centrifuge lysate was then filtered using a 0.45 micron filter and was loaded onto a 5 ml Q column in tandem with a Heparin column that was previously equilibrated with 50 ml of T7 buffer II followed by 50 ml of T7 buffer I (see Table 2.3). After loading, the columns were washed with 500 mLs of T7 buffer I. Following the wash step the Q column was unhooked and a 100 mL NaCl gradient from 0 to 100% T7 buffer II eluted the protein off
the column as a broad tall peak. The fractions containing the T7 polymerase protein were concentrated to a total volume of about 10 mls and were loaded into an injection loop of a superdex 200 26/60 column that was equilibrated with T7 sizing buffer. The T7 polymerase came off in a nice sharp peak. The polymerase fractions were pooled and concentrated to 3.2 mg/ml of protein. The protein was prepared for storage by mixing concentrated T7 polymerase at a 1:1 ratio with T7 sizing buffer that contained 55% glycerol, making the final concentration of T7 polymerase 1.6 mg/ml in T7 sizing buffer that contained 30% glycerol. The protein was flash frozen by pipetting 20 µl beads into liquid nitrogen. Beads were gathered and stored at -80°C.

The activity of T7 polymerase was determined for several steps of the purification process. The concentration of the pooled fractions of post-heparin, concentrated post-heparin, concentrated post-sizing and the flash frozen post-sizing T7 polymerase was determined with the nanodrop spectrophotometer (ThermoFisher) followed by in vitro transcription of tRNA1met. The concentration of protein varied at each purification step with post-heparin as 4 mg/ml, concentrated post-heparin 8 mg/ml, concentrated post-sizing 3.2 mg/ml and flash frozen post-sizing as 1.6 mg/ml. Some protein was lost between the heparin and sizing columns (~ 40%). However it was difficult to observe whether this was due to the loss of contaminating proteins or if the polymerase itself was lost. Activity tests of each protein fraction showed robust polymerization with little to no differences in activity between each fraction.
Table 2.3 T7 purification buffers

<table>
<thead>
<tr>
<th>Buffer</th>
<th>T7 Lysis buffer</th>
<th>T7 buffer I</th>
<th>T7 buffer II</th>
<th>T7 sizing buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycerol</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
</tr>
<tr>
<td>K PO4 pH 8.0</td>
<td>10 mM</td>
<td>10 mM</td>
<td>10 mM</td>
<td>10 mM</td>
</tr>
<tr>
<td>EDTA</td>
<td>10 mM</td>
<td>1 mM</td>
<td>1 mM</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>DTT</td>
<td>1 mM</td>
<td>----</td>
<td>----</td>
<td>1 mM</td>
</tr>
<tr>
<td>β-ME</td>
<td>10 mM</td>
<td>10 mM</td>
<td>10 mM</td>
<td>----</td>
</tr>
<tr>
<td>NaCl</td>
<td>50 mM</td>
<td>50 mM</td>
<td>1 M</td>
<td>100 mM</td>
</tr>
<tr>
<td>Volume</td>
<td>100 mL</td>
<td>1 L</td>
<td>1 L</td>
<td>1 L</td>
</tr>
</tbody>
</table>

Structure solution of E.coli ACTase using PtCl₄ for heavy atom derivatization

In an attempt to determine the x-ray crystal structure of the N-terminus of Ski2, the cytosolic homolog of Mtr4, residues 1-313 of Ski2 were cloned, expressed and purified. The purified protein produced small cubic crystals that were difficult to reproduce (Figure 2-1). Seeding techniques allowed for the reproduction of the small crystals.

A native data set to 2.9 Å was collected using a home source x-ray generator and diffraction collector (see Table 2.4). However, no phasing information was available for structure solution. To produce heavy atom derivatization, a crystal was soaked for 10 minutes in well solution containing 10 mM of potassium tetrachloroplatinate (III) followed by back soaking in well solution for approximately 30 seconds before flash freezing (Joyce et al., 2010; Sun and Radaev, 2002; Sun et al., 2002). The soaked crystal and apo crystals were shipped to the SSRL (Stanford Synchrotron Radiation Lightsource)

---

1 This study was performed in collaboration with Amy Crandall
for data collection on a tunable x-ray beamline. Using the web browser application Web-Ice a fluorescence scan on the derivatized crystal verified the presence of the anomalous scattering Pt in the crystal with an anomalous peak wavelength of 1.07202 Å and a remote wavelength of 0.88557 Å. Full data sets of the derivatized crystal at the remote and peak wavelengths were collected at SSRL. A native data set was also collected at the peak wavelength that diffracted to 2.3 Å.

The indexing, integration and scaling of the data sets were performed using HKL 2000 software (Otwinowski and Minor, 1997). The crystals belong to space group R32
Figure 2-1. Structure determination of ATCase. (A) An image of a small cubic crystal of ATCase is shown. (B) Displayed is a Pt atom coordinated by two cysteines within the 3.5 Å maps derived using SAD phasing strategies. (C) The 3.5 Å maps (left) determined with SAD phasing and PtCl₄ are compared to the 2.3 Å maps (right) determined by molecular replacement with the structure of ATCase (PDB 2HSE). (D) The refined model of ATCase is shown as a twelve subunit complex from the top (left) and side (right) views. Chain A is colored blue and chain B is colored in cyan.

(a=127.3 Å, b=127.3 Å, c=197.7 Å; α=90°, β=90°, γ=120°). To obtain a workable data set, ice rings were removed from the heavy-atom soaked crystal diffraction data by increasing the REJECT fraction value until ice ring spots were no longer picked up by the software. Also it was required that the beam stop shadow be masked out of the data. Anamolous signal was detected using HKL2000 by first scaling with the ANOMALOUS flag on. Next the .sca file was read in, but the ANAMOLOUS flag was turned off.
Anomalous differences between the I+ and I- reflections are detected as $X^2$ values above 1.0, however these differences can also be attributed to noise. Resolution shells that had a $X^2 \geq 2$ were used for the purpose of calculating an anomalous difference Patterson. In the case of ATCase the resolution limit was 3.5 Å.

The AutoSol Wizard of the PHENIX software package was used to identify four Pt atom positions, and generate experimental phases and density modified maps (Adams et al., 2002; Adams et al., 2010; Terwilliger et al., 2009; Zwart et al., 2005). The AutoSol Wizard uses several programs that include HYSS (Grosse-Kunstleve and Adams, 2003), Phaser (McCoy et al., 2007) and RESOLVE (Terwilliger, 2000). The initial model was built into the density maps using the Coot software package (Emsley and Cowtan, 2004). Once an initial trace was built into the electron density, a DALI search (Holm and Rosenstrom, 2010) of the model discovered that the protein structure aligned identically with E. coli ATCase. Using the PHENIX AutoMR package (Zwart et al., 2008), the higher resolution data set and the pdb file 2HSE as a search model maps for ATCase were calculated to 2.3 Å resolution. A model was built using Coot and refinements were performed in PHENIX (Figure 2-1).

**Conclusion**

This chapter provides future researchers with the methodologies employed to probe Mtr4 Structure and function. Inclusive is the molecular cloning, protein expression and crystallography used to obtain Mtr4 crystals that eventually led to the structure solution of the protein. Also contained herein is a description of the biochemical assays and optimization of RNA *in vitro* transcription techniques developed throughout the
course of this work. Lastly, is a narrative of the techniques and programs utilized to determine the structure of ATCase. These methods provide a record of the work that has been accomplished and a comprehensive reference for further structural and biochemical studies with enzymes that act on RNA.

References


CHAPTER 3

THE CRYSTAL STRUCTURE OF MTR4 REVEALS A NOVEL ARCH DOMAIN REQUIRED FOR rRNA PROCESSING$^{1,2}$

Abstract

The essential RNA helicase Mtr4 performs a critical role in RNA processing and degradation as an activator of the nuclear exosome. The molecular basis for this vital function is not understood and detailed analysis is significantly limited by the lack of structural data. Here we present the crystal structure of Mtr4. The structure reveals a novel arch-like domain that is specific to Mtr4 and Ski2 (the cytosolic homolog of Mtr4). In vivo and in vitro analyses demonstrate that the Mtr4 arch domain is required for proper 5.8 S rRNA processing, and suggest that the arch functions independently of canonical helicase activity. Additionally, extensive conservation along the face of the putative RNA exit site highlights a potential interface with the exosome. These studies provide a molecular framework for understanding fundamental aspects of helicase function in exosome activation, and more broadly define the molecular architecture of Ski2-like helicases.

Introduction

RNA surveillance is a fundamental quality control process that prevents aberrantly transcribed or modified RNAs, poorly assembled ribonucleoprotein complexes

$^{1}$Coauthored by Ryan N. Jackson, A. Alejandra Klauer, Bradley J. Hintze, Howard Robinson, Ambro van Hoof and Sean J. Johnson
$^{2}$ EMBO J. (2010) 29, 2205-16
(RNPs), and unneeded RNAs from interfering with normal cellular gene expression and regulation (Houseley and Tollervey, 2009; Lebreton and Seraphin, 2008). Defects in RNA processing have been linked to a variety of disease states, including neurodegenerative diseases and cancer (Bruserud, 2007; Nelson and Keller, 2007). A critical component of nuclear RNA surveillance is the nuclear exosome complex, which contains two exoribonuclease active sites and an endoribonuclease site (Allmang et al., 1999b; Anderson and Wang, 2009; Dziembowski et al., 2007; Houseley and Tollervey, 2009; Jensen and Moore, 2005; Lacava et al., 2005; Lebreton and Seraphin, 2008; Lebreton et al., 2008; Liu et al., 2006; Lykke-Andersen et al., 2009; Schaeffer et al., 2009; Schneider et al., 2009; Vanacova et al., 2005; Wyers et al., 2005).

Although the exosome contains the catalytic activities for RNA processing and degradation, its in vivo activity requires many cofactors, including Mtr4. Mtr4 (for mRNA transport; also known as Dob1) was identified in a genetic screen for yeast mutants that accumulated polyadenylated RNAs in the nucleus (Liang et al., 1996), and was subsequently found to affect rRNA processing and other exosome functions (de la Cruz et al., 1998). Mtr4 may affect some exosome functions in isolation, but is also the largest component of the TRAMP complex, a three protein complex containing a poly(A) polymerase (Trf4 or Trf5), a putative RNA binding protein (Air1 or Air2), and Mtr4 (Anderson and Wang, 2009; Jensen and Moore, 2005; Lacava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005). TRAMP identifies RNA substrates and adds a short poly(A) tail to the 3’ end, thereby promoting their degradation (Anderson and Wang, 2009; Lacava et al., 2005). TRAMP substrates are wide-ranging and include tRNAs,
snoRNAs, snRNAs, ncRNAs, rRNAs, some mRNAs, and cryptic unstable transcripts (Allmang et al., 1999a; Buhler et al., 2007; Houseley et al., 2007; Houseley et al., 2008; Milligan et al., 2005; Reis and Campbell, 2007; van Hoof et al., 2000; Wang et al., 2008). Mtr4 and the exosome are also involved in chromatin remodeling (Houseley et al., 2007; Houseley et al., 2008; Reis and Campbell, 2007; San Paolo et al., 2009) and in normal processing of various RNAs that require 3’ end trimming to reach a mature state (e.g. 5.8S rRNA, snoRNAs, snRNAs, and some mRNAs) (Allmang et al., 1999a; de la Cruz et al., 1998; van Hoof et al., 2000). A homolog of Mtr4, Ski2, is also involved in activation of the exosome, but is located in the cytoplasm and acts primarily on mRNA transcripts (Lebreton and Seraphin, 2008). Mtr4 and Ski2 homologs have been identified in a wide variety of eukaryotes including humans (Anderson and Wang, 2009; Houseley and Tollervey, 2009), indicating that their roles in RNA surveillance, processing and decay are conserved.

Mtr4 is essential for yeast growth (Liang et al., 1996) and provides a critical link between polyadenylation of RNA substrates by Trf4 and degradation by the exosome. A point mutant that disrupts helicase activity of Mtr4 results in the accumulation of polyadenylated RNAs that are no longer removed by the RNA degradation machinery (Wang et al., 2008). In vitro analysis indicates that Mtr4 has RNA dependent ATPase and helicase activity (Bernstein et al., 2008; Wang et al., 2008). The role of Mtr4 in exosome-mediated RNA decay presumably includes unwinding of RNA secondary structure and/or displacement of proteins associated with RNP complexes in order to present a “clean” substrate to the exosome (Houseley and Tollervey, 2009; Lebreton and Seraphin, 2008).
Additionally, it has been suggested that the ATPase activity of Mtr4 and Ski2 may be used to feed RNA substrates into the ring-like structure of the exosome, analogous to ATPases associated with the proteasome (Lorentzen and Conti, 2006; van Hoof and Parker, 1999).

Helicases are ubiquitous and diverse enzymes that are broadly classified into families based on a distinct set of sequence motifs localized in core helicase domains, which are generally the site of nucleic acid and nucleotide binding (Cordin et al., 2006; Jankowsky and Fairman, 2007; Pyle, 2008; Singleton et al., 2007). Mtr4 is designated as a superfamily II RNA helicase belonging to the Ski2-like family of DExH/D RNA helicases (de la Cruz et al., 1999). The most closely related member of this family is the Ski2 protein (38 % identity), from which the family derives its name. More distantly related Ski2-like RNA helicases include Brr2 (RNA splicing) and Slh1 (translation initiation) (Pena et al., 2009).

A molecular understanding of Mtr4 function is significantly limited because no structures are currently available for Mtr4 or Ski2. The only available structures of a Ski2-like RNA helicase are C-terminal fragments of Brr2 (Pena et al., 2009; Zhang et al., 2009), which have limited sequence similarity to Mtr4. Although Ski2-like DNA helicase structures have been characterized more extensively (Buttner et al., 2007; Oyama et al., 2009; Richards et al., 2008; Zhang et al., 2008), the 122 kDa Mtr4 is considerably larger than these DNA helicases (e.g. Hel308 is 79 kDa) and the similarity between these proteins is not apparent beyond the core helicase domains. Only three domains are identifiable in the Mtr4 sequence: the two core helicase domains and an additional C-
terminal domain, designated DSHCT (Dob1/Ski2/HelY C-terminal domain) (Staub et al., 2004), which has no obvious sequence similarity to domains of known structure or function.

Here we report the 3.4 Å crystal structure of full-length Mtr4 from *Saccharomyces cerevisiae*. This structure provides the first view of an exosome activating cofactor, and also represents the first full-length description of a Ski2-like RNA helicase. The structure clarifies which molecular features are shared throughout the Ski2-like helicase family, and highlights features that are unique to Mtr4 and Ski2. Surprisingly, the Mtr4 structure reveals a prominent domain that adopts a novel arch-like appearance. Removal of the arch domain produces a defect in rRNA processing that resembles inactivation of the Rrp6 exonuclease component of the exosome. The data presented here suggest that Mtr4 plays an important role in regulation of exosome activity by employing some mechanisms that are general to the Ski2-like family of RNA helicases and others that are unique to Mtr4 and Ski2.

**Results and discussion**

**Structure determination**

The full length Mtr4 protein from *S. cerevisiae* was expressed in *Escherichia coli* using a construct containing a 6-histidine affinity tag at the N-terminus. The recombinant protein was purified by Ni-chelate, heparin affinity and gel filtration chromatography. Crystallization conditions were identified and optimized to obtain a 3.4 Å native data set and 3.6 Å selenomethionine (Se-met) substituted data set. Since no effective molecular
replacement model was available, the Se-met data was used to solve the phase problem using single-wavelength anomalous dispersion (SAD) methods. The initial model was built using the SAD maps and subsequently refined against the 3.4 Å native data to a final R/R\text{free} of 29.3%/33.3% (Table 3.1).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
\textbf{Data collection} & Mtr4 Se-met & Mtr4 Native \\
\hline
Space group & \textit{P}3\textsubscript{1}21 & \textit{P}3\textsubscript{1}21 \\
Cell dimensions & \begin{tabular}{c}
\textit{a}, \textit{b}, \textit{c} (Å) \\
\textit{\alpha}, \textit{\beta}, \textit{\gamma} (deg)
\end{tabular} & \begin{tabular}{c}
133.4, 133.4, 192.1 \\
90.0, 90.0, 120.0
\end{tabular} \\
& \begin{tabular}{c}
133.5, 133.5, 190.9 \\
90.0, 90.0, 120.0
\end{tabular} \\
Resolution (Å) & \begin{tabular}{c}
30-3.6 (3.73-3.60)\textsuperscript{a} \\
0.120 (0.689)
\end{tabular} & \begin{tabular}{c}
30-3.4 (3.52-3.40) \\
0.059 (0.494)
\end{tabular} \\
\hline
\textbf{Refinement} & & \\
Resolution (Å) & 30-3.4 & \\
No. reflections & 26,097 & \\
\textit{R}_{\text{work}} / \textit{R}_{\text{free}} & 0.293 / 0.333 & \\
No. atoms & \begin{tabular}{c}
Protein \\
Ligand/ion \\
Water
\end{tabular} & \begin{tabular}{c}
6487 \\
20 \\
0
\end{tabular} \\
\textit{B}-factors & \begin{tabular}{c}
Protein \\
Ligand/ion
\end{tabular} & \begin{tabular}{c}
157.9 \\
171.2
\end{tabular} \\
R.m.s deviations & \begin{tabular}{c}
Bond lengths (Å) \\
Bond angles (deg)
\end{tabular} & \begin{tabular}{c}
0.011 \\
1.530
\end{tabular} \\
\hline
\end{tabular}
\caption{Data collection, phasing and refinement statistics (SAD)}
\end{table}

\textsuperscript{a}Values in parentheses are for highest-resolution shell.

One crystal was used for each data set.
The electron density maps are quite clean and continuous, considering the modest resolution of the data (Figure 3-1). Gaps in the current model include 85 residues at the N-terminus and 11 loops of varying length for which no interpretable electron density is observed (Figure 3-1D). The anomalous signal from selenomethionine sites, which constitute ~3% of the Mtr4 sequence and are relatively well distributed throughout the protein, was used to assign amino acid sequence. Alignments of the Hel308 (Buttner et al., 2007; Richards et al., 2008) and Hjm (Oyama et al., 2009) DNA helicase structures to the Mtr4 structure were used to further clarify sequence assignment and connectivity through portions of the structure. Consequently, we were able to confidently model 85% (918 residues) of the Mtr4 sequence as well as four phosphate molecules into the electron density.

**Overall Description**

The crystal structure reveals that Mtr4 is composed of five distinct domains (Figure 3-1). The base of the structure contains four domains (domains 1, 2, 3 and 4) assembled in a circular manner, with a channel of ~12 Å diameter passing through the center. In addition, a fifth domain is a large (265 residue) novel structure that arches over the other four domains. This domain is inserted into the middle of domain 3 and is therefore named domain 3a, or the “arch” domain. A brief description of each of these domains follows.

*Domains 1 and 2 are the core helicase domains.* Domains 1 and 2 are canonical RecA-like domains that are each composed of a central β-sheet surrounded by α-helices (Figure 3-2). As observed in other helicase structures (Cordin et al., 2006), eight
**Figure 3-1.** Structural overview of Mtr4. (A) Stereo view of the Mtr4 structure. (A) Cα trace is shown with rainbow coloring from N- (blue) to C-termini (red). Residue numbers are indicated. (B) Representative stereo view of the 2fσ-fσ electron density map, contoured at 1σ. The (*) corresponds to the same position marked in (A). (C) Ribbon and surface representations of the Mtr4 structure, colored by structural domains (see also Figure 3-2). (D) Mtr4 primary sequence and secondary structure. Helices and strands are numbered as indicated. Dashed lines denote gaps in the model due to lack of electron density. The first strand in domain 1 (β1) is not indicated, but presumably comes from part of the 85 residues at the N-terminus of Mtr4; the precise sequence could not be determined because no electron density is observed connecting it to the rest of the domain.
signature sequence motifs are located in a cleft at the interface of domains 1 and 2 (Figure 3-2A). These motifs are typically involved in nucleotide and nucleic acid binding, as well as hydrolysis of ATP. Domain 1 is an unusually extended beta structure composed of ten strands which, to our knowledge, is the longest sheet currently reported for superfamily 2 helicase structures (Figure 3-2B). Secondary structure predictions of other Mtr4 and Ski2 species suggest that an extended sheet is a common feature of this protein family. The structural or functional role of this extended sheet is unclear, but the additional strands do contribute to the surface area along the base of the structure. Like domain 1, the β-sheet of domain 2 is rather large, containing eight strands. Extending off the second strand (β2) of domain 1 is a loop and short helix that span the cleft and pack with domain 2.

*Domain 3 adopts a winged-helix fold.* An ordered 14 residue loop extends across the face of domain 1, connecting domain 2 to domain 3. Domain 3 adopts a winged helix-like fold composed of helices H21, H22 and H30, followed by a β-hairpin (Figure 3-2C). The fold is structurally similar to the winged helix domains of Hel308 (RMSD=1.97 Å) and Hjm (RMSD=2.04 Å). An additional helix (H31) is positioned C-terminal to the winged helix fold and forms part of the channel wall that runs through the base of the protein.

*Domain 4 resembles the seven helix bundle of Hel308.* The C-terminal domain of Mtr4 is an eight helix bundle comprising residues 912-1073, and corresponds to the previously annotated DSHCT domain (Staub et al., 2004). The Mtr4 structure is the first report of a DSHCT domain. Unexpectedly, the structure closely resembles the seven
helix bundle structures from Brr2 (Pena et al., 2009; Zhang et al., 2009), Hel308 (Buttner et al., 2007; Richards et al., 2008), and Hjm (Oyama et al., 2009). Indeed, the structural similarity between Mtr4 and Hel308 allowed us to identify connectivity between several helices in domain 4 that would have otherwise remained undetermined due to weak electron density in several loop regions.

Figure 3-2. Architecture of Domains 1, 2, and 3 in Mtr4. (A) Domains 1 (blue) and 2 (yellow) are oriented to show the central cleft containing eight signature sequence motifs. The sequence and position of each motif are shown below. (B) The β-sheets for domains 1 (left) and 2 (right) are shown. Domain 1 is a particularly extended sheet composed of 10 strands. (C) Domain 3 adopts a canonical winged helix fold. A large domain (3a or “arch”) is inserted between the second (H22) and third (H30) helices of the winged helix.
Domain 3a – the arch domain – is unique to Mtr4 and Ski2. The most striking and unexpected feature of the Mtr4 structure is a large, 265 residue domain that arches over the rest of the Mtr4 structure and accounts for approximately 25% of the Mtr4 protein (Figure 3-1 and Figure 3-3). This “arch” domain is inserted between the second and third helices of domain 3 (H22 and H30), and replaces the loop that is typically observed in other Ski2-like helicases. The arch domain is composed of two distinct features: (1) the arms and (2) the fist.

The “arm” (H23 and H29) and “forearm” (H24 and H28) are each composed of an ascending and descending helix that form antiparallel left handed coiled coils (Figure 3-3A). Two short loops at the “elbow” are located between the arm and the forearm, allowing the structure to make a sharp ~120° turn. This dramatic bend is primarily responsible for the arch-like appearance of the domain. Several conserved features are observed at the elbow. A proline (Pro839) facilitates formation of the bend, and similar residues are found in the same region throughout Mtr4 species, although the absolute position varies slightly (Figure 3-3A). Hydrophobic packing by several conserved residues (Val637, Val643, Tyr646, Leu840, Leu846 and Tyr853) appears to stabilize the bend. In addition, a potential salt bridge forms between the only invariant residues in this region: Glu640 and Lys856. The low resolution of the data precludes an absolute assignment of side chain conformation, but these residues are clearly positioned in a manner that would allow a salt bridge. Significantly, these residues are only brought into proximity because of the sharp bend at the elbow. A physical interaction between the two residues is currently the best explanation for the absolute conservation observed at these
Figure 3-3. Domain 3a – the “arch” domain. (A) Sequence alignment of the arch domain. Conservation was calculated for Mtr4 sequences using Consurf (Landau et al., 2005), based on 108 Mtr4 sequences (orange, strictly conserved residues; yellow, similar). Conservation was similarly calculated for 42 Ski2 sequences and is displayed below the Mtr4 sequences. Observed (Mtr4) and predicted (Ski2) secondary structure for the arch is indicated. The coloring of the Mtr4 secondary structure corresponds to the stereo figure of the arch shown above. Two types of conservation are observed: residues that play a role in maintaining the overall fold or structure of the arch (▼), and those that have a potential functional role (●). Proline residues that are predicted to influence the bend at the elbow are underlined. (S.cer, Saccharomyces cerevisiae; S.pom, Schizosaccharomyces pombe; N.cra, Neurospora crassa; H.sap, Homo sapiens; M.mus, Mus musculus; D.re, Danio rerio; D.mel, Drosophila melanogaster; C.ele, Caenorhabditis elegans; M.bre, Monosiga brevicollis; A.tha, Arabidopsis thaliana) (B) Comparison of the Mtr4 fist with the L14e ribosomal protein. Structures (top) and topology diagrams (bottom) are shown. Similar features between the structures are colored blue. The central fold of the L14e protein from S. solfataricus (PDB 2joy) is structurally similar to the fist of Mtr4 and was used to guide model building in the fist. A more extensive comparison is shown in Figure 3-4. (C) Domain arrangement of Ski2-like RNA and DNA helicases. Domains 1, 2, 3, and 4 are characteristic of all identified Ski2-like helicases. Sequence and secondary structure analysis indicates that the arch domain is unique to the Mtr4/Ski2 sub-family. The function of DDX60 as an RNA or DNA helicase has not been demonstrated.
sites. The accumulation of conserved features at the elbow suggests that the bend is a genuine feature of the arch domain that cannot be dismissed as a crystallographic artifact. The forearm extends into a globular α-β “fist” that contains a central β-sheet and sits directly above domain 2 (Figure 3-1 and Figure 3-3A). The core of the fist is structurally similar to the ribosomal protein L14e of *Sulfolobus solfataricus* (RMSD=2.72 Å) (Figure 3-3B and Figure 3-4). Both Mtr4 and the L14e protein of *S. solfataricus* adopt a five strand fold followed by a C-terminal helix. Substantial crystal packing is observed along helices H25 and H26 of the fist. In the absence of additional structural data, it is not clear what influence these packing interactions have on the overall fold of the arch domain, but the most likely impact is on the orientation of the fist.

It is important to note that the internal packing features observed in the arch (e.g. non-crystallographic packing) do not necessarily preclude motion at the elbow or other conformational rearrangements in the arch. The loops at the elbow may actually facilitate conformational flexibility. Indeed, given the proximity of the fist to the predicted path of RNA (see below), we anticipate conformational changes in the arch upon binding of substrates or protein co-factors. Rigid-body motions in the arch may also result from conformational changes in other domains (e.g. upon ATP hydrolysis). Such motions, if they occur, are likely to have significant mechanistic implications. Additional crystallographic and biophysical analysis is therefore needed to characterize the potential dynamics of the arch domain.

The arch domain appears to be a novel domain, both in terms of structure and sequence. Comparable arch-like domains have not been observed in other helicase
structures. DALI (Holm et al., 2008) and secondary structure matching (SSM) (Krissinel and Henrick, 2004) searches using the arch as a query structure failed to identify similar structures in the protein data bank (although some similarity exists between the fist of the arch and the L14e protein). Likewise, a BLAST (Altschul et al., 1990) search of the arch

Figure 3-4. Comparison of the Mtr4 fist with the L14e ribosomal protein. Structures (top) and topology diagrams (bottom) are shown. Similar features between the structures are colored blue. (A) The central fold of the L14e protein from S. solfataricus (PDB 2joy) is structurally similar to the fist of Mtr4 (B), and was used to guide model building in the fist. (C) The ribosome bound L14e structure from H. marismortui (PDB 1jj2) adopts a similar fold (Klein et al., 2001). Notably, the primary RNA binding face of L14e protein corresponds to the putative RNA binding interface of the Mtr4 fist.
sequence exclusively yields Mtr4 sequences. We note, however, that Ski2 contains a region of similar length (residues 830-1083), which also has limited similarity to other known sequences. Predicted secondary structure analysis indicates this region in Ski2 adopts a fold that resembles the Mtr4 arch (Figure 3-3A). While the sequence similarity within the arch domain of Mtr4 and Ski2 is low (~34%), it does include several residues that are critical for maintaining the overall structure in Mtr4 (Appendix A). We conclude that, despite the low sequence similarity, Mtr4 and Ski2 both contain an arch-like domain. The domain is not observed in other Ski2-like RNA or DNA helicases, and is therefore unique to the Mtr4/Ski2 subfamily of Ski2-like helicases (Figure 3-3C).

Molecular architecture of Ski2-like RNA helicases

The Mtr4 structure defines the fundamental molecular architecture of Ski2-like helicases. The structure clearly demonstrates that domains 3 and 4 are characteristic features of Ski2-like family members (Figure 3-3C). This observation is particularly significant because sequence analysis failed to identify either region of Mtr4 as similar to other Ski2-like proteins. Domain 3 was not recognized as a winged-helix due to the insertion of the 265 residue arch domain between the second and third helices (H22 and H30) of the motif (Figure 3-1C). However, the structural similarity of this domain to other Ski2-like helicases is strong. Similarly, the eight-helix bundle of the Mtr4 domain 4 is clearly related to that of the Brr2 and Hel308 structures (Figure 3-5C). While our manuscript was under revision, the structure of a DEAH-box RNA helicase, Prp43, was published (He et al., 2010), which also contains domains 1-4. This suggests that domains 1-4 may be a common feature in other helicases beyond the Ski2-like RNA and DNA
helicase family. Each subfamily of the Ski2-like helicases also retains unique architectural features. The most striking distinction between Mtr4/Ski2 and the other subfamilies is the presence of the large arch domain, which accounts for a quarter of the entire Mtr4 protein.

**Model for RNA Binding**

The structural similarity between Mtr4 (excluding the arch domain) and Hel308 allows us to model the likely RNA binding site for Mtr4 (Figure 3-5C). The model was built by aligning the Hel308 DNA-bound structure onto domains 1, 2, 3 and 4 of the Mtr4 structure. Based on this model, it is expected that RNA enters Mtr4 near the interface of the fist and domain 2. Single-stranded RNA then makes a 90° turn to pass through a channel formed by domains 1, 2, 3, and 4, and exits out of the base of the structure. The immediate path of a displaced strand is predicted to extend between domain 2 and the fist, possibly in the direction of the central hole of the arch. However, the trajectory of the displaced strand is unclear, in part because the duplex region likely arises from secondary structures formed by a single strand of RNA, which may preclude large separation of the opposing strands.

The Mtr4 structure reveals all of the general structural components that are required for nucleic acid binding, strand melting and strand translocation (Figure 3-5A). Specific features include: (1) a β-hairpin (residues 521-532 of domain 2) that is involved in melting duplex nucleic acid, (2) the conserved motifs along the interface of domains 1 and 2 are positioned to interact with the phosphate backbone of the unwound strand, and
Figure 3-5. RNA binding model. (A) Nucleic acid from the DNA bound Hel308 structure (PDB 2p6r) was superimposed on the Mtr4 structure to model the likely RNA binding site. Surface (top) and cartoon (middle) representations are shown, colored by domain as in Figure 3-1. Inset images highlight specific structural features observed in Mtr4 that are associated with strand separation and translocation in Hel308, including the β-hairpin (from domain 2) and ratchet helix (domain 4). In addition, extensive interactions are predicted along the fist of the arch. (B) Sequence conservation mapped on the surface of the arch. The color scheme of conservation is the same as in Figure 3-3A. The position of three strictly conserved arginine residues that potentially interact with RNA is indicated (♦). (C) Stereo view of Domain 4. Domain 4 of Mtr4 is an eight-helix bundle and is designated as a DSHCT domain (purple). The domain from Mtr4 is superimposed with the DNA-bound seven-helix bundle of Hel308 (white) to highlight the structural similarity between these two domains. An 18 residue disordered loop in Mtr4 is drawn as a dashed line. Helix H38 of Mtr4 corresponds to the ratchet helix of Hel308.
(3) a ratchet helix. The consistency between Mtr4 and Hel308 (Buttner et al., 2007) in these three conserved structural features along the putative RNA binding path suggests that all Ski2-like helicases likely promote strand displacement and translocation using the same basic mechanisms.

The RNA binding model also predicts protein-RNA interactions that are unique to Mtr4. The arch domain is poised to interact with RNA primarily in the fist. Conserved, positively charged residues (Arg678, Arg774, and Arg800) are located along one face of the fist and are strong candidates for potential interactions (Figure 3-5B). Several large loops, which are disordered in the current structure, potentially interact with the incoming RNA. These loops include: a 15 residue loop (699-713) in the fist, a 29 residue region (361-389) in domain 2, and an 18 residue loop (943-961) in domain 4. Each of these disordered regions contains conserved, positively charged residues that could interact with a negatively charged RNA backbone. We note that the temperature sensitive Cys942Tyr mtr4-1 mutation (Liang et al., 1996) resides in domain 4 at the base of the disordered loop, and potentially affects interactions with incoming RNA.

Localization of the fist near the predicted entry point for RNA and adjacent to the β-hairpin of domain 2 suggests that the arch may influence strand separation or resolution of RNA/RNP structures. The coiled-coil scaffold of the arms of the arch may actually provide the structural rigidity needed to resolve some structures. Alternatively, the arch may regulate RNA access to or navigation through the helicase. In this model, the arch could fill a role analogous to domain 5 of other Ski2-like helicases, which acts as a
“molecular brake” to regulate processivity along various substrates (Richards et al., 2008).

The RNA binding model suggests that the large hole formed by the arms of the arch is probably not the primary interface for RNA binding. Consistent with this model, little conservation of charge or sequence is observed along this surface (Figure 3-3A and Figure 3-5B). The predicted path of RNA binding is also a significant distance away from the winged helix of domain 3. Although winged helix domains are often associated with nucleic acid binding (Gaggiwala and Burley, 2000), it is unclear whether this domain interacts with RNA.

**The bottom surface of Mtr4 is a potential interaction site for the exosome**

In order to identify regions of potential surface interactions, conservation among Mtr4 sequences was mapped onto the Mtr4 structure using Consurf (Landau et al., 2005) (Figure 3-6A). Strong conservation is observed throughout the base of the structure and is particularly concentrated along the bottom surface near the putative RNA exit site. Conservation is less pronounced in the arch domain and is generally limited to one surface of the fist that includes a cluster of highly conserved arginines (also see Figure 3-5B). When Ski2 conservation is mapped onto the Mtr4 structure, a similar pattern of strong conservation is observed along the base of the structure with even more limited conservation in the arch (Figure 3-6B). In contrast to Mtr4 and Ski2, the surface of Hel308 is much less conserved, especially near the nucleic acid exit site (Figure 3-6C).
Figure 3-6. Surface conservation. Evolutionary conservation scores were calculated and mapped onto the protein surface using Consurf (Landau et al., 2005). Conservation is depicted by color gradient ranging from variable (white) to highly conserved (orange). (A) Mtr4 conservation mapped onto the Mtr4 structure. (B) Ski2 conservation mapped onto the Mtr4 structure. (C) Hel308 conservation mapped onto the Hel308 structure (PDB 2p6r). The strong conservation observed along the base of Mtr4 and Ski2 suggests that this is a potential interaction surface. See Appendix A for a more detailed alignment of Mtr4 and Ski2 sequences.
Thus, the conservation along the base of Mtr4 and Ski2 appears to be a unique feature of this subfamily of Ski2-like helicases. This suggests that the conserved surface plays a role that is common to Mtr4 and Ski2, but is distinct from canonical helicase function. One intriguing possibility is that the exosome interacts with Mtr4 (and Ski2) along the base of the structure. This possibility is consistent with the observation that Mtr4 co-purifies with the human exosome (Chen et al., 2001). Such a direct interaction would position RNA emerging from the base of Mtr4 adjacent to the site where RNA enters the exosome. We note that the base of Mtr4 is roughly equivalent in diameter to the entry surface of the exosome.

The current data provide few clues for the location of Trf4-Air2 assembly on Mtr4. Since polyadenylation of RNA substrates is likely to precede interactions with Mtr4, the protein-protein binding site is expected to be nearer the putative RNA entry site rather than at the conserved exit site of Mtr4. The arch domain is an attractive target for Trf4-Air2 assembly because of its proximity to the entry site and its distinctive scaffold-like appearance. However, the lack of conservation in this region yields few clues to potential interfaces.

**The arch is important for exosome function in vivo**

To investigate the function of the arch domain, an Mtr4-archless mutant was constructed by inserting a 4 residue linker between H22 and H30 of domain 3 (Figure 3-7). Thus, the archless mutant essentially resembles domains 1, 2, 3, and 4 of Hel308. Not surprisingly, purified Mtr4-archless protein retains both RNA-dependent ATPase activity
**Figure 3-7.** Construction of the Mtr4-archless mutant. (A) To make the Mtr4-archless mutant, residues 615–878 were replaced with the residues ELST (Glu, Leu, Ser, and Thr). These four residues are similar to those found in Hel308 in equivalent positions (Glu, Ile, Ser and Leu). (B) The structural justification for the Mtr4-archless mutant. Mtr4 is colored by domain as in Figure 3-1. The inset shows that domain 3 of Hel308 (yellow) aligns with domain 3 of Mtr4 and a small portion of the arch domain. The Mtr4 residues that were removed in the construction of the Mtr4-archless mutant are colored white and the four residues that replaced the arch sequence are modeled as a blue dashed line.

To address whether the arch is functionally important, we generated yeast expression plasmids for wild-type and archless Mtr4. A strain containing the Mtr4-archless version is viable, but grows significantly slower than the control wild-type strain (Figure 3-8D). This slow growth phenotype is not simply due to reduced expression of Mtr4-archless, since western blot analysis indicates that wild-type and archless Mtr4 are and ATP dependent helicase activity (Figure 3-8). Furthermore, the activity is comparable to full-length Mtr4 (Figure 3-8B).
expressed at similar levels (Figure 3-8E). Thus, although the arch domain of Mtr4 is not required for \textit{in vitro} ATPase or helicase activity, it is important for proper Mtr4 function \textit{in vivo}.

To analyze whether the growth defect of the Mtr4-archless strain reflected a defect in exosome-mediated functions, we analyzed the processing of 5.8S rRNA and the degradation of the 5’ETS, which are products of a 35S polycistronic precursor (Figure 3-8F). The 5.8S rRNA is generated from a 7S precursor by the exosome and Mtr4 (de la Cruz et al., 1998). The Mtr4-archless strain accumulated a distinct RNA species that is 30 nucleotides longer than the normal 5.8S rRNA (Figure 3-8G). We conclude that the arch is required for exosome-mediated rRNA processing. The 5.8S+30 accumulation phenotype of the Mtr4-archless mutant is similar to that seen in \textit{rrp6}Δ strains (Figure 3-8G). In contrast, point mutations in the core exosome do not lead to the accumulation of this species, but instead cause the accumulation of heterogeneous species that range in length between 7S and 5.8S + 30 (Allmang et al., 1999a). Based on these observations it has previously been suggested that removal of the last 30 nts of the 5.8S precursor specifically requires Rrp6 and cannot be performed by the core exosome (Briggs et al., 1998). The Mtr4-archless phenotype demonstrates that the arch domain is also required for final processing of remaining 30 nts of the 5.8S precursor.

The 5’ ETS is the 5’ most part of the 35S rRNA precursor, which is degraded by the exosome following cleavage from the precursor. Probing a northern blot with a probe for the 5’ETS revealed a pattern in archless that was very similar to that seen in \textit{rrp6}Δ. Both strains accumulate about 4-fold higher levels of the 5’ETS compared to wild type.
Figure 3-8. The arch domain of Mtr4 is required for RNA processing and degradation, but not for helicase or ATPase activity. (A) Displacement of a radiolabeled single-stranded RNA from a 16bp duplex with a 3’-single-stranded overhang by full-length Mtr4 and the Mtr4-archless mutant, as observed on a non-denaturing polyacrylamide gel. Illustrations on the left describe the mobility of duplex and single-stranded RNA through the gel; the asterisk represents the 32P label. Aliquots were taken at the time points indicated after the addition of ATP. Lane 8 displays the complete dissociation of the duplex after heating an aliquot to 95°C, and lanes 7 and 15 show a reaction without ATP after 60 minutes. (B) Time courses for the fraction of displaced (unwound) RNA compared to total RNA is shown. Closed circles indicate Mtr4 activity; open triangles indicate Mtr4-archless activity. A solid line representing Mtr4 and a dashed line representing Mtr4-archless were fit to the data points as a first-order reaction (Wang et al., 2008), allowing for the determination of reaction amplitudes (A = 0.427 ± 0.010, Mtr4; A = 0.364 ±0.016, Mtr4-archless) and observed unwinding rate constants (k’ unw = 0.066 ± 0.005 min⁻¹, Mtr4; k’ unw = 0.058 ± 0.008 min⁻¹, Mtr4-archless). (C) Enhancement of ATPase activity after introduction of RNA for Mtr4 and Mtr4-archless. The rate was determined using a malachite green assay that monitors the rate of release of inorganic phosphate [Pi] from ATP over time. Background values (activity in the absence of RNA) have been subtracted from the total rate. (D) The arch domain is required for optimal growth rate. The indicated plasmids were introduced into a yeast strain that had the MTR4 gene deleted from the chromosome, and that also contained a plasmid encoding wild-type Mtr4 with an URA3 selectable marker. Growth on 5-FOA plates (left) selects for cells that have lost the URA3 plasmid, and thus shows that Mtr4-archless confers a slow growth phenotype. (E) A Western blot probed with anti-Mtr4 antibodies (top) and reprobed with anti-Pgk1 antibodies as a loading control shows that archless and wild-type Mtr4 are expressed at similar levels. (F) The 5’ ETS and 5.8S rRNA are degraded and processed through actions of the nuclear exosome and Mtr4. The 5’ ETS and 7S rRNA precursor are downstream products of 35S rRNA precursor processing events. 5.8S rRNA is generated through processing of 7S and 5.8S+30 intermediates and the 5’ ETS is completely degraded. (G) The 5.8S+30 intermediate accumulates in an Mtr4-archless mutant as well as an rrp6 knockout strain but not in strains that lack Rrp44 exonuclease or endonuclease activity. Shown is a Northern blot probed with a 32P radiolabeled oligo specific for the 5.8S rRNA (top) and re-probed for the RNA subunit of the signal recognition particle (bottom). (H) The 5’ ETS signal is four fold greater in the Mtr4-archless, rrp6 knockout, and rrp44-exo’ strains, but only the rrp44-exo’ strain shows decay intermediates. Shown is the same Northern blot as shown in (G), re-probed with a 32P radiolabeled oligo specific to the 5’ ETS. Figure 3-7 provides the structural justification for design of the Mtr4-archless mutant used in these studies.
For both the 5.8S rRNA processing and 5’ETS degradation assays, the phenotype of archless mtr4 did not resemble that of point mutants inactivating either of the other two catalytic activities of the exosome (i.e. the rrp44-D551N mutant lacking Rrp44 exoribonuclease activity and the rrp44-D171A mutant lacking Rrp44 endoribonuclease activity; Figure 3-8). Thus, our Northern blotting results suggest a possible role of the arch in activation of Rrp6, although they do not rule out the possibility that the arch also affects the activity of Rrp44 on some substrates.

Consistent with a role for Mtr4 in activating Rrp6, a recent paper indicates that the TRAMP complex can stimulate the activity of Rrp6 in vitro in a manner that is independent of the helicase or poly(A) polymerase activities of TRAMP (Callahan and Butler, 2009). Notably, TRAMP enhancement of Rrp6 activity appeared even in a TRAMP complex depleted of detectable levels of Mtr4. However, other cofactors associated with the 5.8S rRNA during processing, including the core exosome, were not present in the in vitro analysis of TRAMP-mediated Rrp6 activation. The mechanism for Rrp6 activation and the role of the Mtr4 arch domain in that process therefore remains unclear. One possibility is that the arch indirectly enhances Rrp6 activity in vivo through other protein-protein interactions (e.g. Trf4 and/or Air2). Alternatively, the arch may assist in making the extra 30 nts of the 5.8S precursor accessible to Rrp6 in vivo. This last possibility is especially interesting in light of recent structural and biochemical data indicating that the distance from the top of the core exosome to the active site of the exonuclease Rrp44 is 30 nts (Bonneau et al., 2009) (i.e. the same as the number left in 5.8S+30). Previous models for RNA processing have not addressed how the core
Figure 3-9. Model for 5.8S rRNA processing. Initial processing of the 7S precursor is performed by Rrp44 in association with the core exosome, leaving a 5.8S + 30 nt RNA. Final processing is performed by Rrp6. Mtr4 is proposed to help make the 5.8S + 30 substrate accessible to Rrp6, possibly by removal of the RNA from the core exosome. While Rrp6 is depicted in association with the core exosome during processing of the 5.8S + 30 substrate, such an interaction may not be required, as previously suggested (Callahan and Butler, 2008).

Exosome would stop once it reaches +30 and how the 5.8S+30 species is released from the core exosome to be targeted to Rrp6. We speculate that Mtr4 could help stop exosomal decay of processed substrates. The Mtr4-archless data further raise the possibility that the arch domain (and/or proteins associated with it) may facilitate the retraction of 5.8S+30 intermediates out of the exosome core to make them accessible for Rrp6 (Figure 3-9).
Conclusion

The full-length Mtr4 structure presented here represents the first complete view of a Ski2-like RNA helicase, and is also the first full-length structure for any RNA helicase of comparable size. The structure reveals a variety of features that suggest that while Mtr4 retains canonical helicase function, it also contains added functionality that extends beyond that of previously characterized helicases. The novel arch domain, in particular, appears to play an important role in processing of RNA substrates. The structural and biochemical data presented here suggest numerous avenues toward a more detailed mechanistic understanding of exosome-mediated RNA processing and decay, as well as a broader view of the activities of Ski2-like helicases.

Materials and methods

Construct design

The full length Mtr4 DNA sequence from *S. cerevisiae* was inserted into a pET 151-D-topo *E. coli* expression vector (Invitrogen) with an N-terminal, TEV-protease cleavable, hexahistidine-tag. Removal of the his-tag by TEV cleavage leaves two additional amino acids at the N-terminus of the native sequence. The archless Mtr4 mutant was constructed by replacing protein residues 615-878 with a 4 residue ELST linker (Figure 3-7).

Protein expression

Mtr4 protein was recombinantly expressed in an *E. coli* BL21(DE3)codon+RIL cell line (Stratagene). Protein expression was induced using an autoinduction protocol
(Studier, 2005). Autoinduction yielded a greater amount of protein per media volume (15-20 mg/L) than traditional IPTG induction methods. Growth at room temperature allowed for greater Mtr4 solubility than that observed for 37°C growth. Cells were harvested at 20,000 rpm and stored at -80°C.

**Protein purification**

Cell lysis was performed by manually breaking up the frozen cell pellet followed by lysozyme treatment and sonication. Ni-affinity, heparin affinity, and gel filtration chromatography techniques were employed to purify the protein. All purification buffers contained 5% glycerol, 50 mM HEPES pH 7.5, and 2 mM beta-mercaptoethanol. Salt concentration was varied in the buffers through the chromatography procedures. The final buffer conditions of purified protein were those described above plus 160 mM NaCl.

**Activity assays**

ATPase activity was determined using a colorimetric malachite green assay that detects the release of inorganic phosphate [Pi] after ATP hydrolysis. The method was adapted from the procedure described by Bernstein, et al (Bernstein et al., 2008) and is described in detail in supplementary materials and methods (appended below).

*ATPase assay* – A colorimetric malachite green assay, adapted from Bernstein, et al (Bernstein et al., 2008), was used to detect the release of inorganic phosphate [Pi] after ATP hydrolysis. Inorganic phosphate release was observed by a change in absorbance at 620 nm.
To prepare the malachite green assay solution, 1 volume of 4.2 % (w/v) ammonium molybdate in 4 M HCl was added to 3 volumes of 0.045 % (w/v) Malachite Green. Reaction mixtures (600 µL) containing 25 mM Tris (pH 7.5), 10 mM Mg Acetate, 2 Mm β-mercaptoethanol, 0.3 µM of protein and 40 nM of ssRNA (5’-GCGUCUUUACGGUGCUUAAAAACAAAAACAAAAACAAAAACAAAA-3’) were initiated by the addition of 2 mM of ATP. At 5, 10, 15, 20 and 30 minutes 120 µL were removed from the reaction and added to 200 µL of malachite green solution. Change in absorbance at 620 nm was observed using a VERSAmax tunable microplate reader (Molecular Devices). The instantaneous reaction rate ([Pi] µM min⁻¹) was calculated by fitting a linear trend line to the absorbance values of time points using KaleidaGraph software. To determine the enhancement of RNA on Mtr4 and Mtr4-archless activity, reactions without RNA were used to obtain background values.

Helicase activity was determined essentially as described by Wang et al (Wang et al., 2008). Briefly, the assay measures displacement of a 32P-labeled 16 nt RNA from a complementary RNA strand containing an additional 25 nt 3’ overhang. Displacement is observed on a non-denaturing polyacrylamide gel. The detailed protocol is described in supplementary materials and methods (appended below).

*Helicase assay* – Helicase activity was determined essentially as described by Wang et al (Wang et al., 2008). Briefly, the assay measures displacement of a 32P-labeled 16 nt RNA from a complementary RNA strand containing an additional 25 nt 3’ overhang. The detailed protocol is described below.
A 41 nucleotide ssRNA that was produced using Ampliscribe T7 flash in vitro transcription kit (Epicentre) was combined with a 16 nucleotide ssRNA ordered from Integrated DNA Technologies. The 16bp RNA was radiolabeled using ATP, [γ²P]-6000Ci/mmol and Kinase Max 5’ end labeling kit (Ambion). The sequence of the 16 bp duplex RNA with a 3’-end 25 base overhang is identical to one used by Wang and colleagues (Wang et al., 2008) and is as follows (duplex region is underlined): (5’-AGCACCCTAAAGACGC-3’ + 5’-GGGUCUUUAAAGGUCCUUAACCAAAAAACAAACAAAAACAAA-3’). Reaction mixtures (40 μL) containing 40 mM NaH₂PO₄ (pH 6.0), 50 mM NaCl, 0.5 mM MgCl₂, 2 mM β-mercaptoethanol, 0.01% NP-40, 0.8 U/μL Anti-RNase (Ambion), and 0.5 nM 32P-labeled duplex RNA substrate were incubated for 5 minutes at 30°C with 800 nM of tr4 or Mtr4-archless protein. Reactions were initiated by adding MgCl₂ and ATP to final reaction concentrations of 2 mM. At designated time points, 8 μL aliquots were removed and quenched with 8 μL of loading dye that contained 1% SDS, 50 mM EDTA, 0.1% bromophenol blue, and 20% glycerol. Aliquots were applied to a 15% non-denaturing poly-acrylamide gel for 1 hour at 100 V. To visualize the radioactivity, gels were wrapped in cellophane and exposed to x-ray film (Kodak). Films were developed and then quantified using Multi Gauge software. Kinetic calculations of helicase observed rate constants (kobs), annealing constants (kann), unwinding constants (kunw) and amplitudes (A) were performed using the equations employed by Wang et al (Wang et al., 2008) (Fraction unwound = A(1-exp(-kobs·t)), kobs = kunw + kann, A = kunw /kunw + kann) and KaleidaGraph software.
**Yeast complementation assay**

MTR4-wild-type or MTR4-archless plasmids containing a LEU2 selectable marker were transformed into an MTR4 deletion strain complemented with a MTR4-wild-type copy plasmid containing a URA3 selectable marker. To test whether the arch was essential for growth, the resulting transformants were grown in Synthetic Complete-LEU (SC-LEU) liquid media overnight at 30°C. The liquid cultures were then serially diluted and spotted onto plates containing 5-Fluoro-Orotic Acid (5-FOA) (to counter select against the URA3 plasmid with MTR4-wild-type), or control plates (SC-LEU). For more details regarding yeast plasmids and methods see supplementary materials and methods.

**Yeast plasmids** – The plasmids for expression of Mtr4-archless (or wild-type Mtr4) in yeast contain 388 bp of MTR4 promoter upstream of the AUG and 264 bp downstream of MTR4 sequence downstream of the stop codon. pAv673 contains wild type MTR4 inserted in pRS416 (a URA3 CEN plasmid; [Sikorski and Hieter, 1989]). Plasmid pAv675 contains wild type MTR4 inserted in pRS415 (a LEU2 CEN plasmid; [Sikorski and Hieter, 1989]). pAv674 contains MTR4-archless inserted in pRS415.

**Yeast methods** – Plasmid pAv675 was introduced into a heterozygote diploid MTR4/mtr4D::KANMX4 strain obtained from Open Biosystems. We sporulated the transformants and obtained haploid progeny spores by the hydrophobic spore-isolation method essentially as described (Rockmill et al., 1991). Isolated spores were germinated to generate the deletion strains complemented with the URA3 plasmid pAv673, resulting
in strains yAv1150 (MatA, ura3-D0, leu2-D0, his3-D1, mtr4D::KANMX4 [MTR4, URA3]) and yAv1151 (Mata, ura3-D0, leu2-D0, his3-D1, met15-D0 mtr4D::KANMX4 [MTR4, URA3]). These strains were transformed with either wild-type Mtr4 or Mtr4-archless plasmids with a LEU2 selectable marker (pAv675 and pAv674, respectively).

To test whether the arch was essential for growth, we grew the resulting transformants in SC-LEU media overnight at 30°C. We diluted these cultures in SC-LEU media to a starting OD600 of 0.2. We grew the cultures until they reached an OD of 0.8, serially diluted them in 96-well plates by a factor of five, and spotted them onto media containing 5-FOA (to counter select against the URA3 plasmid with WT MTR4), or control media (SC-LEU). We incubated these plates at a variety of temperature (room temp to 37°C) (Sikorski and Hieter, 1989), and at each temperature the strain expressing Mtr4-archless grew slower than the wild-type control. Colonies that grew on 5FOA were isolated, and streaked on another 5FOA plate. The resulting strains were used for Western and Northern blot analysis.

**Western blot analysis**

MTR4-wild-type and MTR4-archless strain that grew on 5FOA were grown in YPD, total protein was isolated and analyzed using antibodies against Mtr4 at a 1:5000 dilution. Western blots were stripped and reprobed with antibodies against PGK1 (Molecular Probes) to control for equal loading.
Northern blot analysis

MTR4-wild-type and MTR4-archless strain were grown in YPD, total RNA was isolated, resolved on polyacrylamide gels, and probed with 5' 32P labeled oligonucleotides for 5.8S processing defects (5'TTTCGCTCGTTCTTCATC3'), 5'ETS degradation defects (5'CGACGACAAGCCTACTCG3') and for the RNA subunit of the signal recognition particle (5'GTCTAGCGCGAGGAAGG3').

Crystallization

Crystallization was performed using standard vapor diffusion methods. Crystals were grown at 4°C in 2.4 M ammonium phosphate dibasic, and 0.1 M Tris-base pH 8.5 at a 1:2 protein:well drop ratio, using 10 mg/ml his-tagged Mtr4 protein. Although TEV-cleaved protein is amenable to crystallization in the same conditions, there was no improvement in the diffraction of the crystals. In order to provide phasing data for structure solution, selenomethionine-substituted (Se-met) Mtr4 was expressed in E. coli using a modified autoinduction protocol (Studier, 2005). Se-met purification and crystallization was the same as described for the native protein.

Data Collection and Structure Determination

Crystallographic data were collected to 3.4 Å (native) and 3.6 Å (Se-met) on beamline X29 at the National Synchrotron Light Source (NSLS) (Table 3.1). Data were processed using HKL2000 (Otwinowski and Minor, 1997). The Mtr4 crystals belong to space group P3121 (a=133.4 Å, b=133.4 Å, c=191.9 Å; α=90°, β=90°, γ=120°) and contain one molecule in the asymmetric unit (Matthews coefficient = 3.9; 68% solvent).
Phases were determined by the single-wavelength anomalous dispersion method using Se-met substituted Mtr4. The programs SOLVE (Terwilliger and Berendzen, 1999) and RESOLVE (Terwilliger, 2000), as implemented in the PHENIX software package (Adams et al., 2002), were used to identify selenium positions (30 out of 32 potential sites were identified) and to calculate maps to 3.6 Å. The maps were subsequently extended to 3.4 Å using the native data.

**Modeling and refinement**

The resulting electron density maps were of sufficient quality to manually build an initial backbone trace using Coot (Emsley and Cowtan, 2004). The Hel308 DNA helicase structures (PDB 2va8 and 2p6r) (Buttner et al., 2007; Richards et al., 2008) were used to clarify ambiguous connectivity between secondary structures and to improve the model in domains 1, 2, 3 and 4. The Robetta structure prediction server (Kim et al., 2004) was used to identify homology between the fist of the arch domain and the archaeal L14e protein from *S. solfataricus* (2joy). The L14e structure provided critical guidance for tracing the map in this region. Secondary structure predictions obtained from PSIPRED (Jones, 1999) provided useful guidance for modeling secondary structure. Methionine residues constitute ~ 3% of the protein and are quite evenly distributed throughout the sequence. Consequently, methionine positions identified from the Se-met data provided critical place markers to assign sequence to 95 % of the final structure.

PHENIX (Adams et al., 2002) was used to perform individual b-factor, positional and TLS refinement. Because of the low resolution of the data, extensive secondary structure restraints were employed to maintain proper geometry during refinement. The
final model is a mixed α-helical (50%) and β-sheet (15 %) structure containing 918 out of 1073 total residues and four phosphates. Gaps in the current model include 85 residues of the native sequence at the N-terminus and 11 loops of varying length throughout the rest of the sequence (Figure 3-1). Complete side chains were modeled when supported by the electron density. All other residues were modeled as alanine or serine, as deemed appropriate. The structure was refined to a final $R_{\text{work}}/R_{\text{free}}$ of 29.3% / 33.3%.

Ramachandran statistics calculated using MolProbity (Davis et al., 2007) indicate 74.8% of the residues are in the favored region with 3.8% outliers. Figures were generated with PyMOL (DeLano, 2002).

**Sequence alignment and conservation analysis**

Sequence alignments and conservation scores were calculated with ConSurf (Landau et al., 2005). Using the arch sequence from *S. cerevisiae* Mtr4 as a query, 108 unique Mtr4 sequences were identified for analysis. A similar search with the putative arch sequence from *S. cerevisiae* Ski2 retrieved 42 unique sequences.

**Accession codes**

Protein Data Bank: The atomic coordinates and structure factors for Mtr4 have been deposited with accession code 3L9O.

**References**


CHAPTER 4
CONSERVED FEATURES IN THE MTR4 RATCHET HELIX
REGULATE RNA HELICASE ACTIVITY

Abstract

Mtr4 is a conserved superfamily-2 RNA helicase that activates the turnover, surveillance and processing of nuclear RNAs. Mtr4 is the largest member of the Trf4-Air2-Mtr4 polyadenylation (TRAMP) complex which targets and delivers RNA substrates to the nuclear exosome for 3’-5’ decay. Recent studies have shown that the polyadenylation activity of Trf4 and RNA unwinding activity of Mtr4 within the TRAMP complex is modulated by a sensing mechanism in Mtr4 that detects both length and identity of poly(A) tails. However, it is unclear what structural features of Mtr4 contribute to this sensing function. Analysis of the sequence and recent x-ray crystal structures of Mtr4 reveals significant conservation of ratchet helix residues positioned along the RNA binding path. Although similar conservation patterns are observed throughout Ski2-like and DEAH-box helicases, the identity of conserved ratchet helix residues is helicase specific. While it has been proposed that the ratchet helix is important for unwinding activity, the role of individual residues along the helix in this process has not been investigated. Here we demonstrate that the residue at the 4th turn of the helix (E1033), influences RNA unwinding, RNA functional affinity, and ATPase activities. Finally, when mutated ratchet helix positions confer slow growth phenotypes to *Saccharomyces cerevisiae*. Our findings demonstrate that E1033 is important for helicase activity in vitro, and R1030 and E1033 is important for in vivo Mtr4 function.
The conservation observed in other helicases suggests that members of the Ski2-like and DEAH-box helicase families may utilize ratchet helix residue identity as a general mechanism for regulation of substrate specificity and unwinding rate.

**Introduction**

To maintain correct gene expression in the cell, the integrity of RNA must be tightly regulated through RNA processing, turnover and surveillance pathways (Bernstein and Toth, 2012; Ibrahim et al., 2008; Lebreton and Seraphin, 2008). Several disease states are linked to defects in these RNA quality control mechanisms, including neurodegenerative diseases, congenital diseases and cancer (Astuti et al., 2012; Bruserud, 2007; Fabre et al., 2012; Nelson and Keller, 2007; Staals and Pruijn, 2011). The eukaryotic exosome, which contains both endonuclease and 3’-5’ exonuclease activities, plays a critical role in a wide variety of RNA processing and degradation pathways (Allmang et al., 1999; Lykke-Andersen et al., 2011; Mitchell et al., 1997; Wolin et al., 2012). Regulation of this activity involves multiple protein co-factors including the nuclear Trf4-Air2-Mtr4 polyadenylation (TRAMP) complex (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005). TRAMP facilitates the 3’-end processing of rRNAs, snoRNAs, and snRNAs as well as the degradation of cryptic unstable transcripts (CUTs), aberrant RNAs, antisense RNAs, intronic RNAs, several mRNAs and incorrectly processed RNAs reviewed in (Anderson and Wang, 2009; Bayne et al., 2007; Houseley and Tollervey, 2009; Vanacova and Stefl, 2007; Wolin et al., 2012). TRAMP is also involved in transcriptional regulation, chromatin maintenance and DNA repair (Bayne et al., 2007).
TRAMP is composed of a poly(A) polymerase (Trf4 or Trf5) a zinc knuckle RNA binding protein (Air2 or Air1) and a Ski2-like RNA helicase (Mtr4) that are conserved throughout eukaryotes (LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005). Reminiscent of RNA degradation pathways observed in prokaryotes and organelles (Lin-Chao et al., 2007; Lykke-Andersen et al., 2009), TRAMP promotes 3’-5’ exosomal degradation by adding a short 4-5 nt poly(A) tail to the 3’ end of RNA (Jia et al., 2011; Wlotzka et al., 2011). Depletion or mutation of Mtr4 causes a buildup of polyadenylated TRAMP substrates, demonstrating that Mtr4 is the fundamental link between TRAMP and exosome degradation (Liang et al., 1996; van Hoof et al., 2000). The helicase activity of Mtr4 is proposed to resolve secondary structures and remove proteins associated with the RNA, thus facilitating delivery of single stranded RNA to the exosome (Anderson and Wang, 2009; Houseley and Tollervey, 2009; Lykke-Andersen et al., 2009).

Mtr4 binds poly(A) RNA with higher affinity and with a mechanism distinct from that employed to bind random sequences (Bernstein et al., 2008; Bernstein et al., 2010). Mtr4-substrate interactions are dynamic and are dependent on substrate length and the presence of nucleotide. The ability of Mtr4 to distinguish poly(A) from non-(A) substrates in the absence of nucleotide is enhanced 10-fold as the length of the RNA strand decreases to 5 nucleotides, the minimal Mtr4 binding site (Bernstein et al., 2010). Notably, two studies in vitro by Jia et al demonstrate that not only does Mtr4 show an unwinding preference to substrates with poly(A) tail overhangs one binding site (5-6nt) long (Jia et al., 2012), the polyadenylation activity of TRAMP is restricted by Mtr4 to maintain this optimal tail length in targeted RNAs (Jia et al., 2011). Furthermore, a UV
crosslinking study in yeast recently determined that Trf4 substrates contain an average poly(A) tail length of ~5 nt, supporting the conclusion that poly(A) tail length is regulated \textit{in vivo}. Mtr4 contains a fine tuned mechanism that senses the number and identity of 3’ end poly(A) tracts through a distinct binding mode which modulates the polymerase and unwinding activities of TRAMP. However, it is unclear how Mtr4 senses the length and identity of the sequence, and how this sensing is coupled to unwinding.

Recent crystal structures of Mtr4, including apo and RNA-bound forms, and several related Ski2-like and DEAH-box helicase structures provide insight into the general features employed by these helicase families to bind and translocate along nucleic acid substrates (Buttner et al., 2007; Halbach et al., 2012; He et al., 2011; Jackson et al., 2010; Oyama et al., 2009; Richards et al., 2008; Walbott et al., 2010; Weir et al., 2010; Zhang et al., 2008). Although each helicase exhibits unique features and accessory domains, they all contain a common core structure composed of two RecA domains (Mtr4 domains 1 and 2), a winged helix domain (domain 3) and a 7-8 helix bundle domain (domain 4) (Woodman and Bolt, 2011). In the case of the Ski2-like RNA helicase Brr2, two helicase cores are connected in tandem, with the first core being more relevant to \textit{in vitro} and \textit{in vivo} function (Pena et al., 2009; Zhang et al., 2009). The RecA domains contain conserved sequence motifs that bind nucleic acid, and bind and hydrolyze ATP (Singleton et al., 2007). The Hel308 DNA helicase structure bound to duplex DNA indicates that a beta-hairpin within the second RecA domain facilitates strand splitting as the nucleic acid enters the helicase core (Buttner et al., 2007). The 3’
single stranded nucleic acid then traverses the RecA domains and interacts with domain 4 before exiting the helicase at the base. Multiple interactions are observed with domain 4, particularly along the “ratchet helix” where nucleotides stack with W599 and R592 in a manner that is thought to facilitate DNA translocation (Buttner et al., 2007). Not surprisingly, deletion of domain 4 abolishes helicase activity in Hel308 (Buttner et al., 2007), and the analogous mutation in Mtr4 is inviable in vivo (Holub et al., 2012). Also, point mutations in domain 4 of Mtr4 and other Ski2-like helicases have conferred slow growth phenotypes and loss of in vitro activity (Liang et al., 1996; Pena et al., 2009; Small et al., 2006; Zhang et al., 2009). However, although domain 4 appears to play an important role in helicase activity, a mechanistic description of domain 4 function is lacking, particularly for Ski2-like RNA helicases.

In an effort to better characterize the Mtr4 RNA binding path, we have investigated amino acid residues along the ratchet helix of domain 4. Sequence and structural analysis reveals discrete conservation patterns in Mtr4, Ski2-like and DEAH-box helicases. Mutagenesis studies demonstrate that E1033 plays an important role in sequence recognition and helicase activity. In vivo analysis further underscores the importance of ratchet helix residues for cellular function. These data suggest that residues along the ratchet helix provide a mechanism for regulating helicase activity and modulating sequence specificity.
Results

*Structural analysis of the Mtr4 RNA binding path reveals modes of RNA substrate binding distinct from that observed in Hel308*

In the RNA-bound structure of Mtr4, two molecules are observed in the asymmetric unit (Weir et al., 2010). In both molecules, a 5nt pol(A) RNA interacts with the canonical helicase motifs of the RecA domains 1 and 2 through multiple phosphate backbone interactions, similar to that observed in Hel308 (Figure 4-1A and 4-1B) (Buttner et al., 2007). Domain 4 is positioned on the opposite side of the RNA and interacts directly with the nucleotide bases. The primary base interactions are with E947, located on the end of a loop, and R1026, R1030 and E1033, which occupy one face of the ratchet helix. Notably, each of these base interactions are mediated through hydrogen bonds, whereas the analogous interactions in Hel308 generally involve base stacking (Figure 4-1A and 4-1B). The direct protein-nucleotide base interactions observed in the Mtr4 crystal structure suggest that the function of the ratchet helix may not be restricted to RNA translocation (by analogy to Hel308), but may also involve RNA sequence recognition.

*Ratchet helix residues are conserved in Ski2-like/DEAH-box helicases*

We next examined the conservation of the residues along the ratchet helix for the Ski2-like Mtr4, Ski2, Brr2, Hel308 helicases and the DEAH-box Prp22 and Prp43 helicases (Pena et al., 2009; Walbott et al., 2010; Zhang et al., 2009; Zhang et al., 2008). Clustal W was used to align a diverse set of eukaryotic sequences for each helicase
Figure 4-1. Conserved ratchet helix residues interact with nucleic acid. (A) The DNA-bound Hel308 structure (PDB 2P6R) colored by domains. Inset image highlights the pi stacking interactions of ratchet helix residues W499 and R592 shown as sticks with bases of ssDNA. (B) The RNA-bound Mtr4 structure (PDB 2XGJ) molecule A is colored by domains. Inset helix-bundle domain (domain 4) residues that interact with poly-A RNA are shown as sticks, molecule B aligned residues and RNA bases are colored cyan and light blue respectively. (C) Alignment and conservation scores (calculated using Consurf) of Archeal Hel308 and Eukaryotic Mtr4 ratchet helix sequences. Conservation is colored strictly conserved as orange, to variable as white. Extensive conservation at helical turn 4 is highlighted with an arrow. Alignment of Hel308 includes the sequence crystal structure homologs and 7 other archeal sequences (A.ful, Archaeoglobus fulgidus; S.sul, Sulfolobus solfataricus; P.fur, Pyrococcus furiosus; P.hor, Pyrococcus horikoshii; A.ven, Archaeoglobus fulgidus; F.pla, Ferroglobus placidus; M.the, Methanoseta thermophila S.hel, Staphylothermus hellenicus; T.vol, Thermoplasma volcainum P.aci, Candidatus Parvarchaeum acidophilus; M.kan, Methanopyrus kandleri) Alignment of Mtr4 includes 10 model eukaryotic species (S.cer, Saccharomyces cerevisiae; S.pom, Schizosaccharomyces pombe; N.cra, Neurospora crassa; H.sap, Homo sapiens; M.mus, Mus musculus; D.rer, Danio rerio; D.mel, Drosophila melagonaster; C.ele, Caenorhabditis elegans; M.bre, Monosiga brevicollis; A.tha, Arabidopsis thaliana). (D) Conservation of Ski2-like and DEAH-box helicases are mapped onto a ratchet helix cartoon depicting the observed sequences. The sequence placement of the S1 ratchet of Brr2 was performed using our prediction (see supplementary methods and supplementary figure 1.) Extensive conservation at helical turn 4 is highlighted with an arrow.
(archaeal sequences were used for Hel308) (Larkin et al., 2007). Conservation scores were calculated using the ConSurf server (Ashkenazy et al., 2010). Extensive conservation is observed along the entire ratchet helix for the Ski2-like RNA helicases (Mtr4, Ski2 and Brr2) (Figure 4-1C). Less conservation is observed for Hel308 and the DEAH-box RNA helicases Prp22 and Prp43 (Figure 4-1C and 4-1D). In the case of Hel308, position W599 is the only strictly conserved ratchet helix residue observed to interact with nucleic acid. The conserved residues observed at the N-termini of each helix are involved in interactions with domain 2 and do not interact directly with nucleic acid.

Significantly, no residue along the ratchet helix is universally conserved across all helicases, but conservation patterns are clearly evident. The most striking feature is that the fourth turn of the ratchet helix (counting from the N-terminus) is strictly conserved in a helicase specific manner. Mtr4 and Ski2 always have a glutamate at that position (E1033 in Mtr4; E1247 in Ski2), Brr2 has an arginine (R1107), Prp22 and Prp43 have a glutamine (Q1081 in Prp22; Q622 in Prp43), and Hel308 has a tryptophan (W599).

Among the Ski2-like RNA helicases, we note similar conservation patterns at the second and third turns of the ratchet helix. In the case of Mtr4, both positions are always arginines (R1026 and R1030).

**E1033 is important for Mtr4 unwinding activity**

The interaction of R1030 and E1033 with RNA observed in the Mtr4 structures combined with the strong conservation at each of these positions in Ski2-like RNA helicases suggested that these residues might be important for Mtr4 activity. To assess the role of these residues for Mtr4 helicase activity, we mutated each position in S.
cerevisiae Mtr4 to alanine (R1030A, E1033A). E1033 was also mutated to tryptophan (E1033W) to mimic the sequence observed in Hel308. Unwinding assays were performed with E1033 mutants. Pre-steady-state unwinding assays and calculations were performed using protocols described by Jia and colleagues (Jia et al., 2012). The assay detects the displacement of a $^{32}$P labeled top strand from complementary bottom strand with a 3’ single-stranded extension of 6 nt.

The E1033A protein demonstrated a greater instantaneous unwinding rate constant ($k_{\text{unw}}$) for the polyadenylated substrate (poly(A) at 800 nM protein as compared to wild-type Mtr4 (Figure 4-2A and 4-2B). However, E1033W showed significantly less unwinding activity than wild-type with the same substrate (Figure 4-2B). Instantaneous unwinding rate constants ($k_{\text{unw}}$) at several concentrations were then determined for E1033A and E1033W to obtain a true strand-separation rate constant ($k_{\text{max}}$) (Figure 4-2C and Table 4-1). Compared to wild-type, the E1033A mutation caused an increase in $k_{\text{max}}$ and the E1033W a decrease in $k_{\text{max}}$.

To examine potential RNA sequence effects on the enhanced unwinding of the E1033A mutant, we determined unwinding rate constants for a non-polyadenylated substrate (non(A)) (Jia et al., 2012) and for a CCACCA sequence (Figure 2A). E1033A had a greater $k_{\text{max}}$ than wild type for both the poly(A) and non(A) substrates, with a preference for poly(A) (4-2D). We also observe a weaker functional affinity ($K_{1/2}$) in the E1033A mutant for both the poly(A) and non(A) substrate when compared to wild-type (Figure 4-2E and Table 4-1). We note that we do not observe a greater functional affinity
Figure 4-2. Unwinding assays of ratchet helix point mutants. (A) The three RNA substrates used in this study with identical 16bp duplex and variable 3’end overhangs. (B) Representative time course of fraction unwound values of wild type (WT), E1033A, and E1033W Mtr4 enzymes with poly(A) RNA and PAGE (inset) Mtr4 wild type. Determination of all kinetic constants is described in the materials and methods section see Table 4-1 for all kinetic values. (C) Unwinding rates of poly(A) RNA plotted against enzyme concentration. Error bars represent calculated error from helicase curve fitting calculations. (D) Unwinding curves of wild type Mtr4 and E1033A with different RNA substrates as described in A. (E) Functional affinity ($K_{1/2}$) is plotted against the unwinding rate constant ($k_{max}$) for wild type and E1033A enzyme with poly(A) and random RNA substrates.
Table 4-1. Unwinding and dissociation constants of Mtr4 constructs

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Amplitude</th>
<th>$k_{\text{unw}}$ 800nM</th>
<th>$k_{\text{max}}$ (min$^{-1}$)</th>
<th>$K_{1/2}$ (nM)</th>
<th>$K_d$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.76 ± 0.03</td>
<td>0.32 ± 0.06</td>
<td>0.43 ± 0.04</td>
<td>201 ± 64</td>
<td>9.5 ± 0.5</td>
</tr>
<tr>
<td>E1033A</td>
<td>0.85 ± 0.03</td>
<td>0.58 ± 0.11</td>
<td>0.99 ± 0.08</td>
<td>620 ± 111</td>
<td>7.9 ± 0.5</td>
</tr>
<tr>
<td>E1033W</td>
<td>0.57 ± 0.05</td>
<td>0.15 ± 0.05</td>
<td>0.21 ± 0.07</td>
<td>484 ± 360</td>
<td>10.7 ± 0.5</td>
</tr>
</tbody>
</table>

*Methodologies and equations used to derive kinetic constants are found in the materials and methods section.

or the poly(A) substrate over the non(A) as observed by Jia et al (Jia et al., 2012).

However, this may be due to slight differences in experimental design or a higher degree of error in our data. Regardless, our data clearly demonstrate that the increased unwinding rates for E1033A with the poly(A) and non(A) RNAs come at the expense of functional affinity for the substrate. Surprisingly, neither wild-type or E1033A were able to unwind the CCACCA substrate to any substantial degree, suggesting that the sensing mechanism for substrates may be more complex than simply distinguishing between a poly(A) or non(A) sequence.

**Mutation of E1033 has minimal impact on substrate binding**

To assess whether mutation of E1033 affected substrate binding, we determined dissociation constants ($K_d$) of the RNA substrates with wild type and mutant Mtr4 proteins using an electrophoretic mobility shift assay (EMSA). For each protein, we observed similar binding constants towards poly(A) RNA (Table 4-1). The dissociation constants of non(A) and CCACCA substrates with the Mtr4 proteins showed no distinct differences from those observed with poly(A) (data not shown). Although these
calculated dissociation constants are statistically different from one another the magnitude of the distinction is small, and it is unclear if any true differences in RNA substrate binding are caused by E1033 mutation.

_E1033 influences ATPase enhancement at subsaturating RNA levels_

Previous studies have shown the RNA-dependent ATPase activity of Mtr4 is enhanced differently based on RNA nucleic acid sequence (Bernstein et al., 2008; Bernstein et al., 2010). In order to determine how ratchet helix point mutants influenced ATPase activity, we monitored the release of inorganic phosphate [Pi] from ATP in the presence of Mtr4 proteins and RNA substrates using a colorimetric Malachite Green Assay (Bernstein et al., 2008; Jackson et al., 2010). At sub-saturating levels of RNA (40 nM) wild type Mtr4 showed an enhancement of over two fold with poly(A) substrate, one and a half fold enhancement for the non-A substrate, and no enhancement for the CCACCA substrate (Figure 4-3). At 40nM of RNA neither E1033A nor E1033W (data not shown) demonstrated enhanced ATPase activity. Thus, at subsaturating RNA levels mutation of E1033 disrupts the ability of RNA to enhance ATPase activity. At saturating levels of RNA (400nM) all Mtr4 constructs, including R1030A, showed RNA induced ATPase enhancement (data not shown), indicating that the mutagenesis did not completely disrupt the RNA-dependent ATPase activity but did influence the sensitivity of the mechanism.
R1030 and E1033 are important for Mtr4 function in vivo

After determining that the E1033 residue plays a significant role in Mtr4 helicase activity in vitro, we wanted to explore how mutation at ratchet helix positions affected Mtr4 function in vivo. Mutants *mtr4*-R1030A, E1033A, E1033W and double mutant of *mtr4*-R1030A/E1033A were constructed, serially diluted, and tested for viability at 20, 30 and 37 degrees Celsius. Complementation with plasmid containing wild type MTR4 was

![RNA dependent ATPase Activity](image)

**Figure 4-3.** RNA-dependent ATPase enhancement (A) An EMSA derived fraction bound plotted against Mtr4 concentration for wild type Mtr4 with labeled polyA RNA substrate. The equation used to fit the data is \( \theta = [E]/[E] + K_{d,E} \) where \( \theta \) is the fraction of bound RNA/total RNA and \( K_{d,E} \) = dissociation constant. (B) A bar graph displaying calculated dissociation constants. Observed \( K_d \) is as follows: \( K_{d,WT} = 9.5 \pm 0.5 \text{nM, } K_{d,E1033A} = 7.9 \pm 0.5 \text{nM, } K_{d,E1033W} = 10.7 \pm 0.5 \text{nM} \) (C) RNA-dependent ATPase fold enhancement of wild-type and E1033A proteins.
used as a positive control, whereas \textit{mtr4-archless} and \textit{mtr4-D262A/E263A} mutants were used to demonstrate a slow growth phenotype and an active site knockout respectively. The ratchet helix mutations confer a slow growth phenotype at all temperatures tested, demonstrating that these residues are important for Mtr4 function \textit{in vivo} (Figure 4-4). At 30 degrees, the growth defect appears to be most pronounced for the R1030A mutation. This greater growth defect is surprisingly suppressed by the E1033A mutation as seen in the \textit{mtr4-R1030A/E1033A} double mutant. Interestingly, at 37 degrees growth defects are even more pronounced and the pattern appears to change with E1033W and the R1030A/E1033A double mutant mimicking the R1030A mutant. Regardless of temperature however, the double mutation does not compound the growth phenotype.

\textbf{Figure 4-4.} Growth complementation of an Mtr4 knockout strain by ratchet helix mutants. Shown is observed slow growth phenotypes of the ratchet helix mutants. The R1030A displays the slowest growth out of the helix mutant residues. Mtr4-archless is used as a slow growth control and D262A/E263A as an active site knockout.
observed at single sites suggesting that defect caused by each ratchet helix mutation resides in the same mechanistic pathway.

**Discussion**

Previous studies have shown that Mtr4 preferentially unwinds substrates with a short polyadenylated 3’ overhang (Jia et al., 2012). Kinetic analyses of Mtr4-RNA binding have demonstrated that Mtr4 adopts different conformational states when binding polyadenylated and non-polyadenylated RNA substrates (Bernstein et al., 2010). However, the molecular basis for this differential binding is unclear. In an effort to identify the site of sequence recognition by Mtr4 we carefully examined the RNA binding path observed in the 5nt RNA-bound form of the Mtr4 structure. Several residues were identified in domain 4 along the ratchet helix and a neighboring loop that directly interacts with the RNA bases. Significantly, these interactions are mediated through hydrogen bonding in a potentially sequence-specific manner, as opposed to the non-discriminatory base stacking observed in the Hel308 DNA helicase structure. Sequence conservation analysis further revealed more extensive conservation along the ratchet helix of Mtr4 and other Ski2-like RNA helicases than exists for Hel308 or the more distantly related DEAH/ RHA-box helicases. Although the sequence identity varies for each helicase, several positions along the ratchet helix are strictly conserved, in particular, the residue at the 4th turn of the ratchet helix (E1033 in Mtr4).

Preliminary *in vitro* and *in vivo* analysis of E1033 mutants confirmed that the conserved RNA binding position plays an important role in Mtr4 function. Residue
identity clearly influences unwinding, for three different residues (E1033A, E1033W and wild type) at this position demonstrate three different unwinding profiles. Ski2-like and DEAH/RHA-box helicase share a common four domain scaffold, yet target divergent substrates (Cordin et al., 2012). It is possible that the residue identity at the fourth turn of the ratchet helix may impact substrate specificity and conservation of various residue types is a reflection of targeted substrates.

In the case of Mtr4, the ratchet helix is implicated in the poly(A) sensing mechanism. Indeed, both E1033 mutants showed a loss of functional specificity for poly(A) RNA as well as non(A), suggesting that E1033 may be important for sensing substrate characteristics. Mutation of E1033 disrupts Mtr4 sensitivity to the presence of RNA in subsaturating ATPase enhancement assay, supporting a hypothesis that E1033 helps detect sequence.

Although these studies implicate that E1033 may be involved in sensing substrate features, it is clear that E1033 could not act alone to sense the full length and sequence of the 3’ end. Other residues that lie on the RNA bind path including other ratchet residues are likely involved in the sensing mechanism. In vivo studies showed that ratchet helix residue R1030 is also important for Mtr4 function and current in vitro work with the enzyme will soon elucidate the kinetic parameters of this residue. An interesting result of the in vitro studies is that the double mutant R1030A/E1033A did not cause an additive growth defect demonstrating that both residues are likely acting on the same pathway. This supports the idea that R1030 may also be involved in sensing strand identity.
The sensing mechanism of the poly(A) tail is important to Mtr4 and TRAMP function. However, to unwind and effectively deliver substrates to the RNA exosome a processive mode of substrate binding must also exist in Mtr4 that would allow translocation across a variety of RNAs independent of sequence. It is possible that the differences observed between the DNA bound structure of Hel308 and the RNA bound structure of Mtr4 demonstrates sensing (Mtr4) and processive (Hel308) modes of binding. The large residues of the Mtr4 ratchet helix, including R1030 and E1033 are amenable to pi stacking and may switch from a RNA sensing mode that involves specific hydrogen bonds with base atoms to a processive state in which the residues pi stack with any incoming base. Furthermore it has been shown that Mtr4 loses its preference for poly(A) sequence in the presence of specific nucleotide analogs (Bernstein et al., 2010), suggesting that nucleotide binding may cause conformational changes within the core that cause a shift from each mode of binding. Although this model is feasible and the characteristics of the ratchet helix residues are capable of such a mechanism future experimentation is required to confirm these hypotheses.

A surprising result of this work was that the CCACCA overhang was observed to not unwind to any substantial degree in wild type or the E1033A mutant suggesting that unwinding preferences may not be as straight forward as simply selecting a poly(A) tailed over non(A). Continued studies probing how sequence characteristics of RNA substrates modulate the activity of Mtr4 will great aid in understanding the fundamental question of how are RNA substrates selected for decay.
Here we provide an initial framework for understanding the molecular details of the poly(A) sensing mechanism and the role of ratchet helix residues of Mtr4 in RNA substrate targeting and unwinding.

Materials and methods

Structural analysis and Conservation Scoring of Ski2-like and DEAH-box helicases

The helix-bundle domain (domain 4) of archeal Hel308 (pdb 2P6R) (Buttner et al., 2007) was used as bait in a DALI search (Holm and Rosenstrom, 2010) to find structures containing a helix-bundle domain with associated ratchet helix. Conservation of eukaryotic helicases was determined by multiple sequence alignment of model organisms in Clustal W (Larkin et al., 2007) and conservation scoring with the ConSurf server (Ashkenazy et al., 2010). For the archeal Ski2-like DNA helicase Hel308, 98 archeal sequences were retrieved and scored using the ConSurf server. ConSurf scores of 10, 9, and 8 were considered conserved.

Mutagenesis, Protein Expression and Protein Purification

Point mutants of Mtr4 were made using a modified version of the QuikChange (Agilent) site directed mutagenesis procedure. The expression and purification of Mtr4 and mutant Mtr4 proteins was carried out as performed in (Jackson et al., 2010) No differences in expression or purification were observed between the mutants and wild-type protein. Protein concentration was determined using a NanoDrop spectrophotometer (ThermoFisher) and calculated extinction coefficients.
RNA substrate Design and Purification

The RNA substrates were designed to mimic unwinding substrates used in the recent study by Jia and colleagues (Jia et al., 2011). Three 22 nucleotide ssRNAs (bottom strand), each with a unique 3’ end were incubated independently with a complementary 16 nucleotide ssRNA (top strand) at 95°C for 10 minutes after which samples were slowly annealed to room temperature.

For the unwinding assay and EMSA, the 16 nucleotide top strand was radiolabeled with $\gamma^{32}$P ATP and T4 polynucleotide kinase and quenched by heating to 95°C before annealing. The RNA substrates were purified by native PAGE, gel extraction and ethanol precipitation. All RNAs used in this study were purchased from Integrated DNA Technologies (IDT). The substrate sequences are as follows with duplex regions underlined: R16 (top strand of all three substrates) = 5’AGCACCGUAAAGACGC3’, R22A (polyA overhang) = 5’GCGUCUUUACGGUGCUIUAAAAA3’, R22R (random overhang) = 5’GCGUCUUUACGGUGCUCUGCUGCG3’, CCA = 5’GCGUCUUUACGGUGCUCGCCACCA3’.

Unwinding Assay

Assays were performed essentially as described by Jia et al (Jia et al., 2012). A radiolabeled 16 nucleotide top strand is displaced over time when incubated with Mtr4 and ATP. Reactions were carried out at 30°C in a controlled water bath. The buffer used was 40mM MOPS pH 6.5, 100mM NaCl, 0.5 mM MgCl₂, 5% glycerol, 0.01% NP-40 substitute, 2mM DTT, and 1 U/µl of Ribolock (ThermoFisher). Reactions were allowed
to incubate for 5 minutes with ~0.2 nM RNA (final concentration) and indicated concentration of Mtr4 or Mtr4 mutant protein. Reactions were initiated by the addition of equimolar ATP and MgCl$_2$ at a final concentration of 2mM. At specified time points, aliquots of the reaction were removed and quenched at a 1:1 ratio with buffer containing 1% SDS, 5mM EDTA, 20% glycerol, 0.1% bromophenol blue and 0.1% xylene cyanol. Aliquots were run on a native 15% TBE polyacrylamide gel at 100V for 115 minutes. Radioactivity was visualized as performed previously (Jackson et al., 2010). Gels were wrapped in cellophane and exposed to x-ray film. Film was developed and then quantified using multi-gauge software. Calculations of the observed rate constants ($k_{\text{obs}}$), annealing constants ($k_{\text{ann}}$), unwinding constants ($k_{\text{unw}}$) and amplitudes (A) were performed by fitting the curve to the data collected in triplicate. The equation used was that employed previously (Jackson et al., 2010; Jia et al., 2012; Wang et al., 2008) (Fraction unwound = A(1-exp(-$k_{\text{obs}}$·t)), $k_{\text{obs}}$ = $k_{\text{unw}}$ + $k_{\text{ann}}$, A = $k_{\text{unw}}$ + $k_{\text{ann}}$ / $k_{\text{unw}}$). Curve fits to the unwinding data were also fit as one curve to all data collected in triplicate. The curve fits used are the same as found in (Jia et al., 2012), and are best fits using the equation, $k_{\text{unw}}$=$k_{\text{max}}$, E/[E] + K$_{1/2}$, E where [E] is enzyme concentration, K$_{1/2}$ is functional affinity and $k_{\text{max}}$, E is the unwinding rate constant at enzyme saturation.

**Electrophoretic mobility shift assays (EMSA)**

Reactions were incubated in the identical buffer used in the unwinding assay with ~0.2 nM RNA, equimolar MgCl$_2$ and AMP-PNP (2mM), and varying concentration of enzyme for 15 minutes at 30°C in a controlled water bath. Non-denaturing dye (20% glycerol, 0.1% bromophenol blue and 0.1% xylene cyanol) was added at a 1:1 ratio and
samples were run on a native TBE 15% polyacrylamide gel at 100V for 115 minutes. Radioactivity was visualized the same as the unwinding assays. The dissociation constant ($K_d$) was determined by fitting the data to the equation $\theta = \frac{[E]}{[E] + K_d[E]}$ where $\theta$ is the fraction of bound RNA/total RNA and $K_d[E]$ = dissociation constant.

**ATPase activity assay**

The release of inorganic phosphate [Pi] after ATP hydrolysis was determined using the colorimetric malachite green assay performed previously (Jackson et al., 2010). Reaction mixtures (550 µL) containing the same final buffer components and concentrations as that used in the unwinding assays and EMSAs, 40 nM or 400 nM of RNA substrates, and 300 nM of Mtr4 or Mtr4 mutant protein. Reactions were initiated by the addition of equimolar ATP and MgCl$_2$ (final concentration 2mM). At indicated time points 100 µl were removed from the reaction and added to 200 µL of malachite green solution. Change in absorbance at 620 nm was observed using a VERSAmax tunable microplate reader (Molecular Devices). The absorbance values of a known [Pi] standard were used to convert observed ATPase absorbance values into [Pi] µM. The instantaneous reaction rate ([Pi] µM min$^{-1}$) was calculated by fitting a linear trend line to determined values of time points. Fold enhancement values were calculated by dividing the rate of inorganic phosphate release in the presence of RNA by the rate values determined from reactions with no RNA.

The malachite green solution was prepared as performed previously (Jackson et al., 2010) 1 volume of 4.2 % (w/v) ammonium molybdate in 4 M HCl was added to 3 volumes of 0.045 % (w/v) Malachite Green.
Yeast plasmids

The plasmids for expression of Mtr4-wild-type or Mtr4-ratchet helix mutants contained the same upstream promoter and downstream sequence as used previously (Jackson et al., 2010). The same wild type MTR4 expression plasmids pAv673 (a URA3 CEN plasmid; (Sikorski and Hieter, 1989)), and pAv675 (a LEU2 CEN plasmid; (Sikorski and Hieter, 1989)) were used as in (Jackson et al., 2010). Plasmids expressing ratchet helix mutants are simply point mutants of pAv675.

Yeast methods

A n MTR4 deletion strain of Saccharomyces cerevisiae complemented with a URA3 selectable marker containing MTR4-wild-type copy plasmid was transformed with MTR4-wild-type or MTR4-ratchet helix mutant plasmids. Resulting transformants were grown in Synthetic Complete-LEU (SC-LEU) liquid media overnight at 20°C 30°C and 37°C to test if the mutation caused growth defects. Liquid cultures were serial diluted fivefold and spotted onto control plates (SC-LEU) or 5-fluoro- orotic acid (5-FOA; to counter against the MTR4-wild-type plasmid containing URA3 selectable marker).

Software

Pymol was used to align helicase structures and make figures. Kaleidagraph software was used to generate curve fits to the observed data and rate constants.

References


CHAPTER 5

SKI2-LIKE RNA HELICASE STRUCTURES:
COMMON THEMES AND COMPLEX ASSEMBLIES

Abstract

Ski2-like RNA helicases are large multidomain proteins involved in a variety of RNA processing and degradation events. Recent structures of Mtr4, Ski2 and Brr2 provide our first view of these intricate helicases. Here we review these structures, which reveal a conserved ring-like helicase core that extends beyond the canonical RecA domains to include a winged-helix and ratchet domain. Comparison of apo- and RNA-bound Mtr4 structures suggests a role for the winged-helix domain as a molecular hub that coordinates RNA interacting events throughout the helicase. Unique accessory domains provide expanded diversity and functionality to each Ski2-like family member. A common theme is the integration of Ski2-like RNA helicases into larger protein assemblies. We describe the central role of Mtr4 and Ski2 in formation of complexes that activate RNA decay by the eukaryotic exosome. The current structures provide clues into what promises to be a fascinating view of these dynamic assemblies.

Introduction

Ski2-like RNA helicases are a relatively small family of superfamily 2 helicases (Caprara, 2010; Fairman-Williams et al., 2010; Gorbelenya and Koonin, 1993; Jankowsky and Fairman-Williams, 2010; Singleton et al., 2007). A distinctive feature of these helicases is their large size, spanning 120-225 kDa in S. cerevisiae. Named for their
found member, Ski2 (Anderson and Parker, 1998; Johnson and Kolodner, 1995; Toh et al., 1978; Widner and Wickner, 1993), Ski2-like RNA helicases play important roles in RNA degradation, processing and splicing pathways (Bleichert and Baserga, 2007; Caprara, 2010). As helicases, they process RNAs in a 3′-5′ direction. They generally function as parts of larger complexes and in many cases appear to act as a platform for complex assembly. Several recent studies have given us our first structural view of this family. Here we highlight the common architectural themes and functional implications emerging from these structures. We also discuss the role of Ski2-like helicases in larger complexes, with a particular emphasis on complexes involved in activation of the exosome for RNA degradation and processing.

Ski2-like family helicase structures

In *S. cerevisiae*, four Ski2-like RNA helicases have been identified, including Ski2, Mtr4, Brr2, and Slh1. (Suv3 has previously been categorized as Ski2-like, but recent phylogenetic [Fairman-Williams et al., 2010] and structural [Jedrzezejczak et al., 2011] studies indicate that it more appropriately belongs in a separate subfamily.) Human homologs exist for each protein. The biology of these helicases has been extensively reviewed (Anderson and Wang, 2009; Bleichert and Baserga, 2007; Butler, 2002; Caprara, 2010; Cordin et al., 2012; de la Cruz et al., 1999; Hahn and Beggs, 2010; Houseley and Tollervey, 2009; Lebreton and Seraphin, 2008; Newman and Nagai, 2010; Norbury, 2011; Pyle, 2011; Schaeffer et al., 2010). Briefly, Mtr4 and Ski2 promote RNA decay, Brr2 functions in pre-mRNA splicing, and Slh1 is involved in translation and transcription. The structural data currently available for Ski2-like helicases is
summarized in Table 5-1 and Figure 5-1, and includes nearly complete structures of Mtr4 (Jackson et al., 2010; Weir et al., 2010) (75 N-terminal residues are missing) and an equivalent region for Ski2 (Halbach et al., 2012). Partial structures for Brr2 have also been determined (Pena et al., 2009; Zhang et al., 2009a) that correspond to the C-terminal Sec63 domain from the second of two helicase repeats. The only structure bound to RNA substrates is an Mtr4 complex bound to ADP and a five nucleotide single stranded RNA (Weir et al., 2010). In addition to the RNA helicase structures, several related Ski2-like DNA helicase structures are also available (Buttner et al., 2007; Oyama et al., 2009; Richards et al., 2008; Zhang et al., 2008). As will be discussed below, the structure of Hel308 (Buttner et al., 2007), an archaeal DNA helicase from *A. fulgidus*, includes a partially unwound DNA duplex and has provided valuable insight into unwinding and translocation mechanisms that are likely applicable to the RNA helicases.

**Common architecture of Ski2-like helicases**

The signature feature of superfamily 2 (SF2) helicases is a pair of RecA-like domains arranged on a single polypeptide with N- and C-terminal extensions (Gorbalenya and Koonin, 1993; Singleton et al., 2007). These two domains constitute the “core” helicase domains and provide the motor associated with helicase activity (Jankowsky, 2011; Pyle, 2008). Conserved sequence motifs found within the RecA domains are involved in ATP and nucleic acid interactions. Classification of helicase families is typically based on the precise sequence and arrangement of these motifs. In the case of Ski2-like helicases, sequence analysis has identified two RecA domains containing 12 conserved sequence motifs (Caprara, 2010; Fairman-Williams et al., 2010;
Gorbalenya and Koonin, 1993; Jankowsky and Fairman-Williams, 2010; Singleton et al., 2007). The Mtr4 and Ski2 structures confirm that the RecA domains (designated here as RecA1 and RecA2) adopt structures and sequence motif arrangements similar to that observed in other SF2 helicases (Figure 5-1).

Despite the strong conservation of the RecA domains between Ski2-like helicase family members, it was unclear whether any other similarities existed between family members until determination of crystal structures. The Hel308 structure (Buttner et al., 2007) revealed that a small winged-helix domain and a larger seven helix bundle or “ratchet” domain packs against the RecA domains to form a ring-like structure that accommodates passage of a single strand of DNA. An additional C-terminal helix-hairpin-helix (HhH) domain interacts with the ratchet domain and the exiting DNA strand. Structures of the second Brr2 Sec63 domain (Pena et al., 2009; Zhang et al., 2009a) showed that it resembles the ratchet and HhH domains of Hel308, and additionally contains a Fibronectin type III domain (FN3). Unexpectedly, the Mtr4 (Jackson et al., 2010; Weir et al., 2010) and Ski2 (Halbach et al., 2012) structures showed that they also contain a winged-helix and ratchet domain similar to what is seen in the Hel308 structure. A large insertion in the middle of the winged-helix domain in Mtr4 and Ski2 (the arch domain, discussed below) had prevented prior identification of this domain by sequence analysis. Likewise, although the sequence similarity between the Hel308/Brr2 seven helix bundle and the Mtr4/Ski2 eight helix bundle was not obvious (11% sequence identity/26% similarity), they are clearly structurally related (RMSD=2.7 Å, as calculated by PDBeFold (Krissinel and Henrick, 2004).
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Organism</th>
<th>PDB</th>
<th>Substrate</th>
<th>Domains obs.</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ski2</td>
<td>S. cerevisiae</td>
<td>5VV</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>Brr2</td>
<td>S. cerevisiae</td>
<td>51$\uparrow$</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>$\text{p}$. helicase</td>
</tr>
<tr>
<td>Ski2</td>
<td>S. cerevisiae</td>
<td>51$\uparrow$</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>Mtr4</td>
<td>S. cerevisiae</td>
<td>51$\uparrow$</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>Hel308</td>
<td>A. fulgidus</td>
<td>35</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>P. horikoshii</td>
<td>35</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>P. solfataricus</td>
<td>35</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
<tr>
<td>H. furiosus</td>
<td>35</td>
<td>$\text{ADP}$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\uparrow$, $\downarrow$</td>
<td>$\text{p}$. helicase</td>
<td>unpublished</td>
</tr>
</tbody>
</table>

**Table 5.1.** X-ray crystal structures of Ski2-like helicases.
Figure 5-1. Members of the Ski2-like helicase family share a common four domain helicase core. (A). Schematic representations (top row) highlight the architectural arrangement of the Ski2-like helicases for the Brr2, Mtr4 and Ski2 RNA helicases and the related Hel308 DNA helicase. Representative structures are also indicated (middle and bottom rows). The helicase core is composed of the RecA1, RecA2, Winged Helix (WH) and Ratchet domains (white lettering). Accessory domains are also indicated (black lettering). Brr2 contains two helicase core cassettes. The current Brr2 structures are limited to the second Brr2 Sec63 domain (composed of the ratchet, HhH and FN3 domains). (B) Domain organization for the Ski2-like helicases. Amino acid residue numbers are from S. cerevisiae, except for Hel308 (A. fulgidus) and DDX60 (H. sapiens). Uncharacterized N-terminal regions are not indicated in (A) but are included in (B) as grey boxes. Structures were rendered using PyMol.
A. Ski2-like Helicase Family Structures

DNA

Hel308

RNA

Brr2

Cassette 1

Cassette 2

Mtr4

ji-helix

Ski2

RNA

B. Ski2-like RNA Helicases

<table>
<thead>
<tr>
<th>Helicase</th>
<th>Sec63</th>
<th>Sec63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brr2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sfh1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mtr4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ski2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ski2-like DNA Helicases

<table>
<thead>
<tr>
<th>Helicase</th>
<th>Sec63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hfm1/Mer3</td>
<td></td>
</tr>
<tr>
<td>Hel308</td>
<td></td>
</tr>
<tr>
<td>DDX60</td>
<td></td>
</tr>
</tbody>
</table>
Combined, these structures demonstrate that the fundamental molecular “core” of all Ski2-like helicases is a ring-like four domain assembly of two RecA domains, a winged-helix domain and a ratchet domain (Figure 5-1). Recent structures of DEAH/RHA-box RNA helicases have revealed a similar core structure (He et al., 2010; Kudlinzki et al., 2012; Walbott et al., 2010), indicating that it may be a common feature with useful characteristics for a variety of helicases. This common architecture has been referred to elsewhere as the DExH-box core (Weir et al., 2010). Here we will simply refer to it as the helicase core.

Unwinding Mechanism

Our current understanding of RNA/DNA unwinding and translocation by Ski2-like helicases has been developed primarily from structures of the Hel308 DNA helicase (Buttner et al., 2007), and is detailed in recent reviews (Caprara, 2010; Pyle, 2008; Woodman and Bolt, 2011). Key structural features of the proposed mechanism include a β-hairpin in domain RecA2 that acts as a wedge to split the DNA duplex; nucleic acid backbone interactions with conserved motifs in RecA1 and RecA2; and stabilizing base stacking interactions opposite the RecA domains with the ratchet helix of the ratchet domain. ATP binding and hydrolysis between the RecA domains are believed to cause conformational changes that allow the helicase to step along the DNA backbone using an inchworm mechanism (Tanner and Linder, 2001).

The Mtr4 and Ski2 structures suggest that Ski2-like RNA helicases function similarly to Hel308. All of the key structural features described in Hel308 for unwinding are retained in Mtr4 and Ski2, including the β-hairpin and ratchet helix. Although no
structures of the corresponding region exist for Brr2, mutagenic disruption of the
predicted β-hairpin results in a slow growth phenotype \textit{in vivo} and loss of helicase
activity \textit{in vitro} (Zhang et al., 2009a). No duplex RNA has been crystallized with either
Mtr4 or Ski2, so strand splitting is not directly observed. However, the RNA-bound Mtr4
structure does show backbone interactions with the conserved RecA motifs Ia, Ib, IV and
V. Interactions with the RNA bases are observed along the ratchet helix, although it has
been noted that there are some differences in the mode of RNA binding between the two
crystallographically observed molecules for Mtr4 (Bernstein and Toth, 2012). The
ratchet helix interaction also differs slightly in detail between Mtr4 and Hel308, but it
remains to be determined whether this corresponds to any functional differences between
the two proteins. What is clear is that the ratchet domain is functionally important.
Deletion of the ratchet domain of Mtr4 is lethal (Holub et al., 2012), consistent with the
loss of unwinding activity observed in Hel308 when the ratchet domain is removed
(Buttner et al., 2007). The Mtr4-1 mutation (C942Y), which arose from the initial
genetic screen that identified Mtr4, is located on a loop near the ratchet helix and
produces a growth phenotype in \textit{S. cerevisiae} (Liang et al., 1996). Similarly, mutations
along the putative ratchet helix in Brr2 result in slow growth phenotypes and also exhibit
reduced unwinding activity \textit{in vitro} (Pena et al., 2009; Small et al., 2006; Zhang et al.,
2009a).

It has been proposed that the base stacking interactions mediated by the ratchet
helix provide a mechanism for enforcing directionality during translocation (Buttner et
al., 2007; Pyle, 2008). These interactions, combined with the ring-like shape of the
helicase core, may also confer some degree of processivity to the helicase. However, while Hel308 has been described as a processive helicase (Buttner et al., 2007; Guy and Bolt, 2005), the processivity of the RNA helicases has proven difficult to determine (Jia et al., 2012; Pena et al., 2009).

**Accessory Domains**

The unique characteristics of Ski2-like family members are derived in large measure from accessory domains that decorate the helicase core and provide expanded functionality. Indeed, this appears to be a common theme throughout SF2 helicase families (Fairman-Williams et al., 2010; Pyle, 2011).

**Arch domain**

The most distinctive Ski2-like accessory domain described to date is the Arch domain of Mtr4 and a related domain in Ski2. The domain is a prominent 256-266 amino acid insertion within the winged helix domain that is unique to Mtr4 and Ski2 (Figure 5-1). Originating from the back side of the helicase core, two anti-parallel coiled coils (arm and forearm), each composed of an ascending and descending helix, extend over the helicase core and terminate in a globular region called the fist that in Mtr4 is positioned directly above the RNA entry site. A dramatic bend between the arm and forearm creates the arch-like appearance for which the domain was named.

The fist of Mtr4 adopts a β-barrel fold that contains a KOW motif (Kyrpides et al., 1996) and has therefore been described as a KOW domain (Weir et al., 2010). KOW domains are often associated with proteins that interact with structured RNAs (Steiner et
al., 2002), including the bacterial L24 protein, the archeal L14 protein and the eukaryotic L26 and L27 proteins from the large ribosomal subunit (Kyprides et al., 1996; Selmer et al., 2006; Voorhees et al., 2009; Zhang et al., 2009b). (The L14 protein was actually used to model the fist in the original low resolution maps of Mtr4 (Jackson et al., 2010).) Although sequence conservation is limited on the surface of the fist, and generally throughout the arch, a few conserved positively charged and aromatic residues are found on the surface of the fist which may facilitate RNA interactions (Jackson et al., 2010). In vitro gel shift assays confirm that the fist binds hypomodified tRNA (Weir et al., 2010) and pre-rRNA (Johnson lab, unpublished data). However, the fist does not bind single stranded poly(A) RNA (Halbach et al., 2012), indicating a preference for structured RNA.

Like Mtr4, the Ski2 fist contains a β-barrel, but it lacks a KOW motif and adopts a more rigid conformation by associating directly with the arch forearm via hydrophobic interactions (Halbach et al., 2012). The Ski2 fist appears to be less selective than that of Mtr4, binding to both structured and single stranded RNAs. The surface of the Ski2 fist is more positively charged than Mtr4. Presumably, the different RNA binding properties reflect differences in the types of substrates processed by Mtr4 and Ski2 (Halbach et al., 2012).

While the arch domain is poised to interact with an incoming substrate and may play a role in substrate specificity, it is not required for helicase activity (Jackson et al., 2010). Nor has it been shown to be directly involved in protein-protein interactions. It does, however, play an important functional role. Archless-mtr4 mutants have a slow
growth phenotype in *S. cerevisiae* and exhibit rRNA processing defects (Jackson et al., 2010). Further studies will be required to elucidate the role of the arch domain in Mtr4/Ski2 function.

**Other accessory domains**

Mtr4 and Ski2 both contain N-terminal extensions off of the helicase core that may play functional roles. No structural information is available for this region, with the exception of a 50 amino acid extended β-hairpin in Mtr4 that spans both RecA domains and appears to stabilize the RecA1-RecA2 structure (Jackson et al., 2010; Weir et al., 2010) (Figure 5-1). It is unclear whether a similar structure exists in Ski2 since the corresponding sequence was removed to facilitate crystallization. However, the N-terminus of Ski2 is required for Ski complex formation (see discussion below), suggesting a role in mediating protein-protein interactions.

In Brr2, a helix-hairpin-helix (HhH) domain and Fibronectin type III (FN3) domain flank the C-terminus of the ratchet domain. These three domains are jointly referred to as a Sec63 domain (Ponting, 2000), although the ratchet domain may be more appropriately considered part of the helicase core (Hahn and Beggs, 2010). A similar HhH domain is observed in Hel308 where it interacts with DNA emerging from the helicase core (Buttner et al., 2007) and acts as an autoinhibitory “molecular brake” to repress helicase activity (Richards et al., 2008). FN3 domains are generally associated with protein-protein interactions (Bloom and Calabro, 2009).

Brr2 and Slh1 both contain two copies of the core helicase-Sec63 module, although only the first module is believed to have functional helicase activity (Kim and
Rossi, 1999). The role of the second module is uncertain, but has been suggested to play a role in protein-protein interactions (Liu et al., 2007; Pena et al., 2009; van Nues and Beggs, 2001) and was shown to regulate Brr2 helicase activity (Pena et al., 2009), and may therefore be considered an accessory domain to the Brr2/Slh1 helicase core.

**Helicase Dynamics in Mtr4**

Because multiple structures of Mtr4 are available (apo and ADP/RNA-bound), it is possible to get a sense for the range of motions accessible to the helicase. Since both structures were determined at about the same time, no comparison of the structures has been published to date. Here we describe the major differences between the two structures.

The first observation is that while extensive contacts place restraints on the helicase core, the core domains retain moderate flexibility. The net effect of RNA binding to Mtr4 is a collapse of the entire helicase core around the RNA (Figure 5-2). The RecA domains engage the RNA backbone through the conserved sequence motifs Ia, Ib, IV and V. This binding is accompanied by shifts that bring the RecA domains closer together. On the opposite side of the RNA, the ratchet helix shifts ~ 4.1 Å to interact directly with the RNA bases. Larger motions (up to 5.5 Å) and helical rearrangements are observed throughout the ratchet domain in response to RNA binding (Figure 5-2).

The arch domain is also mobile. This is particularly true for the fist. Compared to the apo structure, rotation occurs in the RNA-bound structure at the “wrist” between the arms and the fist. In one of the molecules observed in the RNA-bound crystal, significant disorder is observed in the electron density which prevented modeling of the
fist, suggesting substantial flexibility in this region (Weir et al., 2010). In contrast, the “elbow” position where the arms make a sharp bend appears to be quite fixed and the arms move as a rigid body (Figure 5-2C). Rotation of the arms with respect to the helicase core occurs at the junction with the winged helix domain. The different arch conformations observed in the apo and RNA bound states are not a function of direct contact with RNA since no interaction is observed in the structure, but may be a result of indirect long-range interactions with the helicase core.

An important conclusion from comparison of the Mtr4 structures is that each domain appears to rotate independently with respect to the other domains (Figure 5-2A). Furthermore, the winged-helix domain is strategically positioned to act as the central hub around which all of the other domains rotate. The winged-helix domain can be thought of as a three-way junction with the RecA domains attached at the N-terminal end, the arch domain inserted in the middle, and the ratchet domain attached to the C-terminus.

Winged-helix domains are commonly utilized to bind DNA and have also been observed in protein-protein interactions (Gajiwala and Burley, 2000). Among Ski2-like helicases, the domain appears to be structurally conserved. However, no general function has been ascribed to this domain other than maintaining the structural integrity of the helicase core (Woodman and Bolt, 2011). We note that an extended 13 amino acid linker located in Mtr4, Ski2 and Hel308 between RecA2 and the winged-helix traverses the RecA domains and places the winged-helix on the opposite side of the helicase from the RNA entry point. Consequently, the winged helix is not expected to interact with RNA
Figure 5-2. Substrate binding induces conformational changes in Mtr4. (A) Comparison of the apo (Jackson et al., 2010) (gray) and RNA/ADP bound (Weir et al., 2010) (color) Mtr4 structures (left panel). Alignment was performed by superimposing the winged helix domains. The cartoon rendition (middle panel) highlights domain motions. Covalent linkages connecting the winged helix to the RecA2, arch and ratchet domains are indicated. The winged helix also stacks with RecA1. The right panel depicts predicted and observed RNA interactions with the RecA, arch and ratchet domains, demonstrating the ability of Mtr4 to interact with every region of the unwinding RNA (i.e. duplex, fork, ssRNA backbone and bases). The winged helix acts as a molecular hub, potentially coordinating motions between the other domains. (B) Comparison of the RecA2 and ratchet domains in the apo and substrate-bound structures. Structures were aligned as in (A). The Arch, WH and most of RecA1 are removed for clarity. The ratchet helix, motifs Ia, Ib, IV, and V and the β-hairpin collapse towards the RNA substrate upon binding. Conformational shifts are depicted with black arrows and labeled with observed distances. Phosphates are colored the same as interacting helicase motifs. (C) Superposition of the arch domains (left) indicates that the arms adopt a rigid structure. Comparison of the arch domain when the winged helix domains are aligned (right) highlights conformational differences upon substrate binding. The right view includes apo (white) and both molecules from the RNA-bound structure (green and dark gray).
in a significant manner. (Minimal interactions are observed between the winged-helix DNA in the Hel308 structure [Buttner et al., 2007]). Instead, it is poised to coordinate and potentially regulate motions between the other domains, each of which interact extensively with RNA.

A detailed understanding of the conformational dynamics employed by the helicase during substrate recognition and processing will require additional structural information. Based on the differential affinity of the Mtr4 arch for various substrates, it seems likely that substrate specific information could be communicated throughout the helicase. For example, given the length and positioning of the arch domain, subtle changes in RNA sequence or structure detected by the fist could have significant conformational and functional effects on the helicase core. Conversely, specific features sensed by the ratchet helix or RecA domains could be easily communicated to the protein surface.

RNA Processing and Degradation

Activation of the exosome by Mtr4 and Ski2

A recent review describes the role of RNA helicases in pre-mRNA splicing, including the integral role of Brr2 in the spliceosome (Cordin et al., 2012). Here we discuss the role of Mtr4 and Ski2 in exosome-mediated RNA decay. The exosome is a large multi-protein ring-like structure involved in the processing and degradation of a wide variety of nuclear and cytosolic RNAs (Butler, 2002; Doma and Parker, 2007; Januszyk and Lima, 2010; Lebreton and Seraphin, 2008; Lykke-Andersen et al., 2009;
Lykke-Andersen et al., 2011). Multiple nucleases are associated with the exosome, including Rrp44 (3’-5’ exonuclease and endonuclease) and Rrp6 (3’-5’ exonuclease, associated with the nuclear exosome) (Butler, 2002; Lykke-Andersen et al., 2011). Exosome activation requires Mtr4 in the nucleus and Ski2 in the cytosol (Schaeffer et al., 2010; Vanacova and Stefl, 2007). Activation generally occurs in the context of larger assemblies, including the Mtr4-mediated TRAMP and NEXT complexes and the Ski2-mediated Ski complex (Anderson and Wang, 2009; Lubas et al., 2011; Schaeffer et al., 2010; Wolin et al., 2012). Mtr4 may also independently activate the exosome for processing of some substrates, such as the 5.8S rRNA (Allmang et al., 1999; Bernstein and Toth, 2012; de la Cruz et al., 1998; LaCava et al., 2005).

How do Mtr4 and Ski2 stimulate exosome activity? One simple explanation is that unwinding of RNA secondary structures and resolution of ribonucleoprotein complexes by Mtr4/Ski2 makes the 3’ end of the RNA accessible to the exosome (Anderson and Wang, 2009; Houseley and Tollervey, 2009). The exosome may also utilize the translocation activity provided by the helicase motor to effectively degrade substrates (Anderson and Wang, 2009). Extensive conservation is observed along the base of Mtr4 and Ski2 where RNA exits the helicase core (Halbach et al., 2012; Jackson et al., 2010; Weir et al., 2010), whereas the base of Hel308 is much less conserved. This suggests that the conservation is related to a function common to Mtr4 and Ski2, but distinct from canonical helicase activity. One possibility is that Mtr4 and Ski2 interact directly with the exosome core, which is roughly similar in dimensions to the base of Mtr4 and Ski2 (Jackson et al., 2010). A second possibility is that other proteins directly
interact with the base of Mtr4 and Ski2 and mediate association with the exosome. Mtr4 and Ski2 association with the exosome has been demonstrated in vivo through a variety of co-immunoprecipitation experiments in yeast (Araki et al., 2001; Peng et al., 2003) and human (Chen et al., 2001; Lubas et al., 2011; Schilders et al., 2007; Tomecki et al., 2010). Although direct interactions with the exosome core have yet to be demonstrated in vitro, human Mtr4 appears to directly interact with the exosome associated components Mpp6 and Rrp6 (Lubas et al., 2011; Schilders et al., 2007).

Maturation of the 5.8S rRNA suggests that the role of Mtr4 in exosome activation may be more complex than simply providing an accessible 3’ end. Conversion of the 7S precursor rRNA to the mature 5.8S rRNA involves exonucleolytic processing by Rrp44, followed by further processing by Rrp6 (Bernstein and Toth, 2012; de la Cruz et al., 1998). Both of these steps require Mtr4. Remarkably, when the arch domain is removed from Mtr4 (“mtr4-archless”), the 7S rRNA is effectively degraded to the 5.8S+30 rRNA precursor (a step associated with Rrp44 activity) at which point processing stalls, similar to what is observed in a Rrp6 knockout (Jackson et al., 2010). This indicates that Rrp6 activity is impaired and further suggests that Mtr4 is able to partition substrates between Rrp44 and Rrp6 in an arch-dependent manner. Recombinant Rrp6 completely degrades a naked 7S rRNA substrate in vitro (Callahan and Butler, 2010). However, a separate study showed that TAP-tag purified Rrp6 from S. cerevisiae cannot degrade 7S rRNA or tRNA substrates unless Mtr4 is present. Furthermore, removal of the Mtr4 fist or the entire arch inhibits RNA degradation (Holub et al., 2012). These data suggest that Rrp6 association with the core exosome (which co-purified with TAP-tagged Rrp6) has an
inhibitory effect on Rrp6 that is overcome by the Mtr4 arch domain. Additional studies are needed to elucidate the precise mechanisms involved in exosome activation.

**Substrate recognition by Mtr4/Ski2-mediated macromolecular complexes**

The fundamental biological problem for exosome dependent decay pathways is distinguishing the correct RNA substrate from all of the other cellular RNAs. Inadvertent removal of needed RNAs or retention of unwanted RNAs is likely to have deleterious consequences for cellular function. Consequently, substrate recognition must be specific and carefully regulated. It is therefore not surprising that a variety of multi-protein complexes have been identified that play a critical role in targeting substrates to the exosome. Significantly, Mtr4 or Ski2 are vital components of these RNA targeting complexes (Figure 5-3). For the most part, a detailed molecular description for these complexes is severely lacking. Our current understanding of these assemblies is summarized below. In each case, important structural questions include: What are the structures of the individual components and how do they assemble to form a larger complex? How does complex formation affect the activity/function of individual components? How are RNA substrates recognized by the complex? What is the trajectory of the RNA through the complex? How are RNAs ultimately delivered to the exosome?

**TRAMP (Trf4/5-Air1/2-Mtr4 polyadenylation complex)**

The best characterized exosome activating complex is the nuclear TRAMP complex (Anderson and Wang, 2009; Bayne et al., 2007; Jensen and Moore, 2005).
**Figure 5-3.** Exosome activating complexes are mediated by Mtr4 and Ski2. Existing crystal structures are indicated (see text for details). Exosome structures also exist (Buttner et al., 2005; Liu et al., 2006; Lorentzen et al., 2007; Lorentzen et al., 2005; Navarro et al., 2008), but are not shown.

TRAMP is a 3-protein complex composed of Mtr4, a poly(A) polymerase (Trf4 or Trf5), and an RNA binding protein containing five zinc knuckles (Air1 or Air2) (Houseley and Tollervey, 2006; LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005). The complex has been described in yeast (Keller et al., 2010; LaCava et al., 2005; Vanacova et al., 2005; Wyers et al., 2005) and human (Lubas et al., 2011), and individual protein components are found throughout eukaryotic species (Anderson and Wang, 2009; Cheng
et al., 2005; Etheridge et al., 2009; Lange et al., 2011; Nakamura et al., 2008; Schmid et al., 2009), suggesting that the complex is found throughout eukaryotes. In *S. cerevisiae*, two forms of TRAMP have been described, TRAMP4 (Trf4, Air2 and Mtr4) and TRAMP5 (Trf5, Air1 and Mtr4) (Houseley and Tollervey, 2006; LaCava et al., 2005). These TRAMP complexes localize differentially in the nucleoplasm (TRAMP4) and nucleolus (TRAMP5) (Dez et al., 2007; Rougemaille et al., 2007; Wery et al., 2009), but they appear to be functionally redundant (Wyers et al., 2005). Deletion of Air1 or Air2 results in only a slight growth defect, whereas the Air1/Air2 double deletion is either synthetically lethal or results in a slow growth phenotype (Inoue et al., 2000; LaCava et al., 2005). Similarly, no growth defects are observed upon deletion of Trf4 or Trf5, but the double deletion is lethal (Wyers et al., 2005). Given the high sequence identity (55% identity for Trf4/5 [Castano et al., 1996]; 45% identity for Air1/2 [Inoue et al., 2000]) and functional redundancy for each TRAMP homolog, it is expected that the TRAMP4 and TRAMP5 complexes will also be similar.

TRAMP adds a short (~5-nt) poly(A) tail to the 3’ end of RNA substrates (Jia et al., 2011; Wlotzka et al., 2011), which is preferentially unwound by Mtr4 (Jia et al., 2012) and subsequently degraded by the exosome. Remarkably, Mtr4 directly modulates the polyadenylation activity of Trf4. In the presence of Mtr4, Trf4 polymerase activity is enhanced until the RNA has 5 adenosines at the 3’ end, at which point Trf4 activity is severely inhibited (Jia et al., 2011). Similarly, Mtr4 helicase activity is enhanced simply by binding to Trf4-Air2 (Jia et al., 2012). It is clear that Mtr4 has the ability to detect the length of a poly(A) tail, which directly influences Trf4 and Mtr4 activity in a TRAMP
context (Jia et al., 2012). A point mutation in the ratchet domain (E947A) disrupts modulation of Trf4 activity by Mtr4 (Jia et al., 2011). In the RNA-bound crystal structure (Weir et al., 2010), this residue is located near one of the RNA bases and potentially helps recognize the poly(A) sequence, but the mechanisms involved in sequence recognition have not been clearly delineated.

Understanding the assembly of the TRAMP complex is an active focus of investigation. Trf4 and Air2 are tightly associated in vitro, remaining in complex above 1M NaCl, while Mtr4 interaction is more dynamic and dissociates from Trf4-Air2 around 500 mM NaCl (Keller et al., 2010; LaCava et al., 2005; Vanacova et al., 2005). A recent crystal structure of the catalytic domain of Trf4 bound to Air2 zinc knuckles 4-5 reveals that these two proteins interact primarily through zinc knuckle 5 and a loop between zinc knuckles 4 and 5 (Hamill et al., 2010). (An NMR structure of Air2 zinc knuckles 1-5 was also recently reported (Holub et al., 2012).) This interaction has been confirmed by mutagenic analysis (Fasken et al., 2011; Hamill et al., 2010; Holub et al., 2012). The site of Mtr4 interactions is less well understood. In vitro pull-down assays from the Conti lab (Weir et al., 2010) and our lab (unpublished data) indicate that the helicase core is sufficient to form a TRAMP-like complex (i.e. the N-terminus and arch domains are not required). Similarly, a domain deletion study shows that TRAMP fails to pull-down with Mtr4 when any of the individual helicase core domains are removed (Holub et al., 2012), although it isn’t clear what effect deletion of individual core domains has on the overall structure or localization. Presumably, the Trf4/5-Air1/2 docking site would be located near the RNA entry site in Mtr4. A few conserved patches on the surface of Mtr4 that are
not found in Ski2 suggest potential interaction regions (Weir et al., 2010), but the location of TRAMP assembly on Mtr4 remains obscure. Recent studies, however, show that the N-terminus of Air1/2 interacts directly with Mtr4 and may provide a bridge between Mtr4 and Trf4/5 (Holub et al., 2012). Earlier yeast 2-hybrid studies indicated that an N-terminal region of Trf4/5 is required for Mtr4 interaction (LaCava et al., 2005). No structures of the relevant regions in Air1/2 or Trf4/5 are currently available.

**NEXT (Mtr4-Rbm7-ZCCHC8)**

Another Mtr4-mediated activator of the exosome is the human nuclear exosome targeting complex (NEXT), which was recently identified by co-immunoprecipitation studies and shown to target promoter upstream transcripts (PROMPTs) for degradation by the exosome (Lubas et al., 2011). Like TRAMP, the NEXT complex contains 3 components including Mtr4, a zinc knuckle protein called ZCCHC8, and a putative RNA binding protein, Rbm7. However, no poly(A) activity is associated with NEXT, indicating that polyadenylation is not required for degradation in this context. In humans, NEXT is localized in the nucleus while TRAMP appears to be confined to the nucleolus. Since Mtr4 is shared by both complexes, it has been suggested that Mtr4 potentially acts as a common scaffold for a variety of protein complexes aimed at shuttling substrates to the exosome (Lubas et al., 2011). It will be interesting to learn whether Mtr4 interacts with each complex (and other yet to be discovered complexes) in a similar manner. For example, does the putative Mtr4-ZCCHC8 interaction resemble that of the TRAMP Mtr4-Air2 interaction?
**FFC (FRQ-FRH complex)**

The FRH (FRQ-interacting RNA helicase) protein from *N. crassa* is a homolog of *S. cerevisiae* Mtr4 (56% sequence identity; 73% similarity) and is an indispensable component of the circadian clock (Baker et al., 2012). FRH forms a 1:2 stoichiometric complex with the Frequency protein (FRQ) that regulates oscillating gene expression (Cheng et al., 2005). The FRQ-FRH complex (FFC) acts within a negative feedback loop that represses transcription of specific genes, including the *frq* allele, by binding and promoting phosphorylation of the white collar complex (WCC) transcription promoter (Baker et al., 2012). FRH is required for the direct interaction between FFC and WCC (Cheng et al., 2005). Moreover, FFC negatively regulates FRQ expression by specifically binding *frq* mRNA and mediating its exosomal decay (Guo et al., 2009).

Mutation of the conserved arginine R806 to histidine allows the formation of the FFC but abolishes interactions with WCC (Shi et al., 2010). Interestingly, the equivalent residue in *S. cerevisiae* Mtr4, R774, is a conserved, solvent exposed residue positioned on the upper region of the fist/KOW motif in the arch (Jackson et al., 2010). Thus, in this system the arch domain appears to play an important role in complex formation.

**Ski complex (Ski2-Ski3-Ski8)**

The Ski complex components are conserved in most eukaryotes including human, and promotes degradation of mRNAs by the cytosolic exosome (Schaeffer et al., 2010). It is composed of Ski2, a tetratricopeptide protein (Ski3), and a WD-40 repeat protein (Ski8) (Brown et al., 2000). The complex was originally described as a trimer (Brown et al., 2000), but subsequent mass spec analysis indicates that it is a tetramer containing two
copies of Ski8 (Synowsky and Heck, 2008). Activation of the exosome also requires an accessory protein, Ski7, that has been shown to physically link the two protein machineries (Araki et al., 2001). A crystal structure of Ski8 reveals a seven-bladed β-propeller that is proposed to function as a scaffold protein (Cheng et al., 2004; Madrona and Wilson, 2004). Yeast 2-hybrid screens and co-immunoprecipitation studies indicate that the N-terminus of Ski2 associates directly with Ski3, while direct interactions with Ski8 or Ski7 are not observed (Wang et al., 2005). While the Ski2 crystal structure provides little information regarding complex assembly (Halbach et al., 2012), forthcoming structures are expected to significantly clarify the architecture of this complex.

**Outlook**

As the above complexes demonstrate, the assembly of different proteins on Mtr4 and Ski2 allow for targeting of a variety of RNA substrates for exosome processing and degradation. The FCC complex also highlights the potential of some complexes to affect protein post-translational modifications. It would not be surprising to discover more complexes that expand the versatility of these helicases. Despite the valuable insight provided by the current structures, we are just beginning to understand how RNA substrates are recognized and interact with Ski2-like RNA helicases. The interaction of these helicases in larger complex assemblies is a critical question that promises to be an active focus of continued research.
References


CHAPTER 6
SUMMARY AND FUTURE DIRECTIONS

Summary

The Ski2-like RNA helicase Mtr4 is a conserved and essential protein that is necessary for proper RNA quality control, RNA processing and gene regulation in eukaryotes. Mtr4 acts as a vital member of large multiprotein complexes, including the TRAMP complex, to activate 3’ end decay of RNA substrates by the exosome. Central to TRAMP function is the Mtr4 poly(A) tail sensing mechanism that regulates both Trf4 polymerase and Mtr4 unwinding activities.

A strong foundation for understanding the role of Mtr4 function in exosome activation and TRAMP regulation has been established by this work. The first x-ray crystal structure of Mtr4 was determined (Chapter 3) (Jackson et al., 2010) which revealed that Mtr4 contains a four domain core (domains include RecA1, RecA2, winged helix and ratchet domain) observed in other Ski2-like and DEAH/RHA-box helicases. The helicase core contains all the features observed in the structure of the DNA helicase Hel308 that is implicated in processive strand unwinding, namely, the conserved motifs of the RecA domains, the β-hairpin and the ratchet helix.

The structural similarity between Mtr4 and other Ski2-like helicases was obscured from sequence analysis by the presence of a large arch-like domain inserted within the winged helix. The arch is poised directly above the putative RNA entry site, suggesting that it may interact with structured regions of RNA. Indeed, the Arch domain binds hypomodified tRNA and pre-rRNA (Figure 6-1). Removal of the arch domain does not
disrupt ATPase, RNA unwinding or TRAMP forming capabilities in vitro (Chapter 3 and data not shown) but does impact exosome activation in vivo. In S. cerevisiae an mtr4-archless mutant displayed impaired growth and the inability to activate some functions of the exosome. A buildup of rRNA processing intermediates 5.8S + 30 rRNA and 5’ ETS was observed in the mtr4-archless strain resembling that of an Rrp6 knockout, suggesting an interaction between Rrp6 and the arch domain.

Conservation analysis demonstrated that the surface of the putative RNA exit site is highly conserved in Mtr4 and its cytosolic homolog Ski2, suggesting a common helicase-exosome interface. Further experimentation is required to explore the possible structural interactions between Mtr4 and the RNA exosome.

A second Mtr4 structure bound to RNA and ADP (Weir et al., 2010) confirmed that Mtr4 RNA binding is analogous to that observed in the Hel308 structure and provided an opportunity for comparison of apo and substrate bound Mtr4 structures. Structure and conservation studies of Mtr4 and other Ski2-like and DEAH/Box helicases identified conserved residues along the ratchet helix likely involved in RNA unwinding (Chapter 4).

Significantly, the residue of the 4th turn of the ratchet helix is identically conserved through all helicases known to contain a ratchet helix, although residue identity is variable helicase to helicase. Mutagenic studies of this position (E1033) in Mtr4 demonstrated that the identity of the residue directly influences in vitro unwinding rate constants, does not impact substrate binding but is involved in RNA dependent ATPase enhancement (Chapter 4). The E1033A mutant displayed an increased, and the
E1033W mutant displayed a decreased unwinding rate constant compared to wild-type Mtr4 on a 16 bp substrate with a 6 nt poly(A) overhang. Interestingly, the functional affinity for the poly(A) substrate was less than that observed in wild-type for both mutants suggesting that E1033 may act to sense features of the poly(A) tail. Further studies of the residues along the RNA binding path are currently underway and are described in more detail below. These studies aim to determine the molecular features and mechanisms utilized by Mtr4 to sense the identity of RNA substrates and translate sequence characteristics into enhancement of Mtr4 unwinding and modulation of Trf4 polymerization.

The Mtr4 apo and substrate bound structures provide the greatest amount of structural information regarding Ski2-like RNA helicases available to date. Within Chapter 5 we have reviewed the structural similarities between Ski2-like RNA helicases and highlight the conformational changes observed in Mtr4 upon substrate binding. Notably, analysis of conformational differences between Mtr4 structures illuminates the role of the winged helix domain as a molecular hub that coordinates domain motions within the helicase during RNA interaction events. Because of its central location, the winged helix domain is a prime target for interactions that might modify/regulate Mtr4 activity such as posttranslational modifications or protein-protein interactions. A common theme among all Ski2-like RNA helicases is their association with large protein complexes. Mtr4 is a member of several exosome activating complexes (TRAMP4, TRAMP5, NEXT and FFC), and the cytosolic homolog of Mtr4, Ski2, forms the exosome activating Ski complex (Chapter 5).
between all these assemblies is a Ski2-like helicase with an arch domain. How these complexes assemble and how Mtr4 activity is employed to activate the exosome remains to be determined through future biochemical and structural studies.

The structural, conservational and biochemical analysis of Mtr4 described in this dissertation has provided a strong molecular framework for studies aimed at exploring Mtr4 activity and its fundamental role in RNA processing and surveillance. However, many molecular details of Mtr4 function remain unanswered. The foremost unanswered questions are: How are RNA substrates that require decay selected over functional RNAs? How does Mtr4 act on RNAs during translocation and/or RNA sensing? What features of Mtr4 are required for assembly into larger complexes? How does Mtr4 activate the exosome? The future directions that should be pursued along with descriptions of existing preliminary data and relevant citations are described below.

**Future directions**

**How does Mtr4 specifically interact with RNA?**

The alignment of the Mtr4 structure the DNA bound Hel308 structure (Chapter 3) provided the first insight of how Mtr4 might bind to an RNA substrate (Buttner et al., 2007). Although the RNA bound Mtr4 structure does reveal some details regarding Mtr4 binding, the single stranded 5 nt long oligo(A) observed in the density only provides binding information for the RecA folds and a few residues of the ratchet domain (Weir et al., 2010). To obtain information regarding strand splitting and interactions of duplex substrate with the arch domain, a crystal structure of Mtr4 with a structured RNA
substrate such as a tRNA or rRNA would be ideal. Efforts to produce Mtr4-RNA crystals have been carried out with some promising outcomes (Figure 6-2).

*Crystallization trials of Mtr4 with structured nucleic acid promise future structures.* In an attempt to produce RNA at the quality and quantity required for crystallization trials, an *in vitro* transcription reaction protocol for tRNA\textsubscript{\textit{i} met}, pre-rRNA and other RNAs was developed (Chapter 2). Several Mtr4 constructs bind these structured RNAs (Figure 6-1) and preliminary crystal screens have been attempted with tRNA\textsubscript{\textit{i} met} and Mtr4 constructs, although no crystals have been observed to date.

The difficulty in obtaining crystals may be due to heterogeneity of the 3’ and 5’ ends. Often heterogeneous ends of RNA substrates inhibit crystal formation and/or growth. The remediation to homogenous substrates greatly improves the possibilities of crystal production with RNA (Doudna, 1997). The Rob Batey and James Kieft labs have created plasmids and optimized a RNA native affinity purification method that produces RNAs with homogenous ends (Batey and Kieft, 2007; Kieft and Batey, 2004). This method incorporates a stem loop into the transcribed RNA that follows a glmS ribozyme used to cleave the RNA in a homogenous manner at the 3’ end. The stem loop is recognized by a His-tagged MS2/MBP protein. Transcribed RNA is purified by incubating the RNA substrate with His-tagged MS2/MBP protein over Ni-resin. The RNA is bound at the stem loop by MS2/MBP and the unwanted products of transcription are washed from the resin. The RNA is eluted by treating the ribozyme with glucosamine-6-phosphate, which induces cleavage. We have recently acquired the
Figure 6-1. The core and fist of Mtr4 bind structured RNA. (A) Mtr4 full length, archless and fist proteins were incubated with $^{32}$P labeled tRNA$_{i\text{met}}$ at concentrations of 0.45 µM, 1.5 µM, 45 µM and 150 µM after which samples were run on a 15% polyacrylamide gel. Binding to Mtr4 is observed as a shift in position on the gel and is indicated left. (B) Samples were prepared and run on a polyacrylamide gel as in A except that a $^{32}$P labeled rRNA substrate was used.
plasmids for the expression and purification of the MS2/MBP protein as well as the transcription vector carrying the stem loop and ribozyme sequence. The method remains to be tested in the lab but promises to yield large volumes of pure RNA with homogenous 3’ ends.

It has been shown by the Conti lab (Weir et al., 2010) and our lab (Figure 6-1) that full length, archless and fist Mtr4 constructs can bind structured RNAs. Crystallization trials with RNA and Mtr4 constructs, including the fist of Mtr4 have been attempted, although many more avenues of trial with Mtr4 constructs and RNA species remain to be probed.

*Mtr4 crystallizes in the presence of nucleic acid.* It is encouraging to note that full length Mtr4 has produced several crystal forms in the presence of RNA (Chapter 2, Table 2-1). The most promising of these preliminary crystals are those of Mtr4 with a DNA hairpin (see Chapter 2 for methods) because they are reproducible, behave reasonably well in cryosolution and diffracts x-rays (10-20 Å). Mtr4-DNA hairpin crystals are grown in 0.1 M SPG (succinic acid, sodium dihydrogen phosphate, and Glycine) buffer at pH 8-9 and 25% PEG 1500 at a 1:1 protein:well solution ratio. Analysis of the crystals showed that degradation was occurring during the crystallization process (Figure 6-2 C). Mass spec analysis showed that that Mtr4 protein within the crystals lacked the N-terminal region and were cleaved between residues 360 and 390 of Mtr4. The region between 360 and 390 is called the loop region because in the structure of Mtr4 it forms a non-structured loop. Mtr4 constructs with the N-terminus and loop removed have crystallized in the presence of hairpin nucleic acid (Chapter 2 Table 2-2). These crystals
Figure 6-2. Mtr4 protein crystal growth in the presence of nucleic acid. (A) An Mtr4 crystal grown in the presence of a 10 nt poly(A) oligomer. Crystallization conditions are displayed below. (B) Crystals grown in the presence of the 37 nt DNA hairpin, conditions below. (C) Gel demonstrating the proteolysis occurring in the Mtr4 + DNA hairpin crystals and drop solution.

grow well at room temperature and 13°C, however, they do not diffract beyond 10 Å.

Further optimization of this crystallization condition may allow for diffraction to higher resolution. The duplex nature of the substrate promises to provide many important details to Mtr4 RNA binding and unwinding function if a co-crystal structure with the hairpin and Mtr4 is determined. The most obvious place to begin optimization is with the nucleic acid substrate. Crystals were obtained primarily with a DNA sequence but subsequent trials demonstrated that in vitro derived RNA hairpin will also promote crystallization.

Modification of the length and sequence of the 3’ overhang would be a good place to begin optimization efforts, followed by modification of the hairpin region. It is possible that the hairpin causes some heterogeneity in the RNA. An approach where the
hairpin is removed and two nucleic acid strands are annealed for crystallization may improve diffraction by eliminating the possibility of multiple annealing options in the substrate. Also, a minimal region of Mtr4 has been defined for crystallization. Further exploration of Mtr4 N-terminal truncations may aid in improving the diffraction quality of the crystals. In summary, Mtr4 does crystallize with nucleic acid (Figure 6-2) and further optimization of the crystallization conditions listed in Table 2.2 in the methods chapter may improve the diffraction needed to obtain structural data of Mtr4 with a structured nucleic acid substrate.

How does Mtr4 couple ATPase activity to unwinding?

Mtr4 possesses RNA dependent ATPase and ATP dependent RNA unwinding activities (Bernstein et al., 2008; Wang et al., 2008). The Hel308 structure bound to a partially unwound duplex DNA (Buttnetn et al., 2007) provides the molecular model for translocation generally accepted in the field for super family 2 helicases containing a beta-hairpin and ratchet helix. The model suggests that Mtr4 works along the backbone of the RNA strand in an inch worm mechanism promoted by motions caused by ATP binding, hydrolysis and nucleotide release. The model also suggests that motions of domain 2 are propagated to domain 4 via surface contacts to provide some processivity to the reaction. The comparison of apo and the RNA/ADP bound structures of Mtr4 allowed for analysis of possible conformations that occur upon substrate binding (Weir et al., 2010). A comparison of the structures allowed for the identification of residues within the core that are likely important for coupling ATPase activity to RNA unwinding.
Several identified residues were mutated, expressed, and purified as performed for wild type protein. Included were residues Q154 and R547 that interact directly with the adenosine ring of ADP. Q154 binds the N6 and N7 positions of the adenosine ring in a bidentate manner and R547 swings about 2 Å to stack with the adenosine ring. The slight motion between domain 1 and 2 caused by R547 stacking with nucleotide may be enough to promote translocation. Indeed, mutation of Q154 and R457 to alanine significantly

![ATPase 400 nM polyA RNA](image)

**Figure 6-3.** The ATPase activity of Mtr4 helicase core mutants. The ATPase activity of Mtr4 wild type (WT), and core mutants was determined using a malachite green colorimetric assay. Following methods described in Chapter 4, reactions containing 300 nM of enzyme and indicated concentration of RNA were initiated with 2 mM ATP/MgCl₂ and change in absorbance at 620 nm was observed over time. Shown is the rate of release of inorganic phosphate in mM min⁻¹. The SS-CC sample stands for the S199C/S205C double mutant.
disrupted the RNA dependent ATPase activity of Mtr4 (Figure 6-3). How this disruption affects unwinding remains to be tested.

Several residues that form the conserved motif Ia were targeted for investigation. These included P200, S199, and S205. Alignment with apo Mtr4 demonstrated that P200 may restrict the rotameric flexibility of the residues of the DExH/D-box motif II. It appears that upon RNA binding P200 shifts towards the D262 and E263 residues, sterically inhibiting the rotamers modeled in the apo structure. To test the hypothesis that the bulk of P200 is required for ATP hydrolysis P200 was mutated to a glycine. The P200G mutation demonstrated no ATP hydrolysis activity (Figure 6-3). Serine residues 199 and 205 that flank either end of P200 are rotamerically modeled pointing at each other with their representative hydroxyl groups. The proximity of these residues appears to differ between the apo and substrate bound structures and it was hypothesized that formation of a disulfide linkage between the residues might constitutively activate or deactivate the ATPase/unwinding activity. The presence of RNA appears to stimulate ATPase activity in the S199C/S205C (SS-CC) double mutant in a reducing environment (2mM DTT) (Figure 6-3). The mutant remains to be tested in an oxidative environment, concurrent with wild type to analyze if the mutation has truly affected the coupling of RNA binding to ATPase activity.

Helicase core residues that may be involved in interactions between domains 2 and 4 were also identified, including E420, K484, and H476. These three residues were mutated to alanines and the H476A mutant showed little unwinding or ATPase activity when tested. E420A and K484A remain to be assayed for activities. The ratchet helix
residue R1026 is also implicated in motion transfer from domain 2 to domain 4 as well as Poly(A) sensing (see below) but mutations of this residue have yet to be made.

Several helicase core mutants have been constructed and provide the future researcher with a wealth of constructs that can be analyzed biochemically. Ideally crystal structures could be pursued with mutants that show interesting biochemical function in an effort to elucidate the detailed molecular mechanism of ATPase and unwinding activities.

**How does Mtr4 sense the poly(A) tail of RNA substrates?**

In Chapter 4 it was shown that residue E1033 regulates the unwinding activity of Mtr4 and may be involved in sensing substrate characteristics. To further characterize the residues that lie along the RNA binding path, a R1030A mutant has been constructed, expressed and purified and is under current kinetic investigation. R1030A displays reduced unwinding compared to wild-type (Figure 6-4B), confirming a functional defect suggested by the *in vivo* studies described in Chapter 4. We aim to determine the unwinding rate constants necessary to evaluate this mutant as well as the R1030A/E1033A double mutant and will hopefully wrap our findings into the results described in Chapter 4. Furthermore, in a search for what sequence characteristics contribute to the sensing mechanism we probed a RNA substrate with a CCACCA 3’ end overhang. Mtr4 unwound this substrate very poorly (Figure 6-4A) suggesting that the mechanism by which Mtr4 recognizes sequence characteristics is more complex than simply selecting poly(A) over non-poly(A). An in depth analysis of several overhangs
Figure 6-4. Probing the substrate sensing mechanism of Mtr4. (A) Wild type Mtr4 displaces a $^{32}$P labeled 16 nt top strand from a poly(A) overhang strand in a robust manner (left). Mtr4 does not displace the top strand from CCACCA overhang strand (right). (B) The R1030A mutant displays slow unwinding at 800 nM concentration compared to other Mtr4 constructs. The curves show the calculated fraction unwound over time of $^{32}$P displacement from a strand containing a poly(A) overhang with E1033A, wild type (WT), E1033W and R1030A constructs. The inset is the displacement observed in R1030A.
consisting of cytosines and adenosines is proposed. The analysis would describe the minimum location and identity of nucleotides in the overhang required for unwinding enhancement. Our discovery that a sequence containing cytosines dissuades unwinding suggests that a mechanism may be present that causes avoidance of certain substrates. The CCA sequence is of particular interest because tRNAs are appended with this feature and it may be a mechanism to avoid selection by Mtr4 and/or TRAMP mediated degradation pathways.

*How does Mtr4 assemble into large exosome activating complexes?*

Mtr4 is the central component to many exosome activating complexes. These complexes are not all equivalent in architecture yet all select specific RNAs for degradation by the exosome. Future studies directed at obtaining crystals structures of these intricate machineries are necessary to understand the fundamental questions of how these complexes assemble and how the individual components act in concert to recognize substrate and activate the exosome. Currently we have constructs of Mtr4 from *S. cerevisiae*, *S. pombe* and *H. sapiens*. Significant effort has been made to obtain TRAMP crystal structures however much can be done to optimize the screens including the use of truncation mutants and the addition of RNA substrates. To date no trials of Mtr4 from *S. pombe* with Cid14 and/or other *S. pombe* TRAMP components has been attempted. Human Mtr4 provides a wealth of opportunities for obtaining a complex structure in that hMtr4 not only interacts with a homologous TRAMP complex but also with the NEXT complex. We currently have DNA for hMtr4 but are having and the DNA for the human
NEXT complex components Rbm7 and Zcche8. One notable observation regarding the NEXT complex is that within the primary manuscript a WD-40 protein was also identified to associate with Mtr4. Interestingly Ski8 from the Ski2 complex is a WD-40 protein. We have expressed and purified Ski8 from *S. cerevisiae* in our lab and it would be interesting to explore if Ski8 can bind Mtr4 and may in fact stabilize complex interactions.

A very interesting Mtr4-mediated complex is the FFC (FRH-FRQ complex) observed in the circadian clock of *N. crassa*. We do not have the DNA for the *N. crassa* homolog of Mtr4, FRH. However, the structure of this complex should be pursued. Circadian clocks are observed throughout all eukaryotes, and while it is unclear if Mtr4 is involved in the circadian rhythms of higher organisms a structure of the FFC would provide needed insight and significant impact to the circadian rhythm field of study.

**Conclusion**

Many opportunities are available for future research with Mtr4 and Mtr4 homologs. Mtr4 from *S. cerevisiae* is amenable to crystallization with nucleic acid and with the recent development of the *in vitro* transcription technique in the Johnson lab the sky is the limit for Mtr4-nucleic acid crystallization trials. A myriad of Mtr4 mutants probing core functions have been made and provide the future researcher with a wealth of avenues for probing the translocation and sensing mechanisms of Mtr4. Lastly Mtr4 assembles into several different protein complexes providing much opportunity for protein complex crystallography and biochemical analysis. It is hoped that this
dissertation can provide a framework and direction for future studies that will answer the important remaining questions of Mtr4 function.

References


APPENDICES
APPENDIX A

Mtr4 and Ski2 Sequence Alignment
<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S. cerevisiae Mr4</th>
<th>S. pombe Mr4</th>
<th>M. oryzae Mr4</th>
<th>N. crassa Mr4</th>
<th>N. scitia Mr4</th>
<th>D. melanogaster Mr4</th>
<th>D. rerio sk12</th>
<th>A. egypri M4r</th>
<th>C. elegans M4r</th>
<th>A. thaliana sk12</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

175
<table>
<thead>
<tr>
<th>Organism</th>
<th>Mtr4</th>
<th>Ski2</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. cerevisiae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. pombe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. crassa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. sapiens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. maritimus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. melanogaster_6551</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. melanogaster_3805</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. gambiae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. elegans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. thalana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. pombe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. crassa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. sapiens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. maritimus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. melanogaster_6551</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. melanogaster_3805</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. gambiae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. elegans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. thalana</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A-1.** Alignment of Mtr4 and Ski2 sequences. An alignment of the Mtr4 and Ski2 proteins from several organisms is shown. Residues that are highlighted in yellow are similar; those highlighted in red are strictly conserved. The numbering above the sequences corresponds to the *S. cerevisiae* Mtr4 protein.
APPENDIX B

Permissions
NPG author licence policy

This publishers’ policy applies to all journals published by the Nature Publishing Group (NPG), including the Nature journals. NPG does not require authors of original (primary) research papers to assign copyright of their published contributions. Authors grant NPG an exclusive licence to publish, in return for which they can reuse their papers in their future printed work without first requiring permission from the publisher of the journal. For commissioned articles (for example, Reviews, News and Views), copyright is retained by NPG.

When a manuscript is accepted for publication in an NPG journal, authors are encouraged to submit the author’s version of the accepted paper (the unedited manuscript) to PubMedCentral or other appropriate funding body’s archive, for public release six months after publication. In addition, authors are encouraged to archive this version of the manuscript in their institution’s repositories and, if they wish, on their personal websites, also six months after the original publication.

For open access content published under a Creative Commons license, authors can replace the submitted version with the final published version at publication.

In all these cases, authors should cite the publication reference and DOI number on the first page of any deposited version, and provide a link from it to the URL of the published article on the journal’s website (see publications A-Z index).

This policy has been developed by NPG’s publishers to extend the reach of scientific communication, and to meet the needs of authors and the evolving policies of funding agencies that wish themselves to archive the research they fund. It is also designed to protect the integrity and authenticity of the scientific record, with the published version on nature.com clearly identified as the definitive version of the article.

NPG recognizes the balance of rights held by publishers, authors, their institutions and their funders (Zwolle principles, 2002), and has been a progressive and active participant in debates about access to the literature. In 2002, NPG was one of the first publishers to allow authors to post their contributions on their personal websites, by requesting an exclusive licence to publish, rather than requiring authors to transfer copyright. NPG actively supports the self-archiving process, and continues to work with
authors, readers, subscribers and site-license holders to develop its policy.

**Manuscript Deposition Service**

To help authors fulfill funder and institutional mandates, authors of original research manuscripts accepted for publication in many journals published by NPG can opt in to NPG's Manuscript Deposition Service.

**Creative Commons licence**

In December 2007, NPG introduced the Creative Commons attribution-non commercial-share alike unported licence for those articles in Nature journals that are publishing the primary sequence of an organism's genome for the first time. In summary, under this type of licence, readers are free to share (copy, distribute and transmit) and remix (adapt) the contribution under these conditions: attribution in the manner specified by the author or licensor; non-commercial — readers and users cannot re-use the material for commercial purposes; and share alike — if readers or other users alter, transform or build upon the work, they may distribute this work only under the same or similar licence to this one. Further details of the licence and of the legal code are available.

A Nature Editorial introducing this service can be read here: comments are welcome at Nautilus, our author blog.

All articles in NPG's open-access journal Molecular Systems Biology are published under a Creative Commons attribution-non commercial-share alike unported licence or a non commercial-no deris licence, at the choice of the authors. Please see the Molecular Systems Biology website for more information, and visit The Seven Stones, the Molecular Systems Biology blog, for online discussion of this policy.

Articles published at Nature Communications under open access and all those published at Scientific Reports are published under the same Creative Commons licences as detailed for Molecular Systems Biology. Please see the licence section on the Nature Communications and Scientific Reports websites for more information.

**PowerPoint slides**

Some Nature journals offer downloadable PowerPoint slides of figures in articles. Subscribers and registrants from licensed institutions (register for free at http://www.nature.com/register) may download slides in PowerPoint format for non-profit, educational use, provided the content is not modified and full
credit is given to the author and journal. Any other re-use requires the permission of Nature Publishing Group (permission can be obtained online by following the ‘Rights and Permissions’ link on the right-hand side of the article’s Full Text (HTML) and Abstract web pages). For images credited to other sources, permission must first be gained from the source in the figure legend or credit line.

**General information**

General information about NPG’s licencing policies can be found at the reprints and permissions website.

NPG’s publishers welcome authors’ feedback and comments on the NPG author licence policy. Please use the email link provided and address your message for the attention of the publisher of the journal in which your paper appears.

Licence to publish forms are available at the website of the journal that is publishing the article.

If your paper is being published in *Nature*, the forms can be found at the journal’s forms and declarations page.

If your paper is being published in any other Nature Publishing Group journal, please see the home page of the journal concerned to navigate to the licence form.

---

**Terms for reuse of archived manuscripts**

Nature Publishing Group (NPG) wishes to clarify the permitted use of NPG-published material that has been archived into academic repositories such as institutional repositories, PubMed Central and its mirror sites, where NPG holds copyright, or an exclusive license to publish. NPG reserves the right to change these usage terms, subject to copyright law and its exemptions.

**Summary**

Users may view, print, copy, download and text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use.

**Conditions of use**

Articles published by Nature Publishing Group (NPG) which are made available through academic repositories remain subject to copyright. Any reuse is subject to permission from NPG. The following restrictions on re-use of such articles apply:
Academic research only
1. Archived content may only be used for academic research. Any content downloaded for text based experiments should be destroyed when the experiment is complete.

Re-use must not be for Commercial Purposes
2. Archived content may not be used for purposes that are intended for or directed towards commercial advantage or monetary compensation by means of sale, resale, licence, loan, transfer or any other form of commercial exploitation ("Commercial Purposes").

Wholesale re-publishing is prohibited
3. Archived content may not be published verbatim in whole or in part, whether or not this is done for Commercial Purposes, either in print or online.
4. This restriction does not apply to reproducing normal quotations with an appropriate citation. In the case of text-mining, individual words, concepts and quotes up to 100 words per matching sentence may be reused, whereas longer paragraphs of text and images cannot (without specific permission from NPG).

Moral rights
5. All re-use must be fully attributed. Attribution must take the form of a link - using the article DOI - to the published article on the journal’s website.
6. All re-use must ensure that the authors’ moral right to the integrity of their work is not compromised.

Third party content
7. Where content in the document is identified as belonging to a third party, it is the obligation of the user to ensure that any reuse complies with copyright policies of the owner.

Re-use at own risk
8. Any re-use of NPG content is at your own risk and NPG accepts no liability arising from such re-use.

http://www.nature.com/authors/policies/license.html#top
Permission Letter

Bradley J. Hintze
Duke University
Department of Biochemistry
Durham, NC, 27710

June 30, 2012

This letter grants my permission to Ryan N. Jackson to use the following publication in part or in full for inclusion in his Ph.D. dissertation.


Sincerely,

Bradley J. Hintze
Permission Letter

A. Alejandra Klauer
University of Texas Health Science Center-Houston
Department of Microbiology and Molecular Genetics
Houston, TX, 77030

June 30, 2012

Ryan N. Jackson
Utah State University
Department of Chemistry and Biochemistry
Logan, UT 84322-0300

This letter grants my permission to Ryan N. Jackson to use the following publication in part or in full for inclusion in his Ph.D. dissertation.


Sincerely,

[Signature]

A. Alejandra Klauer
Permission Letter

Howard Robinson
Brookhaven National Laboratory
Department of Biology
Upton, NY, 11973-5000

June 30, 2012

Ryan N. Jackson
Utah State University
Department of Chemistry and Biochemistry
Logan, UT 84322-0300

This letter grants my permission to Ryan N. Jackson to use the following publication in part or in full for inclusion in his Ph.D. dissertation.

The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. *EMBO J* 29, 2205-2216

Sincerely,

Howard Robinson

Howard Robinson, Ph.D.
Permission Letter

Ambro van Hoof
University of Texas Health Science Center-Houston
Department of Microbiology and Molecular Genetics
Houston, TX, 77030

June 30, 2012

Ryan N. Jackson
Utah State University
Department of Chemistry and Biochemistry
Logan, UT 84322-0300

This letter grants my permission to Ryan N. Jackson to use the following publication in part or in
full for inclusion in his Ph.D. dissertation.

The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. EMBO
J. 29, 2205-2216

Sincerely,

Ambro van Hoof, Ph.D.
EDUCATION

Ph.D. Utah State University, Logan UT; 2012; Biochemistry
B. S. Utah State University, Logan UT; 2005; Biology

RESEARCH EXPERIENCE

Ph.D. Candidate (2007–summer 2012)
Utah State University
Department of Chemistry and Biochemistry
Mentor: Sean J. Johnson, Ph.D.

My dissertation work structurally and functionally characterized the RNA helicase Mtr4 that plays a central role in nuclear 3'-5' exosome-mediated RNA decay and processing. My initial work determined the full length x-ray crystal structure of Mtr4, revealing a novel arch domain that is necessary for proper 5.8S rRNA processing. Currently I am probing the structural features of Mtr4 that are important for identifying poly-adenylated RNA substrates.

M.S. Candidate (2005–2007)
Utah State University
Department of Chemistry and Biochemistry
Mentor: Joseph K. K. Li, Ph.D.

As a M.S. degree candidate I determined the effects of nucleoside analogs on the replication cycle of Blue Tongue Virus using the techniques of animal cell culture, RNA isolation and quantitative Real Time PCR.

Undergraduate Research Associate (2004–2005)
Utah State University
Department of Chemistry and Biochemistry
Mentor: Joseph K. K. Li, Ph.D.

PUBLICATIONS

Ryan N Jackson, A Alejandra Kauer, Lacy Taylor, Anna K Lytle, Lindsey K Lott, Megi Rexhepaj, Ambro van Hoof and Sean J Johnson (2012). Conserved features in the Mtr4 ratchet helix regulate RNA helicase activity. (To be published)

Ryan N Jackson, A Alejandra Klauer, Bradley J Hintze, Howard Robinson, Ambro van Hoof and Sean J Johnson (2010). The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. EMBO J, 29, 2205-16.

RESEARCH TALKS

1st Annual Biochemistry Department Retreat – USU – Logan, UT 2011
Graduate Student Symposium – USU – Logan, UT (*Talk Award) 2010

POSTER PRESENTATIONS

FASEB Summer Research Conference, Nucleic Acid Enzymes, Snowmass, CO 2012
1st Annual Biochemistry Department Retreat – USU – Logan, UT 2011
FASEB Summer Research Conference, Helicases, Steamboat Springs, CO 2011
Intermountain Graduate Student Symposium – USU – Logan, UT (*Poster award) 2011
American Crystallographic Association, 2010 Annual Meeting – Chicago, IL 2010
American Crystallographic Association, 2009 Annual Meeting – Toronto, ON 2009
237th National Meeting American Chemical Society– Salt Lake City, UT 2009
Graduate Student Symposium – USU – Logan, UT (*Poster award) 2009
Graduate Student Showcase – USU – Logan, UT 2008
USU Biology Department Retreat – USU – Logan, UT 2007
Student Showcase – USU – Logan, UT (*Poster award) 2005
19th National Conference on Undergraduate Research – Lexington, VA 2005

RESEARCH FUNDING

Center for Integrated Biotechnology Research Student Grant (CIBR), USU 2007-2008
Undergraduate Research and Creative Opportunities Grant (URCO), USU 2004-2005

DISTINCTIONS

Department of Chemistry/Biochemistry Travel Award – Utah State University 2012
Department of Chemistry/Biochemistry Travel Award – Utah State University 2011
Graduate Student Senate Travel Award – Utah State University 2011
Thomas F. Emery Research Scholar – Dept. of Chem./Biochem. – USU 2011
Travel Award – American Crystallography Association 2010
Graduate Student Senate Enhancement Award – Utah State University 2010
Marjorie H. Gardner Teaching Award – Dept. of Chem./Biochem. – USU 2009
Department of Chemistry/Biochemistry Travel Award – Utah State University 2009
Graduate Student Senate Travel Award – Utah State University 2009
Claude E. Zobell Scholarship Finalist – Utah State University 2006
Graduate Student Teacher of the Year Nomination – Utah State University 2005
Dr. Joseph E. Greaves Memorial Scholarship – Utah State University 2001
John R. Simmons Scholarship – Utah State University 2001

TEACHING EXPERIENCE

Teaching Assistant – Advanced Biochemistry Laboratory (2008–2012, 5 semesters)
I served as an integral part of the design and development of this upper division
Biochemistry course working with Dr. Sean Johnson. After training students in
bioinformatics, molecular cloning, protein expression and purification techniques,
students design, propose and carry out a protein purification experiment. After mastering
protein purification, students are taught analytical techniques such as Michaelis-Menten
kinetics and x-ray crystallography.

Teaching Assistant – Principles of Chemistry I Laboratory – (Fall 2011 and Fall 2007)
I gave laboratory instruction to one section in 2011 and three sections in 2007
with ~20-25 students each working with Dr. Douglas Harris. I monitored student
experiments and graded submitted assignments.

Teaching Assistant – Principles of Chemistry I – (Fall 2009)
I gave recitation instruction to four sections of ~20-25 students working with Dr.
Hubbard. This course covered the first half of general chemistry, which included; the
characteristics of atoms, trends of the periodic table, molecular geometries, and
bonding.

Teaching Assistant – Principles of Chemistry II – (Fall 2008)
I supplemented the lecture material taught by Dr. Lance Seefeldt by teaching four
recitation sections of ~20-25 students. This course covered the second half of general
chemistry, which included; reaction rates, electro, oxidation-reduction, atmospheric,
quantum, and nuclear chemistry. In my recitation sections I taught new material
and provided a forum for answering questions.

Teaching Assistant – Human Anatomy – (Spring 2006 and 2007)
I gave instruction on Human Anatomy in a cadaver lab and held many regular
and special office hours to prepare students for exams. This challenging course taught
by Dr. David A. Anderson covered the major features of gross anatomy such as the
skeletal, muscular, nervous, circulatory, digestive, and reproductive systems. In 2006 Dr.
Anderson allowed me the privilege to guest lecture for 50 minutes on the facial bones
and 50 minutes on the digestive system to the class of ~200 students.

Teaching Assistant – Human Physiology – (Fall 2005 and 2006)
I found this course taught by Dr. David A. Anderson to be one of the most
challenging and rewarding for entry level students. In an effort to provide support to new
students I gave a 50 minute crash course called “How to Succeed in Human
Physiology”. My responsibilities as a teaching assistant included; giving laboratory
instruction to four sections with ~20 students, designing pre-quizzes and pre-exams,
holding regular and special office hours and exam grading. The material covered was
the basic Human Physiology of all organ systems. I was allowed to guest lecture for Dr. Anderson to the large class of 300 students for 50 minutes on muscle contraction in 2006 and the eye in 2005.

**Other Teaching Opportunities**

As a M.S. student in the Biology Department I was often asked to help teach in various ways. I guest lectured twice in Dr. Joseph Li’s Virology course during Spring 2007 teaching details regarding the Reoviridae virus family. I guest lectured on RNA metabolism in General Biochemistry II in Spring of 2006 for Dr. Joan Hevel and oversaw 120 minute Microbiology Laboratory sessions twice in Spring of 2006.

**ACADEMIC LEADERSHIP AND SERVICE**

*Young Scientist Special Interest Group (YSSIG) Executive Committee Member – American Crystallographic Association (2010-present)*

I serve as an executive committee member of the YSSIG to help those in elected office perform their duties to help advance young scientists in the field of crystallography.

*Young Scientist Special Interest Group (YSSIG) Chair – American Crystallographic Association (2009-2010)*

I worked to provide forums for young scientists to present their work and to exchange ideas and information at the annual American Crystallographic Association (ACA) meeting in Chicago. I invited speakers and chaired the Etter Award Symposium at the annual ACA meeting in Chicago that was designed specifically to provide a venue for the presentation of the work of young scientists. I served on the Etter Award Selection Committee to help select the 2011 Etter Award Winner. I managed the affairs of the YSSIG and held elections for the 2011 leadership.

*Young Scientist Special Interest Group (YSSIG) Chair Elect – American Crystallographic Association (2008-2009)*

I created opportunities for young scientists making the transition from education to the workplace. At the annual ACA meeting in Toronto I chaired a professional directions session where a panel of professional crystallographers gave advice and direction to those on the verge of moving into the workplace.
LIST OF REFEREES

Sean J. Johnson, Ph.D.
R. Gaurth Hansen Assistant Professor
Department of Chemistry and Biochemistry
Utah State University
0300 Old Main Hill
Logan, UT 84322
sean.johnson@usu.edu
(435) 797-2089 (phone)

Joan M. Hevel, Ph.D.
Associate Professor
Department of Chemistry and Biochemistry
Utah State University
0300 Old Main Hill
Logan, UT 84322
joan.hevel@usu.edu
(435) 797-1622 (phone)

Christopher P. Hill, Ph.D.
Distinguished Professor and Co-Chair of Biochemistry
Department of Biochemistry
University of Utah
15 N Medical Drive East, Room 4100
Salt Lake City, UT 84112
chris@biochem.utah.edu
(801) 585-5536 (phone)

Additional References (if needed)

Lance C. Seefeldt, Ph.D.
Professor
Department of Chemistry and Biochemistry
Utah State University
0300 Old Main Hill
Logan, UT 84322
lance.seefeldt@usu.edu
(435) 797-3964 (phone)

Alvan C. Hengge, Ph.D.
Professor and Department Head
Department of Chemistry and Biochemistry
Utah State University
0300 Old Main Hill
Logan, UT 84322
alvan.hengge@usu.edu
(435) 797-3442 (phone)

Eckhard Jankowsky, Ph.D.
Associate Professor
Department of Biochemistry
Case Western Reserve University
10900 Euclid Avenue
Cleveland, OH 44106
ej@case.edu
(216) 368-3336 (phone)